2006 Annual Report

Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

(2)

You are holding the 2006
Barcelona Supercomputing
Center Annual Report, which
contains a summary of our
activities as well as a short
description of our organization.

2006 Annual Report

**Barcelona
Supercomputing
Center**
Centro Nacional de Supercomputación

Table of Contents

The year 2006 was marked by significant growth in both Research and Support for the Barcelona Supercomputing Center– Centro Nacional de Supercomputación.

With this new growth, BSC-CNS research departments were able to build upon their already strong foundations in Computer Sciences, Earth Sciences and Life Sciences.

A Year in Perspective

The year 2006 was marked by significant growth in both Research and Support for the Barcelona Supercomputing Center – Centro Nacional de Supercomputación (BSC-CNS). We are particularly happy to report a large increase in our most important resource, our staff, which more than doubled from 63 members at the end of 2005 to 137 members at the end of 2006 and hailed from 18 different countries (Argentina, Belgium, Bulgaria, China, Colombia, Cuba, France, Germany, India, Ireland, Italy, Mejico, Peru, Serbia, Turkey, the United Kingdom, the United States and Spain).

With this new growth, BSC-CNS research departments were able to build upon their already strong foundations in Computer Sciences, Earth Sciences and Life Sciences. The Computer Sciences Department continued to research novel architectures and programming models as well as adapt current hardware and software technologies for supercomputing infrastructures. The Earth Sciences Department further advanced their study of the global and local impact of changes in climate and air quality through their internally developed modeling system. The Life Sciences Department moved forward in their study of complex systems as well as key interactions at the sub-atomic level in hopes of gaining a deep insight into the origin, evolution and function of living organisms. Each of these departments made significant scientific and technological progress which resulted in 103 publications and 26 communications. Moreover, our innovative scientific and technological solutions were applied to real-world problems.

The theme of expansion and growth extended to the MareNostrum supercomputer which is managed by BSC-CNS. In 2006, MareNostrum was upgraded from 4812 processors to 10240 processors which enables MareNostrum to perform more than 94 billion (10^{12}) operations per second. One hundred and twenty six external research teams from the areas of Biomedicine, Health Sciences, Astronomy, Space, Earth Sciences, Physics, Engineering, Chemistry, Material Science and Information Technology performed research using the supercomputer in 2006. MareNostrum continues to be the number one supercomputer in Europe and is ranked fifth worldwide, and as stated in the September 27, 2006 issue of *Fortune*

Magazine, "is definitely the most beautiful [supercomputer in the world]."

BSC-CNS played a key role in the Spain's increase in supercomputing capability in 2006 as the technical coordinator and main node of the Spanish Supercomputing Network (Red Española de Supercomputación or RES). This distributed network of supercomputers was created by the Spanish Ministry of Education and Science in response to the growing demand for supercomputing resources for research purposes. The initial network nodes include BSC-CNS, the Centro de Supercomputación y Visualización de Madrid (CesViMa), the Instituto de Astrofísica de Canarias (IAC) as well as the Spanish Universities of Cantabria, Málaga, Valencia and Zaragoza; all 7 nodes are included in the 28th Top500 list dated November 2006 (www.top500.org).

In 2006, BSC-CNS signed agreements with public institutions and private companies of which the following stand out: INM (Instituto Nacional de Meteorología), CSC (Finnish Supercomputing Center) IBM, Microsoft and Repsol.

The BSC-CNS vision also looks beyond the borders of Spain to Europe. In 2006, BSC-CNS began participation in eight European Commission Sixth Research Framework Programme projects. BSC-CNS is also the Spanish Representative to the HPC European Taskforce Initiative whose main objective is to create a limited number of top-tier world centers as a strategic tool to increase the capabilities of the European Research community.

We believe that our successes in 2006 show that we are on track to fulfilling our mission to contribute to the advancement of science and technology. We are proud of this progress and hope for continued success as we pursue this mission in the future.

Mateo Valero, Director   Francesc Subirada, Associate Director

1

Organizational Structure

Description of the main governing bodies and departments of BSC-CNS.

The figure below describes the different governing bodies and departments that constitute the Barcelona Supercomputing Center – Centro Nacional de Supercomputación (BSC-CNS). The Board of Trustees serves as the main governing body of the BSC-CNS Consortium and consists of representatives from each of the three shareholder institutions in the Consortium: the Ministerio de Educación y Ciencia (MEC), the Departament d'Educació i Universitats de la Generalitat de Catalunya (DEiU) and the Universitat Politècnica de Catalunya (UPC).

Board of Trustees

Chairman of the Board	Francisco Marcellán Español *Secretary-General of Policies on Science and Technology, MEC*	Members (DEiU)	Josep Anton Ferré *Director-General of Universities* Xavier Testar *Director-General of Research* Josep Mª Vilalta *Deputy Dir.-General of Research*
Vice Chairman of the Board	Joaquim Prats *Secretary of Universities and Research, DEiU*	Members (UPC)	Antoni Giró *Rector* Josep Casanovas *Vice-Rector of University Policies* Sebastià Sallent *Director of Fundació I2Cat*
Members (MEC)	Carlos Alejaldre *Director-General of Policies on Technology* José Luís Martínez *Deputy Director-General of Research Projects* Ernest Quingles *Deputy Director-General of Promotion and Technology Infrastructures and Large Installations*	BSC-CNS Director	Mateo Valero *Director, Barcelona Supercomputing Center*
		Secretary of the Board	Francesc Subirada *Associate Director, Barcelona Supercomputing Center*

The Executive Comission is responsible for executing on the plans of the BSC-CNS and monitoring center progress.

Executive Comission

Commission Chairman	Ernest Quingles *Deputy Director-General of Promotion and Technology Infrastructures and Large Installations*
Comission Vice-Chairman	Xavier Testar *Director-General of Research*
Member (MEC)	Ramon López de Arenosa *Department Head, Department Production Technologies and Communications*
Member (DEiU)	Josep Mª Vilalta *Deputy Director-General of Research*
Member (UPC)	Xavier Gil *Vice-Rector of Research*
Member (UPC)	Josep Casanovas *Vice-Rector of University Policies*
BSC-CNS Director	Mateo Valero *Director, Barcelona Supercomputing Center*
Comission Secretary	Francesc Subirada *Associate Director, Barcelona Supercomputing Center*

The Scientific Advisory Board and the Business Board will offer advice on scientific and business matters beginning in 2007. The Scientific Advisory Board will consist of a select group of prominent international scientists while the Business Board will consist of representatives from companies that currently collaborate with and fund the center.

The Access Committee is responsible for all decisions concerning the scientific use of MareNostrum. The members of the Access Committee are selected by the MEC and the Agencia Nacional de Evaluación y Prospectiva (ANEP). The Access Committee is composed of a Core Team and four Expert Panels of prestigious Spanish scientists (external to BSC-CNS).

The four Expert Panels are based on the four-group classification system employed by the Spanish Foundation of Science and Technology (FECYT – Fundación Española de Ciencia y Tecnología): Astronomy, Space and Earth Sciences,

Biomedicine and Health Sciences, Physics and Engineering, Chemistry and Materials Science and Technology. The Expert Panels are composed of high caliber scientists that also have experience in the management of research projects. Each of the four Expert Panels consists of 10 experts and is chaired by a coordinator with the help of an assistant. The Experts Panel can request an ANEP peer review of a project when required.

The Expert Panels prioritize the activities of each of their respective areas and send their recommendations to the Core Team who publishes the list of approved users of MareNostrum on the BSC-CNS website. In 2006, the members of the Access Committee included:

Core Team

Victoria Ley Vega de Seoane	*(Agencia Nacional de Evaluación Prospectiva)*
Pedro de Miguel Anasagasti	*(Universidad Politécnica de Madrid)*
Ramón López de Arenosa	*(Ministerio de Educación y Ciencia)*
Jesús Labarta	*(Barcelona Supercomputing Center-Centro Nacional de Supercomputación)*

Biomedicine and Health Sciences Expert Panel

Coordinator	Alfonso Valencia *(Centro Nacional de Investigaciones Oncológicas)*
Assistant	Manuel Palacín *(Universidad de Barcelona)*

Chemistry and Material Sciences Expert Panel

Coordinator	Agustí Lledós *(Universidad Autónoma de Barcelona)*
Assistant	José María Pitarke *(Universidad del País Vasco)*

Physics and Engineering Expert Panel

Coordinator	Pablo Ordejón *(Instituto de Ciencia de Materiales de Barcelona)*
Assistant	Manuel Laso *(Universidad Politécnica de Madrid)*

Astronomy, Space and Earth Sciences Expert Panel

Coordinator	José María Ibáñez *(Universidad de Valencia)*
Assistant	Vicente Caselles *(Universidad de Valencia)*

2006 BSC-CNS Staff

Director	Mateo Valero
Associate Director	Francesc Subirada

Computer Sciences Director	Jesús Labarta
Computer Sciences Assoc. Director	Eduard Ayguadé
Autonomic Systems and e-Business Platforms Head	Jordi Torres
Researcher	Vicenç Beltran
Researcher	David Carrera
Researcher	Jordi Guitart
Researcher	Kevin Hogan
Researcher	Mario Macias
Resident Student	Ferran Julià
Resident Student	Ramon Nou
Computational Mechanics and Optimization Head	José María Cela
Senior Researcher	Rogeli Grima
Senior Researcher	Guillaume Houzeaux
Senior Researcher	Mariano Vázquez
Researcher	Hadrien P. Calmet
Researcher	Miquel Catalán
Researcher	Raúl de la Cruz
Researcher	Manuel Quero
Researcher	Félix Rubio
Researcher	Xavier Saez
Resident Student	Ane Beatriz Eguzkitza
Computer Architecture Head	Alex Ramírez
Senior Researcher	Francisco Cazorla
Senior Researcher	Adrian Cristal
Senior Researcher	Roberto Gioiosa
Senior Researcher	Osman Unsal
Researcher	Miquel Pericas
Researcher	Francisco Javier Vera
Technical Support	Josep Oriol Prat
Resident Student	Paul M. Carpenter
Resident Student	Milos Milovanovic
Resident Student	Cristian Perfumo
Resident Student	Petar Radojkovic
Resident Student	Suthirta Sanyal
Resident Student	Nehir Sonmez
Resident Student	Srdjan Stipic

Resident Student	Sasa Tomic
Resident Student	Ferad Zyulkyarov
Grid Computing and Clusters Head	Rosa Badia
Researcher	Julita Corbalán
Researcher	Marc de Palol
Researcher	Jorge Ejarque
Researcher	Francesc Guim
Researcher	José María Pérez
Researcher	Ivan Rodero
Researcher	Raúl Sirvent
Resident Student	Pieter Bellens
Performance Tools Head	Judit Giménez
Researcher	Xavier Aguilar
Researcher	Marc Casas
Researcher	Juan González
Researcher	Pedro Antonio González
Researcher	German Llort
Researcher	German Rodríguez
Researcher	Eloy Martínez
Researcher	Xavier Pegenaute
Researcher	Harald Servat
Programming Models Head	Xavier Martorell
Researcher	Jairo Balart
Researcher	Roger Ferrer
Researcher	David Ródenas
Researcher	Javier Teruel
Researcher	Jorge Vaquero
Resident Student	Javier Bueno
Resident Student	Víctor Jimenez
Resident Student	Lluis Vilanova
Storage Systems Head	Toni Cortés
Researcher	Jesús Malo
Researcher	Jonathan Martí
Resident Student	Marta García

Earth Sciences Director	José María Baldasano
Senior Researcher	Pedro Jiménez
Senior Researcher	Oriol Jorba
Senior Researcher	Carlos Pérez
Technical Support	Pedro Gracia
Technical Support	Eugenio López
Resident Student	Eric Benoiston
Resident Student	Juliana González
Resident Student	Patricia Güereca
Resident Student	Mª Teresa Pay
Resident Student	Angel Rincón
Visiting Student	Maria Gonçalves

Life Sciences Director	Modesto Orozco
Computational Genomics Head	David Torrents
Researcher	Carlos Quijano
Researcher	Jesús López
Technical Support	David García
Resident Student	Mª Dolores Fernández
Electronic and Atomic Protein Modelling Head	Víctor Guallar
Researcher	Raúl Alcántara
Researcher	Víctor Gil
Visiting Student	Yolanda Small
INB – Computational Node 2 Head	Josep Lluis Gelpí
Technical Support	Jordi Camps
Technical Support	Aina Martínez
Technical Support	Romina Royo
Resident Student	Alexis Torrano
Molecular Modeling and Bioinformatics	
Visiting Professor	Jiali Gao
Resident Student	Wei Fu
Visiting Student	Josep Ramon Goñi
Visiting Student	Alberto Pérez
Protein Interactions and Docking Head	Juan Fernández-Recio
Researcher	Solène Grosdidier

Researcher	Albert Solernou
Technical Support	Carles Pons
Resident Student	Mathias Wody

Business Development Head	Felipe Lozano
Project Manager	Toni Moreno
Marketing Executive	Renata Giménez
MareNostrum Visitors Manager	Oriol Riu
Mobility Program Manager	Fermín Sánchez
Mobility Program Assistant	Anna Monrós

Human Resources, Finance and Administration Head	Mercè Calvet
Director's Assistant	Neus Jiménez
Associate Director's Assistant	Nuria Sirvent
Administration Assistant	Laura Gutiérrez
Account Officer	Cristina Calonge
Human Resources Manager	Sandra Vargas
Recruiting Assistant	Ricardo de la Cruz

Operations Director	Sergi Girona
System Administration Head	Javier Bartolomé
Helpdesk	Fernando Fernández
Helpdesk	Jordi Inglés
Helpdesk	Carlos Vicente
Security and Networks	Carlos Kishimoto
Security and Networks	Juan Carlos Sánchez
System Administration	Xavier Fustero
System Administration	Sergi Moré
Performance and Planning	Ernest Artiaga
User Support	
User Support	Pere Munt
User Support	Jorge Naranjo
User Support	David Vicente
User SupportAssistant	Asier Roa
Webmaster	Augusto Izquierdo
Maintenance Assistant	Albert Riera

2

Research Activities

The internal scientific departments
and their results during the year 2006.

2.1 Computer Sciences

The Computer Sciences Department focuses on building upon currently available hardware and software technologies.

Jesús Labarta
Computer Sciences
Director

Eduard Ayguadé
Computer Sciences
Associate Director

The Department is involved in the development of the new technologies and adapting them to make efficient use of supercomputing infrastructures. The department researches novel architectures for processors and memory hierarchy and proposes programming models and innovative implementation approaches for these models as well as develops tools for performance analysis and prediction. These research activities are carried out as a part of internal BSC-CNS projects, projects in the Sixth Research Framework Programme of the European Union (FP6) as well as projects with leading-edge companies in the field. The activities carried out in 2006 are presented individually for each of the research groups in the Computer Sciences Department: Autonomic Systems, eBusiness Platforms (eBP), Computational Mechanics and Optimization. (CM), Computer Architecture (CA), Grid Computing and Clusters (GC). Performance Tools (PT), Programming Models (PM) and Storage Systems (SS).

Autonomic Systems and eBusiness Platforms (eBP)

In 2006, the Autonomic Systems and eBusiness Platforms or eBP Group devoted a great deal of time and effort to proposing application servers that adapt to their respective execution environments in addition to researching application characteristics and workload intensity. The group also started three European projects under the IST FP6 (SORMA, BREIN and XtreemOS) as well as a research collaboration with IBM.
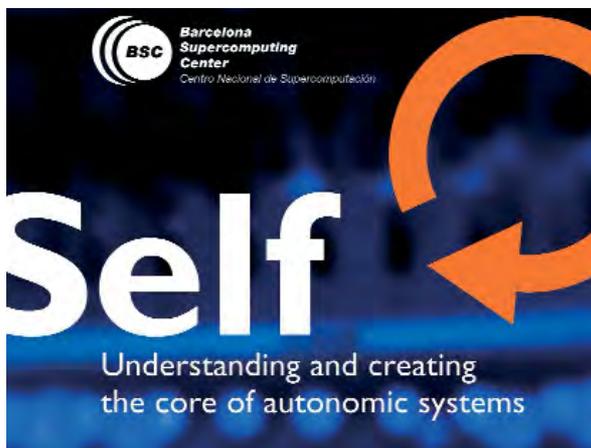
The eBP Group proposed and implemented a new HTTP connector (BSC Hybrid Connector) and tested it with well-known business benchmarks which showed improvement in the performance as well as reducing the tuning complexity. The group pursued this objective as tuning the web container of application servers has been shown to reduce the number of wrongly configured servers as well as introduce performance improvements.

In addition, the group worked on low-level resource management, focusing on intelligent allocation to avoid bottlenecks. The group implemented a prototype of middleware with efficient and autonomic management of heterogeneous workloads (transactional, long running jobs, etc.). Part of the work in this area falls within the scope of the project SORMA, in which the group is building an adaptive middleware based in economic algorithms. The remaining work relates to another project BREIN, in which the group is implementing a resource manager that efficiently allows the sharing of the resources based on semantic rules. The activities in this direction are also supported by the Adaptive Systems SOW with IBM T.J. Watson Labs; in this project, the group made significant contributions in the management of heterogeneous workloads including cluster management and application placement, taking advantage of virtualization techniques.

The group also continued to improve upon the instrumentation and analysis of the BSC-CNS platform with the goal of helping experts to extract knowledge and patterns from complex systems.

Moreover, the group started work in new areas in 2006 such as: using virtualization technology for low-level resource management, proposing logic engines for autonomous systems and considering alternatives such as online simulation or machine learning, and dynamic and adaptive software.

In addition to the above projects and collaborations, the group also participated in the XtreemOS project. The eBP Group contributed to this project in the area of the eBusiness applications. In addition to this project, the group performed research in collaboration with Adaptive Systems which contributed to the LA Grid program, an international multi-disciplinary research community to facilitate collaborative IT research, education and workforce development. Finally, the group joined CoreGrid, a Network of Excellence related to grid technology in the area of self-healing.

Computational Mechanics and Optimization (CM)

The Computational Mechanics and Optimization or CM Group focused on various basic research activities in 2006.

The CM Group focused on making improvements to ALYA, BSC-CNS's parallel computational mechanics platform, and bringing this platform to a mature state. This multiphysics code is capable of efficiently solving coupled problems in parallel taking advantage of the MareNostrum supercomputer architecture. By the end of 2006, ALYA simulated problems such as:

- Incompressible laminar flows (NASTIN module)
- Compressible (transonic and hypersonic) laminar flows (NASTAL module)
- Turbulent flows, using RANS and LES models (TURBUL module)
- Thermal flows, with forced and natural convection (TEMPER module)
- Excitable media electric potential propagation, like cardiac muscular tissue (EXMEDI module)
- Convection-diffusion-reaction transport equations (CODIRE module)
- Basic solid mechanics problems (SOLIDZ module)
- Domain decomposition techniques (DODEME service)

ALYA was used to solve application problems related to flow around high-speed trains, thermal flows in buildings (including the MareNostrum site), cardiac electrophysiology and rotating flows around wind generators. Moreover, ALYA achieved nearly ideal parallelization for explicit schemes (PARALL service coupled with NASTAL and EXMEDI modules).

Runs were performed using up to 2500 MareNostrum processors with 95% scalability.
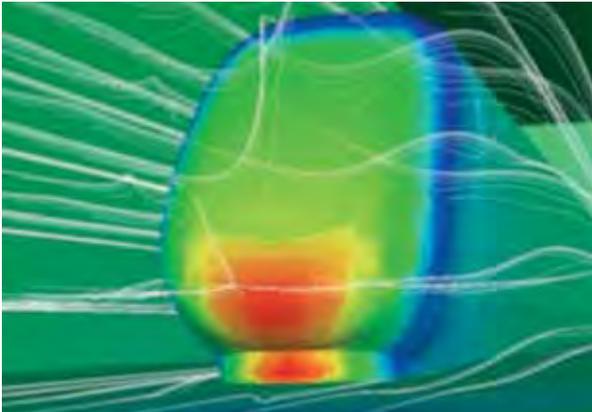
Aeroacustic study for a highspeed train. Hybrid net of 50 millions of elements. Execution with 2,500 processors.

In addition to making significant progress on the ALYA platform, the CM Group collaborated with the "Departament

CENIT project financed by the Spanish Ministry of Industry. This joint project includes research groups such as "Hospital General Universitario Gregorio Marañon" as well as industrial partners such as Phillips and GridSystems. Its main objective is to develop new technological simulation tools for different surgical techniques. In the Flow Simulator project financed by Airbus, the CM Group was responsible for improving the parallel performance of ELSA, a computational fluid dynamics code developed by ONERA (France) that is currently in production at Airbus. Finally, the CM Group participated in the ESSL VMX project funded by IBM. For this project, the group began developing the VMX version of the ESSL scientific library.

Computer Architecture (CA)

In 2006, the Computer Architecture or CA Group addressed such research topics as multithreaded processors, novel processor architectures, transactional memory and accelerators for bioinformatic applications. In addition, the CA Group participated in the FP6 Integrative Project SARC and the Network of Excellence (NoE) HiPEAC and began a collaboration with Microsoft. The research topics and their relationship to CA Group projects are described in detail below.
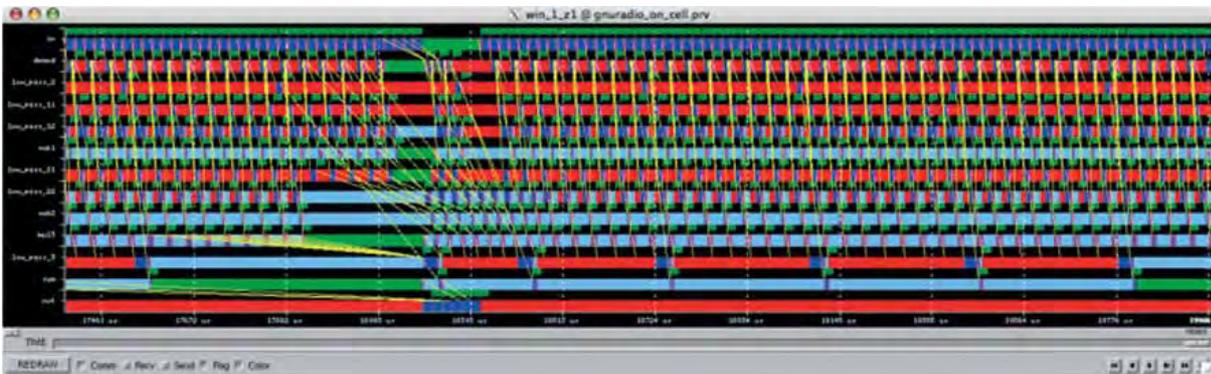
Performance prediction of a Streaming version of GNU FM radio on the Cell Processor.

de Fisica i Enginyeria Nuclear" of the UPC to develop a parallel numerical simulation for the analysis of Coherent Stark nonlinear spectroscopy with chirped pulses optimized to run on MareNostrum.

The CM Group also strengthened links with industry through its technology driven research lines in 2006 and expects to build upon these relationships throughout 2007. In the second year of the OPTIDIS project (funded by the Spanish Government through the National Plan), the CM Group began to apply thermal flows and optimization to problems of energy efficiency in buildings. The objective is to develop computational fluid dynamics tools in the domain of building ventilation. The CM Group participated in the CDTeam, a

The CA Group proposed several middleware (operating system scheduling algorithms) and hardware (resource allocation policies) solutions that attempt to reduce the execution time variability that may exist when running applications in Simultaneous Multithreading (SMT) processors. For example, the CA Group proposed a hardware mechanism that allows the Operating System (OS) to control the resource allocation in an SMT processor coupled with the OS scheduler for real-time systems. The group also developed FAME (fairly measuring multithread architectures), a novel evaluation methodology aimed at accurately measuring the performance of multithreaded processors. FAME re-executes all threads in a multithreaded workload until all of them are accurately represented in the final

measurements taken from the workload. FAME provides much more precise measurements than the previously available solutions which makes it an ideal evaluation methodology for the analysis of novel design ideas implemented in multithreaded architectures.

In the area of novel processor architectures, the CA Group continued research and development on the Kilo-instruction Processor (KiP) architecture. The Decoupled KiP implementation uses multiple cores in a Chip-level Multi-processing (CMP) architecture to allow out-of-order execution of over a thousand in-flight instructions, effectively tolerating memory latencies and significantly accelerating memory-bounded applications. In the SARC project, the CA Group contributed a novel tile-based architecture, where each node is a heterogeneous chip multiprocessor composed of a high-level control processor that manages the application, and a set of specialized accelerators perform the processing. In addition, the group contributed to the definition of the SARC vector accelerator which is targeted at scientific applications and includes novel features like a polymorphic register file and 2-dimensional (matrix) register operations. Finally, the group developed the necessary framework to investigate performance prediction, resource partitioning and design space exploration of heterogeneous CMP architectures such as the nodes proposed in the SARC project.

The CA Group also developed the required infrastructure to perform research on the benefits of transactional memories as well as explore potential implementations. Transactional memory has been identified as a potentially key element of next generation programming models to facilitate the development of parallel applications. On the applications side, the group developed a small benchmark suite written in an imperative language (C/C++) and a second suite of small application kernels written in a functional language (Haskell). As for support tools, the group developed a Software Transactional Memory (STM) implementation, including a source-to-source C translator and a runtime library (libBSC) that provides an easier programming model interface than those previously proposed. In addition, the group began optimizing the Haskell STM library to avoid false conflicts between transactions. Finally, the group started developing a Hardware Transactional Memory simulator based on the M5 codebase that implements both the Transactional Coherence (TCC) model from Stanford and the Log-TM from Wisconsin. The CA Group performed research on transactional memory in collaboration with Microsoft Research.

Finally, the group investigated the possibility of using hardware accelerators for bioinformatic applications. The CA Group evaluated different applications in the bioinformatics field (e.g. pattern matching and protein docking) on processors with SIMD instruction sets, including the Cell Broadband Engine (BE) from IBM / Sony / Toshiba. In particular, the group evaluated the impact of the Pentium 4 SSE2 on FTDOCK using the FFTW library and the impact of PowerPC VMX on the Smith Waterman algorithm. The group also collaborated with the Life Sciences Department to develop its own protein docking application from scratch exploiting SIMD ISAs. In addition, the group developed optimized versions of several sequence alignment applications, using 128-bit SIMD registers to sort patterns, exploiting higher data parallelism. The group then optimized this code for shared memory multiprocessors and for the Cell architecture.

Grid Computing and Clusters (GC)

In 2006, the Grid Computing and Clusters or GC Group focused on two main areas: GRID superscalar and eNANOS. In addition, the group participated in the following projects funded by the European Commission: HPC Europa, CoreGRID, BEinGRID, Brein and XtreemOS.

The CG Group made various improvements to the GRID superscalar. GRID superscalar is a programming and execution environment for computational grids. It is able to parallelize a sequential application at the runtime and the task level. Thus, it is possible to take advantage of applications that are composed of coarse grained tasks (e.g. a simulation, a program, a solver) which are common in bioinformatics, computational chemistry and other scientific fields. In 2006, the group added fault tolerance and componentized the runtime library which allows for local tasks and dependencies between any type of task parameter. The team released a version of the GRID superscalar for the MareNostrum supercomputer that included new scheduling algorithms and the use of a scratch disk. In addition, the group began development of a prototype scheduling module based on ontologies (semantic scheduler).

In 2006, the CG Group also made significant progress in various areas of the eNANOS project. The eNANOS project is based on the idea of having strong low-level support for performing high level scheduling. The eNANOS project implementation includes: fine grain control between scheduling levels, dynamic allocation (MPI+OpenMP jobs) to improve system performance, tracking detailed scheduling and performance information that can be used for future scheduling decisions resulting in efficient scheduling (in terms of slowdown) based on performance prediction. For eNANOS, the group extended the resource broker based in GT3 to support JSDL 1.0. The team extended the eNANOS scheduler and low-level execution framework to manage multiple nodes and implementing scheduling policies based on co-allocation and load-balancing. The group also completed the design and implementation of a progress and performance indicators system as well as the integration of the Palantir meta-information system model with other systems such as GRMS resource management system. At the same time, the group began to explore data-mining techniques for retrieving
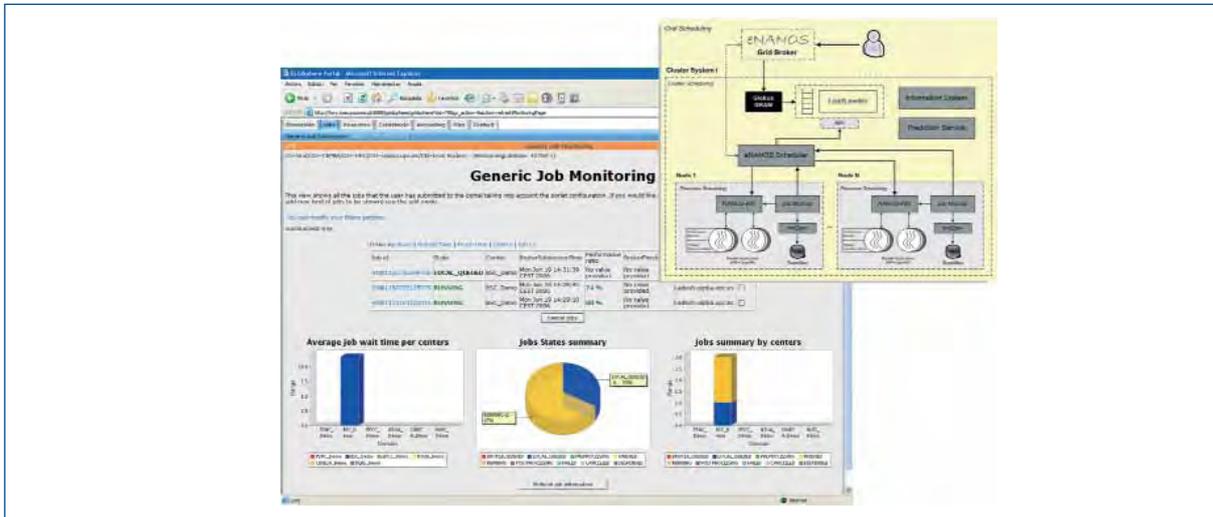
Architecture of eNANOS and its application to the HPC-Europa portal.

information from grid workload logs to be applied in performance forecasting models as a means for improving scheduling. The team investigated the impact of resource load in the performance of the workloads and scheduling decisions with respect to scheduling policies. Furthermore, the team began the implementation of the ISIS simulator framework that will allow to test all these policies in realistic models that includes the modelization of the resource model where the workloads are executed (for example modeling bandwith memory, I/O access, etc.).

The GC Group also participated in several other projects, one of which was the CoreGRID Network of Excellence (NoE). As the research activities of this NoE on GRID are organized into Virtual Institutes, the GC Group was involved in the Institute on Resource Management and Scheduling and Institute on Grid Systems, Tools and Environments. The group worked on the integration of its research with the research of other European groups.

In the HPC-Europa project, the GC Group was involved in the development of a grid portal, with the objective of implementing a Single Point of Access to different grid middleware. To this end, the development of the HPC-Europa Grid Portal, based on the GridSphere portal framework including its GridPortlets mechanism was started. The portal will provide transparent, uniform, flexible and intuitive user access to HPC-Europa resources. This portal will hide the underlying complexity and heterogeneity of these resources and the access to them. In 2006, the group focused on the implementation of the portal and modifications in the middleware to allow the integration of the eNANOS resource management environment.

In the XtreemeOS Project, the group contributed to several Research and Development Workpackages (WP) and also served as the leader of the Application execution management workpackage. The XtreemOS project aims at investigating and proposing new services that should be added to current operating systems to build grid infrastructures in a simple way. XtreemOS targets the widely accepted Linux open source Operating System and extends it to grid with native support for virtual organizations. One of the most important challenges in XtreemOS is the identification of the basic functionalities which are to be embedded in the Linux kernel.

The GC Group began its participation in the BEinGRID project in 2006. This project aims to bring GRID technologies to the industry by means of developing GRID applications in relevant industry sectors. BSC-CNS began leading a Business Experiment (BE14) in this project, in which the GRID superscalar will be integrated with the GridWay metascheduler. An application from the design of new processes and products in the chemistry sector will be implemented on top of these two tools.

In addition to BEinGRID, the group began to participate in the BREIN project in 2006. The objective of this project is to define a new grid architecture with a business-centric approach, that takes into account several business aspects and intra- and inter-organizational interactions. To achieve these goals, artificial intelligence or AI systems, multi-agents and technologies from the semantic web will be applied. In 2006, BSC-CNS was largely devoted to releasing an internal prototype based on autonomic resource management, agents and semantic scheduling.

Finally, the group contributed to LA Grid, an IBM initiative to foster collaborative research and to offer equal opportunities to Latin American students. One of the contributions of BSC-CNS to this project was the specification of a meta-brokering system that permits the support of grid applications with resources located and managed in different domains spanned over a grid computing cyber infrastructure. The meta-broker design in LA Grid is following a distributed model. Multiple meta-broker instances cooperate with each other to achieve grid functionalities using a common API that has been defined. The group also contributed to the Grid Enablement of Hurricane Simulation and Visualization Applications project, where GRID superscalar has been integrated with Transparent Shaping from FIU to offer a grid enabling and autonomic programming environment.

Performance Tools (PT)

In 2006, the Performance Tools or PT Group continued to develop tools for instrumentation, performance analysis and prediction with special emphasis on increasing the scalability and flexibility of work on environments with thousands of processes.

The PT Group focused on producing a light version of the tracing library for instrumenting large runs (many hours) and working on the automatic tracing start and shutdown. They released a new version of the tracing library that only collects information about the areas of computation that are larger than a given size. The new version includes mechanisms for automatic start and shutdown using the collective operations as synchronization points. Throughout the year, the group validated the scalability of the tools with the analysis of traces up to 10,000 processes and instrumenting runs of more than 7 hours. Currently, it is quite common to work with traces of as much as a gigabyte (or even up to tens of GB) that are translated into partial traces of the several hundred megabytes that are loadable with Paraver.

In addition, the PT Group implemented other extensions and functionalities such as the instrumentation of the OpenMP runtime library on MareNostrum, the porting of OMPtrace to the SGI Altix platform using dyninst and the definition of a new trace format for Dimemas. The group worked on using clustering algorithms to identify the structure of traces. This approach currently works with the Dimemas traces and can cluster the Dimemas CPU bursts depending on different duration metrics or simple metrics based on hardware counter information such as cycles or derived metrics like IPC. The results of the clustering can be used to perform projections on partial traces, to identify areas with some outlier behavior, to use as input to load balancing techniques. The other research line is focused on the use of wavelets to identify patterns on the trace that
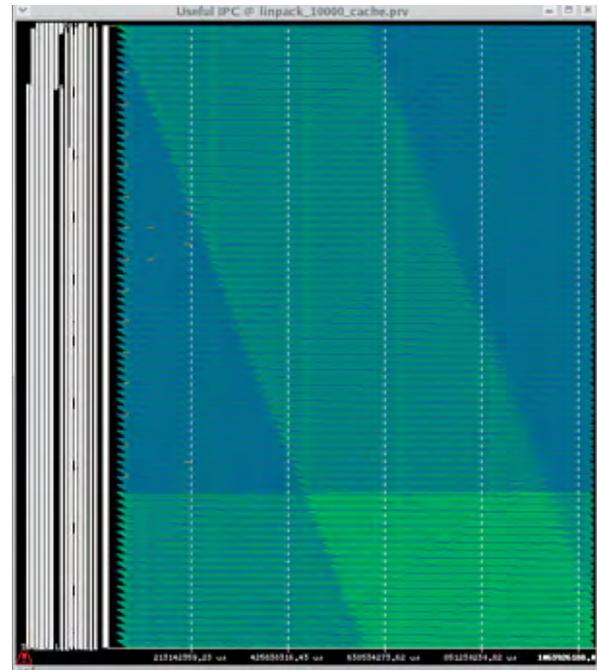
Instructions per cycle (IPC) for a 1700 seconds of a 10.000 processors run of linpack. Gradient from 2.85 (light green) to 2.95 (dark blue).

can be used to automatically drive the analysis. Finally, the group improved upon the basic utilities with new features that allow users to have more flexibility when processing the trace; they also integrated the developed utilities into the visualizer.

In 2006, the PT Group also participated in several national and international projects and collaborations. In the HPC-Europa Project, the PT Group applied the defined methodology to Transnational Access guests running large scale applications as well as main users or codes that run on the different facilities. With respect to the development, the group consolidated the work done during the first half of the project, integrating the utilities to process Paraver traces within the visualizer and extending some functionality. In collaboration with NASA-Ames, the group ported the OpenMP instrumentation to the SGI Altix architecture using the dyninst library as an instrumentation mechanism. The collaboration also included support of the analysis of various applications of interest. Finally, the PT Group worked with IBM to develop a translator from PEM format to Paraver. PEM traces are binary files whose records are defined on one or more XML file. The translator works for any type of PEM trace, although the initial tests were completed only with the MPI instrumentation.

Programming Models (PM)

In 2006, the Programming Models or PM Group focused on improving compiler and runtime infrastructures for the standard OpenMP programming model, development and evaluation for programming models for novel multicore architectures (including the Cell BE) and developing an internal software Distributed Shared Memory (DSM) infrastructure for clusters. In addition, the group participated in projects funded by the European Commission such as SARC as well as research collaborations with industry partners IBM and Microsoft.
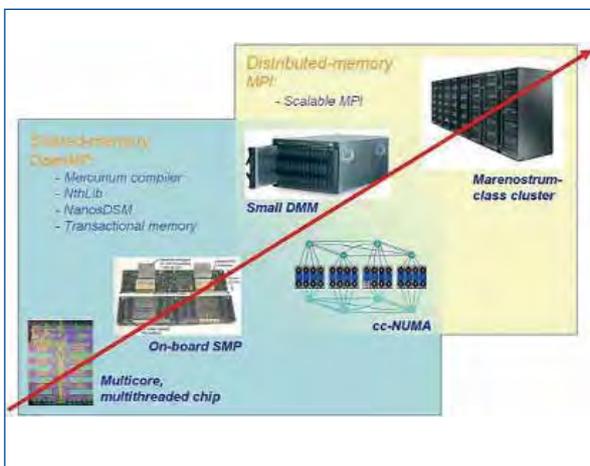
Programming models cover a wide range of architectures.

The group worked on the several scientific problems with respect to the compilation and execution environment for OpenMP. The PM Group improved the Mercurium compiler infrastructure to provide better support for the restructuring of Fortran and C/C++ OpenMP applications. The group worked on a new specification format for the compiler transformations based on new compiler passes written in C++. The code generated by the compiler targets the internally developed OpenMP runtime library (NthLib) which was reengineered during the year. The PM Group also worked on the new tasking execution model to be included in OpenMP 3.0. This work was performed in collaboration with Intel and IBM as well as other research institutions.

In addition to working on OpenMP, the PM Group also continued development on CellSs in 2006. The group began development on CellSs in late 2005, starting from the componentized version of GRID superscalar. The CellSs is a programming model for multicore processors, notably for the Cell BE processor, that is derived from the GRID superscalar programming model for GRID. In 2006, the group proposed and implemented a set of pragma extensions to the C language in the Mercurium compiler. At the same time, the

group implemented the CellSs runtime library, an engine that builds the data dependency graph of the application at runtime based on the annotations provided by the programmer. From this graph, the runtime library is able to schedule the execution of independent nodes to different Synergistic Processing Elements or SPEs in parallel, and all data transfers required for the computations in the SPEs are automatically performed by the runtime library. The group also made the prototype CellSs programming model available for evaluation at different research institutions and performed an extensive benchmark evaluation of the proposal in terms of expression and performance.

The PM Group also began to rethink the industry standard programming models for the emerging multicore architectures, both homogeneous and heterogeneous. The PM Group developed a simple runtime library supporting parallelism on top of the Cell System Processing Units or SPUs in the context of the SARC project. The Mercurium compiler transforms the code in current OpenMP parallel and worksharing constructs to spawn work onto the SPUs. This library also offers primitives to support data transfers between the main memory and the SPUs local stores. In collaboration with IBM Watson, the PM Group proposed extensions to the software cache approach that transparently handles remote memory access from the SPUs. The extensions are based on a novel inspector / executor model that analyzes and forecasts data movement between the main memory and the local memory. The inspector analyzes at runtime the behavior of the applications with respect to memory access. After the inspection, all data access information needed in the code region is collected, and data is moved accordingly.

Finally, the PM Group continued efforts on the NanosDSM runtime system to support the execution of OpenMP shared-memory applications on top of clusters with distributed memory (including MareNostrum). The share-everything approach of the proposal is complemented with added intelligence in the runtime system to learn and predict actions in the memory coherence protocol (presend, preinvalidate and prefetch). The PM Group started a collaboration with IBM Toronto to adapt the IBM XLSMP runtime system to work with the NanosDSM runtime system on clusters of SMPs.

Storage Systems (SS)

The Storage Systems or SS Group focused on two main topics in 2006: file system scalability and I/O for grid systems. In addition, the group participated in the aforementioned project XtreemOS funded by the European Commission. The group carried out an intensive study of the performance of GPFS and Luster in MareNostrum under various kinds of loads. The objective was to detect the strengths and

weaknesses of these two file systems with respect to scalability.

Moreover, the group set-up an infrastructure in order to trace how storage is used by the applications running on MareNostrum. The data that has been gathered will allow for a comprehensive study of the behavior. This work will help those implementing and/or configuring storage systems for large clusters have a better idea of what to expect from real users.

The group also extended the I/O functionality of GRIDsuperscalar by adding a file replica mechanism. On the one hand, the mechanism avoids copying back files when the application is finished while at the same time allowing users to work with replicas previous versions in order to reduce transfer time. In addition, a "pipelined file" mechanism has been implemented, allowing a task to produce a file and other tasks to start consuming it before it has finished writing. The GRIDsuperscalar determines which mechanism to run in a manner that is transparent to the user.

Finally, as part of the XtreemOS project, the group defined the XtreemFS services (a file system for the GRID) should offer to applications in addition to defining the architecture of the system. In particular, the group led the replica-management topics: how and when replicas are created or deleted, monitoring replica usage, etc.

Projects being developed by the Storage System research Group.

The Earth Sciences Department focuses on high-resolution meteorological modeling,

José María Baldasano
Earth Sciences Director

air quality modeling (photochemical and aerosol pollution), climate modeling (feedback with air quality and radiative forcing), and global and regional modeling of desert dust. The department utilizes the computing performance of MareNostrum to study these topics through their internally developed modeling system consisting of models for Meteorology (WRF), Chemistry Transport (CMAQ), Emissions (HERMES) and Mineral Dust (DREAM).

In 2006, this department established a number of national and international collaborations with a diverse range of organizations that include the World Meteorological Organisation, the Goddard Institute for Space Studies-NASA and the National Center for Atmospheric Research in the United States, the University of Athens in Greece, the University of Aveiro in Portugal, the University of Basel in Switzerland, the European Network for Earth System Modeling (ENES), the Observatorio Meteorológico de Izaña (INM), the Centro de Investigación Energéticas, Medioambientales y Tecnológicas (CIEMAT), the Institut de Ciències de la Terra Jaume Almera of the Consejo Superior de Investigaciones Científicas (IJA-CSIC) and the Fundación Centro de Estudios Ambientales del Mediterráneo (CEAM).

The Earth Sciences Department also worked closely with the Support and Computer Sciences Departments in studying the improvements of the parallel performance of their code and models. Also, the department was directly involved in a variety of national sponsored projects with companies such as Gas Natural or Hidrocantábrico.

Air Quality (AQ)

It is a challenge to forecast air pollution events in Europe and the Iberian Peninsula with very high resolution due to the complexity the study area and the non-linearity of the processes governing the dynamic of atmospheric pollutants. In recent years, technological advances have made it possible to incorporate the latest meteorological, transport and chemical research codes into new computer platforms. Moreover, increased computational power as well as access to online data, coupled with advances in the computational structure of air quality models allow for real-time air quality forecasting.

In 2006, the Air Quality Group developed an air quality forecasting tool based on the WRF/HERMES/CMAQ/DREAM modeling system with a high resolution (4 km) for the Iberian Peninsula and 12 km for Europe. This tool provides a better understanding the dynamics of atmospheric pollutants and allows for early communication to the dwellers in areas that exceed air quality thresholds. The group offers a Spain and Europe Air Quality forecasting and assessment service to end-users that takes advantage of this high spatial and temporal resolution of the air quality modeling system. This service is available on the website:
www.bsc.es/projects/earthscience/aqforecast-en/.

As a first step towards the validation of the forecasting system, the Air Quality Group carried out high-resolution annual simulations (specifically for the year 2004) for the Iberian Peninsula (18 km resolution; see figure 1) and nested to Catalonia (2 km resolution) in order to describe the dynamics of air pollution and the relationship between emissions, atmospheric transport, chemistry and deposition. The results were validated against measured data from different air quality networks (EMEP, regional networks, etc.). This analysis and validation was funded by an R+D Grant (CGL2006-08903) from the Spanish Ministry of Education and Science. The group also developed a preliminary version of a high resolution Emissions Model (HERMES, 1 km^2 and 1-hr) for the Iberian Peninsula in order to address for the emission needs of the air quality forecasting.

In 2006, the Air Quality Group began its leadership of CALIOPE, a joint project with other Spanish research groups (CSIC, CIEMAT, CEAM) for improving the accuracy of the forecast and its validation funded by the Spanish Ministry of the Environment. This project also establishes a collaboration with international research groups (University of Aveiro in Portugal,

Laboratoire de Meteorologie Dynamique IPSL in France) to implement the parallel version of their CHIMERE Air Quality Model on the MareNostrum supercomputer. As a result the group performed some preliminary tests for analyzing the parallel performance of the CHIMERE and CMAQ codes when implemented in high performance supercomputing infrastructures.

The Air Quality Group was involved in technology-transfer activities for industries. In projects with Gas Natural, Hidrocantábrico, the group worked on strategies for managing air pollution in large cities (Barcelona and Madrid) by measuring the impact of introducing natural gas vehicle fleets. The team worked on seven different scenarios, focused on the substitution of public transportation vehicles (buses, taxis), freight vehicles and private vehicles with natural gas vehicles. The group also worked on other technology-transfer activities related to the environmental impact assessment (EIA) of combined cycle power plants around Spain and their impact on air quality for several electrical companies.
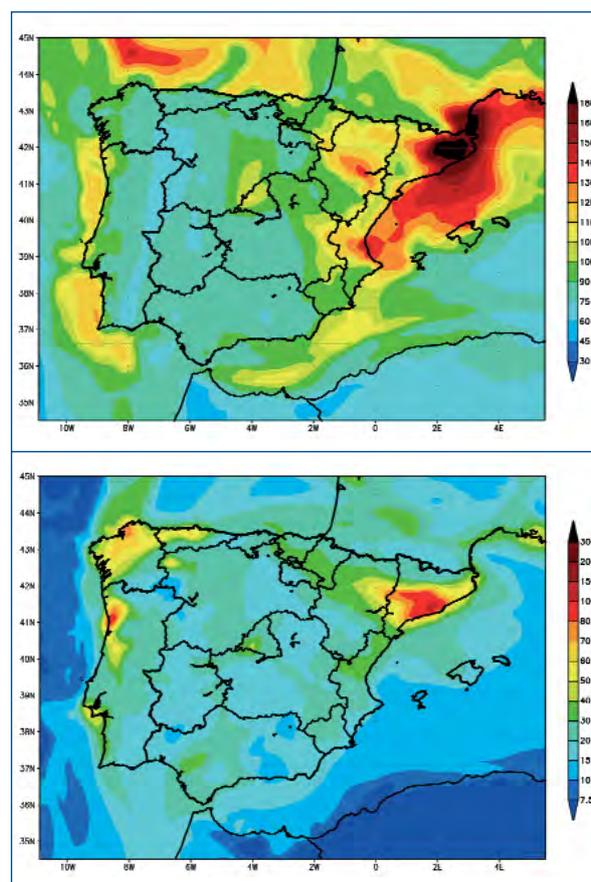

Figure 1. (Above) Ground-level ozone concentration (µg m^{-3}) for the Iberian Peninsula at 12z on August 1st, 2004; (below) Ground-level PM10 concentration (µg m^{-3}) for the Iberian Peninsula at 04z on August 7, 2004.

Climate Modeling (CM)

Most climate simulations using numerical models have emphasized the impact of greenhouse gases on tropospheric climate, but recent advances in atmospheric chemistry as well as increases in computing power have made it possible to develop fully-interactive chemistry-climate models that simulate both tropospheric and middle atmospheric processes. Drawing on the supercomputing capacity of MareNostrum, the Climate Modeling Group began carrying out high-resolution simulations of global climate (resolution of 2x2.5°) as well as regional downscaling over Europe and the Mediterranean region (12-20 km).

The horizontal resolution of coupled Global Circulation Models (GCM) is still too coarse to capture the effects of local and regional forcing in complex areas. The main focus of this research line is the development and testing of regional climate downscaling models for the generation of high-resolution regional climate information from coarse-resolution GCM simulations.

The availability of a world-class computing facility at the BSC-CNS made it feasible to carry out global simulations at 2° x 2.5° resolution and regional simulations at 12-20 km resolution for Europe and the Mediterranean. This allowed the Climate Modeling Group to begin exploring connections between the climate change and the impacts on air quality and extreme events. The group also focused their attention on critical areas such as the Mediterranean which represents a climate change hot spot.

The Climate Modeling Group also implemented the global ModelE by the Goddard Institute for Space Studies (GISS) on MareNostrum as a consequence of the collaboration between the group and the GISS Institute of the NASA. The ModelE constitutes a state-of-the-art global chemistry/climate model (GCM). Past experience suggests that a resolution of at least 2° x 2.5° is necessary to realistically simulate planetary and synoptic atmospheric processes. Therefore, the group carried out simulations with this resolution for the period 1950 - 2000 by using ModelE and the MareNostrum supercomputing facility.

Furthermore, the group also developed and implemented a regional climate model (RCM) based on the WRF / CMAQ / DREAM system for past climate simulations for the Mediterranean Sea and Europe. Emissions by EMEP were linearly disaggregated into a resolution of 20 km for the Mediterranean which was required to capture regional climate features in the Mediterranean and Europe (orographically-forced flows, convective precipitation, air quality, etc.). For this objective, the group performed simulations of the 1960-1980-2000 summertime periods to ascertain the impact of climate change in the trends of extreme events: torrential precipitations, heat waves, drought, air pollution episodes that affect human health, etc (see figure 2).
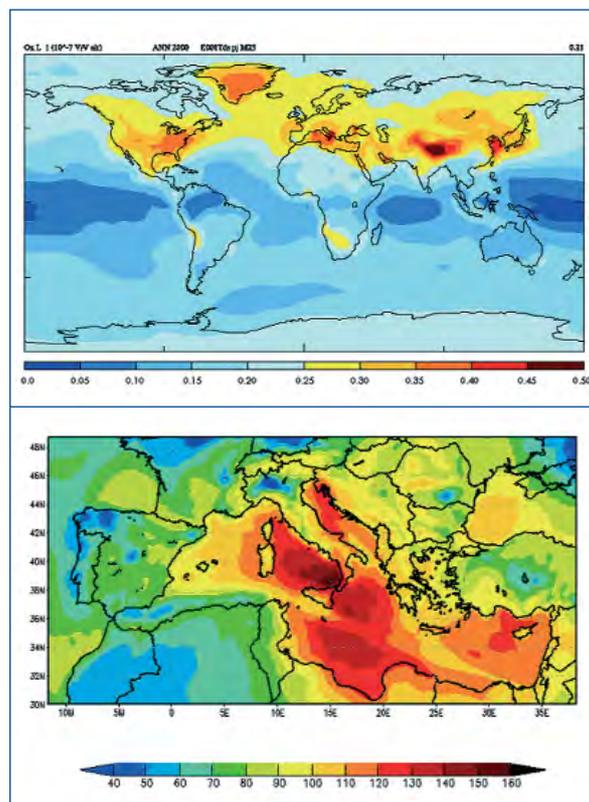

Figure 2. Regional downscaling for the Mediterranean region (above) ground-level ozone global simulations with GISS ModelE; (below) regional simulations with the WRF/EMEP/CMAQ/DREAM modeling system implemented in MareNostrum.

Meteorological Modeling (MM)

The evolution of the atmosphere governs a large number of processes that directly impact society; severe-weather situations lead to high rainfall rates or atmospheric conditions that modify the air quality. The Meteorological Modeling Group or MM Group takes advantage of the computing performance of the MareNostrum super computer to study these areas, focusing on the improvement of the skills scores of their internally developed mesoscale meteorological models for a deeper understanding of the mesoscale phenomena occurring in the atmosphere, paying special attention to the boundary layer processes.

In 2006, the group implemented the existing Weather Research and Forecasting (WRF) Model on the MareNostrum supercomputer. The WRF model is a next-generation mesoscale numerical weather prediction system

designed to serve both operational forecasting and atmospheric research needs. It features multiple dynamical cores and a software architecture that allows for computational parallelism and system extensibility. WRF is suitable for a broad spectrum of applications across scales ranging from meters to thousands of kilometers. The MM Group used this model to provide the meteorological inputs for air quality modeling in the domain of the Iberian Peninsula (4 km resolution) and Europe (12 km; see figure 3) as a part of the previously mentioned CALIOPE project. Additionally, they performed and evaluated high-resolution meteorological simulations for a selected year (2004) for the Iberian Peninsula (18 km resolution) and nested to Catalonia (2 km resolution).

The group also used the WRF model to study the extra-tropical storm, Delta. On November 28-29th, 2005 an extratropical storm affected the Canary Islands causing significant damage related to high sustained wind and intense gusts over some islands of the archipelago. Delta was the twenty-sixth tropical or subtropical storm of the 2005 Atlantic hurricane season. It represents an unusual meteorological phenomenon for that region, and its impacts were underestimated by the different meteorological forecasts during the previous days of the arrival of the low near the Canary Islands. The MM Group established a collaboration with the "Observatorio Izaña" of the "Instituto Nacional de Meteorología" to describe the local effects of the flow that were observed over the Canary Islands during this event using high-resolution mesoscale meteorological simulations. In addition the group analyzed the role of the complex topography of the Canary Islands. The WRF model was applied at 9, 3 and 1 km horizontal resolution.

The current WRF software framework (WSF) supports two dynamic cores: the Advanced Research WRF Model (ARW) and the nonhydrostatic Mesoscale Model (NMM), and in 2006, the MM Group began a performance study of both cores of the WRF model. The purpose of the benchmark study is to analyze the computational performance and scaling of the WRF model on the MareNostrum supercomputer environment as a diagnostic of the current performance of the model under a supercomputer architecture such as MareNostrum as well as determine how to improve its performance. The benchmark, developed in collaboration with the Support Group and the Computer Sciences Department of the BSC-CNS, focuses on both dynamical cores.

Finally, the group performed an analysis of the mean flows affecting different coastal Mediterranean sites through a backtrajectory clustering methodology. The aim of the work was to summarize and quantify the flow characteristics of the region with the application of multivariate statistical techniques. Cluster analysis was used to group trajectories according to wind speed and direction in order to describe

the main flows arriving at the Mediterranean basin and to identify the origin of the air masses affecting the region.
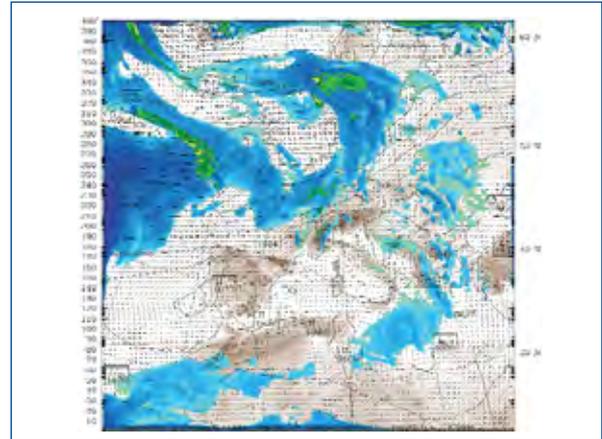

Figure 3. High-resolution meteorological forecast for Europe and the Mediterranean basin with a 12km resolution (sea-level pressure, cloud cover and rainfall).

Mineral Dust (MD)

A large portion of atmospheric particulate matter is derived from arid regions of the Earth and is distributed all over the globe, especially desert mineral dust from the Saharan desert. The impact of mineral dust on air quality, climate and ecosystems represents a major scientific issue. One of the topics in which the European Commission has shown a greater concern is the necessity of assessing the influence of mineral dust emissions on air quality at the regional scale in the European Union.

The Mineral Dust Group or MDGroup provides daily operational forecasts of mineral dust for the Euro-Mediterranean and the East Asia region based on the Dust Regional Atmospheric Model (DREAM), designed to simulate the atmospheric cycle of mineral dust. DREAM has reached a level of delivering reliable operational dust forecasts capable to predict all the major dust events. This service is available in the website: www.bsc.es/projects/earthscience/DREAM. In June of 2006, the Atmospheric Research of National Technical Aerospace Institute (INTA) and the Atmospheric Observatory in Izaña of INM used the DREAM model in the context of the TROMPETA campaign (Tropical Monitoring Phase in the Atmosphere).

The MD Group worked with the parallel NMM meteorological core in order to introduce the dust equations in a parallelized and non-hydrostatic version of the atmospheric model driver. This, in turn, increased the model's spatial and temporal resolution for operational use. This work has been carried out by the support of the National Institute of Meteorology (INM)

and the Earth Science Institute Jaume Almera (CSIC) and the funding of the project CGL2006-11879 funded by the Spanish Ministry of Science and Education.

The group, in collaboration with Naples University and the Technical University of Catalunya (UPC), used aerosol extinction vertical profiles for evaluating the skills of the model to forecast the dust vertical distribution in the region. The Saharan Dust events were measured by two lidar systems in the western and central Mediterranean during a two-year period (March 2001– March 2003).

Some of the research topics cover the definition of dust climatology over the Mediterranean and the Canary Islands. The group performed a model simulation for a specific twenty-year period (1987-2006) at 0.3 º x 0.3 º resolutions in order to analyze the monthly, seasonal and year-to-year variation of the atmospheric dust load, surface concentration, deposition and the frequency and duration of the events over the regions (see figure 4). The group also explored the links of the simulated dust parameters to the variability of the large-scale circulation and possible positive trends of dust in the last years over the Mediterranean.

Lastly, the Mineral Dust Group studied the strong dust storms originating in western and northwestern China affecting Beijing, causing high particulate matter concentrations over the city. The dust events were monitored by continuous aerosol lidar (during its intense phase) and Sun photometric (AERONET) observations in collaboration with international groups of research. The aerosol vertical profiles obtained by these lidar and the AOT values from the Sun photometer were used for the first time to validate the DREAM dust forecast model data over the Beijing region.
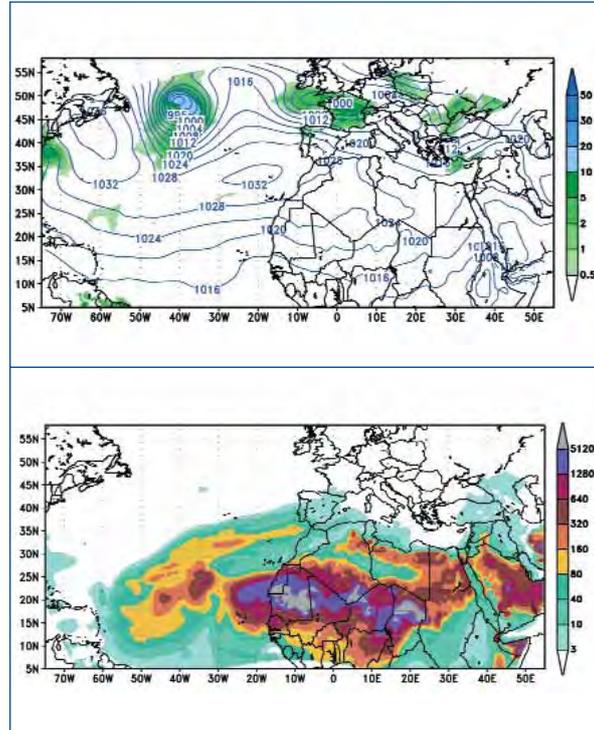

Figure 4. Climate Atlantic DREAM simulations: (above) Precipitation and mean sea level pressure and; (below) lowest model level dust concentration ($\mu g\ m^{-3}$) at 12z on February 14, 1997.

2.3 Life Sciences

The primary goal of the Life Sciences Department is to gain a deep insight into the origin, evolution and function of living organisms using theory and computation.

Modesto Orozco,
Life Sciences Director

The department's focus goes from a global analysis of living entities understood as complex systems to detailed studies of key interactions at the sub-atomic level. Overall, the department's objective is to make theory and simulation one of the driving forces behind the advancement of research in life sciences.

Computational Genomics (CG)

The Computational Genomics or CG Group strives to uncover the molecular processes and evolutionary forces that shape genomes with time, and that have lead to modern genomic structures and current biological diversity. The large availability of complete genomes and the ongoing efforts of the community to obtain more and more genomic DNA sequences make this the perfect moment to carry out comparative analyses to achieve these goals. Taking advantage of this sequence data and the computing power of MareNostrum, this group performs large-scale comparisons between metazoan genomes in order to identify, classify and understand the function and evolution of their biologically relevant regions. The following is a description of the status of the different ongoing projects of the group for 2006, as well as the specific goals of each of these projects.
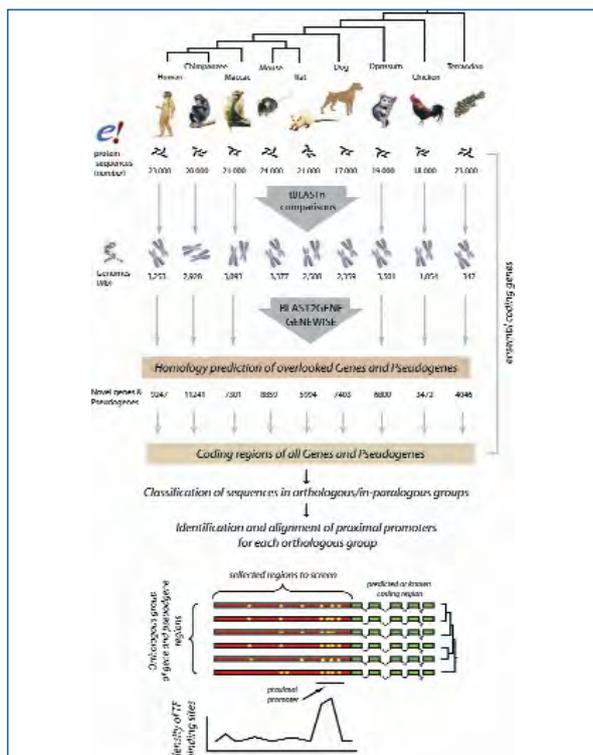

Figure 1. Schematic representation of the complete procedure for the identification and preliminary analysis of vertebrate promoter regions.

One of research activities of the CG Group is the study of genomic regions (proximal promoters) that dictate the expression pattern of genes within organisms. The main goal of this activity is to measure the patterns and rates of sequence divergence of these regions in order to understand how gene expression changes over time in different species and within species after gene duplication.

The first step on the path to achieving this goal was to obtain a representative and comprehensive collection of vertebrate functional and non-functional promoters (i.e. those of duplicated pseudogenes) and to classify these promotors according to their phylogenetic relationships (i.e. in orthologous / paralogous groups). The lack of publicly available large databases of vertebrate promoters provided the incentive for the CG Group to develop an automatic pipeline to detect genes and pseudogenes of all selected genomes which allows the extraction of their proximal regulatory regions (see Figure 1). The CG Group identified genes and pseudogenes in ten large vertebrate genomes that were chosen for their quality and as representative of the major vertebrate classes.

The group then began classifying these genes and pseudogenes according to their relationships in terms of homology, i.e. into orthologous groups. This procedure demands a great deal of computing power because it requires the pairwise comparison of all sequences to identify which sequences are related as well as to evaluate the evolutionary distance of all identified relationships. For each of the orthologous groups identified, the group began constructing phylogenetic trees and extracting the upstream regulatory region, which will later be analyzed in collaboration with the group of Roderic Guigó, at the "Centre de Regulació Genòmica" (CRG) in Barcelona.

The group obtained its first results in 2006; these results will lead to the following manuscripts and databases: (a) a database of orthologous groups of vertebrate genes and pseudogenes, and (b) their corresponding promoter regions; (c) the evaluation of the patterns of divergence of promoters in vertebrates among organisms, and (d) after gene duplication; and finally, (e) the identification of events of positive selection associated to changes in gene expression.

The group also devoted much effort to the analysis of the digestion and assimilation of the proteins in the diet of vertebrates which will constitute a pilot study for broader analysis of other biological systems. The group began an exhaustive compilation of physiological, biochemical and medical information about intestinal amino acid transport, in combination with genomic comparative analyses and theoretical simulations. The goal and major steps of this project were to first determine which molecules (i.e. amino acid and peptide transporters) are involved in the absorption of amino acids in the intestine and then to identify and evaluate the corresponding genes in all available vertebrate genomes to detect episodes of duplication, silencing and modification that can be correlated to physiological or morphological changes. Thus, the group adapted and applied the procedure used for the detection of pseudogenes to the analysis of promoters (see above). The group also constructed a physiological model that, on the basis of classical biochemical experiments and

biomedical data, would explain the net flux of amino acids of the diet from the intestinal lumen to the blood (figure 2a). The analysis of all the genes in vertebrates of one of the families involved in the absorption process led to the detection of a significant number of variations as shown in the phylogenetic tree of these sequences (figure 2b). In collaboration with the group of Manuel Palacín, "Universitat de Barcelona", the group plans to characterize the function of this gene in 2007, which may shed light on the ancestral functions of this family of transporters.

On top of this, the group began to plan theoretical simulation studies in collaboration with the National Institute of Bioinformatics (INB) led by Josep Lluis Gelpí. This group had already modeled the influx and efflux of amino acids in different cell types. As the kinetic properties of nearly all of the transporters considered are already described in the literature, the goal will be to model the absorption of amino acids in the diet and predict different behaviors and responses of the system according to different diets (i.e. different proteins ingested). The group will continue these studies into 2007 as this model could have an important impact on various fields, including for example for the treatment of patients with nutritional deficiencies that suffer diseases such as the Lysinuric Protein Intolerance (LPI) where one of these transporters is affected.
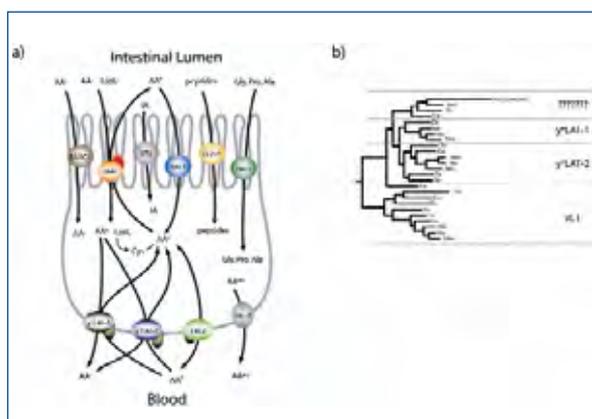

Figure 2. a) Schematic model of the transporters involved in the flux of amino acids from the diet into the blood. AA⁻, anionic; AA⁺,cationic; AAᵃʳᵒ, aromatic amino acids; CssC, Cystine. b) Phylogenetic subtree of the light subunits of heteromeric amino acid transporters (LSHAT) showing the presence of a novel member in some vertebrate species. Hs, *Homo sapiens*; Mm, *Mus musculus*; Gg, *Gallus gallus*; Xt, *Xenpus tropicalis*; Dr, *Dario rerio*.

In addition to the projects mentioned above, the group began several collaborations with different groups in Europe including EMBL in Heidelberg for the preliminary analysis of metagenomic data, Kyoto University in Japan and "Consejo Superior de Investigaciones Científicas" (CSIC) in Madrid for the analysis of gene duplicates in vertebrates to evaluate the duplicability and the functional spectrum explored by different lineages as well as an internal collaboration with Professor Modesto Orozco and Ramon Goñi of the MMB Group for the analysis of structural basis of vertebrate promoters and their capability for de novo prediction of regulatory regions, and Sandra Peiró and Antonio Garcia at the "Institut Municipal d'Investigació Mèdica" (IMIM) in Barcelona for the identification of genes regulated by the Snail transcription factor through the analysis of their promoter regions.

Electronic and Atomic Protein Modeling (EAP)

The Electronic and Atomic Protein Modeling or EAP Group moved from Washington University to BSC-CNS in July 2006 in order to continue the group's research in the following three areas: i) electron transfer and catalysis in heme proteins, ii) protein energy landscape exploration and iii) ligand dynamics and induced fitted docking.

Numerous studies show the importance of the heme propionates in regulating the metal center redox potential. Most of these systems share the catalytic putative species, a highly reactive iron-oxo species known as Compound I (Cpd I), very difficult to characterize experimentally. Recently the group presented QM/MM results indicating the delocalization of the third unpaired electron into the heme propionates. The group studied the implication of this electronic delocalization into the oxidation mechanism in ascorbate peroxidase. Figure 3, panel B, clearly indicates the spin delocalization as a result of the electron transfer from the ascorbate substrate to the iron metal center. Thus, the theoretical modeling of Cpd I supports its high oxidative nature.

By using mixed QM/MM methods, the group can activate/deactivate the electronic (in addition to the atomic) description of different regions in the enzyme (the quantum region). Panel A in Figure 3 displays the spin density when ascorbate is excluded from the quantum region. Thus, eliminating the quantum description of the ascorbate substrate eliminates the possibility of an electron transfer from this species to the iron metal center. In this way the group can obtain an intermediate, extremely difficult to detect experimentally, of the electron transfer pathway. The results clearly indicate the direct involvement of the propionate lone pairs into the electron transfer pathway. This is the first direct evidence of the involvement of the heme propionate substituents into a cytochrome oxidation process.

With respect to Hemoglobin allostery, the group continued to collaborate with Professor Thomas G. Spiro in seeking to understand the mechanism behind hemoglobin cooperativity. Initial results indicate a differential in binding affinity between the T and R structures, in agreement with experimental results. The
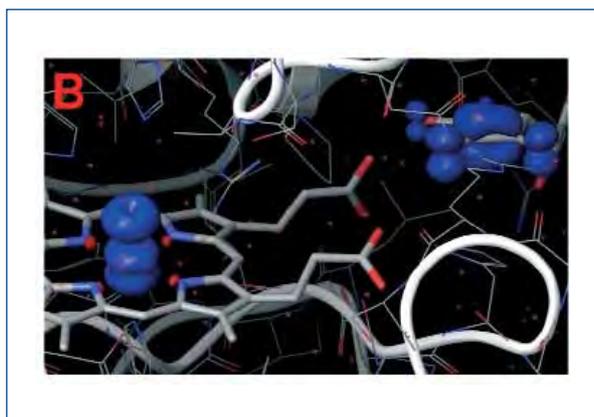
Figure 3. Spin density for putative active species, compound I, in Ascorbate Peroxidase. panel A: quantum region includes the heme group, the oxo ligand and the proximal histidine; panel B the quantum region also includes the substrate ascorbate. Main pictures in each panel show the quartet spin state; the inset on the right bottom corner in panel A shows the doublet spin state.

theoretical simulation, at an atomic resolution, indicated the main contributors to this differential binding, and interestingly, the larger effects are located in the inter chain contacts.

In addition, the group developed a sampling algorithm to map the protein energy landscape known as PELE (Protein Energy Landscape Exploration) which combines a steered stochastic approach with protein structure prediction algorithms. Initial applications showed its potential in projecting the migration dynamics of ligands in protein. The methodology is based on consecutive iterations of three moves (steps): a localized perturbation, a side chain sampling and a minimization. Along 2006 the group has added new capabilities to the PELE program. The initial perturbation step now includes the possibility of sampling the backbone movement through carbon alpha restrained displacement obtained from an

elastic network model. The group is also adding a coarse grained model by means of the UNRES force field. The goal for 2007 will be to couple the elastic network and the coarse grain force field with the all atom description of PELE to accomplish a multiscale level sampling technique, capable of larger collective protein sampling. Application to gating mechanism for ligand entry and escape will follow.

The group also began a development project to adapt PELE to a rigid ligand docking algorithm, GLIDE. This project is performed in collaboration with the Schrodinger software company (www.schrodinger.com) which will co-finance a postdoctoral researcher in 2007.

The capabilities of the new algorithms for ligand/protein landscape exploration (PELE, paragraph above) are being used in several applications. In collaboration with Professors Denis L. Rousseau and Syun-Ru Yeh at the Albert Einstein College of Medicine (New York City, USA) the group
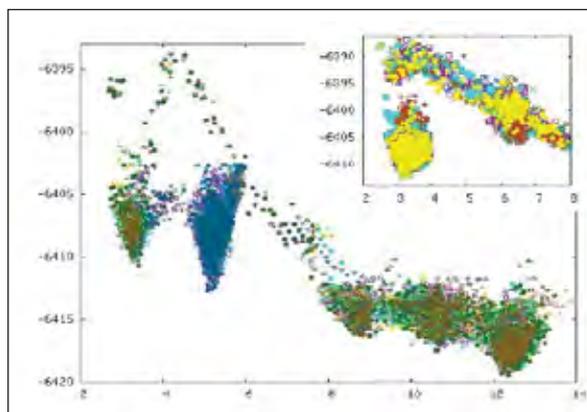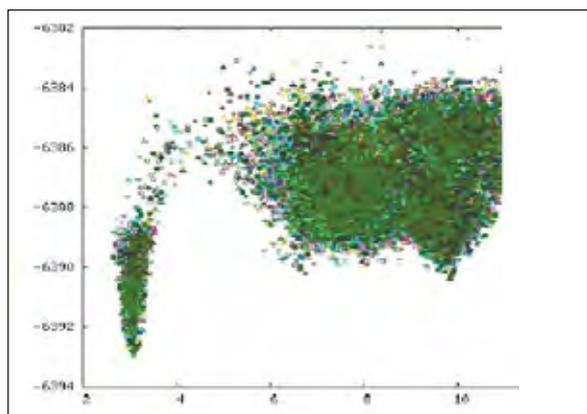

Figure 4. WT (bottom) and G8F (top) mutant energy landscape for the ligand escape dynamics in trHbO. The X axis represents the Fe-C distance and the Y axis the energy differences, in Å and kcal/mol units, respectively. The inset in panel right shows a longer run (more sampling) of the active site escape barrier.

continued studying the enzymatic and ligand migration mechanisms in several truncated hemoglobins. Initial results in trHbO have indicated the importance of the Trp G8 residue in the active site ligand dynamics. Figure 4 indicates the energy landscape for a carbon monoxide migration dynamics (escape from the active site to solution) in the WT (bottom panel) enzyme and the G8F mutant (top panel). In the mutant there is a much deeper active site binding of the ligand, corresponding to Fe-C distances of ~3 Å, and a significantly lower escape barrier, in agreement with experimental observations in the Yeh lab.

A joint project with professors Ernest Giralt (PCB, Parc Científic de Barcelona) and Ricardo Perez Tomas (UB, Universitat de Barcelona) focusing on the design of kinases inhibitors was started in 2006. Using PELE the group seeks to understand the binding dynamics of different inhibitors. Using QM/MM techniques the group is also seeking to understand the enzymatic mechanism of the enzymes, with the goal of characterizing the catalytically important protein residues. The combination of these two methods should facilitate the *a la carte* design of mutant-resistant inhibitors, a central objective in cancer research. The main idea is to design novel inhibitors whose binding is driven by interactions with those residues critical for normal enzymatic function. The studies will center on the de-regulated tyrosine kinase, *Bcr–Abl*, and in several other kinases in the PI3K-Akt apoptotic pathway. Initial studies have been initiated with AT514, a natural inhibitor discovered in the lab of BSC-CNS collaborators.

National Institute of Bioinformatics (INB) Computational Node 2

The National Institute of Bioinformatics (INB) is a research facility created by Genoma España Foundation with the aim of giving support to Bioinformatics groups that participate in Spanish Genomic and Proteomics projects. The Institute has a nodal structure distributed among the most important bioinformatics research groups in Spain. BSC-CNS's Life Sciences Department hosts the computational node of INB, whose purpose is to provide access to generic and project-related biological databases and to develop web services and applications covering a broad range of analysis software.

The INB Node 2 Group focuses on 2 areas. The group provides external project support and helps in the implementation and optimization of users' applications in the supercomputing environment. The group also develops web services and databases and is responsible for creating Biomoby web services and managing biological databases.

External Project Support

In 2006, the INB Node 2 Group worked with Professor Alfonso Valencia group at the "Centro Nacional de

Investigaciones Oncológicas" (CNIO) to adapt a variety of software dealing with the prediction of protein function and protein-protein interactions to the MareNostrum architecture.

The INB Node 2 Group optimized Funcut, a software which does automatic annotation of protein function. Predictions have been made for a set of protein sequences from the Protein Family database (PFAM) for protein sequences sets from species Streptomyces and Arabidopsis. The group's optimization led to the reduction of execution time by a factor of 5.

TSEMA is a software application that is used to predict physical interactions between protein superfamilies using the information in their phylogenetic trees. The INB Node 2 Group optimized this application to make it run 1000 times faster than the original version. This translates to the completion of 44,000 program executions using only the 3000 CPU hours assigned to the project.

The INB Node 2 Group also started a project (allocated 10,000 CPU hours) that focuses on the optimization and benchmarking of several docking algorithms for the prediction of protein-protein interactions so as to take advantage of the MareNostrum architecture. In the first phase of this project, the group installed and compiled RosettaDock on MareNostrum in order to avoid unexpected execution problems common to the PowerPC architecture.

In the area of Comparative Genomics, the INB implemented an automated pipeline involving massive TblastX calculations on MareNostrum (Fig 5) for Professor Roderic Guigó of the "Centre de Regulació Genómica" (CRG). The predictive power of many bioinformatics tools make use of the similarities and differences between whole genomes detected by TblastX. Gene prediction can take advantage of the different conservation patterns between coding and non-coding DNA sequences.
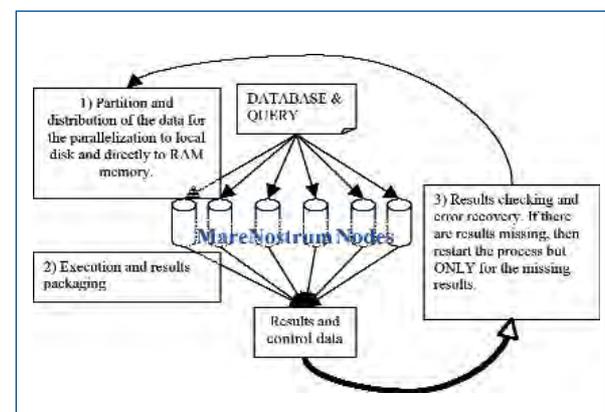

Figure 5. Pipeline for TBlastX calculations.

The INB Node 2 Group also designed a pipeline to run on MareNostrum to align EST sequences to the human genome for the purpose of finding genes for Professor Eduardo Eryas at the "Institut Municipal d'Investigació Mèdica" (IMIM) in Barcelona.

Expressed Sequence Tags (ESTs) allow for the identification of exon sequences in genomes. However, the heterogenic nature of experimental information available requires a complex combination of alignment and sequence identification calculations. Exonerate is a program for finding exonic structures over a genome, given a set of ESTs. When searching for exonic structures, its computational cost grows exponentially along with the size of the input data, so it can't be run using a whole genome as input and set of ESTs for given species. In order to solve this problem, the group developed a two-step execution routine:

In the first step Exonerate is run on the genome without the exon-modeling feature which is faster as it only performs simple alignments. For each EST, the 10 best alignments are saved.

In the second step, Exonerate is run on the genome with exon-modeling feature in a restricted search space in the following fashion: i) The genome is partitioned; ii) the ESTs are clustered into files that contain the set of ESTs that aligned against each one of the genome partitions performed in step (i); iii) Exonerate is executed on each partition with the exon structure using the corresponding sets of ESTs from (ii). This pipeline will be applied to other genomes in the future.

Lastly, the INB Node 2 contributed to work on Protein Domain prediction for Professor Baldo Oliva's group at the "Institut Municipal d'Investigació Mèdica" (IMIM) in Barcelona. The group adapted a protein domain predictor, DomHunt, to supercomputer execution,designing a pipeline allowing for the analysis of thousands of protein structures in parallel. The group also perfomed a massive BlastP search between Uniref series and a NR protein database. Results of these calculations will be available at the beginning of 2007, after the expected upgrade of the INB web server.

Web Services and Database Development
INB computational node was previously located at the UPC and moved to BSC-CNS in the early part of 2006. Part of the initial work during this year has been the technical migration to the new hardware and software environment. This phase included:

- Installation of external programs: EMBOSS, HMMER, NCBI Blast, WU Blast, Mirror.
- Installation of perl modules: MOBY, BioPerl, SOAP, LibXML and many others.

- Modification of source code according to hardware requirements and execution policy of BSC-CNS. This work was done in close collaboration with the systems administration group of BSC-CNS.
- Implementation of databases in the new storage system: EMBL, GenBank, UniProt, TrEMBL, SwissProt, PDB, FSSP, DSSP, HSSP. These applications were offered for internal BSC-CNS users (including MareNostrum projects) and as external web services.

Database	NCBI Blast	WU Blast	FASTA
SwissProt	✓	✓	✓
TrEMBL	✓	✓	✓
UniProt	✓	✓	✓
PDB	✓	✓	✓
NR	✓	✓	✓
RefSeq_aa*	✓	✓	✓
RefSeq_dna*	✓	✓	✓
GenBank*	✓	✓	✓
EMBL*	✗	✗	✗

*RefSeq, GenBank and EMBL were formatted by divisions.

The group also developed and adapted the following webservices to the MareNostrum architecture:

Protein motifs:
- Pratt: Search for patterns conserved in a set of protein sequences.
- PPsearch: Search query sequence for protein motifs, rapidly compare query protein sequence against all patterns stored in the PROSITE pattern database.
- Ps_scan: Search an amino acid query sequence against Prosite and scan a protein sequence for the occurrence of a pattern.
- Interpro: Query an Amino Acid sequence against InterPro database.

Retrieval and parsing services:
- Retrieval services: retrieve an entry from Prosite (getEntryfromProsite) and retrieve a documentation entry from Prosite (getDocumentationFromProsite).
- Parsing services: parse the matrix or rules or pattern of a motif (parseMatrixFromMotif, parseRulesFromMotif, parsePatternFromMotif).

Emboss-based services.
- Composition: coderet, embossdata, cups, codcmp.
- Physico-chemical properties: antigenic, charge, octanol, pepinfo, pepwindow, pepwindowall.
- Protein motifs: patmatdb, patmatmotifs.

All of these new services have been documented and registered at the INB and main Biomoby catalogues. The complete documentation is available at the node website (inb.bsc.es). As the content of the site is automatically generated from registry information, it is consistent and up-to-date with respect to the Biomoby catalogue. Example workflows illustrating most common usage of the services are also available.

Molecular Modeling and Bioinformatics (MMB)

The objective of the Molecular Modeling and Bioinformatics or MMB Group is to understand the behavior of living organisms by means of theoretical models, whose roots are anchored in the basic principles of physics and chemistry. With this general aim, the group works with different methodologies, from mining of biological databases to classical dynamics and quantum chemistry calculations.
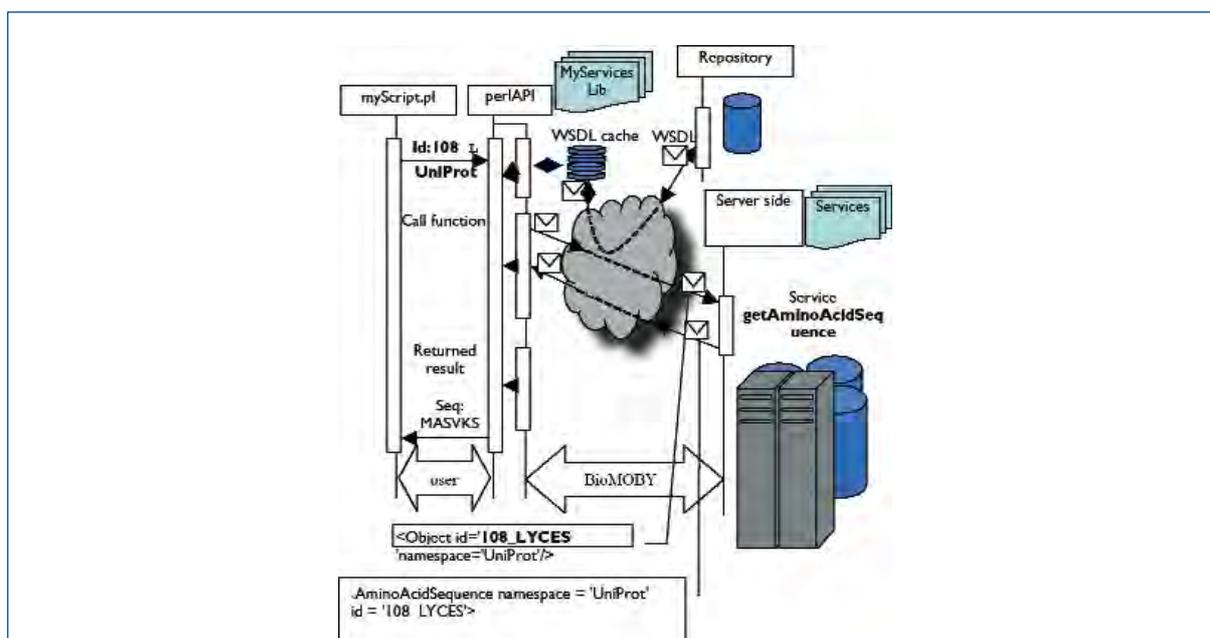
Figure 6. BioMOBY API Webservices Architecture.

Finally, the group developed two new utilities that were implemented as Perl modules to be added to the standard Biomoby distribution.

The MobyLite API consists of a set of perl modules and scripts that have been built to allow programmatic interaction with Biomoby web services for non-trained Perl programmers not familiar with biomoby webservices. In the past, using the web services within bioinformatics applications required a deep understanding of the standard. The Mobylite API builds the necessary Perl modules automatically from the Biomoby registry which is completely transparent to the users.

Moreover, the API hides the Biomoby functionality in a set of normal Perl modules that can be used for a normal Perl programmer without any knowledge of the standard, even thought that underlying machinery uses it (Figure 6).

The Service Testing utility consists of a set of scripts has been developed to allow Biomoby service providers to perform automatic tests of their services, statistics and check results accuracy.

The use of this wide range of methodologies allows the group to explore a wide range of problems, from enzyme reactivity to genome analysis.

This group has a long trajectory in the study of small model systems, whose study can provide the clues for a better understanding of the behavior of much more complex biological molecules. Almost a decade ago, the group realized that such studies were simple in the gas phase but very difficult in aqueous solution, which led the group to develop methods to describe solvent; some of these methods are currently considered the "state of the art" in the field.

In 2006, the group continued working on the development and improvement of methods to study condensed phases of small model systems by introducing dispersion contributions, a methodology that has been successfully applied to complex chemical processes. The group also made progress in transferring the developed methodology to the macromolecular world, notably on the applications side, it completed the first ab

initio atlas of nucleobases-nucleobase stacking. The group also worked quite intensively with non-natural derivatives of nucleobases which can be incorporated in DNA.

In 2006, the group also continued collaborating with Professor Darío A. Estrín's group at the University of Buenos Aires in the analysis of mechanisms of ligand diffusion and reaction of Mycobacterium Tuberculosis truncated-hemoglobin-N. This enzyme is responsible for NO detoxification, and its over-expression is responsible for the resistance of Mycobacterium Tuberculosis to macrophage action. In 2006, the group characterized the main diffusion pathways for NO and O2 inside the enzyme, locating the bottlenecks and control points of the process. The group also participated in a project in collaboration with a pharmaceutical company that led to the definition of the three dimensional recognition properties of a new family of P38α MAP Kinase inhibitors which are now included in 2 clinical trials as a powerful anti-inflammatory compound.

A growing interest in understanding the general aspects of protein structure and flexibility has led the group to develop MODEL (Molecular Dynamics Extended Library; see Figure 7), the largest database of molecular dynamics simulations that currently contains dynamic information for approximately 1300 proteins.

The analysis of this massive database of trajectories provides a complete picture of protein dynamics. These macromolecules are found to be melted solids and the essential dynamics of the proteins is clearly printed in their three dimensional structure. Hydrogen bonds are the stiffer interactions of proteins while the saline bridge appears to be surprisingly labile.

The group is also interested in obtaining information on key biological processes through the analysis of available biological databases. Thus, the group focused on the following: i) analysis
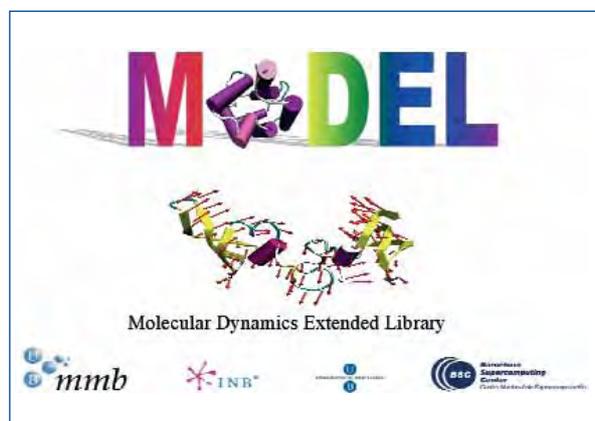
Figure 7. Main page of the MODEL webserver and database (mmb.pcb.ub.es/MODEL).

of genomes to determine DNA with unusual structures and ii) development of predictive tools to determine the pathological character of Single Nucleotide Polymorphisms (SNPs).

With respect to analyzing genomes with unusual DNA structures, the group completed the mapping of Triplex Targeted Sequences (TTS) in the human genome in 2006. Their results confirmed that these triplexes are many times more common in the human genome than in background models (see Figure 8), and that they are over-concentrated in promoter regions, especially the regions of genes related to the control of cellular functions. These results open exciting possibilities for triplex strategies such as antigene therapies, as triplex formation at promoters are known to knock-out or knock-down genes.

In 2006, the group also continued to make progress in the analysis of the trends that determine when a mutation is
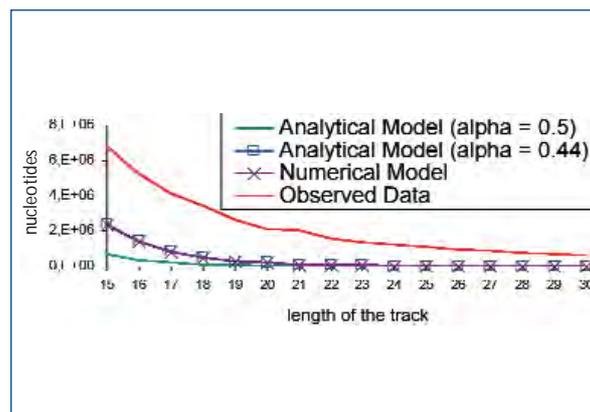
Figure 8. Representation of the number of nucleotides in TFOs of different lengths compared with two background models.

neutral or pathological. The group made improvements to the PMUT server (mmb2.pcb.ub.es:8080/PMut/) which is one of the well-known references for prediction of the pathological nature of SNPs. Furthermore, the group performed a comparative analysis on neutral, pathological and correlated mutations and found that each individual mutation of the correlated pair is characterized by a low pathogenicity index; this explains the maintenance of individual mutations before being stabilized by complementary mutations.

The group also focused their work in another direction on both methodological aspects of analysis of nucleic acids (NA) molecular dynamics trajectories and the detailed study of several unusual conformations. In particular, the group explored the existence of short fragments of Hoogsteen antiparallel DNA embedded in long B-DNA duplexes. Extended MD simulations demonstrated that not only is the

parallel Hoogsteen a stable structure, but that the H-B junction does not introduce dramatic distortions to the overall structure of the helix. The group also focused on quadruplex DNA (G-DNA) and found that G-DNA is stable enough to survive full vaporization in the presence of monovalent ions. The group confirmed experimentally (available experimental information supports our hypothesis) that G-DNA will be the first biopolymer to be stable in gas phase conditions, opening unexpected biotechnological applications for DNA quadruplexes.

The group is interested in anti-gene therapy. The idea behind this approach is to target a duplex DNA by means of triplex forming oligonucleotides (TFO), which binds to the major groove of the duplex generating a triplex. Once the triplex is formed the gene is knock-out, especially when triplex formation occurs at the promoter region. Multiple groups are putting much effort in designing TFOs leading to more stable triplexes. However, very often the formation of the triplex is hampered by that of a tetraplex. Part of the work done during 2006 has been focused in designing nucleobases derivatives which avoid the formation of the TFO-tetraplex. The simulations suggest that excellent results can be obtained with 8-aminoguanine, a result that was later confirmed experimentally.

Protein Interactions and Docking (PID)

Protein-protein interactions are essential for most biological processes in living organisms. The Protein Interactions and Docking or PID Group aims to understand the mechanism behind the formation of specific complexes between proteins and to develop computational tools to make structural, kinetic and thermodynamic predictions. The group's ultimate goal is to design small-molecule interaction regulators of biomedical interest. Thus, the group takes advantage of the parallel computational capabilities of BSC-CNS to work on algorithmic solutions, based on physico-chemical principles.

In 2006, the PID Group developed and optimized computational tools for characterizing and understanding protein-protein interactions. In particular, they worked on developing docking algorithms. As an important aspect of this development, the group optimized the most up-to-date scoring functions for reevaluating rigid-body docking poses, which yielded excellent results in earlier tests. The docking algorithms are being continuously evaluated both by in-house benchmarks, and also externally through the Critical Assessment of PRediction of Interactions (CAPRI) competition. The group's docking algorithms achieved a high rate of success in rounds 4 and 5 of the competition, making good predictions on all targets (Figure 9). The group implemented their docking algorithms in a modular program called pyDock. The program was parallelized in MareNostrum and was subjected to massive benchmarking and comparison with other programs. Moreover, in 2006, the group started to integrate their scoring function into a Grid search procedure using Fast Fourier Transfer (FFT) algorithms to speed up calculations.

Another important development has been the optimization of pyDock to include sequence conservation or mutational data as distance restraints, which dramatically improves the predictive accuracy of the docking results. The group plans
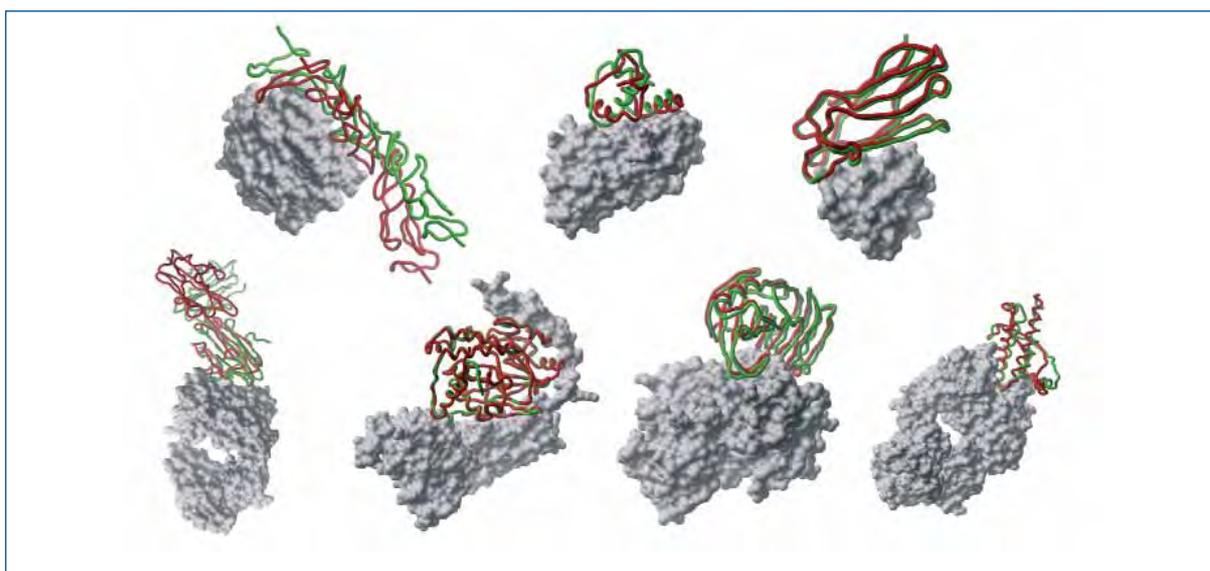

Figure 9. Predicted interaction models (in red ribbon) for protein-protein complexes of the CAPRI experiment (http://capri.ebi.ac.uk) obtained by applying our docking algorithms. Ligand RMSD values with respect to the crystallographic structures (in green ribbon) start from as low as 0.6 Å.

to explore the possibility of introducing flexibility based on MD or NMA analysis.

In 2006, the group also began extending their powerful binding site prediction tools, such as Optimal Docking Area (ODA) and Normalized Interface Propensity (NIP) to the identification of hot-spot residues at binding interfaces. The *NIP* values were able to predict experimental hot spots with 79% accuracy (Figure 10). The tools were also applied at proteomic scale to the study of large protein families, such as the carboxypeptidases. Moreover, the group also implemented the ODA method together with other solvation descriptors into a web server for determining Fractional contributions to Solvation in proteins (FSolv) in collaboration with the INB Group.

The group aims not only to perform state-of-the-art academic research on protein docking, but also to make computational tools that are useful for the biomedical scientific community. In order to accomplish this, it is essential to collaborate with other experimental laboratories on specific protein-protein systems not only to obtain feedback on the tools but also on the methods upon which the tools are based. Throughout 2006, the group collaborated with other laboratories to model protein-protein complexes involved in electron transfer reaction, such as the ones formed between ferredoxin-NADP+ reductase and the proteins flavodoxin and ferredoxin, or the complex between cythochrome c oxidase and the proteins plastocianin and cyt c6. Regarding other practical docking problems, the group also provided rationale about
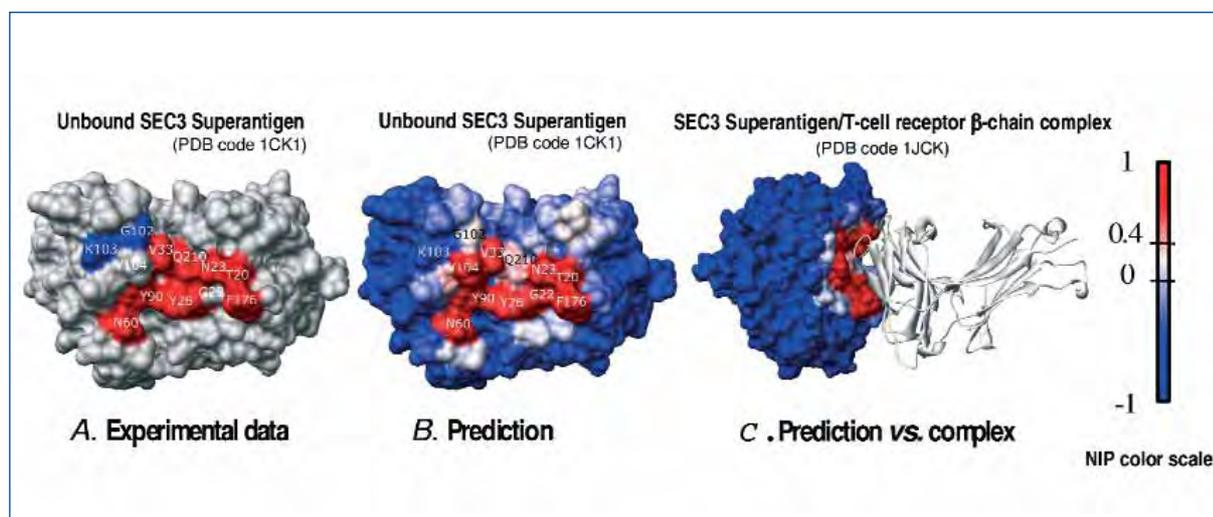
Figure 10. A) Residue contribution to binding energy ($\Delta\Delta G$) upon mutation to alanine in the complex between SEC3 super antigen and T-cell receptor ß-chain. Hot-spot residues are shown in red ($\Delta\Delta G > 1$kcal/mol), while in blue are shown residues with $\Delta\Delta G < 1$kcal/mol; and in grey, residues with not available $\Delta\Delta G$. B) Predicted hot spots are shown in red (*NIP* > 0.4), while in blue are shown residues with a *NIP* < 0.4. C) Complex view, SEC3 super antigen residues are coloured according to their *NIP* values whereas the T-cell receptor , ß-chain is represented as a grey ribbon.

In addition, the group applied physical principles to study the energy landscape of the process of association between two proteins. A deep understanding of protein-protein association is essential, not only for accurate modeling and prediction of complex structures, but also for designing small-molecule inhibitors of protein-protein complexes. Interestingly, the group found good rigid-body docking minima in the electrostatic cases, regardless of the position of the initial conformations with respect to the near-native structures. These minima could represent the first approach of the two interacting proteins, and the group plans to pursue this area further in 2007.

interaction energy and structural requirements on challenging multi-protein systems such as the bacteria multi-drug pump, the protein complexes of the plant immune system, the Bin1 type membrane curvature-inducing proteins; the interaction between helicase and ribonuclease, the possible interactions of the CK2ß, protein, etc. The group is also interested in small-molecule binding to proteins, with the ultimate goal of integrating protein-ligand and protein-protein docking procedures for future drug design studies.
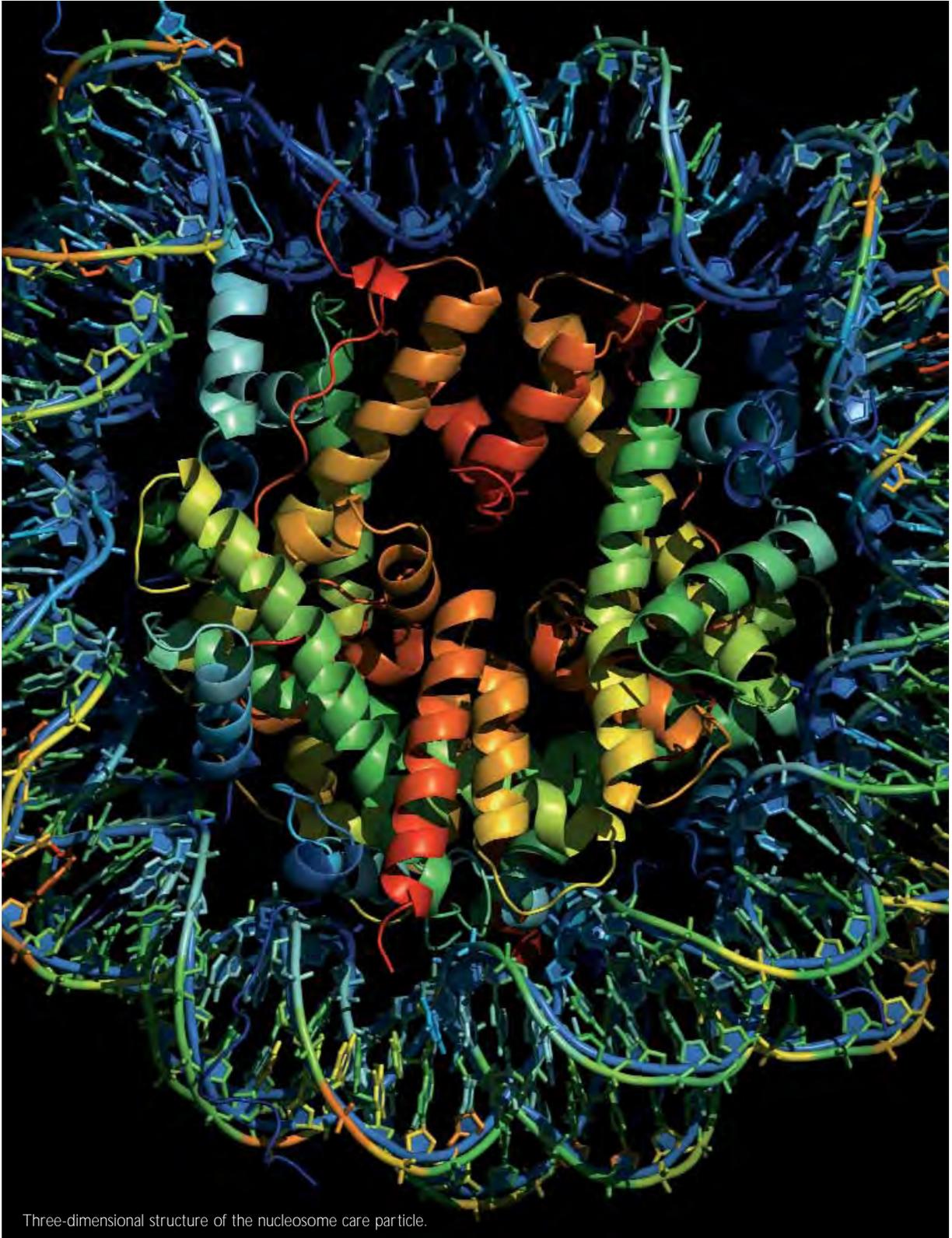
Three-dimensional structure of the nucleosome care particle.
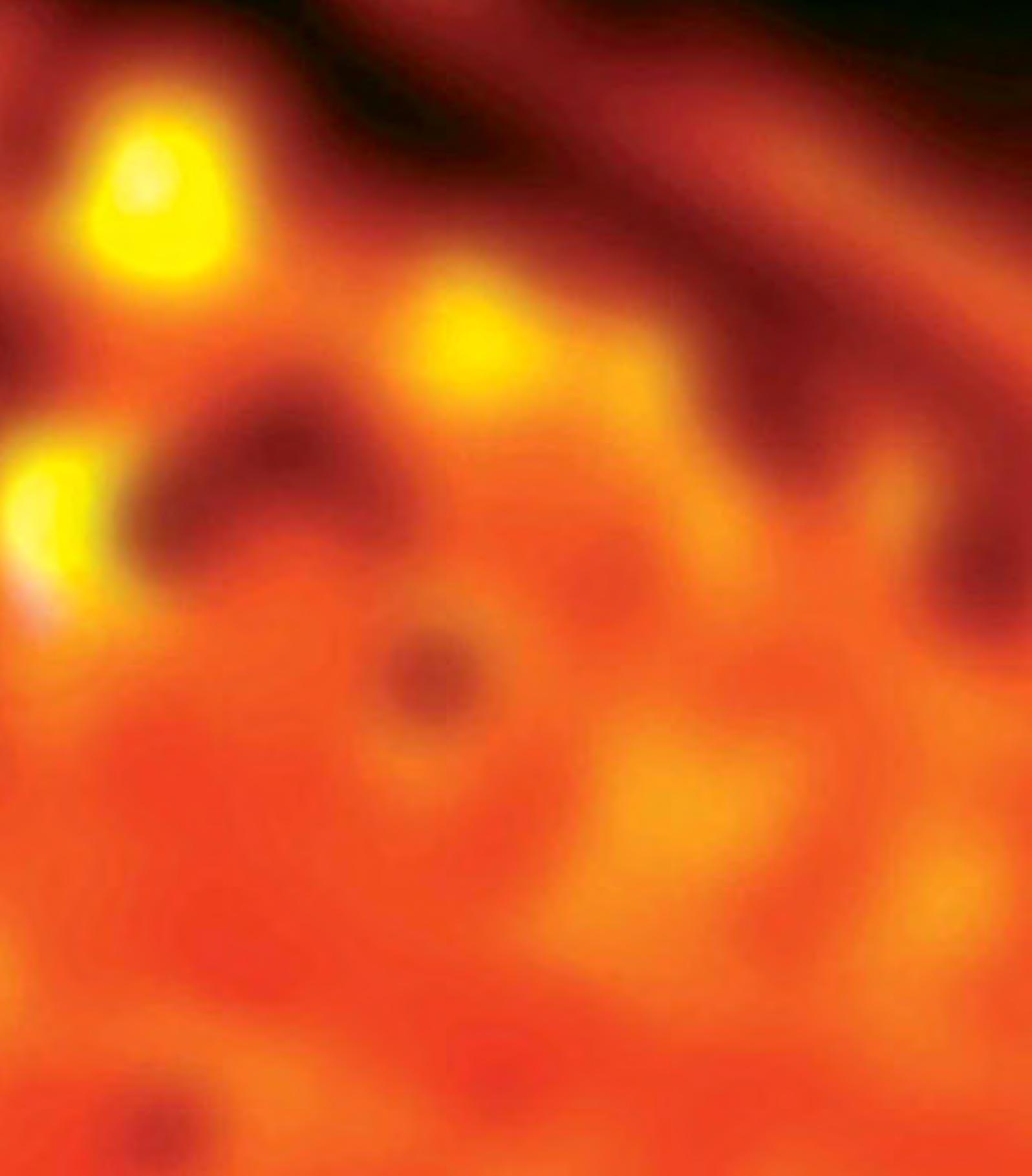
Research
Results

Publications and communications.

3.1 Computer Sciences

Publications

Journals and Book Chapters

• Rosa M. Badia, Raul Sirvent, Jesus Labarta, Josep M. Perez. "Programming the GRID: An Imperative Language Based Approach". Book chapter in Engineering the Grid, Section 4, Chapter 12, January 2006.

• F. Cazorla, P. M.W. Knijnenburg, R. Sakellariou, E. Fernandez, A. Ramirez and M. Valero. "Predictable Performance in SMT processors: Synergy Between the OS and SMTs". IEEE Transactions on Computers, vol. 55(7) IEEE Computer Society 785-799, July 2006.

• R. Codina and G. Houzeaux. "Implementation Aspects of Coupled Problems in CFD involving time". Promuval. In press, October 2006.

• J.J. Costa, T. Cortes, X. Martorell, E. Ayguade, and J. Labarta. "Running OpenMP application efficiently on an everything-shared SDSM". Journal on Parallel and Distributed systems (JPDC), Volume 66, Issue 5, pp. 647-658 , June 2006.

• F. Courty, T. Roy, B. Koobus, M. Vázquez and A. Dervieux. "Error analysis for P1-exact schemes". Int. J. Num. Meth. Fluids, Submitted paper, October 2006.

• Dervieux, M. Vázquez, L. Hascoet and B. Koobus. "Optimization loops for shape and error control". CFD Journal, Submitted paper. Also published in Post-SAROD Indo-French Workshop, Bangalore, December 12-13 in "Recent Trends in Aerospace Design and Optimization", B. Uthup, S.P. Koruthu, R.K. Sharma, P. Priyadarshi Eds. pages 363-373, Tata-McGraw Hill, New Delhi, 2005, October 2006.

• G. Houzeaux and J. Príncipe. "A Variational Subgrid Scale Model for Transient Incompressible Flows". Int. J. Comp. Fluid Dyn. Submitted, October 2006.

• T. Morad, U. Weiser, A. Kolodny, M. Valero and E. Ayguadé. "Performance, Power Efficiency and Scalability of Asymmetric Cluster Chip Multiprocessors". IEEE CAL, Computer Architecture Letters, Volume 5, number 1, pp.14-17, June 2006.

• D. Ródenas, X. Martorell, E. Ayguadé, J. Labarta, G. Almási, C. Cascaval, J. Castaños, and J. Moreira. "Exploiting Multilevel Parallelism using OpenMP on a Massive Multithreaded Architecture". Journal of Embedded Computing, Special issue: Issues in embedded single-chip multicore architectures, Vol 2(2), pp.141-155, ISSN1740-4460, April 2006.

• M. Vázquez, A. Dervieux and B. Koobus. "A methodology for the shape optimization of flexible wings". Engineering Computations, 23. 4., January 2006.

• J. Vidal, M. March, Ll. Cerdá, J. Corbal and M. Valero. "A DRAM/SRAM Memory Scheme for Fast Packet Buffers". IEEE Transactions on Computers, Vol. 55 No. 5, pp. 588-602, May 2006.

International Conferences and Symposiums

• J. Alastruey, T. Monreal, V. Viñals and M. Valero. "Speculative Early Register Release". ACM International Conference on Computing Frontiers, Ischia, May 2-5, 2006, May 2006.

• Rosa M. Badia, Raül Sirvent, Marian Bubak, Wlodzimierz Funika, Cezary Klus, Piotr Machner, Marcin Smetek. "Performance monitoring of GRID superscalar with OCM-G/G-PM: integration issues. 2nd CoreGRID Integration Workshop". Integrated Research in Grid Computing. Krakow (Poland), October 2006.

• Jairo Balart, Marc Gonzàlez, Xavier Martorell, Eduard Ayguadé and Jesús Labarta. "Runtime Address Space Computation for SDSM Systems". The 19th Int. Workshop on Languages and Compilers for Parallel Computing (LCPC 2006). New Orleans, Louisiana, November 2006.

• J.Caubet, K. Hogan, R. Nou, J. Labarta and J. Torres. "Supporting the Introduction of Autonomic Computing in Middleware". Hursley, England. Engineering Conference VII, IBM Academy of Technology Conference IBM, October 2006.

• Alejandro Duran, Roger Ferrer, Juan José Costa, Marc Gonzàlez, Xavier Martorell, Eduard Ayguadé, and Jesús Labarta. "A Proposal for Error Handling in OpenMP". International Workshop on OpenMP (IWOMP 2006), University of Reims Champagne Ardenne. June 2006.

• M. Farrera, T. Cortes, J. Labarta, and G. Almasi. "Scaling MPI to short-memory MPPs such as BG/L". 20th ACM International Conference on Supercomputing (ICS-06), Cairns, Australia, June 2006.

• J. L. Gonzalez, T. Cortes. "Evaluating the Effects of Upgrading Heterogeneous Disk Arrays". International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS 2006), Calgary, Canada. (Best paper award), August 2006.

• Elzbieta Krepska, Thilo Kielmann, Raül Sirvent, Rosa M. Badia. "A service for reliable execution of Grid applications". 2nd CoreGRID Integration Workshop. Integrated Research in Grid Computing, October 2006.

• Róbert Lovas, Raül Sirvent, Gábor Hermann, Josep M. Pérez, Rosa M. Badia, Péter Kacsuk. "Dynamic workflows in the service-oriented P-GRADE portal using GRID superscalar. 2nd Austrian Grid Symposium". Innsbruck., September 2006.

• Xavier Martorell, Marc Gonzàlez, Alejandro Duran, Jairo Balart, Roger Ferrer, Eduard Ayguadé, and Jesús Labarta. "Techniques Supporting threadprivate in OpenMP". 11th Int. Workshop on High-Level Parallel Programming Models and Supportive Environments (HIPS 2006), (in conjunction with IPDPS2006), Rhodes Island, Greece. April 2006.

• R. Nou, J. Guitart and J. Torres. "Simulating and Modeling Secure Web Applications". Lecture Notes on Computer Science (LNCS) Vol. 3991, pp. 84-91. 6th International Conference on Computational Science (ICCS'06), University of Reading, UK, May 28-31, 2006.

• Ramon Nou, Jordi Guitart, David Carrera, Jordi Torres. "Experiences with Simulations - A Light and Fast Model for Secure Web Applications". pp. 177-186, 12th International Conference on Parallel and Distributed Systems - Volume 1 (ICPADS'06), 2006, Volume: 1, On page(s): 177- 186, ISSN: 1521-9097, ISBN: 0-7695-2612-8, July 2006.

• R. Nou, F. Julià, D. Carrera, K. Hogan, J. Caubet, J. Labarta, J. Torres. "Monitoring and analyzing a Grid Middleware Node". Proc. 7th IEEE/ACM International Conference on Grid Computing 2006, Barcelona, September 2006. ISBN 1-4244-0344-8, September 2006.

• M. Pericás, R. González, A. Cristal, D. Jiménez and M. Valero. "A Decoupled Kilo-instruction Processor". IEEE HPCA, International Conference on High Performance Computer Architecture, Austin, February 2006.

• T. Ramírez, M. Pajuelo, O. Santana and M. Valero. "Kilo-instruction Processors, Runahead and Prefetching". ACM International Conference on Computing Frontiers, Ischia, May 2-5, 2006.

• T. Ramírez, M. Pajuelo, O. Santana and M. Valero. "A Simple Speculative Load Control Mechanism for Energy Saving". MEDEA Workshop: "MEmory performance: DEaling with Applications, systems and architecture", Held in conjuction with PACT 2006 in Seattle, USA, September 2006.

• A.M. Rosolen, S. Ordas, M. Vázquez and A.F. Frangi. "Numerical schemes for the simulation of three-dimensional cardiac electrical propagation in patient-specific ventricular geometries". Egmond aan Zee (The Netherlands), European Conference on Computational Fluid Dynamics - ECCOMAS CFD 2006, September 2006.

• F. Sánchez, E. Salami, A. Ramírez and M. Valero. "Performance Analysis of Sequence Alignment Applications". IISWC, IEEE Internacional Symposium on Workload Characterization, San José, USA, October 2006.

• O. Santana, A. Falcón, A. Ramírez and M. Valero. "Branch Predictor Guided Instruction Decoding". IEEE PACT Parallel Architectures and Compiler Techniques, Seatle, September 2006.

• Raül Sirvent, Rosa M. Badia, Natalia Currle-Linde, Michael Resch. "GRID superscalar and GriCoL: integrating different programming approaches". 2nd CoreGRID Integration Workshop. Integrated Research in Grid Computing. Krakow (Poland), October 2006.

• B. Slamat, D. Nicolaescu, A. Veidenbaum and M. Valero. "Fast Speculative Address generation and Way Caching for Reducing L1 data Cache Energy". IEEE ICCD Internation Conference on Computer Design, San Francisco, USA, October 2006.

• J. Vera, F. Cazorla, A. Pajuelo, O.J. Santana, E. Fernández and M. Valero. "A Novel Evaluation Methodology to Obtain Fair Measurements in Multithreaded Architectures". MoBS-2, Workshop on Modeling, Benchmarking and Simulation, In conjunction with ISCA. Boston, June 2006.

National Conferences

• J. Alonso, J. Guitart and J. Torres. "Scheduling Connections using System-Level Thread Priorities in Web Applications". XVII Jornadas de Paralelismo. Albacete, Spain, September 18-20, 2006.

• Rosa M. Badia, Pieter Bellens, Vasilis Dialinos, Jorge Ejarque, Jesús Labarta, Josep M. Pérez and Raül Sirvent. "The GRID superscalar project: Current status". XVII Jornadas de Paralelismo. Albacete (Spain), September 2006.

• C. Boneti, F. Cazorla, M. Valero. "Improving EDF for SMT processors". XVII Jornadas de Paralelismo, Albacete, Spain, September 2006.

• I. González, O.J. Santana, A. Pajuelo, M. Valero. "Implementando recuperaciones precisas en procesadores con consolidación fuera de orden". XVII Jornadas de Paralelismo, Albacete, Spain, September 2006.

• K. Kedzierski, F. Cazorla, M. Valero. "Analysis of multithreading capabilities of current high-performance processors". XVII Jornadas de Paralelismo, Albacete, Spain, September 2006.

• J. Vera, F.J. Cazorla, A. Pajuelo, O.J. Santana, E. Fernández, M. Valero. "Looking for novel ways to obtain fair measurements in multithreaded architectures". XVII Jornadas de Paralelismo, Albacete, Spain. September 2006.

3.2 Earth Sciences

Communications

• I. González, O. J. Santana, A. Pajuelo and M. Valero. "A First Glance at the Implementation of Precise Recoveries in Out-of-order Commit Processors". ACACES 2006, Poster Abstracts. Advanced Computer Architecture and Compilation for Embedded Systems, Láquila, pp. 9-11. Academia Press, ISBN 90 382 0981 9 , July 2006.

• K. Kedziersky, F.J. Cazorla and M. Valero. "Analysis of Simultaneous Multithreading Implementations in Current High-Performance Processors". ACACES 2006, Poster Abstracts. Advanced Computer Architecture and Compilation for Embedded Systems, Láquila, pp. 113-116. Academia Press, ISBN 90 382 0981 9, July 2006.

• M. Moreto, A. Ramírez and M. Valero. "Reducing Simulation Time". ACACES 2006, Poster Abstracts. Advanced Computer Architecture and Compilation for Embedded Systems, Láquila, pp. 233-236. Academia Press, ISBN 90 382 0981 9, July 2006.

• M. Pericás, A. Cristal, R. González, F.J. Cazorla, D.A. Jiménez and M. Valero. "Boosting ILP&TLP with the Flexible Multi-Core (FMC)". Poster Abstracts. Advanced Computer Architecture and Compilation for Embedded Systems, ACACES 2006. Láquila pp. 125-128. Academia Press, ISBN 90 382 0981 9 , July 2006.

• M. Pericás, R. González, A. Cristal, and M. Valero "Overcoming the Memory Wall with D-KIPs". 1st International Summer School on Advanced Computer Architecture and Compilation for Embedded Systems (ACACES 2005), (Láquila, Italy). Academic Press, ISBN 90 382 0802 2. July 2006.

• T. Ramírez, M. Galluzzi, A. Cristal, and M. Valero "Different Approaches using Kilo Instruction Processors". 1st International Summer School on Advanced Computer Architecture and Compilation for Embedded Systems (ACACES 2005), (Láquila, Italy). Academic Press, ISBN 90 382 0802 2., July 2006.

• Mateo Valero. "Kilo-instruction Processors: Overcoming the Memory Wall". University of Irvine at California, February 2006.

• E. Vallejo, M. Galluzzi, A. Cristal, F. Vallejo, R. Beivide, P. Stenström, J.E. Smith and M. Valero. "Chip Multiprocessors with Implicit Transactions". ACACES 2006, Poster Abstracts. Advanced Computer Architecture and Compilation for Embedded Systems, Láquila, pp. 167-170. Academia Press, ISBN 90 382 0981 9 , July 2006.

Publications

Journals and Book Chapters

• Baldasano J.M. "Cambio climático y necesidades energéticas. Estrategía Mediterránea de Desarrollo Sotenible". Editor: Generalitat Valenciana: 61-81. 2006.

• Baldasano J.M., P. Jiménez P., O. Jorba, E. López, R. Parra and C. Pérez "Dynamics of air pollution in the North-Western Mediterranean Basin (Catalonia) in summer by using high-resolution air quality modeling". Contribution to Science 3 (2):167-180. 2006.

• Balis D., V. Amiridis, S. Kazadzis, A. Papayannis, G. Tsaknakis, S. Tzortzakis, N. Kalivitis, M. Vrekoussis, M. Kanakidou, N. Mihalopoulos, G. Chourdarkis, S. Nickovic, C. Pérez, J.M. Baldasano and M. Drakakis "Optical characteristics of desert dust over the East Mediterranean during summer: a case study". Annales Geophysicae, 24: 807-821. 2006.

• Güereca P., S. Gassó, J.M. Baldasano and P. Jiménez "Life cycle asessment of two biowaste management systems for Barcelona, Spain". Resources, Conservation and Recycling, 49: 32-48. 2006.

• Jiménez P., O. Jorba, R. Parra and J.M. Baldasano "Evaluation of MM5-EMICAT2000-CMAQ performance and sensitivity in very complex terrains: high-resolution application to the northeastern Iberian Peninsula". Atmospheric Environment, 40, 26: 5056-5072. 2006.

• Jiménez P., J. Lelieveld and J.M. Baldasano "Multi-scale Modelling of air pollutants Dynamics in the North-Western Mediterranean Basin during a typical summertime episode". Journal of Geophysical Research 111, D18306, doi:10.1029/2005JD006516. 2006.

• Parra R., P. Jiménez, J.M. Baldasano "Development of the high spatial resolution EMICAT2000 emission model for air pollutants from the north-eastern Iberian Peninsula of Catalonia, Spain". Environmental Pollution 140/2, 200-219, doi:10.1016/j.envpol.2005.07.021. 2006.

• Pérez, C., P. Jiménez, P. Jorba, M. Sicard and J.M. Baldasano. "Influence of the PBL Scheme on High Resolution Photochemical Simulations in an Urban Coastal Area over the Western Mediterranean". Atmospheric Environment 40, 27: 5274-5297. 2006.

• Pérez C., S. Nickovic, M. Sicard, F. Rocadenbosch, V. E. Cachorro, G. Sappe and J.M. Baldasano "A Long Saharan Dust over the Western Mediterranean: Lidar, Sun Photometer Observations and Regional Dust Modelling". Journal of Geophysical Research 111, D15214, doi:10.1029/2005JD006579. 2006.

• C. Pérez, S. Nickovic, G. Pejanovic, J.M. Baldasano and E. Ozsoy "Interactive Dust-radiation Modeling: A Step to improve Weather Forecast". Journal of Geophysical Research 111, D16206, doi:10.1029/2005JD006717. 2006.

• J.L. Rubio, F. Garcia Novo, J.M. Baldasano, R. Martín Mateo "Estrategia mediterrània de desarrollo sostenible". Generalitat Valenciana (3/2006) – ISBN 84-7579-844-6, Valencia, pp 213.

• Sicard M., C. Pérez, F. Rocadenbosch, J. M. Baldasano, D. García-Vizcaino "Mixed-Layer Depth Determination in the Barcelona Coastal Area From Regular Lidar Measurements: Methods, Results and Limitations", Boundary-Layer Meteorology, 119, 1 (Apr): 135 - 157, DOI 10.1007/s10546-005-9005-9. 2006.

International Conferences and Symposiums
• Baldasano J.M., Jiménez P., Jorba O., Parra R., Arévalo G., López E., Pérez C., Gassó S., Gonçalves M., Gracia P., Benoiston E., Monroy O., Passola O., "High-resolution air quality modeling in the Iberian Peninsula: applications and developments in a supercomputing framework". Advanced Atmospheric Pollution Modelling Workshop, Hatfield, UK. March 22, 2006.

• Jiménez P., Jorba O., Baldasano J.M. "Modeling the dynamics of air pollutants over the Iberian Peninsula under typical meteorological situations". 28th NATO/CCMS International Technical Meeting on Air Pollution Modeling and Its Application, Leipzig, Germany, pp. 282-290. May 15-19, 2006.

• Jiménez P., Jorba O., Pérez C., Baldasano J.M. "High-resolution modeling of atmospheric pollution by aerosols in the Iberian Peninsula: annual cycle simulations". Seventh International Aerosol Conference (IAC), St. Paul, Minnesota, USA. September 10-15, 2006.

• Jiménez P., Jorba O., Pérez C., Baldasano J.M. "High-resolution annual-cycle simulations of aerosols in the Western Mediterranean in a supercomputing framework". Global Environmental Change: Regional Challenges. Earth System Science Partnership (ESSP) Open Science Conference. World Meteorological Organisation (World Climate Research Programme) and China Meteorological Administration, Beijing, China. November 9-12, 2006.

• Jiménez-Guerrero P., Jorba O., Pérez C., Baldasano J.M. "Annual high-resolution simulations and evaluation of particulate matter in the Iberian Peninsula in a supercomputing framework". ACCENT/GLOREAM Workshop. Paris, October 11-13, 2006.

• Kishcha P., Alpert P., Shtivelman A., Krichak S.O., Joseph J.H., Kallos G., Katsafados P., Spyrou C., Gobbi G.P., Barnaba F.,

Nickovic S., Pérez C. and Baldasano J.M. "Forecast errors in dust vertical distributions over Rome". 28th NATO/CCMS International Technical Meeting on Air Pollution Modelling and its Application. Leipzig, Germany. May 15-19, 2006.

• Pérez C., Nickovic S., Pejanovic G., Özsoy E., and Baldasano J.M., "Impact of Saharan dust on numerical weather forecasts". European Geosciences Union EGU 2006. Vienna, Austria. April 2006.

• Sicard M., F. Rocadenbosch, A. Hénon, C. Pérez, A. Rodríguez, C. Muñoz, D. García Vizcaino, A. Comerón, J.M. Baldasano "Evidence of discrepancies between columnar-averaged Lidar ratios measured by Sun Photometer and Lidar by means of a Raman Lidar in Barcelona". 23rd International Laser Radar Conference (ilrc2006) Nara, Japan, July 24-28, 2006.

National Conferences
• Baldasano J.M. "L'Entrada en vigor del protocol de Kyoto: prespectives espanyola i catalana". Aula d'Ecologia 2005. Editor J. Terrades. Ajt. de Barcelona-UAB. 33-40. 2006.

• Baldasano J.M. "Inventario de emisiones de España para modelizar la calidad del aire con un alto nivel de desagregación". V Seminario Español de Calidad del Aire, Santander, October 16-18, 2006.

• Caballero C., S. Gassó, J.M. Baldasano y P. Güereca "Una propuesta de indicadores genéricos para la evaluación ambiental estratégica". X Internacional Congreso on Project Engineering, Valencia (España), September 13-15, 2006.

• Jimenéz-Guerrero P., O. Jorba, C. Pérez and J.M. Baldasano "Sistema de Pronóstico de la Calidad del Aire en el Barcelona Supercomputing Center-Centro Nacional de Supercomputación". V Seminario Español de Calidad del Aire, Santander, October 16-18, 2006.

• Jiménez-Guerrero P., O. Jorba, C. Pérez y J.M. Baldasano "Sistema de pronóstico de la calidad del aire en el BSC-CNS". 6° Congreso Internacional de Química de la Anque: "Química y Desarrollo Sostenible", Puerto de la Cruz, Tenerife (España), December 5-7, 2006.

• Jorba O., P. Jiménez and J.M. Baldasano "Simulación meteorológica anual con elevada resolución espacial para aplicaciones de calidad del aire". XXIX Jornadas Científicas de la Asociación Meteorológica Española: Aplicaciones Meteorológicas-7° Encuentro Hispano-Luso de Meteorología: Meteorología y Eficiencia Energética. Pamplona (España), April 24-26, 2006.

• Jorba O., Jiménez P., Pérez C., Baldasano J.M. "Requerimientos que deben exigirse a los modelos de calidad

del aire para cada problema o zona de España". V Seminario Español de Calidad del Aire, Santander, October 16-18, 2006.

• Martin F. et al. "Grupo de Trabajo sobre Modelización de la Calidad del Aire: Resumen y conclusiones más relevantes". V Seminario Español de Calidad del Aire, Santander, October 16-18, 2006.

• Pérez C., O. Jorba, P. Jimenéz-Guerrero and J.M. Baldasano "Sistema de Pronostico de Eventos de Polvo Sahariano en el Mediterráneo y las Islas Canarias en el Barcelona Supercomputing Center-Centro Nacional de Supercomputación". V Seminario Español de Calidad del Aire, Santander, October 16-18, 2006.

Communications

• Baldasano J.M. "Modelos para entender la Calidad del Aire". CosmoCaixa, Barcelona March 2, 2006.

• Baldasano J.M. "La Co-incineración de Residuos. Ventajas y Limitaciones". Divulgação Pública do Relatório de Actualização dos Processos de Co-incineração de Resíduos em Articulação com os CIRVER, Ministério do Ambiente, do Ordenamento do Território e do Desenvolvimento Regional, Oporto (Portugal) March 3, 2006.

• Baldasano J.M. "El Protocolo de Kioto y su influencia en España". Jornadas "Aspectos Jurídicos del Medio Ambiente". Univ. de Burgos, Fac. de Derecho, Burgos March 22, 2006.

• Baldasano J.M. "El Cambio Climático. Evidencias Científicas". II Curso de Ciencia y Sociedad. Univ. de Alicante. April 3, 2006.

• Baldasano J.M. "El Cambio Climático. Importancia y Evidencias de su Realidad". Encuentro del Medio Ambiente y de la Coorperación: Realidades del Cambio Climático, XIV Edición. MEDITERRÁNEA-Univ. Rovira i Virgili, April 6, 2006.

• Baldasano J.M. "Impact of the Kyoto protocol on the electricity suply industry in Europe: the Spanish example". KAM-Together Meeting, ATEL. Yverdon-les Bains (Switzerland), May 16, 2006.

• Baldasano J.M. "Modelización del ciclo atmosférico del polvo mineral desértico y su influencia en la meteorología y la calidad del aire". Campus de Excelencia 2006. Fuerteventura (Spain), May 24, 2006.

• Baldasano J.M. "Un Model Energètic Insostenible II: Les Emissions de Diòxid de Carboni (CO2) I EL Canvi Climàtic". Energia i Sostenibilitat, Universidad de Barcelona. Barcelona, July 3, 2006.

• Baldasano J.M. "Revisión histórica de la legislación sobre Residuos Industriales. Gestión de suelos contaminados y residuos industriales. situación actual en el campo de Gibraltar", XXVII Edición de los Cursos de Verano de la Universidad de Cádiz en San Roque, dtor. J.M. Quiroga, July 14, 2006.

• Baldasano J.M. "Contaminación atmosférica: fuentes y contaminantes, evaluación de la calidad de aire, modelización de la dispersión atmosférica". Presente y Futuro de la Ingeniería Ambiental. Universidad de Oviedo, dto. M.M. Mahamud, July 17, 2006.

• Baldasano J.M. "Una visión sistémica global del futuro del mundo en que vivimos ¿deterioros irreversibles?" Curso de Verano, Universidad Complutense de Madrid (UCM): "El horizonte posible del 2100: la difícil perpetuación de la especie humana", dtor. R. Tamames. El Escorial August 7, 2006.

• Baldasano J.M. "Una visión sistémica global del futuro del mundo en que vivimos ¿deterioros irreversibles?". Curso de Verano, Universidad Complutense de Madrid (UCM): El horizonte posible del 2100: la difícil perpetuación de la especie humana, dtor. R. Tamames. El Escorial, August 7, 2006.

• Baldasano J.M. "De los grandes depósitos en masa a los depósitos de balas plastificadas." Gestión de Residuos Municipales en Catalunya. Barcelona, September 27-28, 2006.

• Baldasano J.M., Pérez, C., "Atmospheric mineral dust forecast modeling system in the Barcelona Supercomputing Center: current and future activities". WMO Sand and Dust Store Meeting, Shanghai (China), 30 October–1 November, 2006.

• Baldasano J.M. "Evolució en l'Ús de l'Energia a Catalunya". Jornada sobre Canvi Global: Efectes i Estratègies de Gestió. IEC (Institut d'Estudis Catalans-IEC). Barcelona, November 30, 2006.

• Jorba, O. "Descripción del Área de Ciencias de la Tierra del Barcelona Supercomputing Center". Seminario FIB, February, 2006.

• Pérez, C. "Modelización del ciclo atmosférico de polvo mineral desértico y su influencia en la meteorología y la calidad del aire". Seminario UB, April, 2006.

• Pérez C., S. Nickovic, G. Pejanovic, E. Özsoy, J.M. Baldasano "Interactive dust-radiation modeling: a step to improve weather forecasts". Earlinet Workshop, Granada (Spain), October 9-11, 2006.

3.3 Life Sciences

Publications

Journals and Book Chapters

• Ahmed Altun, Victor Guallar, Richard Friesner RA, Sason Shaik and Walter Thiel. "The effect of heme environment on the hydrogen abstraction reaction of camphor in P450(cam) catalysis: A QM/MM study". Journal of the American Chemical Society, 128: 3924-3925, 2006.

• Bidon-Chanal A., Martí M.A., Crepo A., Milani M., Orozco M., Bolognesi M., Luque F.J. and Estrin D.A. "Ligand-Induced Dynamical Regulation of NO Conversion in Micobacterium tubercolosis Truncated Hemoglobin-N" Proteins-Structure Function and Genetics, 64: 457-465. 2006.

• J.R. Blas, J.M. López-Bes, M. Márquez, J.L. Sessler, F.J. Luque and M. Orozco "Exploring the dynamics of Calix[4]pyrrole: The effect of solvent and fluorine substitution". Chemistry an European Journal. In Press. 2006.

• Blundell T.L., Fernandez-Recio J. "Brief encounters bolster contacts". Nature (in press, News & Views). 2006.

• Bolanos-Garcia V., Fernandez-Recio J., Allende J.E., Blundell T.L. "Identifying interaction motifs in CK2beta - a ubiquitous kinase regulatory subunit?" Trends Biochem. Sci. In press. 2006.

• O. Carrillo and M. Orozco. "GRID-based molecular dynamics. A Tool for genomic-scale simulations of channels in proteins". Proteins-Structure Function and Genetics. In Press 2006.

• Casal E., Federici L., Zhang W., Fernandez-Recio J., Priego E.M., Miguel R.N., DuHadaway J.B., Prendergast G.C., Luisi B.F., Laue E.D. "The Crystal Structure of the BAR Domain from Human Bin1/Amphiphysin II an Its Implications for Molecular Recognition". Biochemistry. In press. 2006.

• Chelliah V., Blundell T.L., Fernandez-Recio J. "Efficient Restraints for Protein-Protein Docking by Comparison of Observed Amino Acid Substitution Patterns with those Predicted from Local Environment". J.Mol.Biol. 357, 1669-1682. 2006.

• Cubero E., Luque F.J. and Orozco M. "Theoretical study of the Hoogsteen<->Watson-Crick junctions in DNA"'. Biophysical Journal, 90: 1000-1008. 2006.

• Curutchet C., Orozco M., Luque F.J., Mennucci B., Tomasi J. "Dispersion and repulsion contributions to the solvation free energy: Comparison of quantum mechanical and classical approaches in the polarizable continuum model" Journal of Computational Chemistry, 27: 1769-1780. 2006.

• Federici L., Di Matteo A., Fernandez-Recio J., Tsernoglou D., Cervone F. "Polygalacturonase inhibiting proteins: players in plant innate immunity?" Trends Plant Sci. 11, 65-70. 2006.

• C. Ferrer-Costa, M. Orozco and X. de la Cruz "Characterization of compensates mutations in terms of structural and physico-chemical properties" Journal of Molecular Biology, 365, 249-256. 2006.

• Fuentes-Cabrera M., Lipkowski P., Huertas O., Sumpter B.G., Orozco M., Luque F.J., Wells J.C., Leszczynski J. "Aromaticy induced changes in the electronic properties of size-expanded DNA bases. The case of xC". International Journal of Quantum Chemistry, 106: 2339-2346. 2006.

• Goñi J.R., Vaquerizas J.M., Dopazo J., Orozco M. "Exploring the reasons for the large density of triplex-forming oligonucleotide target sequences in the human genome". BMC Genomics, 7: 63-72. 2006.

• Victor Guallar, Andrjez Jarzecki, Richard A. Friesner, and Thomas H Spiro. "Modeling of ligation-induced helix/loop displacements in myoglobin: Toward an understanding of hemoglobin Allostery". Journal of the American Chemical Society, 128: 5427-5435, 2006.

• Victor Guallar and Brett Olsen. "The role of the heme propionates in heme biochemistry". Journal of Inorganic Biochemistry 100: 755-760, 2006.

• O. Huertas, J. Poater, M. Fuentes-Cabrera, M. Orozco, M. Sola and F.J. Luque. "Local aromaticity in natural nucleobases and their size-expanded benzo-fused derivatives" Journal of Physical Chemistry B. In Press. 2006.

• López de la Osa J., González C., Gargallo R., Rueda M., Cubero E., Orozco M., Aviñó A., Eritja R. "Destabilization of quadruplex DNA by 8-aminoguanine" ChemBioChem , 7: 46-48. 2006.

• Meyer T., Ferrer-Costa C., Pérez A., Rueda M., Bidon-Chanal A., Luque F.J., Laughton C.A. and Orozco M. "Essential Dynamics: A Tool for Efficient Trajectory Comprenssion and Management". Journal of Chemical Theory and Computation ,2: 251-258. 2006.

• J. Muñoz-Muriedas, X. Barril, J.M. López, M. Orozco and F.J. Luque "A hydrophobic similarity analysis of solvation effects on nucleic acids" Journal of Molecular Modelling. In Press. 2006.

• Noy A., Meyer T., Rueda M., Ferrer C., Valencia C., Pérez A., De la Cruz X., López-Bes J.M., Luque F.J., Orozco M. "Datamining of molecular dynamics trajectories of nucleic

acids" Journal of Biomolecular Structure & Dynamics ,23: 447-455. 2006.

• Noy A., Meyer T., Rueda M., Ferrer C., Valencia A., Perez A., de la Cruz X., Lopez-Bes J.M., Pouplana R., Fernandez-Recio J., Luque F.J., Orozco M. "Data mining of molecular dynamics trajectories of nucleic acids". J Biomol Struct Dyn. 23, 447-456. 2006.

• A. Noy, A. Pérez, C. Laughton and M. Orozco. "Theoretical study of large conformational transitions in DNA: The B??A conformational change in water and ethanol/water". Nucleic Acid Res. In Press. 2006.

• M. Orozco, A. Noy, T. Meyer, M. Rueda, C. Ferrer, A. Valencia, A. Pérez, O. Carrillo, J. Fernández-Recio, X. de la Cruz and F.J. Luque. "Data mining of molecular dynamic trajectories of nucleic acids". In "Computational Studies of RNA and DNA". Challenges and Advances in Computational Chemistry and Physics. Vol 2. J.Sponer and F.Lankas (Eds). Springer. Dordrecht. pp 121-146. 2006.

• A. Pérez, I. Marchán, D. Svozil, J. Sponer, T.E. Cheatham, C.A. Laughton and M. Orozco. "Refinement of the AMBER force-field for nucleic acid simulations. Improving the representation of ·/Á conformations". Biophys J. In Press. 2006.

• M. Rueda, C. Ferrer, T. Meyer, A. Pérez, J. Camps, A. Hospital, J.L. Gelpí and M. Orozco. "A consensus view to Protein Dynamics". Proc. Natl. Acad. Sci. USA. In Press. 2006.

• Rueda M., Luque F.J., Orozco M. "G-DNA can maintain its structure in the gas phase". Journal of the American Chemical Society, 128: 3608-3619. 2006.

• Small YA, Guallar V, Soudackov AV, et al. "Hydrogen bonding pathways in human dihydroorotate dehydrogenase" Journal of physical chemistry B, 110: 19704-19710, 2006.

• R. Soliva, J.L. Gelpi, C. Almansa, M. Virgili and M. Orozco. "Dissection of the recognition properties of p38 MAP Kinase. Determination of the binding mode of a new pyridinyl-heterocycle inhibitor family". Journal of Medicinal Chemistry. In Press. 2006.

• Soteras I., Lozano O, Gómez-Esqué A., Escolano C., Orozco M., Amat M., Bosch J., Luque F.J. "On the origin of stereoselectivity in the alkylation of oxazolopiperidone enolates" Journal of the American Chemical Society ,128: 6581-6588. 2006.

• Sponer J., Jurecka P., Marchan I., Luque F.J., Orozco M., Hobza P. "Nature of base stacking. Reference quantum chemical stacking energies in ten unique B-DNA base pairs steps".Chemistry-A European Journal, 12: 2854-2865. 2006.

• Suyama M., Harrington E., Bork P., Torrents D. "Identification and analysis of genes and pseudogenes within duplicated regions in the human and mouse genomes". PLoS Comput Biol. Jun 30;2(6):e76. Epub 2006 May 16, 2006.

• Suyama M., Torrents D., Bork P. "PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments". Nucleic Acids Res, 34(Web Server issue):W609-12, July 2006.

• D. Talavera, A. Morreale, A. Hospital, C. Ferrer, J.L. Gelpi, X. de la Cruz, T. Meyer, R. Soliva, F.J. Luque and M. Orozco. "A fast method for the determination of fractional contributions to solvation in proteins". Protein Science. 15, 2525-2533. 2006.

• D. Talavera, C. Vogel, M. Orozco, S.A. Teichmann and X. de la Cruz. "The (In)dependence ofalternative splicing and gene duplication". PLOS Comp.Biol. In Press. 2006.

• Yraola F., Garcia-Vicente S., Fernandez-Recio J., Albericio F., Zorzano A., Marti L., Royo M. "New Efficient Substrates for Semicarbazide-Sensitive Amine Oxidase/VAP-1 Enzyme: Analysis by SARs and Computational Docking". J.Med.Chem 49, 6197-6208. 2006.

Nexus II building.
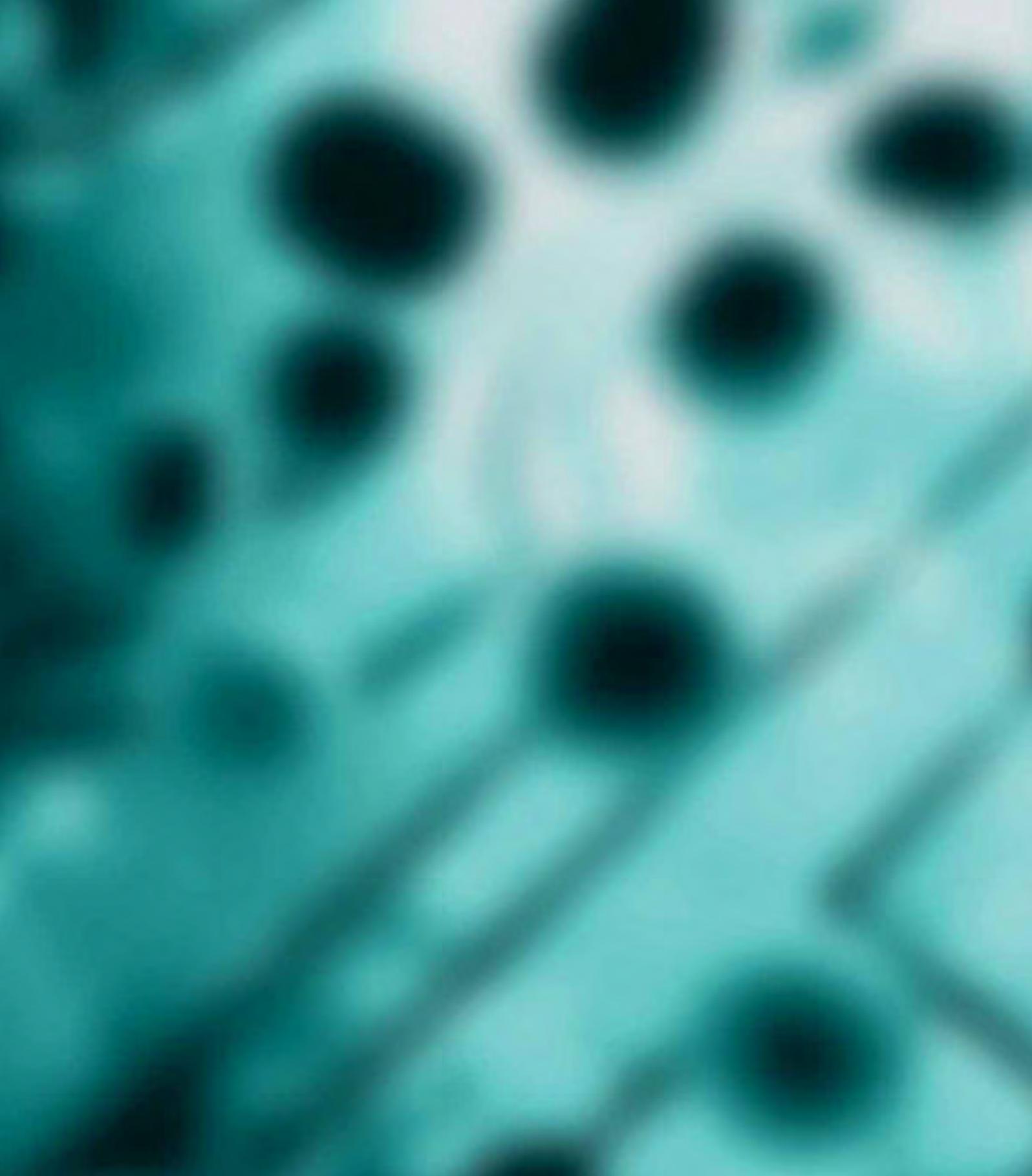
Support

Set of support activities to the scientific community and industry.

4.1 Business Development

In 2006, the Business Development Team worked on the promotion of the center as well as on the management of proposals projects funded by public calls as well as by sponsored research contracts.

Felipe Lozano
Business Development Head

A detailed description of the activities performed by this department during 2006 is structured into the following three sections: project management, technology transfer and corporate image.

Project Management Activities

In 2006, BSC-CNS participated in an increasing number of funded projects and grants. As a consequence, the project management activities also intensified.

International Activities

At an international level, the Barcelona Supercomputing Center has already participated in an impressive number of activities. In 2006, BSC-CNS was active in the Sixth Framework Programme of the European Commission and participated in several activities to prepare for the transition to the Seventh Framework Programme, launched in December 2006.

In 2006, BSC-CNS participated in eight FP6 projects: DEISA, eDEISA, HPC-EUROPA, SARC, BREIN, BEinGRID, XtreemOS and SORMA. DEISA and eDEISA are infrastructure projects. HPC-EUROPA includes a transnational access activity as well as research and development. SARC, BREIN, BEinGRID, XtreemOS and SORMA are primarily R&D projects.

BSC-CNS began its participation in the DEISA and HPC-EUROPA Projects prior to 2006 and began participating in the eDEISA project in 2006 to complement some of the activities of the DEISA project. The SARC project was transferred from UPC and started in January 2006.

During the first half of 2006, proposals for BREIN, BEinGRID, XtreemOS, SORMA projects that had passed the threshold in the last call of the Information and Communication Technology program entered into the negotiation stage and soon after became active projects. These projects helped the BSC-CNS achieve a ranking as the eighth Spanish institution in terms of funding received in the last call of the Sixth Framework Programme; BSC-CNS was only surpassed by larger institutions such as Telefónica I+D, ATOS Origin, Technical University of Catalonia (UPC), Technical University of Madrid (UPM) and the Technical University of Valencia (UPV). BSC-CNS has achieved particularly good results in the specific objective related to grid technologies; it ranked among the top 3 Spanish institutions in terms of funding achieved.

The following is a brief description of the projects and activities in which BSC-CNS participated in 2006. Some of the more technical aspects of projects SARC, BREIN, BEinGRID, SORMA, XtreemOS and HPC-EUROPA are described in the section devoted to Research Activities.

DEISA Project / eDEISA

DEISA is a consortium of leading national supercomputing centres in Europe that are coordinating their actions to jointly build and operate a distributed terascale supercomputing facility.

BSC-CNS joined the DEISA Project (www.deisa.org) in May 2005. The project has secured funds until May 2008 and has received funds for extended activities through eDEISA (extended DEISA).

During 2006, the BSC-CNS participated in the following activities related to the DEISA and eDEISA Projects: management through participation on the Executive Committee, dissemination, service activities and life science joint research activities.

In addition, four proposals were submitted to the DEISA Extreme Computing Initiative (DECI) from Spain (the total number of proposals submitted by the 12 partners was 41) of which one was selected for execution.

HPC-Europa Project

HPC-Europa is a consortium of six leading High Performance Computing (HPC) infrastructures, including BSC-CNS, and five centers of excellence. This consortium aims to provide advanced computational services in an integrated manner to the European research community. The culture of co-operation fostered by the consortium's joint research and networking activities generates critical mass for computational science. It enables researchers working in any eligible country in Europe to visit a participating research institute to carry out a collaborative visit of up to 3 months' duration and to gain access to some of the most powerful High Performance Computing (HPC) facilities in Europe.

In 2006, BSC-CNS welcomed 35 researchers through the European Commission's HPC-Europe Transnational Access program. These researchers visited and used BSC-CNS facilities. The following graphs offer additional information about the distribution of visitors for 2006, for the whole Consortium and for each HPC Center:

PER TA SITE

Physics
Mathematics
Life sciences
Human Sciences
Engineering
Earth sciences
Computing
Chemistry
Astrophysics

FOR CONSORTIUM

- Astrophysics 0%
- Chemistry 25%
- Computing 0%
- Earth sciences 2%
- Physics 32%
- Engineering 22%
- Mathematics 18%
- Human Sciences 0%
- Life sciences 1%

PER TA SITE

- Undergraduate
- Postgraduate (student with first University degree or equivalent)
- Postdoctoral researcher
- Experienced researcher (professional researcher, lecturer or higher)

FOR CONSORTIUM

- Undergraduate 3%
- Experienced researcher (professional researcher, lecturer or higher) 35%
- Postgraduate (student with first University degree or equivalent) 40%
- Postdoctoral researcher 25%

Generally speaking, all BSC-CNS visitors were satisfied with their visits both in terms of administrative as well as for technical issues.

HPC-Europa Transactional Access Meeting

BSC-CNS organized the third HPC-Europa Transactional Access Meeting (TAM) to take place on June 14-15, 2006. This annual "in person" meeting of the Scientific Users' Selection Panel (SUSP) was held in the afternoon of June 15

in order to allow the panel members to attend TAM. The TAM started on June 14th with keynote speeches by Professor Mateo Valero, Director of BSC-CNS, and Professor Carme Rovira, of the Parc Cientific de Barcelona.

During the meeting, 43 users presented their work which they had undertaken during their HPC-Europa visit. Twenty eight users presented talks in themed parallel sessions, while 15 presented posters (including one joint poster presented by 2 visitors to EPCC). In addition to the presentations, there was a panel session allowing visitors to share their experiences and give feedback to the HPC-Europa consortium and make suggestions for improvements to the program. This was followed by technology demonstration sessions showcasing the current results of NA2, NA3, JRA1 and JRA2 and emphasizing their relevance to the Transnational Access visitors. Finally, all users participated in a guided visit of MareNostrum. This opportunity was highly appreciated by everyone as most users do not usually get the chance to see the supercomputing facilities which they use. In total, 111 individuals attended the meeting of which 43 were HPC-Europa users whose attendance was funded by the program, with another 7 self-funded participants. The remaining attendees consisted of 41 members of the HPC-Europa consortium, 8 members of SUSP, 11 BSC-CNS staff members and local scientific hosts (including the 2 keynote speakers) as well as the EC Project Officer.

The following researchers joined the HPC program in 2006:

Viktorya Aviyente
Bogazici University (Turkey)

Heribert Reis
Nacional Hellenic Research Foundation (Greece)

Kamila Reblova
Academy of Sciences of the Czech Republic (Czech Republic)

Bernd Mohr
Forschungszentrum Jülich (Germany)

Patrizia Vignolo
Scuola Normale Superiore (Italy)

Massimo Coppola
CNR-Consiglio Nazionale delle Ricerche (Italy)

Andreas Glotis
NCSR "Demokritos" (Greece)

Adrian Florea
"Lucian Blaga" University of Sibiu (Romania)

Arpad Gellert
"Lucian Blaga" University of Sibiu (Romania)

Pietro Vidossich
SISSA (Italy)

Manuel Perucho
Max-Planck-Institut für Radioastronomie (Germany)

Rolf Oettking
Technische Universitaet Ilmenau (Germany)

Felix Lippold
University of Stuttgart (Germany)

Alberto Bertoldo
University of Padova (Italy)

Sinan Akpinar
Firat University (Turkey)

Ana Sofia Pinto
University of Porto (Portugal)

Dimitri Komatitsch
University of Pau (France)

Jacek Wojdel
Delft University of Technology (Netherlands)

Alexey Zaikin
Postdam University (Germany)

Jonas Tölke
Technical University of Braunschweig (Germany)

Nelido Gonzalez-Segredo
AMOLF (Netherlands)

Theo Ungerer
University of augsburg (Germany)

Jennifer Chan
Imperial Collage London (United Kingdom)

Francesco Ancilotto
University of Padova (Italy)

Monika Golebiowska
University of Gdansk (Poland)

Anna Sadowska
University of Gdansk (Poland)

Sofiane Benhamadouche
Electricité De France (EDF R&D) (France)

Noelia Faginas Lago
University of Perugia (Italy)

Jacek Stolarczyk
University College Dublin (Ireland)

Nadine Utz
University of Freiburg (Germany)

Gian Pietro Miscione
Universita' degli Studi di Bologna (Italy)

Tomek Rog
Helsinki University of Technology (Finland)

HPC Europa publications

The following is a list of publications of the HPC-Europa visitors to the BSC-CNS facilities:

• L. Vintan, A. Gellert, Adrian Florea, M. Oancea, "Understanding Prediction Limits through Unbiased Branches", Lecture Notes in Computer Science, vol. 4186-0480, pp. 483-489, Springer-Verlag, ISSN 0302-9743, Berlin Heidelberg, 2006.

• M. Oancea, A. Gellert, Adrian Florea, L. Vintan, "Analyzing Branch Prediction Contexts Influence", Advanced Computer Architecture and Compilation for Embedded Systems (ACACES 2006), ISBN 90 382 0981 9, L'Aquila, Italy, 2006.

• Adrian Florea, A. Gellert, "Memory Wall – A Critical Factor in Current High-Performance Microprocessors", (Florea Adrian final report) - to appear in HPC-Europa Science and Supercomputing report 2006.

• A. Gellert, Adrian Florea, "Finding and Solving Difficult Predictable Branches", (Gellert Arpad final report) - to apear in HPC-Europa Science and Supercomputing report 2006.

• Perucho, M., Martí, J.Mª, Hanasz, M., "Nonlinear stability of relativistic sheared planar jets." Astronomy and Astrophysics, 443 863-881 (2005).

• Perucho, M., Lobanov, A., Martí, J.Mª, P.E. Hardee. "The role of Kelvin-Helmholtz instability in the internal structure of relativistic outflows. The case of the jet in 3C 273". Astronomy and Astrophysics, 456, 493-504 (2006).

• A. Bertoldo. "FEMS: An Adaptive Finite Element Solver". International Workshop on Parallel and Distributed Scientific and Engineering Computing (PDSEC) at the International Parallel and Distributed Processing Symposium (IPDPS), Long Beach, California USA, 2007. To be published.

• A. Bertoldo. "An Adaptive Parallel Solver for Finite-Element Applications". Ph.D. Thesis, University of Padova, Italy, December 2006.

Initiatives towards the Seventh Framework Programme of the European Commission

BSC-CNS joined several initiatives to smooth the path toward FP7. The most relevant initiatives were the HPC European Taskforce and the technology platforms NESSI and INES.

HPC European Taskforce

The HPC European Taskforce initiative (HET) deserves special mention for its efforts to develop a European strategic

approach to high-performance computing, concentrating the resources in a limited number of top-tier world centers in an overall infrastructure connected with associated national, regional and local centers to form a scientific computing network to utilize the top-level machines. This overall architecture will respond both to Capability (high-performance) and Capacity Computing (high-throughput) needs. Different machine architectures will fulfill the requirements of different scientific domains and applications. The structure can be represented as a pyramid, where local centers constitute the base of the pyramid, national and regional centers constitute the middle layer and the high-end HPC centers constitute the top layer.

NESSI and INES Technology Platforms

The NESSI platform (www.nessi-europe.com) aims to provide a unified view for European research in Services Architectures and Software Infrastructures that will define technologies, strategies and deployment policies fostering new, open, industrial solutions and societal applications that enhance the safety, security and well-being of citizens. BSC-CNS participates in this platform as a member.

The INES Technology Platform is Spain's NESSI counterpart which hopes to produce a strategic research agenda for the software and services field in Spain, and therefore increase the competitiveness of the Spanish industry. BSC-CNS participates in this platform as a member of the Steering Committee.

National Activities

BSC-CNS also participated in a significant number of national and regional calls. In 2006, BSC-CNS was actively involved in six R&D projects of the Spanish National Plan: two in Life Sciences, two in Earth Sciences and one in Computer Sciences. The center received three grants from the Spanish Ministry of Science and Education to participate in international activities (technology platforms and other initiatives): Networked European Software and Services (NESSI), Global Monitoring for Environment and Security (GMES) and HPC European Taskforce. In addition, the center received three grants from the Spanish Ministry of Science and Education (MEC) to improve its infrastructure with respect to: 1) increased storage capacity, 2) fire protection and 3) expansion of the Spanish Supercomputing Network plan. Moreover, BSC-CNS received from this ministry matching funds to participate in seven of the European projects where BSC-CNS is a partner. The center also hosted two *Ramón y Cajal* researchers. Finally, BSC-CNS coordinated a project funded by the Spanish Ministry of the Environment to monitor the air quality in the Iberian Peninsula (CALIOPE).

Since the start of 2006, BSC-CNS ran a national access program called ICTS, with the objective to take advantage of experience of the knowledge in supercomputing and eScience from BSC-CNS. This program allowed 10 researchers to access

BSC-CNS facilities during 2006; their expenses were covered by the Spanish Ministry of Science and Education. Selection meetings are held every four months.

The following researchers joined the ICTS program during the year 2006:

Ramón Beivide Palacio
Universidad de Cantabria

Eduardo W. Vieira Chaves
Universidad de Castilla-La Mancha

Carlos Quijano San Martín
Universidad Autónoma de Madrid

José Miguel Alonso
Universidad del País Vasco

Pere Miró Ramírez
Institut Català Investigació Química

Rubén Darío Costa Riquelme
Universidad de Valencia

Sebastián Reyes Ávila
Universidad de Castilla-La Mancha

Rubén Muñoz González
Universidad Autónoma de Madrid

Carlos Marrero
Instituto Nacional de Meteorología

Ismael Marín Carrión
Universidad de Castilla-La Mancha

Technology Transfer

In 2006, BSC-CNS continued collaborating with technology many of the transfer projects that had started 2005. It also began work on new projects with such industry partners as IBM, Microsoft, Repsol and Desafío Español Copa América 2007.

Several projects in Computer Sciences continued the research lines that UPC and IBM started 5 years ago through their joint initiative, CEPBA-IBM Research Institute (CIRI). In 2006, the IBM Innovation Initiative at the Barcelona Supercomputing Center (I3@BSC) was established as the first European center for open Linux application scaling. This collaboration combines the power of advanced research and development with the practicality of commercial implementation to solve some of the industry's most challenging problems.

During 2006, I3@BSC conducted basic research in the areas of high performance computing and associated computer architectures, including proof of concept testing for system level and applications software on a large scale cluster system such as MareNostrum.

The I3@BSC collaboration enabled and optimized scientific computing libraries for the Power Architecture and Vector Multimedia Extensions (VMX). The collaboration also resulted in efficient scalable performance analysis tools for large scale supercomputers that were integrated the BSC Performance Toolkit with the IBM High Performance Toolkit in order to accelerate application development and performance analysis. Following research work initiated in 2005, I3@BSC increased its focus on emerging heterogeneous multicore architectures like the Cell Broadband Engine. This includes the development and distribution of critical opensource software components like the Linux Operating System, Software Development Kit (SDK), compiler and toolchain for Cell BE (see: www.bsc.es/projects/deepcomputing/linuxoncell/). Building on this operating system and middleware work, I3@BSC started a new project to develop efficient fine-grain programming models like OpenMP and Cell Superscalar for systems based on the Cell BE processor. The main objective of this research area is to hide complex architectural features from the programmer, while at the same time enabling the generation of portable efficient code.

The aim of the project is to cooperate with Bull in its Lustre technology, where BSC-CNS offers its expertise on storage systems.

BSC-CNS initiated a project with Gas Natural that consists of studies of the impact of introducing natural gas vehicles on the air quality for cities Barcelona and Madrid.

BSC-CNS started working with Airbus to improve its eLSA code (fluid dynamics) which was initially developed by ONERA in France and is currently used worldwide by Airbus. The design of the next generation chip multiprocessor will clearly impact how computers look ten years from now and beyond. BSC-CNS worked and continues to work with Microsoft to explore potential chip designs for general purpose and the personal mass computer market. At this time, the project focuses primarily on architectural ideas, with an understanding of the importance of total system integration with software.

The BSC-CNS began working with Repsol YPF in the Kaleidoscope Project. The focus of this project is to optimize and execute in MareNostrum, as well as in cell processors, the Reverse Time Migration (RTM) codes used in next generation seismic imaging technology. It is expected that these codes will accelerate and streamline oil and gas exploration by several orders of magnitude compared to current industry standards.

The BSC-CNS continued its collaboration with Desafío Español Copa América 2007 to develop advanced software algorithms that improve the design of competition sailboats.
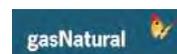
In addition, the BSC-CNS will help to obtain a meteorological model based in a neuronal network that will allow a real-time forecasting.

The objective of the INYPSA S.A. agreement was to perform an impact assessment study on the air quality from the combined cycle power station in Arrúbal (Logroño).

The objective of the SGS TECNOS S.L. agreement was to perform an impact assessment on the air quality study of the from combined cycle power stations

The objective of the BCNEcologia agreement was to perform a simulation for the Air Quality Plan of Catalonia.

Memorandums of Understanding

During 2006 the BSC-CNS signed various Memorandums of Understanding:

- CSC (Finnish Supercomputing Centre). The objective of this agreement was for both centers to act as a hub in their respective countries in order to bring the respective scientific communities together.

- INM (Instituto Nacional de Meteorología). The objective of this agreement was to implement, disseminate and validate the operational prediction of the northafrican dust transport in the Iberian Peninsula as well as to perform modeling, detection, follow-up and characterization studies of atmospheric material.

Corporate Image

In addition to functioning as a tool for scientific research, MareNostrum also served as a tool for promoting science in society.

Over the course of 2006, BSC-CNS received 2,500 visitors from national and international centers, including universities, research centers, industry and non-profit organizations.

MARENOSTRUM VISITS

Individual visits 5%
Associations 4%
Congresses 10%
Companies 25%
Universities 56%
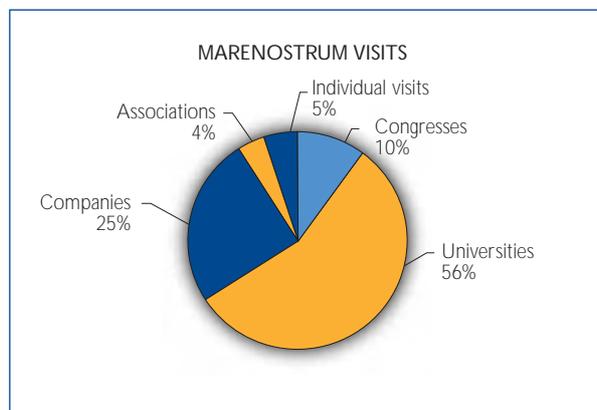
Conferences

BSC-CNS members participated in the following conferences and courses in order to disseminate BSC-CNS activities:

International

- HPC Europa. 3rd Transnational Access Meeting (TAM), HLRS, University of Stuttgart (Germany)
- HPC Europe, 2nd Workshop on High Performance Computing, Oxford, UK, September 24th-27th 2006
- Jornadas de Bioinformática, Zaragoza, November 2006.
- ISC 2006, Dresden, Germany, June 2006.
- SC2006, Tampa, November, USA, November 2006.
- "Kilo-instruction Processors: Overcoming the Memory Wall". Distinguished lecture. University of Irvine at California, February 19, 2006.

National

- Cercle de la FIB, Facultad de Informática de Barcelona. UPC, Barcelona, February 2006.
- Forum Tecnologías de la Información. UPC, Barcelona, March 2006.

BSC-CNS in the Media

BSC-CNS was mentioned in various media articles throughout the year. In 2006, BSC-CNS was mentioned 233 times in written media including national and international newspapers, newsgroups and magazines, and all articles and press releases are available on the BSC-CNS website www.bsc.es. The center was primarily mentioned in print; however, this was followed by TV as shown in the graph below:

MEDIA

TV 8%
Radio 4%
Print 88%
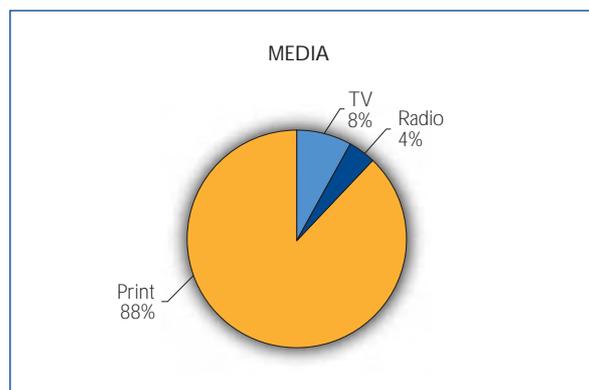
Events Organized by BSC-CNS

In 2006, BSC-CNS organized and attended the following congresses, seminars and workshops:

- TAM 2006, HPC Europa, June 2006.
- Seminar Automatic Differentiation and the Tapenade AD Tool: applications in CFD, Mauricio Araya - Projet Tropics - INRIA Sophia-Antipolis, BSC-CNS, March 22nd 2006.
- Cluster/Grid 2006, Barcelona, September 2006.

The Administration Department includes the following groups: Human Resources, Finance and Administration Support.

Mercè Calvet,
Human Resources, Finance
and Administration Head

The year 2006 was marked for the growth of the activities of the Support department. At the end of the year, the BSC-CNS staff counted with a total of 137 members. In terms of annual accounts, the total operating budget increased in 414 %.

Human Resources

In 2006, both the Scientific and Support staff structures grew significantly. This led to the consolidation of the six original departments into a new organization consisting of 15 research and 6 support groups, with a total of 137 staff members. The total headcount of the organization included 87 permanent staff members and 50 dedicated to specific projects. BSC-CNS initiated the recruitment process 11 times and conducted it based on the principles of ability, publicity and open competition.

PERSONNEL STRUCTURE

Support Staff 22%

R&D Staff 78%
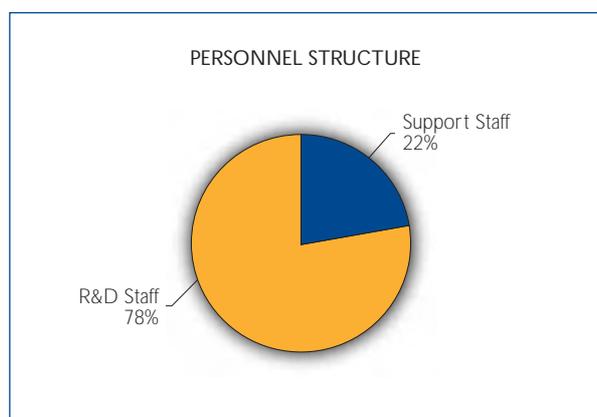
BSC-CNS staff included shared staff from other public institutions such as the Technical University of Catalonia (UPC) and the Institute de Recerca de Barcelona - Parc Científic de Barcelona -Universitat de Barcelona (IRB-PCB-UB); this staff must adhere to the established rules and agreements for each of the centers to which he or she is affiliated. BSC-CNS hopes to welcome staff from additional institutions in 2007.

BSC-CNS also welcomed high level scientific personnel from special human resources public programs such as the *Ramon y Cajal* Program, the ICREA Program and other personnel training research programs sponsored by various Spanish Ministries.

At the same time, the BSC-CNS Fellowship program welcomed applicants from relevant scientific disciplines to participate in several European research projects and collaborations with international industry such as IBM and Microsoft. These fellowships were offered for periods of one year, renewable for the duration of the project. At the end of 2006, BSC-CNS had 25 student researchers associated to several research projects.

The following is a breakdown of BSC-CNS staff:

DEPARTAMENT	HEADCOUNT 2005	HEADCOUNT 2006
Director, Associate Director	2	2
Human Resources, Administration and Finance	4	7
Operations	11	16
Business Development	5	6
Earth Sciences	4	12
Life Sciences	9	24
Computer Sciences	28	70
TOTAL	63	137

The setting up and launching of the center in 2006 led to a high level of activity and a significant increase in staff as well as an increase in the complexity of the organizational structure

PERSONNEL EVOLUTION

Management, Administration & finance, Operations, Business development, Earth Sciences, Life Sciences, Computer Sciences

2005   2006

STAFF INCREASE

120%
100%
80%
60%
40%
20%
0%
          2005   2006

R&D Staff
Support Staff
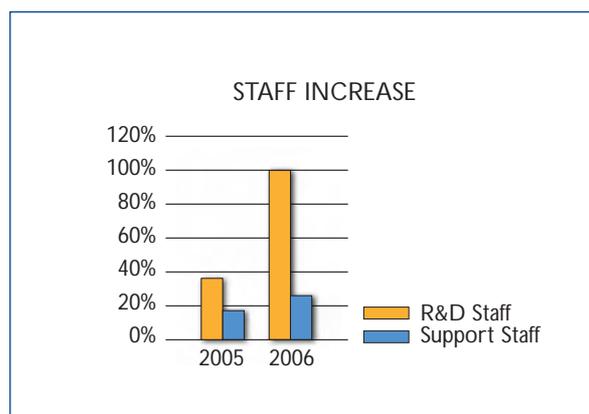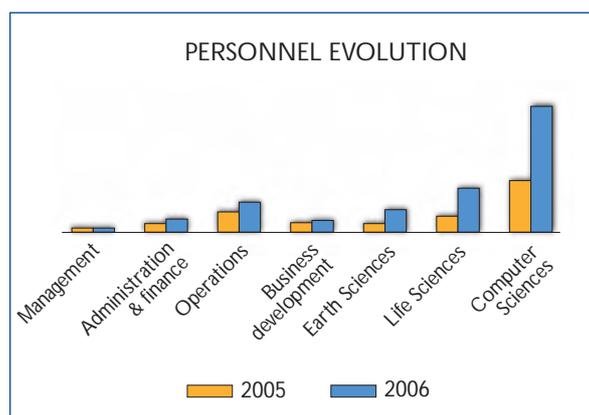
Finance

The following is a detailed description of the financial accounts for 2006. The annual accounts were drawn up following the accounting principles laid out in the General Plan of Public Accounting.

The operating budget of the BSC-CNS Consortium for the fiscal year 2006 (see table below) included expenses incurred from the fulfillment of its financial obligations and income from contributions of the administrations and organizations that compose it, as well as income derived from agreements reached with private organizations.

Income

The Consortium income recognized for fiscal year 2006 came from public administration contributions from the Ministry of

Expenses

The expenses of the BSC-CNS Consortium in the fiscal year 2006 include the upgrade of MareNostrum, financed through a special budget. The current expenses include the office space rental, security services, maintenance and cleaning services, telephones and networking, legal services, marketing, insurances and power. In addition, BSC-CNS acquired the furniture, fixtures and fittings, and the necessary computer equipment.

EXPENSE	AMOUNT	INCOME	AMOUNT
Personnel	2,925,938	Current Income	
		•Ministerio de Educación y Ciencia	3,717,000
Investments	16,663,083	•Generalitat de Catalunya	2,183,000
Current Expenses	2,256,705	Competitive income	
		•Private Companies	1,058,493
		•European commission	2,446,224
		•Ministerio de Educación y Ciencia	2,695,615
R & D Applied Reserves	6,457,285		
		Capital Transfers	2,675,679
		Financial Liabilities	13,527,000
TOTAL	28,303,011	TOTAL	28,303,011

Education and Science (MEC), Generalitat de Catalunya (GdC) and European Commission (EC), as well as from agreements, contracts or other collaborative agreements with private organizations. The center was also granted a loan to finance the upgrade of MareNostrum as well as create the Spanish supercomputing network. The loan will be repaid by 2015 through increases to the initial budget allocations from MEC and GdC. Furthermore, the consolidated budget of BSC-CNS included the assignment of its own resources coming from the previous year's provisions and financial results obtained in 2005.

EXPENSES 2006

Current expenses 8%
Investments 59%
R&D Applied reserves 23%
Personnel 10%
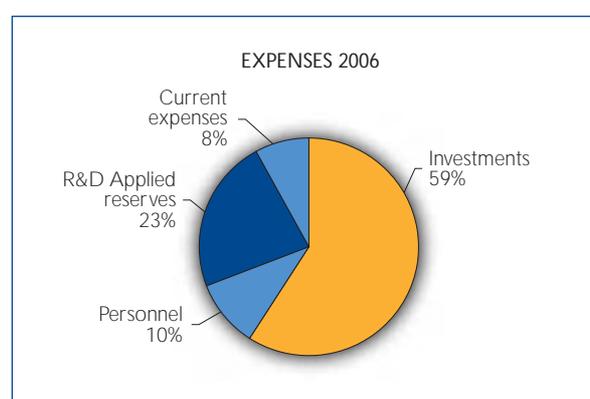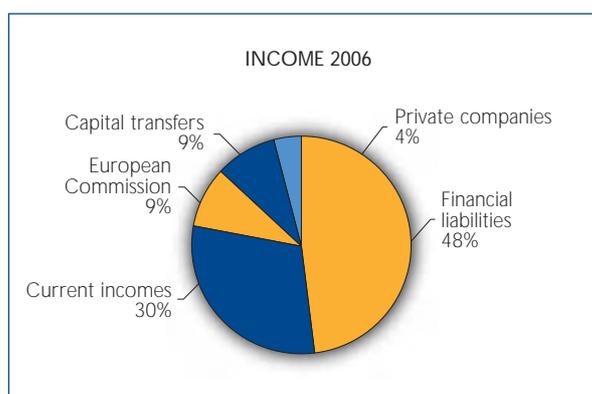
In 2006, BSC-CNS invited 14 public calls in order to accept 10 supply tenders, 3 services tenders and one construction tender. All of the acquisitions were made following the legal procedures established by the law regulating contracting in public administrations, and all contracts have been opened to public tenders.

INCOME 2006

Capital transfers 9%
European Commission 9%
Private companies 4%
Financial liabilities 48%
Current incomes 30%

Administration Support

The Administration Support Group is responsible for various support activities including the organization of official meetings and events, travel planning, space allocation, supply orders and execution of the tender process as well as providing general administrative support and reception services.

In 2006, the group handled the logistics for all official meetings and including regular Access Committee meetings, Executive Commission meetings as well as the meetings of the Board of Trustees. This included arranging for meeting spaces, catering services and travel-related assistance. The group coordinated the Recruitment Committees which were formed in order to find high quality candidates from Spain and abroad. Moreover, the group was responsible for receiving visitors for national and international mobility programs, representatives of various institutions and companies in addition to invited speakers from all over the world.

The group also played a significant role in finding and adapting office space to accommodate the various research groups. They worked with the contractors to arrange the spaces, move equipment and furnish offices; the team also provided facilities management services.

In 2006, BSC-CNS rented space in addition to acquiring space from the Technical University of Catalonia (UPC) according to the established collaborative agreement. BSC rents Nexus II and Nexus I buildings which are managed by the Zona Franca Consortium, and the combined rented area is 1441 square meters. UPC transferred the University Chapel where MareNostrum is located to BSC-CNS. This space required a great deal of renovation in order to create the necessary office space for staff directly responsible for the physical operation of the supercomputer. The total usable space available in the newly renovated Chapel is now 1865 square meters. As the BSC-CNS building is on the UPC campus, UPC agreed to provide some communal services, such as campus coordination and general maintenance. BSC-CNS took out several insurance policies which guarantee coverage with respect to civil responsibility, damage or injury to third parties or UPC property.

At the end of the 2006, the Administration Support Group organized the first all-hands meeting of BSC-CNS employees that took place at Codorniu cellars situated in San Sadurni d'Anoia (Penedés, Barcelona). One hundred and ten staff members participated in the event. Department and General Directors explained the activities and results obtained during 2006 as well as their objectives for 2007. The BSC-CNS staff was provided the opportunity to ask questions about the BSC-CNS organization as well as to propose ideas to improve the center.

BSC-CNS staff

4.3 Operations

The Operations Department consists of the System Administration and User Support Groups.

Sergi Girona,
Operations Director

The System Administration Group's responsibilities include security, resource management, networking and helpdesk services. This group ensures that the BSC-CNS systems are available 7 days a week, 24 hours a day. The service is guaranteed via an automatic monitoring mechanism that reports any events and reacts automatically to most technical problems. The main task of the User Support group is direct user support, and group competencies include programming models, libraries, tools and applications. In 2006, the Operations Department focused on the following four areas:

· Operation and upgrade of MareNostrum
· Preparation of new High Performance Computing (HPC) systems
· Technical definition related to the Spanish Supercomputing Network (RES, Red Española de Supercomputación)
· User Support

The department also contributed to European projects DEISA / eDEISA integrating into the DEISA infrastructure. And in HPC-Europa providing support to different visitors including knowledge and support in HPC systems.

MareNostrum

BSC-CNS hosts MareNostrum, one of the most powerful supercomputers in Europe. MareNostrum is managed by the Operations Department which takes care of its availability, security and performance. The department also supports the scientists that use MareNostrum as well as helps them to improve their applications and ultimately provide better research results.

System Architecture

MareNostrum is a Linux system based on the BladeCenter architecture with PowerPC processors and a Myrinet interconnection. These four technologies integrate into an architecture and design that will significantly impact the future of supercomputing. MareNostrum was upgraded in September 2006 in order to meet the high demand of Spanish scientists. This system upgrade resulted in an increase in power consumption to 740KWh which thus necessitated a modification to the site facilities as well as the air conditioning systems and power supply.

The system upgrade included the following improvements:
- Increased peak performance to 94.21 Teraflops
- 10240 IBM 970MP Power PC 2.3GHz processors (2560 dual 64-bit processor blade nodes)
- 20TB of main memory
- 370TB of disk storage
- Interconnection networks: Myrinet, Gigabit Ethernet, 10/100 Ethernet

MareNostrum has 44 racks and occupies a space of 120m². The following is a floorplan that indicates the distribution of MareNostrum:

Blade Center Racks

MareNostrum has 31 racks dedicated to performing calculations. These racks have a total of 10240 IBM 970MP Power PC 2.3GHz processors and 20TB of total memory. Each rack is composed of 6 blade centers. Each rack has 336 processors and 672Gb of memory with an estimated peak performance of 3.1Tflops.
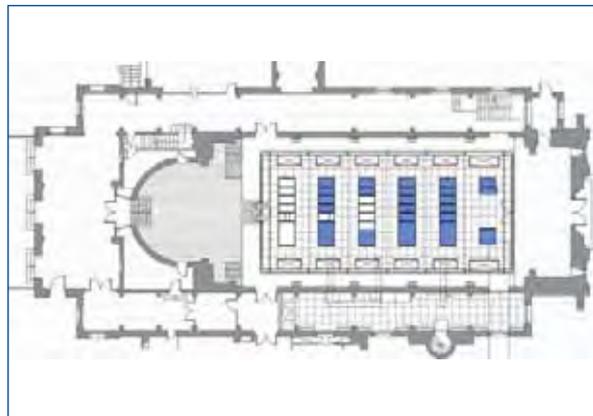
Blade Center

The JS21 nodes are grouped in a blade center which contains 14 dual processors with a total of 56 processors.

Each blade center has two redundant power supplies. If one of the power supplies were to fail, the other power supply would keep the blade center running. Each blade center also has a switch for the Gigabit interconnection network.

Blade centers   Storage servers   Gigabit switch
Myrinet racks   Operation rack   10/100 switches
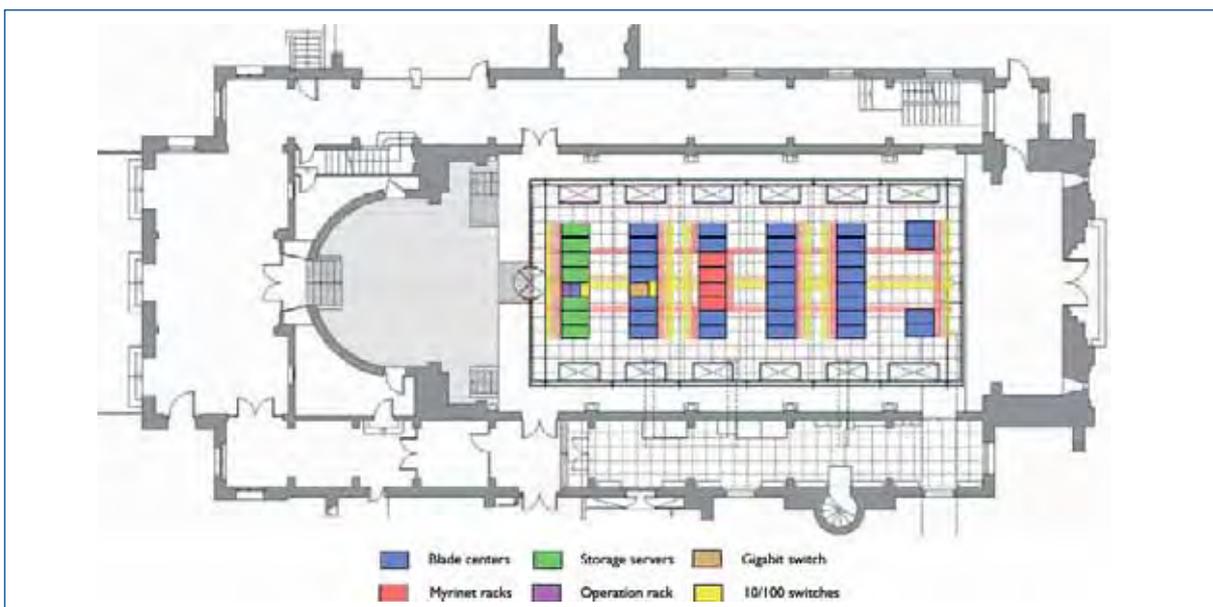
JS21 Server Blade

Each blade center has 14 JS21 server blades. Each of these nodes has two 2.3 GHz 970MP PowerPC processors with 8Gb of shared memory between both processors and a local SAS disk of 36Gb.

Each node has a M3S-PCIXD-2-I Myrinet network card that connects it to the high speed network and the two connections to the Gigabit network.

970MP PowerPC Processor

The 970MP PowerPC processor has a 64-bit architecture designed for general use (the Mac G5 contains this type of processor). It is a superscalar processor with vector extensions SIMD (VMX) whose design is based on the high provision Power4 processor.

The new 970MP PowerPC processor is based on a 90nm technology that works at a frequency of 2.3 GHz. It is ready to launch a maximum of 4 instructions per cycle and start up to 200 in-flight instructions.
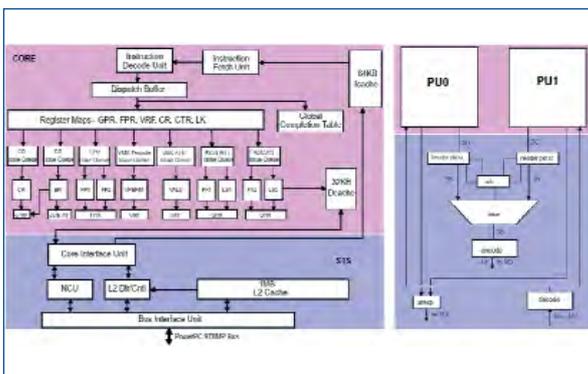
Myrinet Racks

The 2560 JS21 blade nodes are connected through a Myrinet high speed network. The different nodes are connected via fiber optic cables.

Four of the 44 racks in MareNostrum are dedicated to networking which allow for the interconnection of the different nodes connected to the Myrinet network. These four racks are located in the center of the space and each node has its own fiber optic cable. The network elements connect the different cables allowing the interconnection from one point to another from the different nodes.

The total Myrinet interconnection network consists of the following:
- 10 Clos256+256 switches
- 2 Spine 1280 switches
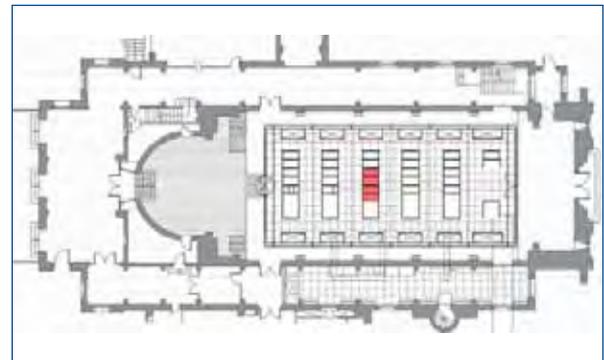
Storage Servers

In addition to the 36GB of capacity of the local disk of each node, MareNostrum has 20 storage servers arranged in 7 racks. These have a total of 560 disks of 512GB and each one provides a total capacity of 280TB of external storage. These disks are working with GPFS (Global Parallel File System) which offers a global vision of the file system and also allows a parallel access.

The 2560 nodes access the disks through the Gigabit interconnection network.

Each of the 20 storage nodes has two p615 nodes to handle disk requests, a FAStT100 controller and one EXP100 unit. Each p615 node has both a Gigabit and 10/100 Ethernet connection.
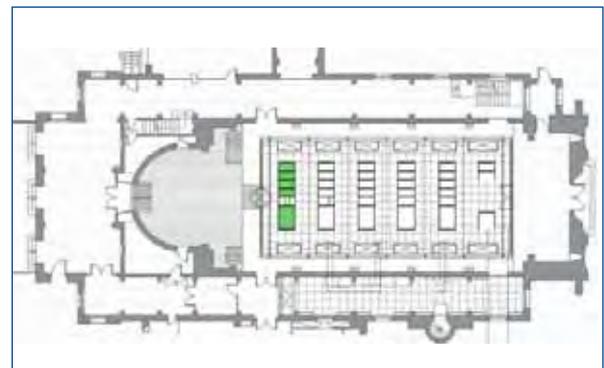
Operation Rack

One of the racks serves as the operation rack that manages the system. This rack is located in the machine console and is composed of the following:

- 1 7316-TF3 monitor
- 2 p615 nodes
- 2 HMC 7315-CR2 consoles
- 3 asynchronous remote nodes
- 1 BladeCenter IO chassis
- 4 Cisco 3550 switches
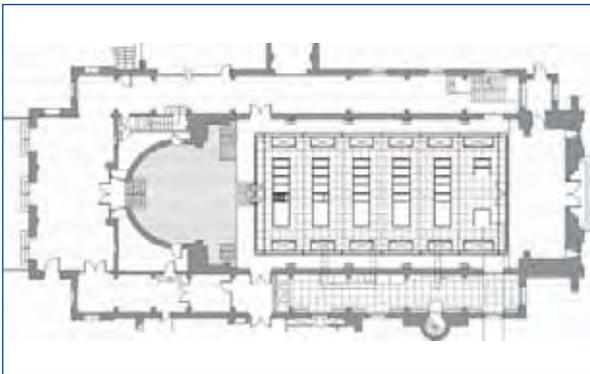
Gigabit Switches

One of the racks of MareNostrum is dedicated to the interconnection of the Gigabit network and the Ethernet 10/100 network. The rack includes the following:

- 1 E600 Force10 Gigabit Ethernet switch - The switch contains a total of 7 slots, 6 of them have cards 10 / 100 / 1000 Base-T of 48 ports offering a total of 288 ports Gigabit Ethernet / IEEE 802.3
- 4 Cisco 3550 48-port Fast Ethernet switches

Preparation of New HPC systems

BSC-CNS upgraded its HPC services with new equipment in order to meet the needs of BSC researchers and external users. In 2006, the following systems were made available:

Tape Library

Due to the high demand for storage capacity, BSC-CNS installed and configured a new SL8500 library with a total capacity of 3 petabytes (one thousand terabytes) without compression. This equipment allows for the back-up of all compute systems at BSC-CNS and includes HSM (Hierarchical Storage Management). Using HSM, the MareNostrum file system (240 Terabytes) has additional space in the library with a total capacity of 1PB.

The SL8500 library is equipped with LTO3 technology, and will be migrated to LTO4 as soon as it becomes available. The complete solution includes SAMFS Servers, Tivoli Servers and 31.4TB of diskspace for caching purposes.
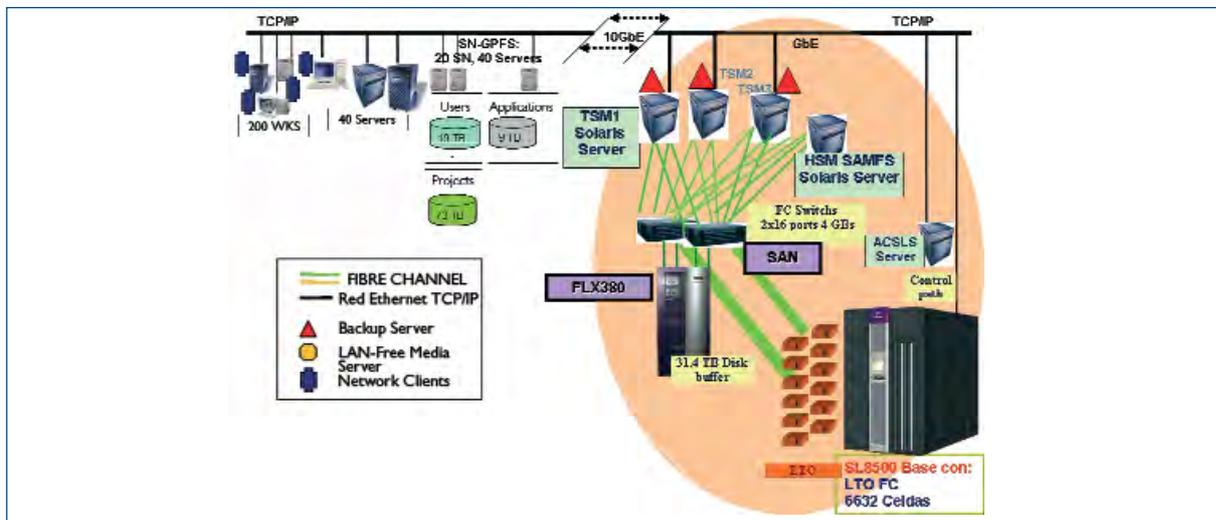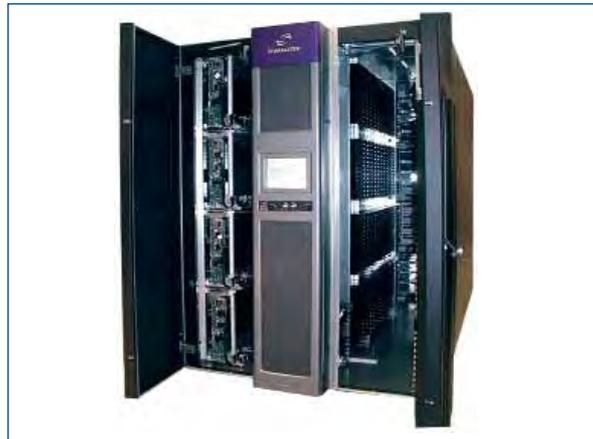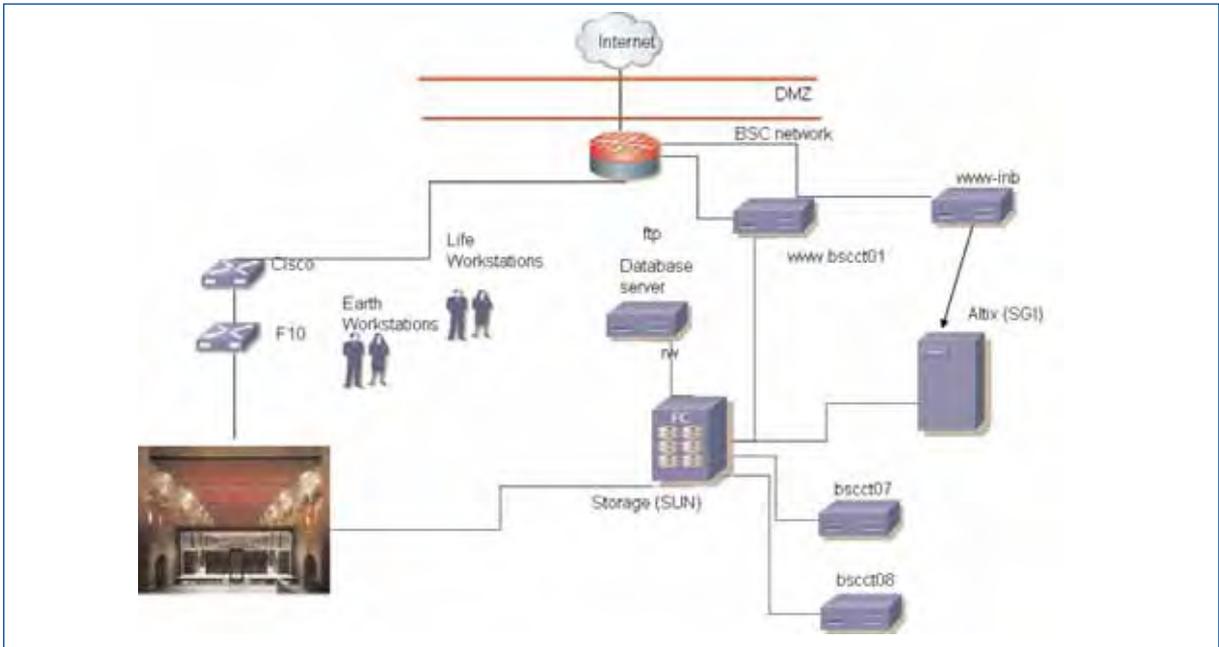
Shared Memory System

In the final quarter 2006, the Operations Department installed an Altix system with 128 cores and 0.5TB of main memory. This is largely due to the fact that specific pre-processing and post-processing analysis requires a large memory system and BSC-CNS's programming models research groups require this equipment for their research activities.

DataBase Storage and Servers

A new system for DataBase storage and handling was installed at BSC-CNS to serve the Life Sciences and Earth Sciences departments. This new equipment is composed of a StorageTek 6540 system with current capacity of 32TB in 120 disks (300 GB FC 10krpm), two Brocade 4100 switches (32 ports and 4Gb each), and 2 Sun Fire X4200 servers. This configuration is completely redundant and is illustrated in the figure above.

User Support

The User Support Group provides assistance with all aspects of scientific computing. This assistance includes general user support, writing and porting serial and code parallelization for the supercomputer, development of scripts for ease of use as well as assistance with software packages, tutorials on specialized topics or programs, user training, code optimization parallel model building support and assistance with developing code for the supercomputers.

In 2006, the User Support Group logged 6708 requests which can be divided into the following categories:
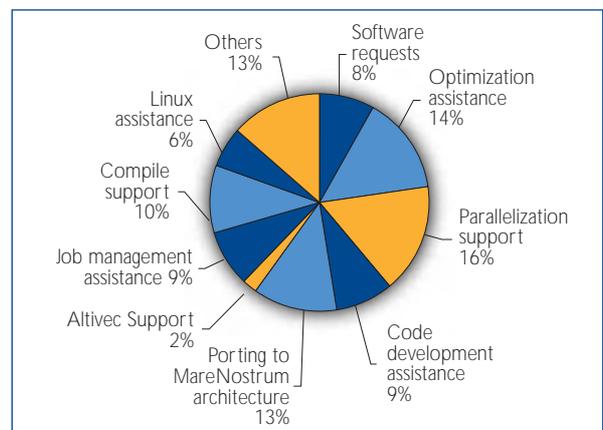
User Support Requests Classified by Topic
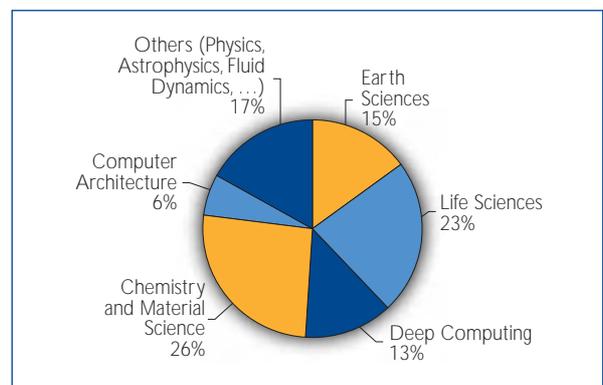
User Support Requests Classified by Area

Accounting

The system use is determined as follows:

$$\left(\frac{total\_cpu\_hour\_used}{total\_hours} \right) x100$$

MareNostrum was online throughout 2006, with the exception of a maintenance period and the time required for the system upgrade and the Linpack execution. Using the remaining part of the year as the basis for calculation, the observed system utilization is approximately 80%.
In addition to BSC-CNS internal groups, more than 120 external groups accessed MareNostrum. These groups can be classified into the following scientific areas:
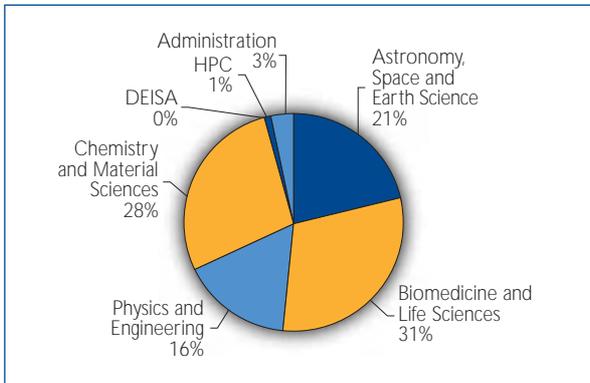
Total utilization of MareNostrum Classified by Scientific Area
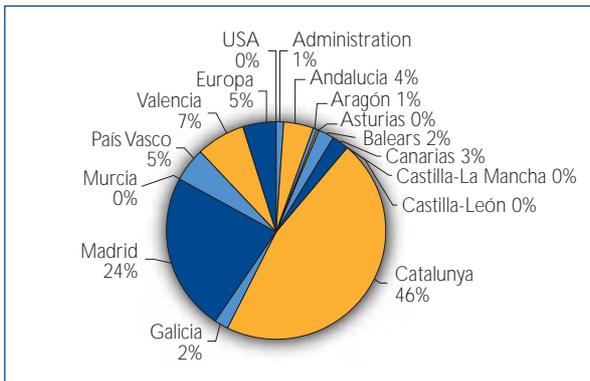
External utilization of MareNostrum Classified by Location

Top500 List

In November 2006, MareNostrum was ranked the most powerful supercomputer in Europe and ranked the fifth most powerful in the world on the Top500 list (www.top500.org) with a 62.63 TFlops/s Linpack performance.
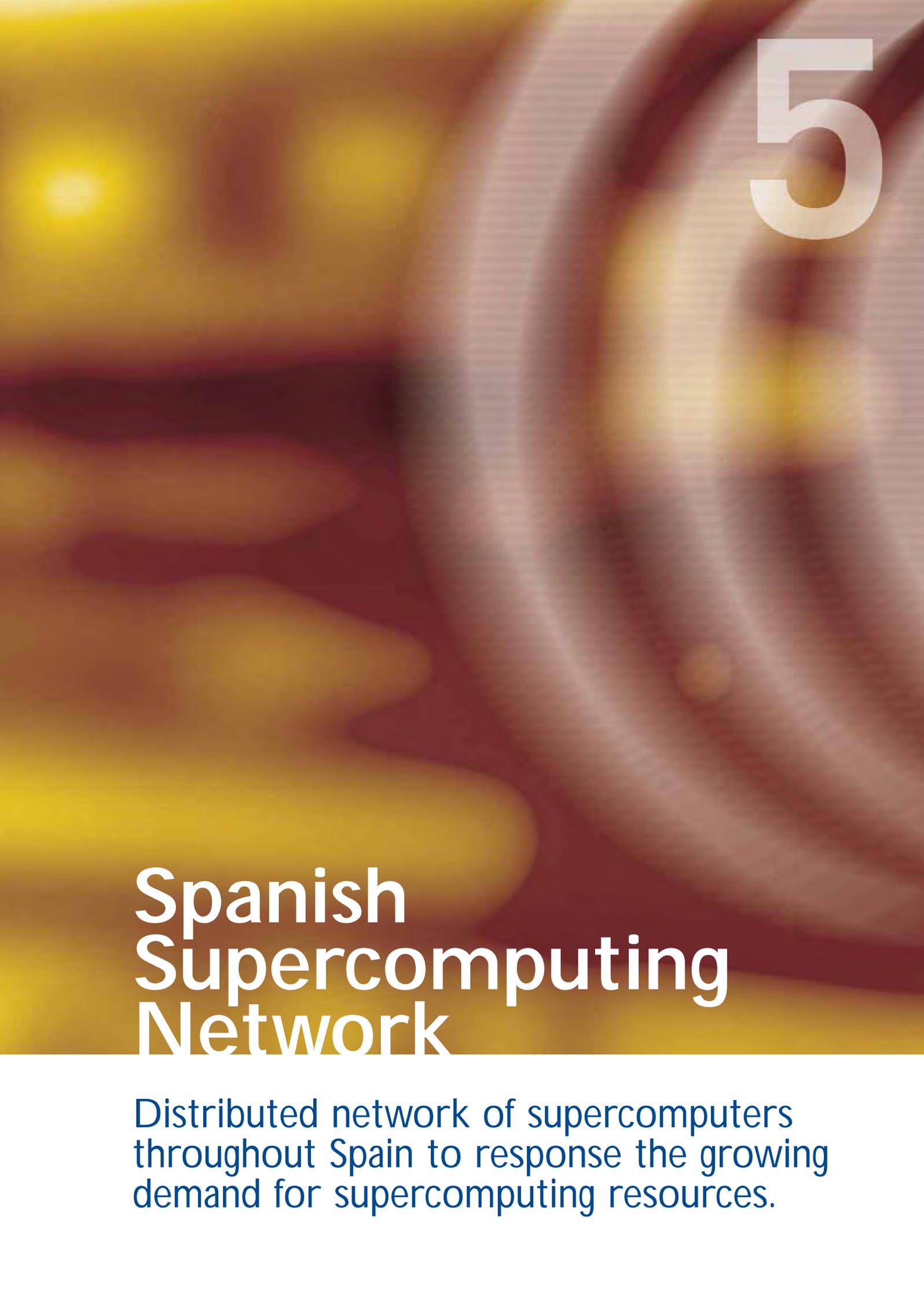
MareNostrum

Spanish Supercomputing Network

5

Distributed network of supercomputers throughout Spain to response the growing demand for supercomputing resources.

In November of 2006, the Spanish Ministry of Education and Science created the Spanish Supercomputing Network (Red Española de Supercomputación or RES). This distributed network of supercomputers throughout Spain was created in response to the growing demand for supercomputing resources for research purposes. In addition to providing the scientific community with broad access to computing capacity, the mission of RES is to increase collaboration between theoretical, experimental and computer scientists. The total capacity of RES is more than 133 millions hours per year, and it is expected that RES will play a key role in the production of scientific results in Spain in the future.

The initial network nodes are located at the BSC-CNS, at the Centro de Supercomputación y Visualización de Madrid (CesViMa), at the Instituto de Astrofísica de Canarias (IAC) as well as at the Spanish Universities of Cantabria, Málaga, Valencia and Zaragoza. All systems made the November 2006 Top500 list. The supercomputer of the CeSViMa is ranked among the top ten computers in Europe and is ranked 34th in the world. The other supercomputers of the Spanish Supercomputing Network are ranked 412th, 413th, 415th, 416th and 417th, respectively.

The following is a brief description of each supercomputer:

- ### BSC-CNS:
  10240 processors, PPC 970, 2.3 GHz, 20Tbytes of main memory, Ranked 5th on the Top500 list.

- ### CeSViMa:
  2408 processors, PPC 970, 2.2GHz, 4.7Tbytes of main memory, Peak performance of 21,19 TFlops, Ranked 34th on the Top500 list.

- ### Universities:
  Universidad de Cantabria
  Universidad de Málaga
  Universidad de Zaragoza
  Universidad de Valencia
  Instituto Astrofísico de Canarias (IAC): each with 512 processors, PPC970, 2.2GHz, 1Tbyte of memory, Peak performance of 4.5 TFlops, Ranked 412th to 417th on the Top500 list.

BSC-CNS serves as the technical coordinator of RES in addition to hosting one of the initial nodes. Access to RES by the various Spanish research groups will be determined by an Access Committee consisting of independent well-known Spanish scientists.

RES - Red Española de Supercomputación

MareNostrum Users

List of research results from all research areas for the year 2006.

This section provides a summary of many of the projects of the external researchers that have accessed MareNostrum in 2006. It includes a list of resulting papers that were published before the printing of this Annual Report. The researchers and their projects have been classified in terms of the knowledge areas of the Access Committee.

Biomedicine and Life Sciences

- José María Carazo - Centro Nacional de Biotecnología
"Revealing different functional states of the ribosome through maximum-likelihood classification of electron microscopy images"

- Robert Castelo - Universitat Pompeu Fabra
"Simulation and learning of large biomolecular networks using graphical Markov models"

- Joaquín Dopazo - Centro de Investigación Príncipe Felipe
"From genome to phylome: towards a complete catalogue of all the human orthologous and paralogous genes"

- Ernest Giralt – Universidad de Barcelona
"Evolutionary Algorithms and de Novo Peptide Design"

- Roderic Guigó – Centre Regulació Genòmica
"Use of TBLASTX to find regions of homology among multiple large-size full genomes"

- Alfonso Jaramillo – Ecole Polytechnique
"Computational design of a new function into a Thioredoxin protein using an automated procedure"

- Francisco Javier Luque – Universidad de Barcelona
"Ligand exit pathways in Mycobacterium tuberculosis truncated haemoglobin"

- Sara Marsal - Institut de Recerca Vall d'Hebron
"Whole genome epistatic modelling in rheumatoid arthritis with constructive induction through parallel computation. Study of genetic combinations associated with diagnosis and disease progression"

- Baldomero Oliva - Universitat Pompeu Fabra
"Aligned profiles with multiple environments and methods: Selection of the best parameters"

- Juan Jesús Pérez González – Universitat Politècnica de Catalunya
"Molecular dynamics simulations of G-protein coupled receptors"

- Marta Riutort – Universitat de Barcelona
"Multigenic and phylogenomic approach to the origin and diversification of Lophotrochozoan Bilaterians"

- Carme Rovira i Virgili – Parc Científic de Barcelona
"Activation of glycoside hydrolase enzymes by specific carbohydrate-protein interactions"

- Carme Rovira i Virgili – Parc Científic de Barcelona
"Unravelling the mechanism of drug activation in Mycobacterium tuberculosis catalase-peroxidase by means of first principles Car-Parrinello simulations."

- Alfonso Valencia - Centro Nacional de Biotecnología
"Docking methods benchmarking and protein evolutionary information integration"

- Alfonso Valencia - Centro Nacional de Biotecnología
"Systematic Evaluation of Prediction of Protein Pairings Between Interacting Families based on Protein Coevolution."

- Jordi Villà Freixa – Institut Municipal d'Investigació Mèdica
"ByoDyn: parameter estimation of multicellular models of gene regulatory networks"

- Jordi Villà Freixa – Institut Municipal d'Investigació Mèdica
"Parameter estimation and optimal design in computational systems biology"

- Wolfgang Wenzel - Research Center Karlsruhe (Germany)
"De-novo protein folding and structure prediction with free-energy forcefields"

Astronomy, Space and Earth Sciences

- Irene Arias - Universitat Politècnica de Catalunya
"Large-Scale Simulations of Dynamic Fracture"

- Marc Balcells - Instituto Astrofísico de Canarias
"Galaxy Transformations through Interactions, Mergers and Accretion"

- Emilio Cuevas - Instituto Nacional Meteorología (OAI)
"MM5 and WRF-ARW Sensibility Study in Tropical Storm Delta Meteorological Environment"

- Pablo Fosalba – Universitat Autònoma de Barcelona
"Large numerical simulations for dark-energy surveys"

- Rolando R. García - University Corporation for Atmospheric Research
"Simulation of changes in atmospheric climate and chemical composition 1950-2050"

- Enrique García Berro - Universitat Politècnica de Catalunya
"Smoothed Particle Hydrodynamics in Astrophysics"

- Víctor Homar – Universitat de les Illes Balears
"Application of ensemble prediction techniques to high impact weather episodes in the Western Mediterranean el Mediterráneo Occidental (ENSEMBLE)"

- Jose María Ibáñez – Universidad de Valencia
"Stability of extragalactic jets"

- Fernando Moreno Insertis – Instituto Astrofísico de Canarias
"Eruptive phenomena in the atmosphere of the sun and cool stars"

- Josep María Solanes – Universitat de Barcelona
"Investigation of the DIffuse LIght COmponent in compact groups of galaxies (IDILICO)"

- Nikolaos Stergioulas - Aristotle University of Thessaloniki (Greece)
"R-mode Frequencies for Relativistic Stars in LMXBs"

- Romain Teyssier. Commissariat à l'Énergie Atomique (CEA). HORIZON Consortium
"HORIZON@MareNostrum: Galaxy formation in a cosmological context"

- Jordi Torra i Roca – Universitat de Barcelona
"Gaia: Simulation of Telemetry Stream"

- Álvaro Viudez Lomba – CSIC
"Generation of inertia-gravity waves in geophysical vortex interactions"

- Gustavo Yepes – Universidad Autónoma de Madrid
"Cosmological Simulations of Large-Scale Structure formation in the Universe"

Physics and Engineering
- Juan Acebrón - Universitat Rovira i Virgili
"Probabilistic domain decomposition for partial differential equations"

- Jose Luis Álvarez Pérez - Universitat Politècnica de Catalunya
"Numerical Simulation of Emissivity of Soils and Ocean"

- Jean Claude Andre - Le Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique (CERFACS) (France)
"Large Eddy Simulation of an industrial gas turbine"

- Richard J. Duro – Universidad de Coruña
"Automatic design of wind turbine blades"

- Jordi Faraudo – Universitat Autónoma de Barcelona
"Molecular Dynamics Simulations of Biomimetic Membranes"

- Domingo Giménez Cánovas – Universidad de Murcia
"Parallel routines optimization and applications"

- Vicent Giménez Gómez – Universidad de Valencia
"Monte Carlo numerical computations of the properties of hadrons"

- Antonio González-Arroyo - Universidad Autónoma de Madrid
"Particle Production and Field Generation at Preheating after Inflation"

- Pilar Hernández – Universidad de Valencia
"Non-perturbative aspects of QCD in flavour physics"

- Javier Jiménez Sendín - Universidad Politécnica de Madrid
"DNS in turbulent channels"

- Bruno Julià Díaz - Universitat de Barcelona
"Dynamical Coupled-channel Analysis of Excited Baryons"

- Thilo Knacke - Technische Universitaet Berlin (Germany)
"Numerical simulation of airframe noise"

- Enrique Lomba - Instituto de Química Física Rocasolano, CSIC
"Ab initio simulation of amorphous SiC and C"

- Enrique Lomba - Instituto de Química Física Rocasolano, CSIC
"Dynamics of Molten SeTe alloys in the semiconductor-metal transition region"

- William Marnane - University College Cork (UK)
"Development and Charaterizaton of new LDPC codes, encoders and decoders"

- Monty Newborn – McGill University, Montreal (Canada)
"Parallel Automated Theorem Proving"

- Asensi Oliva – Universitat Politècnica de Catalunya
"Direct Numerical Simulation of Turbulent Flows"

- Ignacio Pagonabarraga – Universitat de Barcelona
"Lattice Boltzmann simulations of complex fluids"

- Assumpta Parreño – Universitat de Barcelona
"Study of SubAtomic Interactions through Lattice Quantum Chromo Dynamics on Mare Nostrum (SAIL)"

- Josep María Porta – Universitat Politècnica de Catalunya
"Isolating configuration spaces of polycyclic robots and molecules by extended bound smoothing"

- Jose E. Roman – Universitat Politècnica de València
"Scalable Eigensolvers in the SLEPc Library"

- Greg Thain - Wisconsin University (USA)
"Condor on MareNostrum"

- Iñigo Toledo – Desafío Español Copa América
"Desafio Espanol Challenge 2007 – America-s Cup Yacht CFD Research"

- Daniella Tordella - Politécnico de Turín (Italia)
"On the role of turbulent energy and integral scale gradients in turbulent diffusion and intermittency"

- Markus Uhlmann – CIEMAT
"Direct Numerical Simulation of Turbulent Flow with Suspended Solid Particles"

Chemistry and Material Science

- Carlos Alemán – Universitat Politècnica de Catalunya
"Dynamics of an Artificial Muscle based on Conducting Polymers and Calix[4]arene Scaffolds"

- Carlos Alemán – Universitat Politècnica de Catalunya
"Nanotechnology: De Novo Design using the Self-Assembly of Biomolecules"

- Pedro de Andrés – CSIC
"Molecular adsorption on metallic and oxide surfaces"

- Andrés Arnau – Universidad del País Vasco
"Adsorption of alkanethiols on Au(111) at different coverages up to one monolayer"

- Carles Bo – ICIQ
"Structure of polyoxometalates in solution: from single anions to nano-capsules"

- Maria José Caturla Terol – Universidad de Alicante
"Modeling materials properties at the nanoscale: from defects to nanodevices"

- Albert Cirera – Universitat de Barcelona
"Ab initio computation of the effects of vacancies for transparent electronics in In_2O_3"

- Avelino Corma – CSIC
"Selective hydrogenation of nitroaromatics catalyzed by gold"

- Jordi Faraudo – Universitat Autònoma de Barcelona
"Molecular Dynamics Simulations of Electrostatic Interactions between Macromolecules and/or Macroions in Aqueous Solution"

- Ernesto García - Universidad del País Vasco
"Dynamics of the $OH+CO->H+CO2$ reaction at low collision energy"

- Gregory Geneste – Ecole Centrale Paris (France)
"Ab initio simulation of ionic conductors for solid oxide fuel cells"

- Manuel Maria González Alemany – Universidad de Santiago de Compostela
"First-principles calculations for the electronic structure of semiconductor nanowires"

- Eduardo Hernández – ICMAB-CSIC
"Electronic structure of materials: thermal effects and ab initio thermodynamics"

- Francesc Illas – Universitat de Barcelona
"Scalability of the VASP code and applications to heterogeneous catalysis"

- Francesc Illas – Universitat de Barcelona
"Towards understanding the molecular mechanisms of heterogeneous catalysis"

- Zbigniew Lodziana - Polish Academy of Sciences, Krakow (Poland)
"Quantum calculations of electron correlations and the surface properties of magnetite"

- Nuria López – ICIQ
"Hydrogen production from the degradation of ammonia on Ir nanostructured crystals: structure sensitivity"

- Feliu Maseras Cuni – Institut Català d'Investigació Química
"A DFT computational approach to supramolecular catalysis"

- Sílvia Picozzi - CNR-INFM (Italy)
"Ab-Initio Approach to Multiferroics Based on Perovskites: Exchange Mechanisms and Chemical Trends"

- Josep María Poblet – Universitat Rovira i Virgili
"Modelling of Polyoxometalates"

- Ramon Reigada Sanz – Universitat de Barcelona
"Molecular Dynamics study of the effect of Cholesterol in the Cell Membrane"

- Antonio Rodríguez Fortea - Universitat Rovira i Virgili
"First-principles molecular dynamics simulations of the formation mechanisms of polyoxometalates with low nuclearities"

- Ángel Rubio – Euskal Herriko Unibertsotatea
"Spectroscopic properties of biomolecules, nanostructures and extended systems"

- Eliseo Ruiz – Universitat de Barcelona
"Magnetic Properties of large Single Molecule Magnets: Mn_{25} and Mn_{32}"

- Enrique Sánchez Marcos - Universidad de Sevilla
"Hydration of pharmacologically interesting $Pd2^+$ and $Pt2^+$ complexes"

- Miquel Solà Puig - Universitat de Girona
"Theoretical Study of the reaction mechanism of human superoxide dismutase"

- César Tablero – Universidad Politécnica de Madrid
"Study of the nonradiative recombination"

6.1 Biomedicine and Life Sciences

José María Carazo

Centro Nacional de Biotecnología

Revealing different functional states of the ribosome through maximum-likelihood classification of electron microscopy images

In principle, cryo-electron microscopy allows visualization of flexible macromolecular assemblies in their various functional states. However, the inability to classify noisy projection data from structurally heterogeneous samples has posed serious limitations on this technique. To solve this problem, the team is developing a novel classification method based on maximum-likelihood principles. Preliminary results indicate that this method is capable of separating heterogeneous data into structurally distinct classes without the need for prior knowledge about the variability in the sample. The team's current investigations focus on the general applicability of this approach. If this approach would prove generally applicable to heterogeneous electron samples, this would greatly expand the applicability of cryo-electron microscopy to the numerous, flexible macromolecular assemblies that perform many crucial tasks in the living cell, such as the decoding of the genetic material and the defence against bacterial or viral infections.

Thanks to the increasing supercomputing capabilities such as MareNostrum, it is now feasible to apply maximum-likelihood based classification approaches to the tremendous problem of classification of structurally heterogeneous cryo-electron microscopy data sets. These data sets may contain up to billions of data points, and the proposed approach may take up to several months of CPU to separate a single data set. Therefore, these developments would not be feasible without the availability of supercomputers such as the MareNostrum.

Robert Castelo

Universitat Pompeu Fabra

Simulation and learning of large biomolecular networks using graphical Markov models

The team is working in implementing a parallel version of the software qp for learning multivariate models of biomolecular networks from microarray data. Currently, the PC implementation does not allow to work with more than a few hundred variables and it is expected that the parallel version running in MareNostrum will enable the team to work with thousands of them and then tackle problems like learning a gene regulatory network for thousands of genes that might be involved in a particular biological pathway.

The long-term goal of this research is to obtain an implementation that would allow to learn the biomolecular network of the entire human gene set, thus coping with about 30,000 random variables.

The team is still in the process of developing the software in MareNostrum but a succesful end of this task will allow to tackle biomolecular network modeling problems of unprecedented complexity.

Random undirected graph with a scale-free distributed connectivity on 50 vertices. The probability distribution represented in this graph contains many partial correlations that can be only uncovered with multivariate models. In microarray data vertices may represent genes, and edges co-expression relationships where the actual complexity easily scales up to thousands of vertices representing genes under study.

Joaquín Dopazo
Centro de Investigación Príncipe Felipe

From genome to phylome: towards a complete catalogue of all the human orthologous and paralogous genes

The team has reconstructed, for the first time, the collection of evolutionary histories of all genes encoded in the human genome. This represents a significant step forward towards the understanding of the evolution of our species and provides important insigths into the evolution of our genes.

The human phylome is also useful since it allows establishing orthologs (evolutionary related) of human genes in model species, thus facilitating the transfer of knowledge obtained from experiments performed in model species such as mouse or fruit fly to human genes. Moreover, knowledge on the evolution of genes can be used to elucidate the function of uncharacterized genes, by detecting patterns of co-evolution that are common to genes involved in a given biological process.

Thanks to the increasing supercomputing capabilities such as MareNostrum, it is feasible to make an extensive use of multiple sequence alignments, parameter estimation, phylogenetic reconstruction to reconstruct the evolution of more than 20.000 individual genes.
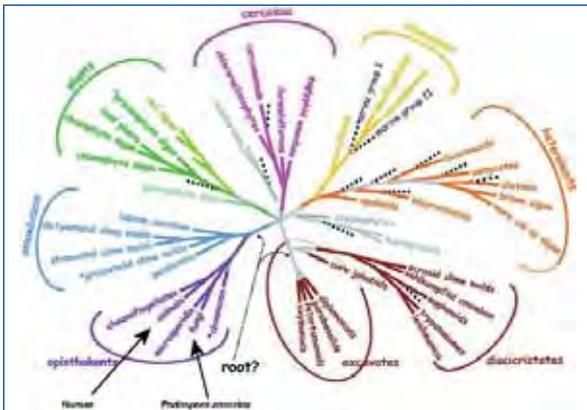
Ernest Giralt
Universidad de Barcelona

Evolutionary Algorithms and de Novo Peptide Design

The project is developing the program "Evolutionary structure de *Novo* Peptide Design Algorithm" (ENPDA). Given a user-defined target protein surface, ENPDA is able to design peptides able to recognize this surface, i.e., bind to it with good binding energies. In particular, the project focus falls in the implementation of novel features into ENPDA such as its capacity to design variable length peptides, reliable and fast docking strategies, and new parallelization architectures to achieve higher speed-ups. In addition, and as part of testing problems, ENPDA was used to design peptide inhibitors for prolyl oligopeptidase, an important enzyme related to several neuropathologies such as schizophrenia, bipolar disorder, bulimia nervosa, etc.
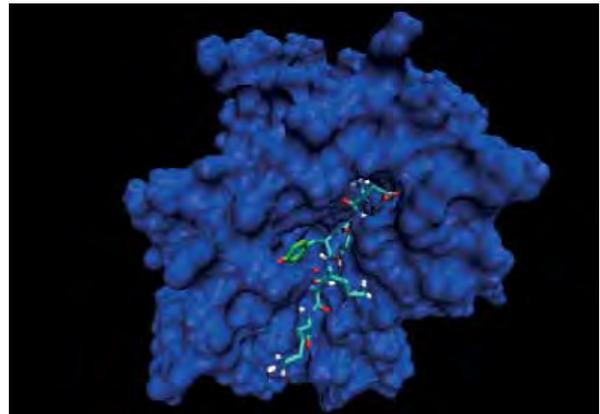
Structure of the best peptide (EKYKEK) proposed by ENPDA that recognizes MHC H-2Kb surface. ENPDA evaluated 9500 different peptides to propose this final sequence.

Roderic Guigó
Centre Regulació Genòmica

Use of TBLASTX to find regions of homology among multiple large-size full genomes

The team is working on the analysis an annotation of many eukaryotic genomes. In particular the project contributes towards the delineation of the gene sets in these genomes. This is the first essential step to translate the sequence of the genome into biologically relevant knowledge. Knowledge that can in turn, be applied to biomedicine and biotechnology. Indeed, genes code for proteins, the basic building blocks of all living organisms.

The team requires MareNostrum to run the computationally intensive algorithm TBLASTX to find regions of homology between two full genomes (i.e. Mouse vs. Human). Running TBLASTX to find potentially conserved regions between different species is an essential component of the homology-based gene prediction tool SGP which has been developed by the group. Generally, TBLASTX comparison among species with large genomes, such as Human and Mouse, would require 7-10 days on a 20-25 processor grid. However, MareNostrum´s processing speed and specially parallelization potential allows the same work to be performed in approximately 12 hours (on 256 CPUs and excluding queuing time). This is important because it allows executing TBLASTX to find homologies among many different genomes at the same time, in a reasonable time frame, which would be impossible to achieve locally or using a smaller computer grid. This, in turn, allows developing and improving SGP for large-genome species to a level which would be unattainable by using a smaller grid.

Alfonso Jaramillo
Ecole Polytechnique (France)

Computational design of a new function into a Thioredoxin protein using an automated procedure

The team is developing a new systematic approach to the problem of introducing enzymatic activity into an arbitrary inert protein scaffold, with the aim of designing new proteins with targeted function. This will have important applications in areas like medicine (treatment of neurodegenerative diseases, cancer, autoimmune diseases and development of peptide-based vaccines), biotechnology (biosensors, biocatalysts, reduction of waste products and toxicity, synthetic biology) and basic science. The computing resources of MareNostrum are well suited for computational protein design and have provided with extensive CPU and memory crucial for this work. It's been allowed to explore a larger combinatorial space and therefore improve the final design. In addition, to use MareNostrum has pushed to improve the methodology in order to fully benefit from MareNostrum's enormous capabilities.

The team has designed a thioredoxin mutant that preserves its original function (general disulphide oxidoreductase) and adds an esterase activity (previously absent) with histidine as nucleophile and p-nitrophenyl-acetate as substrate. Only three mutations were needed and the design was experimentally validated by collaborators in Univ. Granada (Dr. Jose Manuel Sanchez-Ruiz).

The methodology developed is able to scan a larger sequence and conformational space than has ever been possible thanks to the supercomputing facilities of BSC. The team has succesfully incorporated activity and stability (protein folding) in the combinatorial optimization procedure, and therefore the team is able to include mutations to charged residues without destabilizing the protein. This modelization of catalysis is based purely on physicochemical principles, which introduces a novel way to design enzymes.
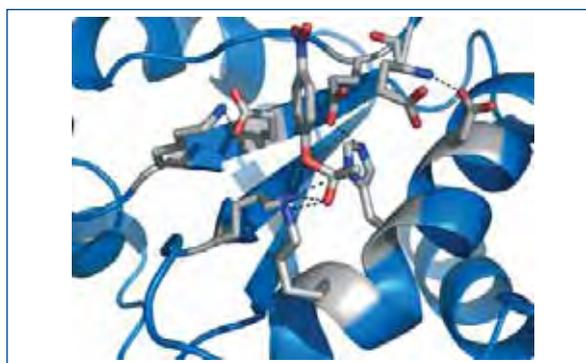
Design of a p-nitrophenyl aceteate esterase active site in thioredoxin with histidine as nucleophile.

Francisco Javier Luque
Universidad de Barcelona

Ligand exit pathways in Mycobacterium tuberculosis truncated haemoglobin

Mycobacterium tuberculosis, the causative agent of human tuberculosis, is forced into latency by nitric oxide produced by macrophages during infection. In response to nitrosative stress *M. tuberculosis* has evolved a defence mechanism that relies on the oxygenated form of "truncated haemoglobin" N (trHbN), formally acting as NO-dioxygenase, yielding the harmless nitrate ion. X-ray crystal structures have shown that trHbN hosts a two-branched protein matrix tunnel system, proposed to control diatomic ligand migration to the heme, as the rate-limiting step in NO conversion to nitrate. Extended molecular dynamics simulations (0.1 μs), employed here to characterize the factors controlling diatomic ligand diffusion through the apolar tunnel system, suggest that O_2 migration in deoxy-trHbN is restricted to a short branch of the tunnel, and that O_2 binding to the heme drives conformational and dynamical fluctuations promoting NO migration through the long tunnel branch. In turn, this is achieved by modulating the conformation of the residues TyrB10 and GlnE11, which regulate the dynamical behaviour of helices B and E. The simulation results suggest that trHbN has evolved a dual-path mechanism for migration of O_2 and NO to the heme to achieve the most efficient NO detoxification.
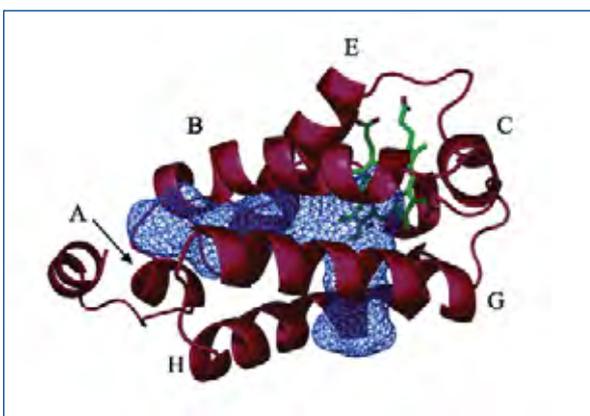
Structure of trHbN showing the two branches of the tunnel system.

Sara Marsal
Institut de Recerca Vall d'Hebron

Whole genome epistatic modelling in rheumatoid arthritis with constructive induction through parallel computation. Study of genetic combinations associated with diagnosis and disease progression

Whole Genome Association Studies (WGAS) offer a new approach to gene discovery that is unbiased with regard to presumed functions or locations of causal variants. In this study, the team uses this *ab initio* strategy to identify non additive genetic interactions *(epistasis)* associated with susceptibility to rheumatoid arthritis (RA) in the Spanish population. The group has genotyped 600 patients with arthritis and 200 hypernormals controls which includes data on 317.000 single-nucleotide polymorphisms (SNPs) for each individual.

The analysis of epistatic interactions was performed by a parallel implementation of *Multifactor Dimensionality Reduction* and *information theory-based Interaction Analysis* on MareNostrum.

A true interaction, also called epistasis or synergy, is observed when a combination of genes yields more information about the disease than the sum of the information from each individual gene. The difference in gained "extra" information is a measure of the synergy of genes and is used to identify disease-associated genes that manifest themselves only in the context of other genes.

Using the computational resources at MareNostrum, the team has analyzed data from a preliminary WGAS which includes information on over 317.000 SNPs and 800 individuals. The team has identified a list of candidate SNPs, appearing in two-way and three-way interactions, that are planned to be test on data on a new set of patients.

Discovering interactions among disease-causing genes in large WGAS data sets is an inherently combinatorial task. By definition, there is no short-cut but to look at each combination of two, three or more genes and calculate their interaction. Although various heuristics can be applied to reduce the set of combinations, the number of combinations to explore remains extremely large (in the range of 10^{12}). Only the vast computational power available at MareNostrum enables to analyze and discover gene interactions in a reasonable amount of time.
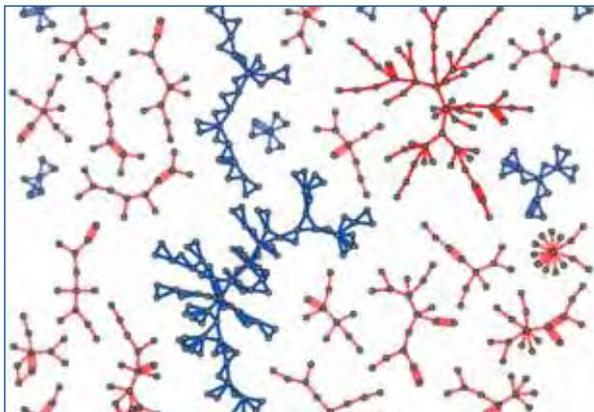
(Coming from page 85). An intricate web of interacting disease-associated genes appears when we connect genes discovered to act in two (red edges) and three-way (blue edges) interactions.

The method of alignment uses two variants of COMPASS, which have been here analysed: it is compared recall and specificity for 21 tested relationships between both variants. The parameter combination can be set given more importance to recall or specificity. Furthermore, it will be possible to select a default combination that will take both concepts in consideration.

Baldomero Oliva
Universitat Pompeu Fabra

Aligned profiles with multiple environments and methods: Selection of the best parameters

The comparison of multiple sequences might provide clues about the structure and function of a protein family. Besides, it has been shown the advantages of using sequence profiles to recognise distant homologies between proteins. Sequence-profile or profile-profile alignment is important not only for protein annotation but also for comparative modelling. Due to this importance, several methods have been designed in order to obtain profiles of multiple sequence alignments. In spite of these advances, many errors are still found on the alignment of sequences in comparative modelling. Accurate sequence alignments are crucial in the regions of 20-30% pairwise sequence identity, where most often methods fail to correctly align these sequences. The team has developed a method that combines all the improvements obtained from 13 different methods of profile-profile alignment. In order to test the best parameters involved in the profile alignment method, the team uses 336 folds (SCOP classification).

The results are obtained for the first group of parameters selected by comparing the project's alignments with gold standard alignments in order to know the best parameter combination for the new profile-profile alignment method. As an example, image shows the calibration between the two variants of the COMPASS method. The team expects the software to be usefull to improve sequence alignments and, thus, to help on comparative modeling. The other part of the project is developing a database that will contain profiles for almost 3 million proteins, a prediction of their domains, and interactions among them. The whole database of domains and the server for profile-profile alignment is intended to serve as a powerful tool for further research on the proteomics field. The team expects to provide the scientific community with a fast and versatile tool for the research on the proteomics field. The database, with millions of precomputed protein alignments, will be accessible by the scientific community via a web server.

This one fast access to the results can save a lot of time and computational resources to the users, while the profiles can be used for computing more reliable protein alignments. Additional aspects of interest of this project are the development of a method for predicting protein domains and protein interactions.

To make the tests, it's necessary to do about 13.3e6 runs with an average duration of 50 minutes. Running one process after the other implies more than a thousand years, so it would have been impossible without MareNostrum processor capability. Furthermore, the RAM memory requirements couldn't have been fullfilled without the BSC's support.

Juan Jesús Pérez González
Universitat Politècnica de Catalunya

Molecular dynamics simulations of G-protein coupled receptors

The study carried out involves a comparative analysis of a set of five 16 ns MD simulations of rhodopsin embedded in a DPPC lipid bilayer using different simulation protocols. The simulations were carried out using periodic boundary conditions on a system consisting in a molecule of rhodopsin, 197 lipid molecules and 15,817 water molecules. Calculations were carried out considering two different approximations for the treatment of long-range electrostatics: on the one hand, the use of a cutoff and on the other, the PME method in combination with two different sampling ensembles: the canonical (NVT) and the constant pressure (NPT). The results suggest that special care must be taken when selecting the statistical ensemble and the type of treatment for long-range electrostatics. The use of NPT is necessary for the bilayer equilibration and for the process involving the bilayer rearrangement around the protein. The use of the PME method provides a better description of lipids and ions but with a higher computational cost.
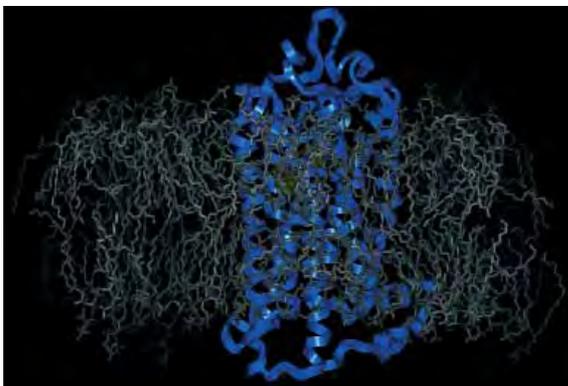
View of the system subject of the calculations. Water molecules are removed for clarity.

Marta Riutort
Universitat de Barcelona

Multigenic and phylogenomic approach to the origin and diversification of Lophotrochozoan Bilaterians

Constructing robust phylogenies is central to spot the extant organism closest to the first animal ancestor and to understand how all present day types of animals evolved. Morphology has been unable, after more than 150 years, to solve these problems. Therefore, the group propose a multigenic approach, identifying and using sequences of multiple genes. The team will concatenate these genes and analyze them using the latest and more sophisticated phylogenetic inference methods. The team's main aims are to find the present day closer relative to the ancestor of all bilateral animals and to resolve the order of appearance and relationships of all the animal phyla. Identifying the extant most basal bilaterian will be instrumental through functional analyses and the use of transgenic organisms to understand how the transitions from the original ancestor to all present day body plans (types of animal phyla) took place.
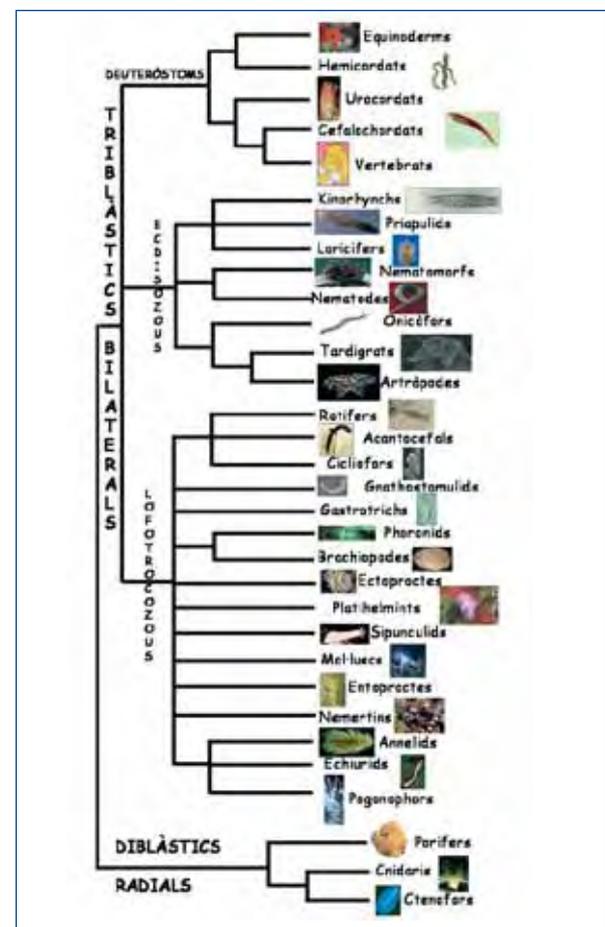
Carme Rovira i Virgili
Parc Científic de Barcelona

Activation of glycoside hydrolase enzymes by specific carbohydrate-protein interactions

The important roles that carbohydrates play in biology and medicine have stimulated a rapid expansion of the field of glycobiology. The structures of glycoside hydrolases, the enzymes involved in the degradation of glycosidic bonds of carbohydrates, are now being elucidated in molecular detail. It is becoming increasingly accepted that only a complex interplay of electronic/structural properties can explain the perfect match between the substrate and the enzyme. In the project the team addresses these problems by means of QM/MM molecular dynamics simulations.

The group has investigated the structure and dynamics of the enzyme-substrate complex of Bacillus 1,3-1,4-ß-glucanase, one of the most active glycoside hydrolases used in food industry. The calculations, performed by means of ab initio molecular dynamics simulations, show that the substrate adopts a distorted conformation that favours the catalytic mechanism of the enzyme. In a subsequent study, the team has rationalized the occurrence of distorted sugar conformations in all the available PDB structures of glucoside hydrolase Michaelis complexes. The project shows that these enzymes recognize the most stable distorted conformers of the substrate which are also the ones better prepared for catalysis. This suggests that the factors governing the distortions present in these complexes are largely dictated by the intrinsic properties of a single glucose unit. The results allow to predict the conformation of the substrates that will favour an efficient catalysis in these enzymes.

Thanks to the increasing supercomputing capabilities such as MareNostrum, it is feasible to do projects that involve a large number of atoms to be treated quantum mechanically. Only using a large number of processors the team can obtain the results on a timely way which is what makes the project feasible.
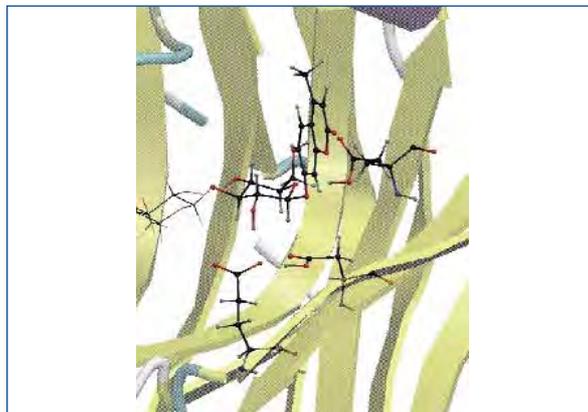
Close view the catalytic site of Bacillus 1-3,1-4-b-glucanase in complex with the substrate obtained by ab initio QM/MM computer simulations.

Publications:
X. Biarnés, J. Nieto, A. Planas, C. Rovira. "Substrate distortion in the Michaelis complex of Bacillus 1,3-1,4-ß-glucanase. Insight from first principles molecular dynamics simulations". Journal of Biological Chemistry. 281, 1432-1441 (2006).

X. Biarnés, A. Ardèvol, A. Planas, C. Rovira, A. Laio, M. Parrinello, "The conformational free energy landscape of ß-D-glucopyranose. Implications for substrate preactivation in ß-glucoside hydrolases". Submitted.

Carme Rovira i Virgili
Parc Científic de Barcelona

Unravelling the mechanism of drug activation in
Mycobacterium tuberculosis catalase-peroxidase by means
of first principles Car-Parrinello simulations

Catalase-peroxidases (CatPs) are bifunctional heme-
dependent enzymes belonging to the class I peroxidase family.
Despite the sequence homology, CatPs are the only members
of the family displaying catalatic activity, which dominates the
peroxidatic reaction. The Mycobacterium tuberculosis CatP is
a target of pharmacological interest, as it activates isoniazid
(INH), a drug used to treat tuberculosis. In particular, when
the enzyme decreases the activity, the resistance to the drug
increases. Therefore, the precise knowledge of the enzymatic
mechanism is of fundamental importance to understanding
both the activation process of the drug and the significance of
variations in the enzyme (mutations) which may be directly
responsible for INH-resistance. The project is intended to
model the enzyme mechanism and find the most likely
binding site for the INH drug in order to understand the
drug activation mechanism. To this aim, the project performs
several simulations with varying conditions (e.g. protonation
state of the residues in the binding pocket) in order to see
how they could affect the enzymatic activity through changes
on the electronic structure of the reactive species. The
analysis is performed in both monofunctional catalases (e.g.
Helicobacter pylori catalase) and bifunctional catalases
(CatPs). The intermediate reaction for monofunctional
catalases basically shows the typical oxoferryl pophyrin
radical, i.e. mostly delocalized on the porphyrin ring. However,
slight changes in the radical location are found for other
catalases, including CatPs, which could be related to
differences in enzyme activity. The calculations are performed
using the so-called Car-Parrinello molecular dynamics method,
as implemented in the CPMD program. The project usually
runs on 64-256 MareNostrum processors.

Thanks to the increasing supercomputing capabilities such as
MareNostrum, it is feasible to do projects that involve a large
number of atoms to be treated quantum mechanically. Only
by using a large number of processors can the project obtain
the results on appropriately, which is what makes the project
feasible.
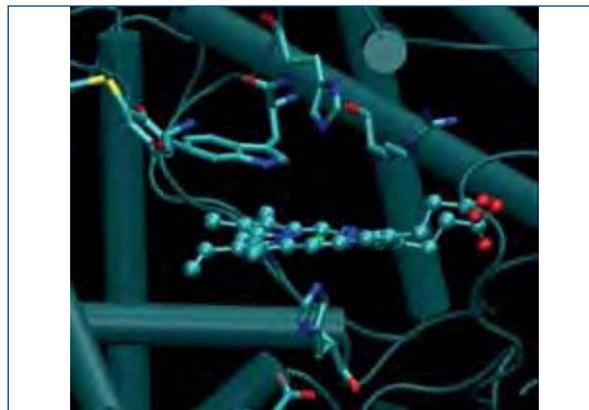
Molecular view of the active centre of the catalase-peroxidase enzyme.
The heme group is shown in ball and stick and the catalytic residues are
shown as sticks.

Publication:
C. Rovira, M. Alfonso, X. Biarnés, I. Fita, P. Loewen, "A first
principles study of the binding of formic acid in catalase
complementing high resolution X-ray structures". Chemical
Physics. 323, 129-137 (2006).

Alfonso Valencia
Centro Nacional de Biotecnología

Docking methods benchmarking and protein evolutionary information integratio

This project aims to conduct a large scale evaluation of several protein docking programs. This data will be very useful to design a rationale in order to integrate these results efficiently.

This information is plan to be used to further develop an efficient protocol for the integration of these methodologies with other methods built to predict protein interaction sites. These prediction methods analyze different aspects of the evolutionary history of interacting proteins such as conservation, co-evolution, among others.

The successful achievement of the aforementioned goal will provide a new tool for the prediction of structural protein-protein interaction models. This will obviously increase the available knowledge to understand most of the main cellular processes.

Protein interactions constitute the basement of most cellular processes. The study of such interactions is crucial to guide our understanding of cell systems. From the experimental side, the large-scale application of proteomic methods is contributing remarkably to increase the wealth of knowledge since those approaches are providing overviews of complete protein interaction networks.

In addition to the global picture, obtaining detailed structural information about these complexes is very important to understand the biochemistry behind the dynamic behaviour. Unfortunately, the number of structurally determined complexes that have been deposited in Protein Data Bank (PDB) is still rather small. Subsequently there is a necessity to develop alternative strategies to deal with this problem. In the absence of highly efficient High Throughput initiatives for X-ray determination, the most straightforward solution is to develop docking algorithms which might reproduce the physical interaction of proteins in protein complexes.

Since the group has been developing methods to predict protein interactions extracting evolutionary information from multiple sequence alignments, the team is interested in analysing how these sequence-based methods might work to score protein docking predictions. The rationale relies on the fact that if evolutionary information can be used to predict successfully interaction partners, then perhaps it could contribute in the successful prediction of protein interfaces between proteins which are known to interact.

While some success is obtained in this aforementioned combined docking-sequence procedure, there are obvious drawbacks with this approach. The obvious one refers to the proper docking methods as they provided a restricted number of solutions based on their own scoring schema. This consequently produces a small space of solutions to search. A second caveat is the lack of any global benchmarking of both the docking programs and the scoring methods used in protein-protein docking.

Nonetheless, as docking methodologies are becoming widely available, it should be possible to broaden the sample space by employing a range of different docking methods.

In this framework, the team aims to undertake a large-scale benchmark of two types of docking methodologies; the physical-based docking programs and the sequence-based methods. This will enable to evaluate the different docking algorithms using a range of target examples and to integrate a multiple sequence alignment-based approaches as filters of the docking models at this level.

Since the scoring methods were developed for predicting protein interactions, the team would also be interested in developing a protein-protein prediction pipeline integrated with additional docking data. Then, relevant evolutionary information for predicting interactions partners obtained from structures is taking into account to score physical region of interactions for docking solutions between proteins that are predicted to interact. To build a benchmarking system, the team started with an initial list of eight programs. However, due to a number of problems related to software licenses and hardware architecture incompatibilities, the analyses had to be limited to two different docking programs (FT-Dock and RosettaDock). Then, run the analyses over a set of 84 complexes extracted from PDB. Despite this, the team has obtained a library of about 110.000 docking predictions for each complex that are currently being evaluated. As a result of this evaluation the team will decide whether to further integrate these programs with these evolution-based methods or not.

The supercomputing capabilities of MareNostrum make possible to obtain a large scale evaluation of several computationally intensive protein docking algorithms.

Alfonso Valencia
Centro Nacional de Biotecnología

Systematic Evaluation of Prediction of Protein Pairings Between Interacting Families based on Protein Coevolution

This project aims to test a new method for predicting protein interaction specificities using multiple sequence alignments of the interacting protein families. This method will allow to enhance our knowledge about those protein families that have been partially characterized, but in which concrete protein interactions for each of their members remains unclear. This is a common problem in biological studies involving eukaryotic organisms that have sets of large protein families essential for the development of relevant processes (i.e. Ras/Ras effectors, chemokines/chemokine receptors). The team's method relies on the fact that protein interacting phylogenetic trees are more similar than expected.

Therefore, this method explores the mapping possibilities space in order to maximize the similarities among the distance matrices obtained from phylogenetic trees.

The advance in the last years regarding the study of molecular machineries such as complexes or metabolic pathways has been remarkably fruitful at the organism level. These studies are based on both experimental and computational techniques that allow increasing our knowledge on interesting features. Those important questions range from global issues such as the structure, evolution and behavior of the cell machinery, to more detailed and specific questions like protein and protein complex functions as well as their roles in cellular processes and diseases.

In this framework, the team has developed a method that predicts the protein interaction mapped in two protein families -known to interact with each other- based on the idea that the real mapping would be the one maximizing the similarity between the corresponding phylogenetic trees.

This method assumes that interacting protein families tend to have similar phylogenetic trees. The rationale behind this has been shown before and was previously used to deal with this problem. A feasible hypothesis to explain the relationship between interaction and tree similarity states that interacting partners are somehow forced to adapt to each other. This process of co-adaptation should lead to correlated evolutionary histories, which in turn should be reflected in a tree similarity higher than expected.

The team's aim was then to provide a new fully available research tool (a server) to produce some insights about protein interaction specificity in interacting protein families. The tool makes use of the aforementioned principles.
It must be mentioned that due to the combinatorial nature of the problem, the exhaustive exploration of all possible mappings between two sets of proteins (when searching the one maximizing the tree similarity) is unfeasible. To circumvent this problem, the team has implemented a Monte Carlo approach to perform a guided exploration of the space of solutions. This will avoid the exhaustive exploration of all the possibilities. As an improvement from previous methodologies, the method has been designed to deal with different number of sequences per families, which is a general limitation of the previous methods.

As a general rule, obtaining a proper set for large scale analyses is computationally intensive. In this regard, one of the team's main concerns was to test the ability of this approach to solve this optimization problem. To do so, the team has analyzed different parameter settings on large test sets. The group performed the tests on two different data sets, the first one was a set of 115 multiple sequence alignments containing groups of orthologous proteins. Alignment pairings were based on experimentally known protein interactions from Escherichia coli. The second set was composed by 150 domain multiple sequence alignments from the Pfam database. In this case the alignments were paired if they contain different domains of the same proteins.

In both cases the predictions were compared to self-pairings runs (i.e both alignments being the same). As a result of these evaluations the team established the optimal settings for this algorithm and also has learnt the main limitations of the methods. The group found that most of the wrong pairings were due to ambiguities in the phylogenetic trees of the studied families. For example, if two proteins have the same evolutionary distances to rest of the family's proteins (they are identical or they have diverged at the same ratio from their splitting point) the co-evolutionary principle cannot solve this ambiguity. In a similar way, low informative trees (very symmetrical) lead to ambiguous assignments. Moreover, when the right pairing was not the best solution the automatic method couldn't get it.

In order to improve the reliability of the tool the team has calculated previously the Shannon's entropy for each of the trees and advice the user of possible low-entropy cases. The team also display the pairs that can't be distinguished and the team has also implemented an interface allowing manual testing of alternative pairings.

The supercomputing capabilities of MareNostrum make possible to obtaine a large scale evaluation of an heuristic approach for exploring all the possible pairing among putative protein interaction parners among two protein families.

Jordi Villà Freixa
Institut Municipal d'Investigació Mèdica

ByoDyn: parameter estimation of multicellular models of gene regulatory networks

Integrating multilayer biological/biomedical information is setting the ground for what is known as biomedical informatics. This field is partially based on systems biology approaches, where the research focuses on systemic rather than reductionist views. Within this framework, ByoDyn is an open source computational package aimed at studying the dynamical behaviour of small to massive biochemical networks (input as SBML files). The model can be simulated, the sensitivity of the system with respect to the parameters can be analysed and, more relevantly for the project, kinetic parameters can be estimated using experimental time course data. The project carried out on the MareNostrum aims at developing a parallel framework for the optimization techniques implemented in the programme. Thus, MPI tools are being developed to parallelize the critical aspects of the hybrid optimization algorithms used to find the optimal sets of parameters that fit a given set of experimental data.

The research has focused on the calibration of biological models of the Notch/Delta multicellular gene regulatory network and of the complete set of models in the BioModels database at the EBI. Starting from experimental data for key nodes of the networks and mathematical models of their interaction the team solves the so-called inverse problem using an algorithm that aims at minimizing the Euclidean distance between the experimental measurements and the simulated results. New algorithms are being implemented and tested in the computer software ByoDyn to improve the performance of the optimization protocol, being the ultimate goal of this algorithmic improvement the analysis of the parameter sets controlling highly complex biochemical networks.

The team's research is focused on the calibration of biological models of gene regulatory and metabolic networks. After building a kinetic model using a set of non-linear ordinary differential equations (ODEs) the team uses experimental time series measurements from available key nodes of the model. The experimental data is then included into an inverse problem solver algorithm that aims to minimize a fitness function that describes the distance between the experimental measurements and the results of the simulation.

The team has implemented an improved strategy for the inverse problem solution that succeeds to lower the fitness function three orders of magnitude lower than previous methods. It is based on a hybrid combination of local and global parameter estimation routines.

The development of this type of new algorithms is a great step forward towards the calibration of much complex biological gene regulation networks.

Plans for future developments are the improvement of the specific parameter estimation approaches. With respect to global methods the team will incorporate multivariate statistical tools to steer the optimization path at the high dimensional space. As for the local optimization module, faster convergence routines are under development.

The inmediate outcome of the ByoDyn software is to be used in the context of the new EU-funded STREP project BioBridge, which aims at developing a simulation environment able to relate omics and clinical information on COPD in an integrated way.

Given the number of parameters that need to be inferred in the networks, the estimation of their optimal values makes use of high performance computing resources. However, the team does not base the results just on simple parameter sweep, which can be done in a typical highly distributed computational environment like the Grid. Instead, the team is developing the communication framework that will allow the parameter estimation tools to make extensive use of tightly coupled parallelization environments like the one provided in MareNostrum. These new paradigms are being developed in the framework of the new QosCosGrid EU-funded project.
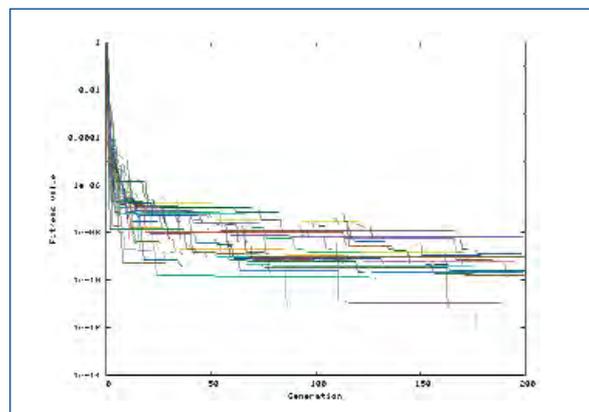
Evolution of the optimization. Vertical axis shows the value of the fitness function the team wants to minimize. Horizontal axis represents the number of iterations. Different starting points have been chosen. After 200 iterations the best algorithm arrives to a solution of fitness value of 1e-12, three orders of magnitude better than previous methods.

Publications:
"General report on the requirements for casting each use cases into a quasi-opportunistic framework WP 3: CS Simulations on the Grid (use case scenarios)". The QosCosGrid consortium. Manuscript in preparation.

Adrian L. Garcia-Lomana, Gianni de Fabritiis, Gina Abelló, Berta Alsina, Jordi Villà-Freixa. "Parameter estimation in complex gene regulatory networks during chicken embryo development with BYODYN". Manuscript in preparation.

Adrian L. Garcia-Lomana and Jordi Villà-Freixa "Strategies for reverse engineering of an oscillatory gene regulatory networor: the repressilator model". Manuscript in preparation.

Àlex Gómez, Adrián L. Garcia-Lomana, Jordi Villà-Freixa "Parameter estimation for biochemical models in a variety of parameter landscapes". Manuscript in preparation.

Jordi Villà Freixa
Institut Municipal d'Investigació Mèdica

Parameter estimation and optimal design in computational systems biology

The team's research has focused on the calibration of biological models of the Notch/Delta multicellular gene regulatory network and of the complete set of models in the BioModels database at the European Bioinformatics Institute (EBI). Starting from experimental data for key nodes of the networks and mathematical models of their interaction the team solves the so-called inverse problem using an algorithm that aims at minimizing the Euclidean distance between the experimental measurements and the simulated results. New algorithms are being implemented and tested in the computer software ByoDyn to improve the performance of the optimization protocol, being the ultimate goal of this algorithmic improvement the analysis of the parameter sets controlling highly complex biochemical networks.

Integrating multilayer biological/biomedical information is setting the ground for what is known as biomedical informatics. This field is partially based on systems biology approaches, where the research is focussed on systemic rather than reductionist views. Within this framework, ByoDyn is an open source computational package aimed at studying the dynamical behaviour of small to massive biochemical networks (input as SBML files). The model can be simulated, the sensitivity of the system with respect to the parameters can be analysed and, more relevant for the project, kinetic parameters can be estimated using experimental time course data. The project carried out in the BSC aims at developing a parallel framework for the optimization techniques implemented in the program. Thus, MPI tools are being developed to parallelize the critical aspects of the hybrid optimization algorithms used to find the optimal sets of parameters fitting a given set of experimental data.

Wolfgang Wenzel
Research Center Karlsruhe (Germany)

De-novo protein folding and structure prediction with free-energy forcefields

Predicting the three dimentional structure of a protein, given the sequence of amino acids is one of the central problems in biophysical chemistry. The group has developed an effective free energy forcefield PFF02 which locates the native state of a protein at its global minimum. The team has also developed a scalable evolutionary algorithm for the search of the minimum. Here PFF02 has been used as well as evolutionary algorithm to reliably predict the three dimentional structure of proteins at an all atom level.

Three different proteins have been folded, Tryptophan zipper (1LE0), E6 Binding Protein A (1EDK) and Engrailed Homeodomain (1ENH) protein using the modified free energy forcefield (PFF02) and a scalable evolutionary algorithm. While tryptophan zipper is a beta hairpin protein, protein A and the engrailed homeodomain are up-down and orthogonal helical bundles respectively. The results indicate the modified forcefield is more versatile in handling proteins of different kinds. Secondly the scalable evolutionary algorithm has proved to be a better strategy for larger size proteins overcoming the freezing problem of basin hopping simulations. Tryptophan zipper (14 amino acids) folds in just 11 minutes using 128 processors, Protein A (56 amino acids) using 256 processors for 72 hours and Engrailed homeodomain (54 amino acids) in 30 hours using 512 processors. These results encourage the team to predict the tertiary structure of even larger proteins and to also understand the thermodynamin properties of the folding process. The scalability of the algorithm allows the usage of even larger number of processors for the ambitious studies.

Thanks to the increasing massively parallel supercomputing facilities such as MareNostrum, it is feasible to make an extensive sampling of the protein energy landscape using evolutionary algorithm to locate the global minimum of the energy landscape of various proteins. The combination of large number of processors available at MareNostrum along with the scalability of evolutionary algorithm makes it an ideal combination for protein structure prediction.
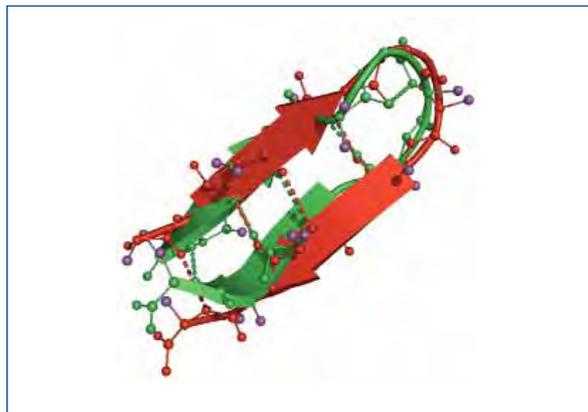
Overlay of predicted (red) and native (green) structures for Tryptophan Zipper (1LE0) and Engrailed Homeodomain (1ENH) protein.

6.2 Astronomy, Space and Earth Sciences

Irene Arias
Universitat Politècnica de Catalunya

Large-Scale Simulations of Dynamic Fracture

Dynamic fracture is ubiquitous in nature, yet its fundamental physics are poorly understood, particularly when it comes to crack branching and fragmentation processes. The goal of the project is to further understand the physics of high-energy dynamic fracture and massive fragmentation through large scale finite element simulations. Supercomputing is required to resolve in a single computation the enormous scale disparity between the processes at the crack tip and the macroscopically relevant sample sizes. In the first set of runs on MareNostrum, a scale disparity of 4 orders of magnitude has been reached. This allows the team to extract preliminary, physically meaningful scaling laws that characterize this complex phenomenon.

The dynamic fracture and fragmentation of brittle solids exposed to high energy loadings has attracted much attention due to its technological and scientific interest, as well as its intriguing complexity. It governs very diverse phenomena such as the response of brittle solids to impact and explosive loadings, quarry blastings or meteorite showers. Although ubiquitous in nature, the fundamental physics of dynamic fracture are still poorly understood, particularly when it comes to crack branching and fragmentation processes. Consider for instance how a light bulb breaks into peaces when it is dropped, or the intricate crack patterns in the windshield of a car when impacted by a small object. It seems very difficult to reproduce such an event; indeed, the shapes, sizes and number of pieces of glass will not be exactly the same even if we try hard to perform a controlled laboratory experiment. Despite this chaotic nature, there seems to be some underlying well-defined physical mechanisms that prevent the bulb from breaking into just two or three peaces, or become completely pulverized. The structured branched crack patterns of the windshield suggest even strongly the presence of fundamental physical mechanisms.

A large body of experimental work has shown that indeed this is the case. By statistical analysis of fragmentation experiments, researchers have demonstrated that the fragments size distributions follow robust power-law statistics in a very wide range of sizes (several orders of magnitude). In addition, the exponents in these power-laws are independent of the details of the experiment such as the material or the loading conditions. In the field of statistical physics, power-laws are very important objects because they often reveal universal physical laws, which are self-similar and scale-invariant. However, the experimental challenge of dynamically observing high-speed fragmentation events has limited these investigations to the post-mortem scenario. Computer simulations appear as an ideal tool to analyze the physics of the dynamical process leading to the well-characterized post-mortem statistics. In particular for applications, it is very important to understand how energy is dissipated in fragmenting brittle materials. Such a computer simulation is more easily stated than performed, as a method capable of describing the nucleation of cracks, their dynamical propagation, and complex changes in crack topology such as branching and merging, is required. In addition, the huge scale disparity between the small-scale physical processes at the crack tip and the macroscopic events at the sample scale must be resolved in a single simulation. This inevitably leads to supercomputing.

The goal of the project is to investigate the physics of high-energy dynamic fracture and massive fragmentation through large-scale finite element simulations. These simulations at MareNostrum will allow to understand the full dynamical process, rather than just the postmortem picture. The statistical analysis of the acoustic emission signals generated in simulations of the dynamic fracture and fragmentation of a plate subjected to high initial strain rates reveals a self-similar behavior also in time. The implications of these preliminary results are under investigation.

Given the extreme difficulty of performing full field, real-time diagnostics on massive dynamic fracture and fragmentation experiments, computer simulations emerge as the only viable approach to these phenomena. The huge scale disparity, from the dimensions of the system to the characteristic length-scales of the fracture processes, together with the long simulation times required, inevitably leads to supercomputing platforms such as MareNostrum. The team is resolving four orders of magnitude in the scale disparity, which is essential if meaningful conclusions are to be extracted.
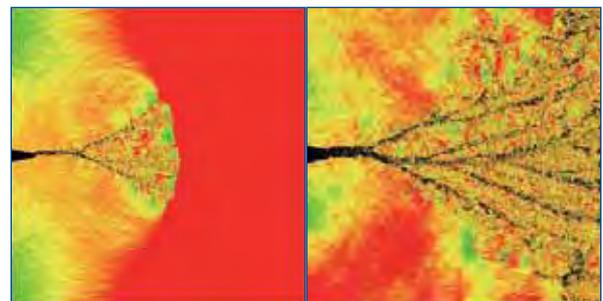
Dynamic fracture process of a 2D brittle plate subjected to an initial strain rate in the vertical direction: dynamic crack propagation, branching and fragmentation.

Publication:
I. Arias, J. Knap and M. Ortiz, "On energy dissipation through massive crack branching", 2006. In preparation.

Marc Balcells
Instituto Astrofísico de Canarias

Galaxy Transformations through Interactions, Mergers and Accretion

Collisions and subsequent merging of galaxies were very common in the first half of the Universe, and still occur today. They shape galaxies, and affect the rate at which galaxies form stars and grow. With N-body methods the project simulates the gravitational evolution of merging galaxies; predict observable properties of such galaxies, and validate or rule out specific hypotheses. Much of the work deals with the so-called minor mergers, in which a big galaxy swallows a 10- to 50-times smaller companion. Recent efforts focus on the kinematic structure of the nuclei of merged galaxies.

The team has used MareNostrum to run N-body gravitational models of galaxy mergers at high spatial resolution. The team's goal is to investigate specific aspects of the hypothesis that elliptical galaxies may be the result of merging disk galaxies. This problem has been addressed many times in the past decades. However, results have systematically suffered from poor resolution due to limitations in computational power. Gravitational systems differ from other many-particle systems such as liquids and gases in that gravitational forces have infinite range; moreover, the dynamic range of parameters such as mass density is also extremely high in galaxies. This demands a sophisticated mapping of the forces as well as the mass distribution. Poor modelling of the mass density distribution of galaxy cores may lead to poor rendering not only of the properties of those cores but of the entire galaxy. In this context, the team has studied the line-of-sight velocity distributions (LOSVDs) of stars inside galaxies formed by mergers. It had long been recognized that LOSVDs were potentially important diagnostics on the types of mergers that lead to elliptical galaxies, owing to the characteristic asymmetries observed in LOSVDs of real ellipticals, which are defined by steep leading edges and longer trailing wings. Previous studies with N-body merger models found that mergers led to the exact opposite behaviour: an extended leading wing and a steep trailing edge.

The team used MareNostrum to increase the number of particles in the models by a factor 5, and then 10, with respect to previous studies. These high-resolution merger models yield a LOSVD asymmetry with the same sign as that observed in real elliptical galaxies. The difference with previous models is interpreted as due to the lower resolution of those models. Low-resolution models yield a higher proportion of box-orbits, box orbits show no net rotation, and naturally yield steep trailing sides in the LOSVDs. The high-resolution of the MareNostrum models more accurately mimics the gravitational potential of galaxy cores; such 'cuspy'

cores prevent box orbits, and instead favor z-tube orbits. The latter yields the type of LOSVD asymmetry as observed in real galaxies. In collaboration with V. Quilis (U. Valencia) the team has investigated the emission of gravitational radiation produced in encounters of dark matter galactic halos.

Gravitational radiation will become one of the main targets for astrophysical research in the near future with the deployment of experiments such as LISA, LIGO, CLOVER and others whose aim is to detect such radiations. Although compact objects such as black holes are the most promising targets for such experiments, the gravitational radiation emitted by the dark matter of galaxies in galaxy encounters may provide an unambiguous signature of the presence of such elusive matter. The simulations performed show that the frequencies of the gravitational waves produced by galaxy mergers could be detected by some of the planed future missions, and that the amplitude is comparable to that coming from primordial waves originated in the early Universe.
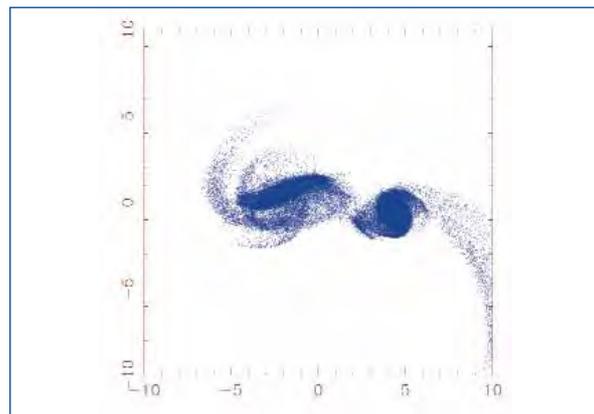
Shown after a first close approach, both members in this unequal pair of orbiting disk galaxies (3:1 by mass) grow strong distortions. The tail-like arcs are due to resonant reactions of the two disks to the gravitational potential of the other galaxy; stars in the tails will eventually fall back onto the central body. The galaxies will merge into an object that resembles an elliptical galaxy. Thanks to the superb spatial resolution of this MareNostrum simulation (the model uses over 10^6 particles) the team obtained an accurate prediction of the velocity distribution of stars in the galaxy nuclei.

Publications:
Gonzalez-Garcia A.C., Balcells M., Olshevski V.S., 'Line-of-sight velocity distributions of elliptical galaxies from collisionless mergers', 2006, MNRAS, 372, 78

Quilis V., González-García A.C., Font, J.A., 'Gravitational Waves from Galaxy Encounters', PRD, submitted.

Emilio Cuevas
Instituto Nacional Meteorología (OAI)

MM5 and WRF-ARW Sensibility Study in Tropical Storm Delta Meteorological Environment

The proposed activity for MareNostrum is to test the sensibility of MM5 and WRF-ARW models in the meteorological environment associated with the Tropical Storm Delta and validate the results using the available INM observational data. Access to MareNostrum would allow this sensibility analysis to be conducted in an efficient manner in a reasonable time-frame which would not be possible on other computing platforms. In addition, the experience of the Barcelona Supercomputing Center in MM5 and WRF-ARW would contribute to the success of this exercise. As a result of the sensibility analysis the optimal configuration of MM5 and WRF-ARW and their boundary conditions will be selected.
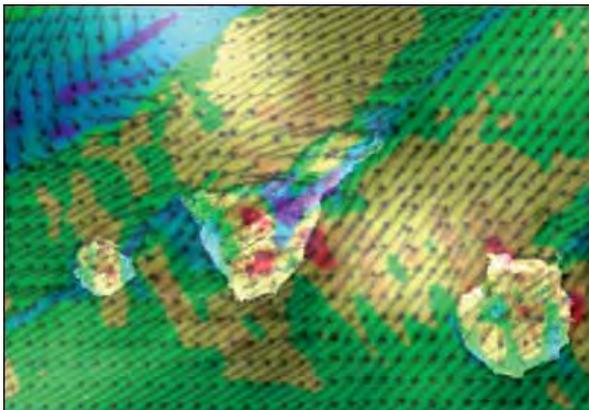
Preliminary Wind Simulation using WRF-ARW model in Tropical Storm Delta conducted by the Earth Sciences Department Team of BSC

Pablo Fosalba
Universitat Autònoma de Barcelona

Large numerical simulations for dark-energy surveys

The evolution and ultimate fate of the universe is determined by the so-called dark-energy, which represents about 75% of the total energy-density of the universe today. IEEC (ICE-CSIC) has undertaken a long-term project to understand the properties of this form of energy through its imprints on the large-scale distribution of matter and radiation in the universe.

In the project, the team is developing the largest N-body simulations to date to build mock galaxy catalogues for the new generation of surveys, such as the Dark-Energy Survey in which the group participates along with other Spanish and international institutions.

The project involves simulating 10,000 million dark-matter particles filling the entire observable universe and following its gravitational evolution immediately after the big bang to present day. This poses a serious computational challenge. For this project, the MareNostrum supercomputer used up to 1000 computers in parallel to achieve 1 TeraByte of memory and several Terabytes of disk storage, which is a requirement for developing these cutting-edge numerical simulations.

The universe is observed to expand at an accelerated rate, what is thought to be caused by the so-called dark-energy, a form of energy effectively acting as a gas "pressure" that dominates gravity and determines the fate of the universe.

The team is carrying out the largest dark-matter simulations to date to understand the role of dark-energy in the evolution of the universe. In particular, the group proposes to measure the effect on the spatial and time distribution of matter, galaxies and clusters of galaxies, and how this information will be used by upcoming galaxy surveys, such as DES and WFMOS, to measure dark-energy properties with unprecedented precision.

The lightcone simulation produces a realization of how the observer would see the universe. Here we present a figure of the closest volume (64 Mp/h size) to the observer who is in the right bottom corner.

Rolando R. García
University Corporation for Atmospheric Research

Simulation of changes in atmospheric climate and chemical composition 1950-2050

This project consists of two parts:

(1) Simulation of the period 1950-2003. The comparison of model results and observations, in particular for the period since 1980, which is very well covered by both conventional and satellite observations, is both a test of model performance and a means of interpreting the physical/chemical mechanisms that have led to the changes observed from 1950 to date.

(2) Prognostic simulation up to 2050. This simulation covers the years 1980-2050, with the period 1980-2000 being used as a control against observations, and will be used to assess the future state of the atmosphere as ozone-depleting compounds are reduced (leading to ozone recovery), while greenhouse gases continue to increase (leading to further warming in the troposphere and cooling in the middle atmosphere).
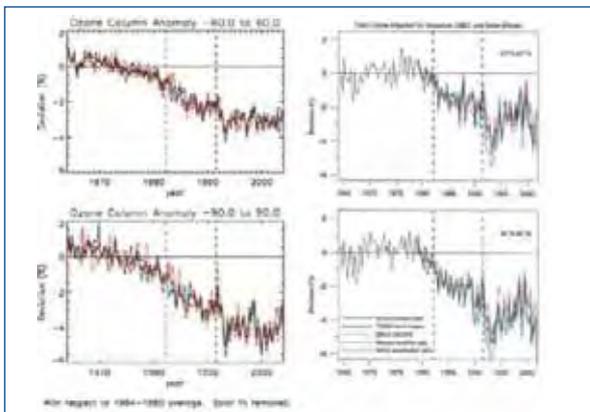
Evolution of zonal-mean ozone column anomalies (percentage change averaged over ±60° and ± 90°) for each member of the model ensemble (left two panels) compared with the anomalies derived from various observational datasets (right two panels). Both model results and data are smoothed with a 3-month running average, and the percentage anomalies are calculated with respect to the average column values for the period 1964–1980. The dashed lines in each panel denote the dates of the eruptions of El Chichón and Mt Pinatubo.

Enrique García Berro
Universitat Politècnica de Catalunya

Smoothed Particle Hydrodynamics in Astrophysics

The team has computed the gravitational wave emission arising from the coalescence of binary systems composed of a white dwarf and a neutron star. In order to do so, the team has followed the evolution of such systems using a *Smoothed Particle Hydrodynamics* code. Special attention has been paid to the detectability of the emitted gravitational waves. Within this context, the project shows which could be the impact of individual merging episodes for LISA. In another field of research the team has studied the coalescence of binary white dwarfs. Binary white dwarfs are assumed to coalesce as the result of the emission of gravitational waves. Therefore, the merger of white dwarfs is the final destiny of the evolution of a large fraction of binary systems. However, due to the very long timescales associated to the emission of gravitational radiation such events are rare. The expected galactic rate is ~ 10^{-2} yr^{-1}. The merger of double white dwarfs has also been invoked as a plausible scenario for the formation of thermonuclear supernovae, for explaining the origin of some massive halo white dwarfs and for explaining the anomalous spin down rate of some pulsars. Here the team reports theoretical and observational evidences that the massive white dwarf GD 362 is likely the result of the merger of a binary white dwarf.

The coalescence of a two binary systems have been computed. The first one is composed of a white dwarf and a neutron star. The second one is double white dwarf binary system. As a result of the calculations performed at the BSC there a couple of scientific papers currently submitted for publication in the most prestigious journals in this field and it's expected to submit a third one shortly. The first case studied (the merger of a white dwarf and a neutron star) is the very first calculation of this process, whereas the second one (the merger of a binary white dwarf) was previously computed using a small number of SPH particles and, hence, the resolution was very poor. These two simulations are leading examples of the capabilities of the so-called particle methods in which the fluid is described by means of a collection of particles of finite size. Thus these methods are grid-less and are especially adequate to describe intrinsically three-dimensional astrophysical processes, such as the ones the team is interested in. Moreover both kinds of binary systems are powerful emitters of gravitational wave radiation. The detection of gravitational waves is a formidable task given their intrinsic weakness and the current generation of terrestrial detectors and future space-borne observatories, like the joint ESA/NASA mission LISA, sorely need reliable templates of gravitational wave emission to which compare the huge amount data needed to detect them. It is worth

noting that the detection of gravitational waves is the last of the predictions of Einstein's Theory of General Relativity which still today remains to be observationally confirmed by direct means. Consequently, the detection of gravitational waves is not only of inherent interest but it will also open a new (and probably the last) observational window in modern astrophysics.

The extensive simulations carried out at the BSC-CNS premises have allowed the team to use a large number of SPH particles to describe the fluid, which otherwise would be impossible to achieve with the current computational resources. This, in turn, has allowed the team to obtain a better insight of the three-dimensional dynamics of the coalescence and to obtain very clean gravitational wave radiation templates.
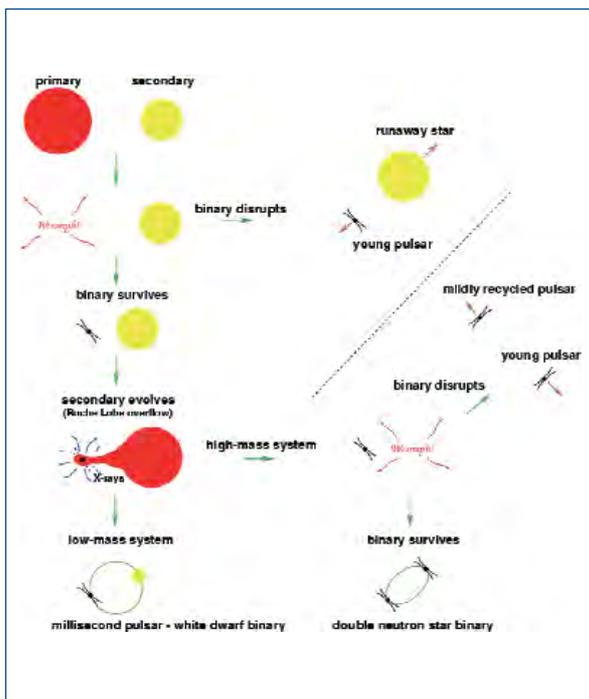
Evolutionary channels

Publications:
"The merging of white dwarf and neutron star systems: gravitational radiation", Journal of Physics, 2006.

Víctor Homar
Universitat de les Illes Balears

Application of ensemble prediction techniques to high impact weather episodes in the Western Mediterranean el Mediterráneo Occidental (ENSEMBLE)

ENSEMBLE is a research project dealing with the challenge of high impact weather prediction in the Western Mediterranean. The project suggests using ensemble prediction systems (EPS) to tackle this problem in the Western Mediterranean. The generation, interpretation and application of short-range mesoscale EPS is still currently explored. ENSEMBLE suggests exploring a set of EPS generation techniques focusing on extreme events in the Western Mediterranean, with special emphasis on techniques that use sensibility fields from lineal adjoint models. As an additional benefit from the proposed project, a climatology of high impact weather sensitivity fields will be built.

Jose María Ibáñez
Universidad de Valencia

Stability of extragalactic jets

The study of the stability of extragalactic jets is a powerful tool to understand the physics and derive the physical parameters involved in extragalactic jets from Active Galactic Nuclei. Work on this topic has been mainly done in two dimensional simulations, where high numerical resolutions have been shown to be needed in order to reproduce successfully the growth of small instabilities. These requirements make very difficult fully three dimensional simulations without the use of high performance computers. Full parallelization of a three dimensional hydrodynamic code with MPI libraries has been performed in the Barcelona Supercomputing Center (BSC). This will allow not only to study the stability of relativistic jets in three dimensions but also to perform simulations on the long term evolution of jets in their motion through interstellar and intergalactic media.
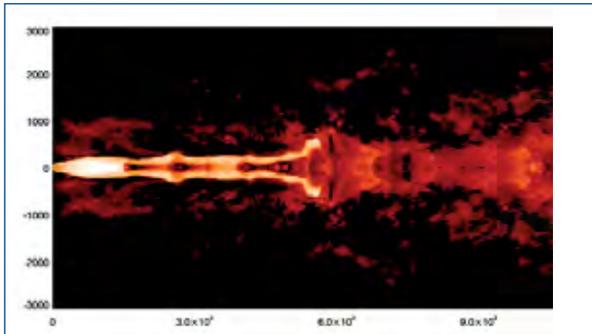
Snapshot of Lorentz factor from a long term simulation at t = 7:2 106 yrs. The color scale stands for the values of the Lorentz factor, ranging from 1 (black) to 5 (white). Coordinates are in parsecs.

Fernando Moreno Insertis
Instituto Astrofísico de Canarias

Eruptive phenomena in the atmosphere of the sun and cool stars

In the project the team is using three different, large parallel magnetohydrodynamic codes adequate to run in massively parallel architectures. The team implemented in the MareNostrum one of those codes. The implementation involved a number of optimizations; it also included extensive scaling tests. The conclusion, for large computational grids, like the ones needed for these experiments, is that code shows excellent scaling in the MareNostrum up to 512 CPUs.
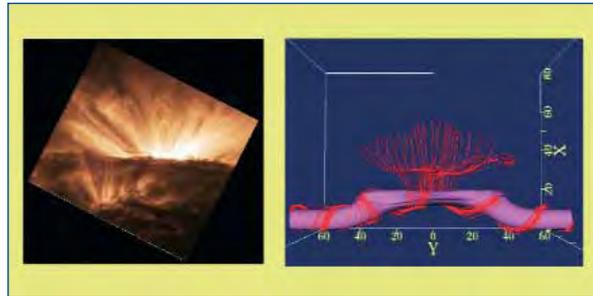
(left) image taken by the solar space mission TRACE showing very hot gas that delineates magnetic field lines in the solar corona (outer part of the solar atmosphere).
(right) snapshot from a numerical simulation of magnetic flux emergence into the solar atmosphere showing the magnetic field line distribution; the uppermost field lines, located in the corona, strongly resemble the satellite image.

Josep María Solanes
Universitat de Barcelona

Investigation of the DIffuse LIght COmponent in compact groups of galaxies (IDILICO)

The IDILICO project is an international collaboration involving researchers from Universitat de Barcelona, Universitat Rovira i Virgili, and Instituto de Astrofísica de Andalucía in Spain and Observatoire Astronomique de Marseille-Provence in France. IDILICO is a basic research project specifically oriented towards studying the role of gravitational interactions on the evolution of galaxies in the highest density regions of the universe, known as compact groups of galaxies.

Preliminary simulations performed with the MareNostrum supercomputer have shown that compact groups of galaxies (CGGs) of the local universe might come from small galaxy systems that are just now collapsing, and whose dynamic evolution is mainly driven by mergers that selectively destroy intermediate mass galaxies. This scenario could provide a sensible explanation for both the observed bimodal nature of the luminosity function of CGGs and the origin of the large amount of diffuse visible light found recently in the intergalactic medium of some of these systems.

The IDILICO project aims to use the superb capabilities of the MareNostrum supercomputer to seek an explanation for the origin of the bimodal luminosity function characteristic of low-velocity dispersion compact groups of galaxies. The team also expects to gain insight into the causes of the surprisingly large amount of diffuse visiblelight detected in the intergalactic medium of some of these systems.

The group will attack these problems by following in detail the gravitational collapse of an small group of galaxies with N-body simulations that, for the first time ever, will comprise up to 100 million particles in total. This will allow to model the simultaneous evolution of dozens of cosmologically consistent disk and elliptical galaxies with a broad range of luminosities embedded on a background of dark matter.

The scientific objective of the IDILICO Project is specifically oriented towards studying the role of gravitational interactions on the evolution of galaxies in the highest density regions of the universe, known as compact groups of galaxies.

The activity program is focused on the numerical study of galaxy and larger structure formation in a cosmological framework by means of demanding and innovative simulations that require large, highly scalable parallel computational resources. Preliminary simulations done with the MareNostrum supercomputer appear to support the idea that the compact groups of galaxies observed in the local universe descend from small galaxy systems (~30 initial members) that are just now collapsing, and whose dynamical evolution has been mainly driven by mergers that cause a selective destruction of the galaxies. This scenario could provide a sensible explanation for both the observed bimodal nature of the luminosity distribution in these systems and the origin of the large amount of diffuse visible light found recently in the intergalactic medium of some of them.

Supercomputers such as MareNostrum are fundamental to deal with complex, innovative, challenging simulations in the field of Astronomy that would not be possible without the BSC infrastructure. The highly specialized nature of this machine will boost in a few years high-level expertise in the areas of computational astrophysics, applied mathematics and computer science, while helping to build at the same time a strong theoretical knowledge in Astrophysics and Cosmology.

NASA's Hubble Space Telescope optical image of a compact group of galaxies known as Seyfert's Sextet, representative of the sort of galaxy systems studied by the IDILICO project. The galaxies are so tightly packed together that gravitational forces are beginning to rip stars from them and distort their shapes in a slow dance of destruction that will last for billions of years. In fact, one of the sextet members isn't a galaxy at all but a long "tidal tail" of stars (below, right) torn from one of the galaxies. The small face-on spiral with the prominent arms (center) is also excluded from the action: it is a background galaxy almost five times farther away than the other four aligned by chance with the group.

Publications:
Toribio, M.C., Athanassoula, E., Bosma, A., García-Gómez, C., Lambert, J.C., del Olmo, A., Perea, J., and Solanes, J.M. "IDILICO: Studying the properties of compact groups of galaxies with a supercomputer". Proceedings of the XLIst Rencontres de Moriond, XXVI Astrophysics Moriond Meeting: 'From dark halos to light', eds. L. Tresse, S. Maurogordato, & J. Tran Thanh Van (Gif sur Yvette: Éditions Frontières), (2006)

Solanes, J.M., Athanassoula, E., Bosma, A., García-Gómez, C., Lambert, J.C., del Olmo, A., Perea, J., and Toribio, M.C. "Supercomputer Simulations of Compact Groups of Galaxies: the IDILICO project" 'Nuevos Retos de la Astrofísica Española', Proceedings of the VII Reunión Científica de la SEA, in preparation (2006)

Nikolaos Stergioulas
Aristotle University of Thessaloniki (Greece)

R-mode Frequencies for Relativistic Stars in LMXBs

The project's aim at computing precise oscillation frequencies for r-modes in a large number of rotating neutron star models. This survey will constrain the possible range of frequencies of gravitational waves from the r-mode instability in accreting neutron star binary systems (LMXBs). The results will be used in ongoing data-analysis effort at the new gravitational-wave observatories GEO600, LIGO, and VIRGO. Precise oscillation frequencies are computed for r-modes in a large number of rotating neutron star model. This survey constrains the possible range of frequencies of gravitational waves from the r-mode instability in accreting neutron star binary systems (LMXBs), which has been found to be a promising source of gravitational waves. The results will be of high significance in the currently ongoing data-analysis effort at the new gravitational-wave observatories GEO600 and LIGO (and soon VIRGO). If successful, the detection of gravitational waves from r-modes in LMXBs will be a significant breakthrough in both experimental and theoretical physics and it will open a new observational window for studying physics in strong gravitational fields.
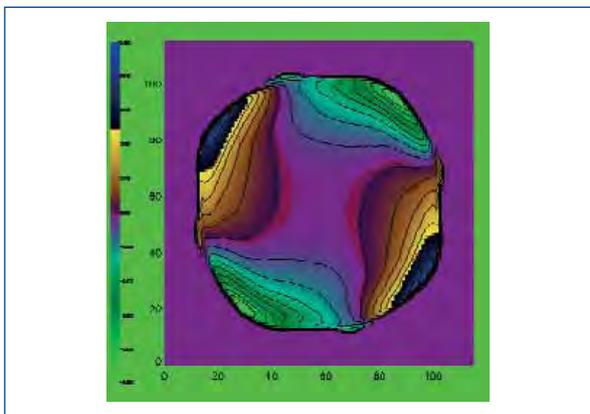
Contour plot of the r-mode velocity field in the equatorial plane of a rapidly rotating relativistic star.

Publication:
Tziampazlis, V, Stergioulas, N., Font, J.A., "R-mode Frequencies for Relativistic Stars in LMXBs", in preparation.

Romain Teyssier
Commissariat à l'Énergie Atomique (CEA).
HORIZON Consortium

HORIZON@MareNostrum: Galaxy formation in a cosmological context

The Horizon Project is a French consortium gathering around 20 scientists in 6 different institutes. More information can be found at www.projethorizon.fr. It has been funded by a three years Grant starting January 2006 by the French Research Foundation (ANR). The scientific objective is specifically oriented towards studying galaxy formation in a cosmological framework. Its transverse and federative nature will however allow developing in a few years high-level expertise in parallel and distributed computing, in database management and virtual observations, in applied mathematics and computer science, and build in the same time a strong theoretical knowledge in astrophysics.

The simulation the Horizon team performs on MareNostrum follows the evolution of dark matter, whose dynamics is describe by the Vlasov-Poisson equations thanks to a "Particle Mesh" scheme, and the evolution of baryons, the gaseous component, whose dynamics is described using the Euler equations. This last component is also subjected to dissipative processes, such as shock heating and atomic radiative cooling/heating.

The Horizon team was able to visualize the data, discovering an impressive level of details. They reminded the team of the first very large N body simulations, such as the one performed in 2000 by the Virgo consortium, except counting with a precise description of gas dynamics, star formation and other complex physical processes. The large-scale density distribution (shown in Figure 1), is typical of self-gravitating Gaussian random fields, with a filamentary structure often referred to as the "Cosmic Web". It is worth mentioning that these filaments are in quasi pressure equilibrium, so that, in this respect, the MareNostrum simulation can be considered as "converged". These filaments constitute the "Lyman alpha" forest that appears in the distant quasars' spectral energy distribution. The current epoch of the simulation (redshift 4) is however slightly to young to be able to compare directly with observed quasars' spectra, so the team will need to wait for further snapshots to do so.

At smaller scales, it is possible to discover in this simulation one very important aspect in galaxy formation, namely the accretion flows around forming galaxies. Figure 2 shows a zoom on one massive halo where several galaxies are visible as dense disc-like features in the upper image. In the same figure, lower images, one can see the gas temperature, with galaxies appearing as cold knots in the image. Clearly visible around the

halo are elongated filaments, often referred to as "cold stream", which are feeding the central galaxies with fresh gas. This gas will be processed into stars by the sub-cell model for star formation occurring in the galactic discs. In between these filaments, one sees hot shocks, typical of large haloes such as the one shown in the image. What can be seen here is the transition from the "cold mode" to the "hot mode" of accretion, bimodality believed to be at the origin of the "blue" versus "red color" dichotomy in the present-day galaxy population. The team has been able to observe such a phenomenon in the MareNostrum simulation, because of a) the large box size, allowing for "rare events" to occur in the simulated volume and b) the good shock-capturing properties of the AMR scheme (Godunov method) that allows for a accurate treatment of shock heating. Another interesting plot is shown in Figure 3. It is a mass weighted-histogram of the gaseous component in the 7 hydrogen density versus temperature plane. This histogram is widely used in cosmology to describe the thermal history of baryons in the universe. One clearly shows 3 different phases: a cold and diffuse background, in the lower left part of the diagram, also known as "the intergalactic medium", which correspond to a large reservoir of fresh gas from which galactic discs accrete new material. On the lower-right part of the diagram, the group find the galactic discs gas, heated by supernovae and that appears as a tight polytropic relation. Star formation occurs within this phase, using 2 different routes: a standard, quiescent mode above 0.1 H/cc and a starburst mode, above 100 H/cc. The third phase is the hot, shockheated phase, shown as a fuzzy halo in the upper part of the plot. The MareNostrum simulation, thanks to its spectacular size, allows for a very high statistical quality in this plot.

Publications:
Teyssier, Fromang & Dormy, "Kinematic dynamos using constrained transport with high order Godunov schemes and adaptive mesh refinement", Journal of Computational Physics, 218, 44, 2006

Fromang, Hennebelle & Teyssier, , "A high order Godunov scheme with constrained transport and adaptive mesh refinement for astrophysical magnetohydrodynamics", Astronomy and Astrophysics, 457, 371, 2006

"The Universe in a Box", Liberation, June 17th, 2006
"Computational Cosmology", Les Houches School of Physics 2006, "The Fabrics of Space Time".

J. E. G. Devriendt1, C. Pichon2, R. Teyssier3, D. Leborgne3, and D. Aubert4, "The luminosity budget of the high-redshift universe", Astronomy & Astrophysics manuscript no. MN c ESO 2006 November 30, 2006.
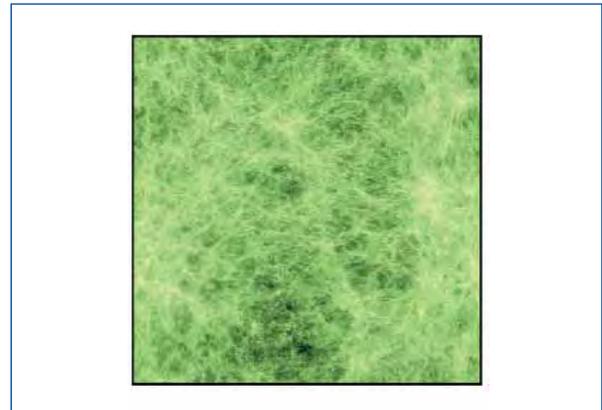
Figure 1. Projected gas density at redshift 4 for the whole simulated volume. The image size is 204x2048 pixels and clearly shows the filamentary structure of the intergalactic medium.
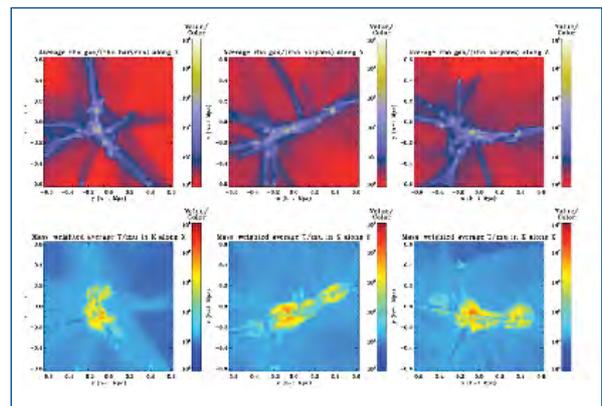
Figure 2. 3D view of the one of the largest halo in the MareNostrum simulation. Upper row is the gas density, while the lower one shows the gas temperature.
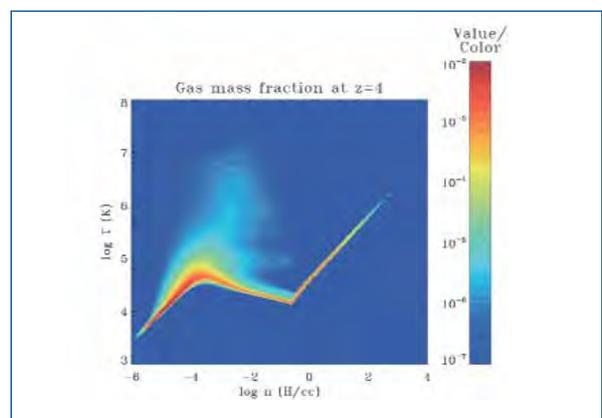
Figure 3. Density-temperature diagram in the MareNostrum simulation showing the various gas phases with unprecedented statistical quality.

Jordi Torra i Roca
Universitat de Barcelona

Gaia: Simulation of Telemetry Stream

Gaia is an ESA mission that aims to observe a thousand million stars with unprecedented accuracy. The UB team is the main developer of two simulation tools for Gaia: GASS (which outputs a realistic telemetry stream) and GOG (which generates intermediate and catalogue data).

Thanks to MareNostrum, the team has been able to generate the 5 years of science telemetry of a sample of 4 millions of stars. The resulting 700GB of raw data have made possible the test of the data processing system of Gaia. The complete catalogue of a thousand million stars has been simulated with MareNostrum as well, thanks to which the team can validate the several steps of the data reduction system.

The Gaia System Simulator (GASS) is a software package developed in Java by the UB team, in collaboration with the Meudon Observatory (Paris). It uses complex models for the Gaia instruments, its attitude, and its orbit, as well as for the stars to be observed – including very complex models of the Galaxy and the interstellar extinction. At the end, the main result is a very realistic telemetry stream, equivalent to the one that will be received from Gaia by the ground station. Afterwards, these simulated data can be used for testing and validating the data reduction systems that are currently under development, which shall be used in the final mission as well.

The team has simulated the complete mission (5 years) with GASS in MareNostrum, generating just a limited set of stars – 4 million instead of a thousand million. It has led to 700GB of raw telemetry data, which have been transferred to the European Space Astronomy Center (ESAC) that ESA has in Villafranca (Spain). The Gaia team working there is developing and testing the complex data reduction system that will generate the final catalogue of the mission from the telemetry data. This is an extremely complex system which requires huge amounts of simulated telemetry data for being tested. It would be impossible to test the system without telemetry data, and hence GASS (and MareNostrum) plays a crucial role here.

On the other side, the team has also developed (in collaboration with other research institutes of Europe) a simulation tool called GOG, or Gaia Object Generator. This software, also developed in Java, makes use of the same Universe models than GASS for simulating the millions of stars and stellar objects that Gaia will observe. The difference is that GOG outputs intermediate data and catalogue data, that is, the results of the data reduction system at different processing stages – instead of the raw telemetry data. In this way, the results of the complex system can be validated being developed and tested at ESAC, as well as of an initial data treatment system

being developed and tested at the UB. The complete set of one thousand million stars has been simulated using MareNostrum, which has been used to validate the Universe model.

The development of the Gaia simulation and data processing systems is an evolving task based on development cycles of 6 months of duration. Each new cycle will lead to additional complexity in the models used for the simulation, leading to an even higher degree of realism. It will obviously increase the computing power requirements. Furthermore, in the case of GASS, the size of the testing data sets will increase until the complete mission (a thousand million stars and 5 years) will be simulated at the end. That will lead to an impressive set of requirements, both on computing power and output size (estimated on 50TB), for which a facility like MareNostrum will play a crucial role.

Due to the complex models used by GASS and GOG, the processing requirements of both simulators are huge. As an example, the generation with GASS of 3 seconds of realistic data in a very crowded area of the Galaxy can require about 10 hours in a personal computer. This, fortunately, can be solved thanks to large supercomputing facilities like MareNostrum. Similarly, the GOG simulator requires massive processing requirements.
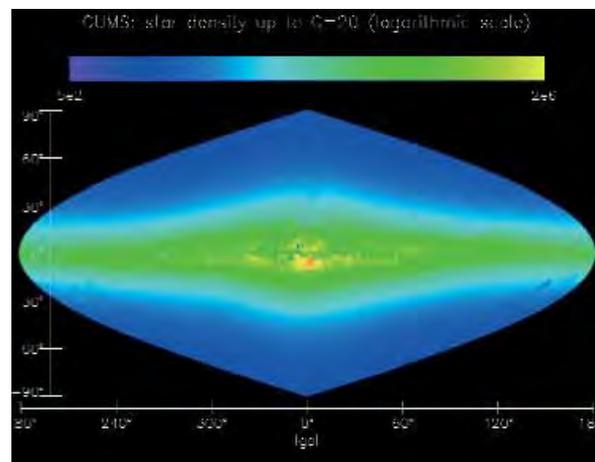
Star density in galactic coordinates, obtained with the GOG simulator.

Publications:
Isasi, Y., Masana, E., Luri, X., "Gaia System Simulator (GASS) overview", GAIA-C2-TN-UB-YI-001 (2006)

Masana, E. , Luri, X, Babusiaux, C., "Proposal for the definition of the Gaia Object Generator (GOG)", GAIA-C2-SP-UB-EM-001(2006)

Luri, X., Masana, E., Altamirano, P., Isasi, Y., Fabricius, C., "The Gaia mission simulator". In: Highlights of Spanish Astrophysics IV, ed by F. Figueras, J.M. Girart, M. Hernanz & C. Jordi (Springer, Barcelona 2006)

Álvaro Viudez Lomba
CSIC

Generation of inertia-gravity waves in geophysical vortex interactions

The team carried out about 50 numerical simulations using their three-dimensional numerical model initialized with different vortices configurations to evaluate how the cyclones and anticyclones interact and spontaneously emit inertia-gravity waves. Numerical resolutions with 256x256x256 and 512x512x512 grid-points were required to correctly simulate simultaneously two different scales of motion: the large scale of the vortical flow and the small scale of the inertia-gravity waves.

The spontaneous generation and propagation of short-scale inertia-gravity waves (IGWs) during the merging of two initially balanced (void of IGWs) baroclinic anticyclones was numerically investigated. The IGW generation was analyzed in flows with different potential vorticity (PV) anomaly, numerical diffusion, numerical resolution, vortex aspect ratio, and background rotation. The vertical velocity and its vertical derivative were used to identify the IGWs in the total flow, while the unbalanced flow (the waves) is diagnosed using the optimal PV balance approach. Spontaneous generation of IGWs occurs in all the cases, primarily as emissions of discrete wave packets. The increase of both the vortex strength and vortex extent isotropy enhances the IGW emission. Three possible indicators, or theories, of spontaneous IGW generation were considered, namely, the advection of PV, the material rate of change of the horizontal divergence, and the three-dimensional baroclinic IGW generation analogy of Lighthill sound radiation theory. It was found that different mechanisms for spontaneous IGW generation may be at work. One mechanism is related to the advection of PV, with the IGWs in this case having wavefronts similar to the PV isosurfaces in the upper layers, and helical patterns in the deep layers. Trapped IGWs are ubiquitous in the vortex interior and have annular wavefronts patterns. Another mechanism is related to the spatially coherent motion of pre-existing IGWs, which eventually cooperate to produce mean flow, in particular larger-scale horizontal divergence, and therefore larger-scale vertical motion, which in turns triggers the emission of new IGWs.

The numerical simulation of the spontaneous generation of inertia-gravity waves in three-dimensional baroclinic rotating and stably stratified flows is an important step for the improvement of the short range atmospheric and oceanic numerical forecasting.

The possibility of using a large number of grid points in numerical simulations has made possible to correctly simulate hydrodynamical processes in the atmosphere and oceans that involve two different spatial and time scales, like the spontaneous generation of inertia-gravity waves (short scales) during the interaction between cyclones and anticyclones (large scale).
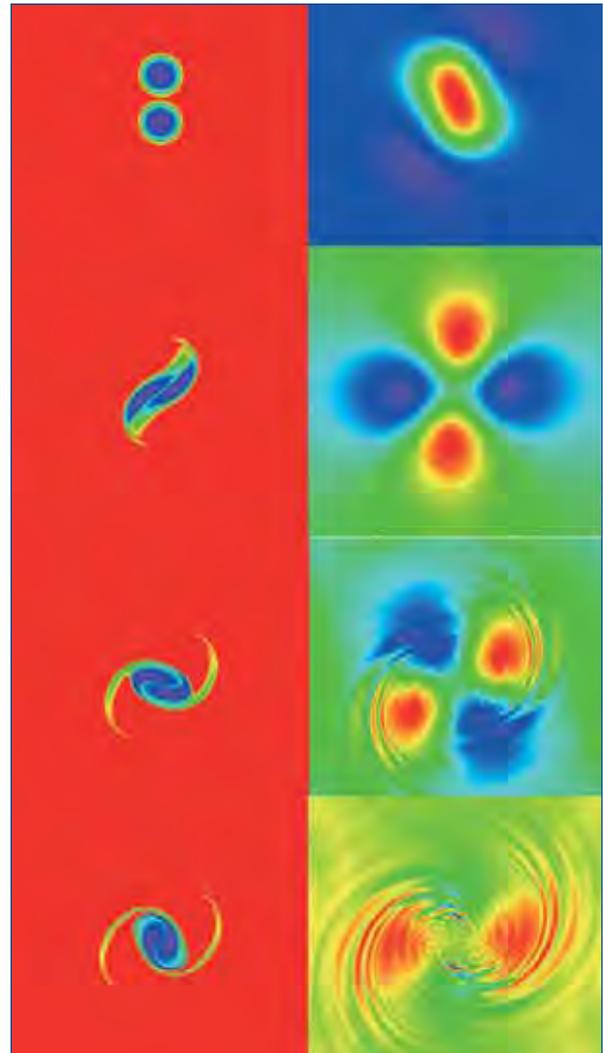
Time evolution of potential vorticity at the surface (left) and vertical velocity at mid depth (right) at time t=1, 5, 7 and 8 days. The images show the spontaneous generation of inertia-gravity waves during the fusion of two oceanic anticyclones. Three dimensional numerical grids of 256x256x256 and 512x512x512 grid points were used to correctly simulate the two spatial scales (the short wavelength and the long spiralling) involved in this phenomenon.

Publication:
Pallàs-Sanz, E. and Viúdez, A. 2007: Spontaneous generation of inertia-gravity waves during the merging of two baroclinic anticyclones. Journal of Physical Oceanography (in press).

Gustavo Yepes
Universidad Autónoma de Madrid

Cosmological Simulations of Large-Scale Structure formation in the Universe

The scientific objective of the project is to understand how galaxies and larger structures are formed in the Universe from cosmological initial conditions and to determine the relevant physical processes responsible for the formation and evolution of the objects we see in the universe today and in past times.

Accordingly, the project employs state-of-the art numerical codes (called GADGET) that simultaneously take into account the gravitational evolution of density disturbances in an expanding universe and the hydrodynamic and shorter scale processes acting on the visible component of the matter: the baryonic gas.

The codes are fully parallel (with MPI) and can treat a large number of particles representing the collisionless dark matter and the collisional baryonic gas fluid. They also have a module to include all the extra physical processes (non-adiabatic) acting on the baryon component related to electromagnetic and nuclear interactions. The multiphase nature of the ISM is also taken into account, as well as chemical enrichment modelling. This makes the code very useful for studies of cosmic structure formation on different scales, ranging from large-scale distributions in which the non-adiabatic effects have less importance, to the formation of galaxies themselves, in which a very detailed modelling of the baryonic processes is required.

The first application of GADGET in MareNostrum has been able to simulate the largest-known SPH simulation of large-scale structure formation. The so called MareNostrum Universe simulation consists of more than 2 billion particles (half dark matter and half SPH) distributed in a 500 /h Mpc comoving box. The project used 256 nodes (512 processors) of MareNostrum.

Clusters of Galaxies

Publications:
Stefan Gottlöber, Gustavo Yepes, Christian Wagner & Raúl Sevilla "The MareNostrum Universe. Proceedings of XXVI Astrophysics MoriondMeeting". La Thuile, 2006.

Stefan Gottlöber, Gustavo Yepes, Christian Wagner & Raúl Sevilla. "Baryon distribution in the Mare Nostrum Universe". Proceedings of DSU 2006 The Dark Side of the Universe. Madrid 2006.

Gustavo Yepes, International Conference: Bernard's Cosmics Histories: "From Primordial Fluctuations to Cosmological Structures. The MareNostrum Universe. The biggest SPH simulation up to now". Valencia 2006.
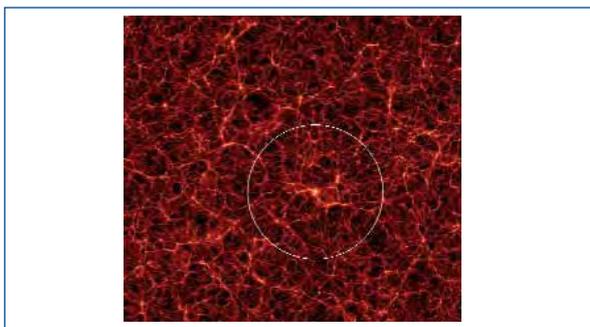
6.3 Physics and Engineering

Juan Acebrón
Universitat Rovira i Virgili

Probabilistic domain decomposition for partial differential equations

An algorithm has been developed recently capable to numerically solve partial differential equations (PDEs) by probabilistically induced domain decomposition methods. This project aims at developing efficient numerical methods to solve a number of partial differential equation problems, arising from Engineering and Physics.

The team is planning to run an algorithm on a variety of different and increasingly complex partial differential equations of interest in Physics and Engineering. The following plan was done:

1. Testing scalability for complex elliptic problems. The algorithm already tested for solving simple elliptic partial differential equations has been analyzed its performance for a quite general elliptic problem. In practice, being now he domain more complex, it constrains to use smaller time step in the Monte Carlo part of the algorithm, which has increased the expected overall CPU time.

2. Complex domains. There exists already a PPD code suited for complex domains (polygonal domain), where the subdomains are now discretized by finite elements instead of finite differences. The associated matrix of the linear algebra problem becomes denser, increasing correspondingly the CPU time needed to solve iteratively the linear system. It is planned to test the scalability of the new code comparing it to the performance of the classical DD based on PARMS.

For both systems the group has tested the scalability. Since intercommunication heavily affects the performance of the classical domain decomposition methods, using MareNostrum (connected through a fast low-latency network called Myrinet) is essential to check the real advantages offered by the algorithm developed for complex problems.

Jose Luis Álvarez Pérez
Universitat Politècnica de Catalunya

Numerical Simulation of Emissivity of Soils and Ocean

The field of Earth Observation from space is a fundamental field for the study of the environmental conditions of our planet. There are a great number of sensors and techniques which make it possible to assess from space different aspects of the Earth natural surfaces, either marine or over land. One class of sensors is called passive since they observe the spontaneous electromagnetic emission of natural targets due to their finite temperature. This allows the team to characterise different relevant parameters such as soil moisture of ocean salinity. The main objective of any sensor designed to monitor the Earth surface from the space of from an aircraft is to characterise the target and to provide geophysical information concerning natural processes and the human impact on them. In the context of this study the team aims at the numerical interpretation of radar remote sensing data collected over land and ocean surfaces. So far the team has focused in the study of ocean surfaces. The first task has been to model these surfaces considering different sea state cases considered. There are two components in the sea profile: the sea wave field due to the wind and the swell component. At this point the work has taken into account wind sea wave fields rather than swell, via a so-called Pierson-Moskowitz shaped spectrum with directional spreading. The long waves, that belong to the main wave energy contribution within the spectral shape, modulate the emissivity effects due to the capillary waves or sea surface ripples. Therefore the emissivity response must be affected by the specific ocean wave spectral features. Once the target is build the team has developed a code to compute the emissivity of the 2-D surface in terms of its roughness, caused by the wind, and its electrical permittivity, dependant upon the salt content of the water. The method in use relies on the use of a conjugate gradient method twined with the so-called physics based approximation. The difficulty of the problem rests on the fact that a huge matrix equation must be solved many times to encompass a variety of ensembles. Several approximations have been made in the past to achieve analytical results in a number of asymptotic situations. The main result so far when comparing the simulated, quasi-exact results obtained by the methods implemented in the code with the asymptotic formulae reported in the literature is the inability of the latter to properly describe the so-called coherent component of the radiation and to comply with the energy conservation principle. Once the algorithm is fully operational the objective of the project is to be able to provide researchers in this field and Earth observation data users with a comprehensive database of ocean and land emissivities at a variety of frequencies and scenarios (wind, salinity, soil moisture) against which real data from space- and airborne can be tested.

Polarisation	|Rqp|	|Rqp (coh)|	|rqp(incoh)|
hh	0.690	0.262	0.07
vh	0	0	0.0028
vv	0.682	0.259	0.070
hv	0	0	0.0027

COHERENT AND INCOHERENT REFLECTIVITIES CONTRIBUTING TO THE EMISSIVITY FOR AN ELECTRICAL PERMITTIVITY $\varepsilon = 72{:}08 + 66{:}59j$, CORRESPONDING TO A SALINITY OF 30 PPS AND A ROGHNESS GIVEN BY $k\,\sigma = 0.5$.

The use of a parallel machine to solve the linear equations posed by the electromagnetic problem of the emissivity by a 2-D rough surface is mandatory. No single CPU is able to deal with the amount of data and computational time required to invert the matrix. There are other instrumental computations that serve that end, basically Fourier transformations, which can be used massively at MareNostrum and that are beyond the capabilities of any conventional, non-parallel computer.
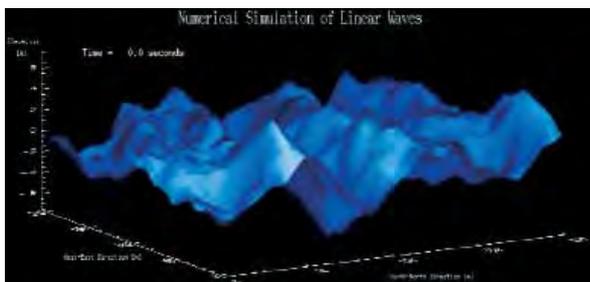
The figure represents a piece of ocean surface as simulated by our software running on the MareNostrum platform.

Jean Claude Andre
Le Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique (CERFACS)

Large Eddy Simulation of an industrial gas turbine

Recent advances in computer science and highly parallel algorithms make LES (Large Eddy Simulations) an efficient tool for the study of complex flows. The available resources allow today to tackle full complex geometries that can not be installed in laboratory facilities. Processes controlling flame ignition and propagation in helicopter combustion chambers are critical phenomena which control the performances of these systems: being able to start an helicopter engine at high altitude and low temperatures is critical. The prediction of these phenomena has long been out of reach of CFD (Computational Fluid Dynamic tools). The capacities of LES tools have been combined with the power of new computers to tackle this problem: one of the most advanced LES solvers developed jointly by CERFACS and IFP, the so-called AVBP code, is run on a high-resolution mesh, in order to compute ignition and flame propagation. Iso-surfaces of the reaction rate and cross-section of the velocity vector (coloured by temperature) within an helicopter combustion chamber: the high-resolution computation exhibits much more smaller-scale features and show the fine-scale structure of the flame. Computation of non stationnary reacting flows remains a huge challenge. For now LES (Large Eddy Simulations) have proved their capability in predicting stationnary flows. The purpose here is to assess LES in highly non-stationnary situation, ie extinction phases in a combustion chamber, using a high resolution mesh. A comparison is made with an equivalent calculation performed on a much coarser mesh in order to analyse subgrid-scale and combustion models. The results show that mean values describing the flows trend fit well, in particular the mean temperature temporal evolution. Discrepancies appear at the end of the calculation. This might be caused by the increased influence of small scales. This discrepecies are expected since a larger portion of the small scale phenomena are resolved in the finer mesh than in the coarse mesh. Nevertheless the same overall behavior is captured in both cases. Thanks to the increasing the supercomputing capabilities such as MareNostrum, it is feasible to perform calculations of non-stationnary flows on huge configurations, such as complex combustion chamber geometries.

Computation done with 1,242,086 cells (230,118 nodes).
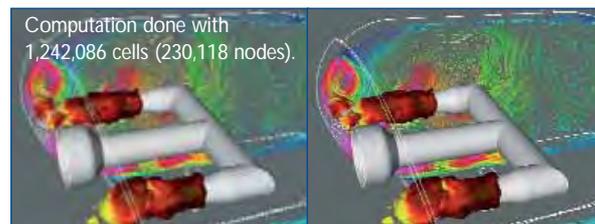
Richard J. Duro
Universidad de Coruña

Automatic design of wind turbine blades

An automatic design environment for wind turbines has been developed with the aim of obtaining the optimal windmill blade profile for a specific location. This system has two main parts: a search module based on a macroevolutionary algorithm, and an evaluation module based on an aerodynamic simulator. The search module tries to find the best blade profile with an evolutionary process where each individual represents a possible blade profile, whereas the evaluation module is in charge of computing the efficiency of each one of those blades.

In this work the project developed a system to produce automatic designs using evolutionary algorithms and simulators. For that purpose, a general decision module with easily adjustable parameters has been created. This module is independent of the simulator and can be used to solve many different design problems.

As a particular test problem, the team used the system in order to design a wind mill blade having a maximum diameter of 4.5 m. and for a wind velocity that follows a Weibull distribution with an average velocity of 7 m/s and a value for its form parameter of 2. A theoretical maximum power, called the Betz limit, can be easily obtained for any windmill, so this is used as a control parameter for the fitness function. Therefore, the fitness parameter is obtained as the calculated power multiplied by wind velocity probability. A blade is defined by providing its distributions of chord, twist and airfoil shape along the span. In this case, 30 sections have been settled, considering the rest of the sections as interpolations of these (in the evaluation process, a total number of 150 sections are considered). The aerodynamic simulator evaluates any blade profile using potential flow methods combined with boundary layer theory and the results are corrected with a neural network to take into account effects due to the detachment of the flow. The final efficiency value is calculated, using the blade element method, as an integral of all sections that compose the blade.

The macroevolutionary algorithm used had a population of 40960 individuals distributed in 32 races. The analysis of the fitness landscape generated by the fitness function lead the team to use this kind evolutionary algorithm. The evolutions were run in the MareNostrum at the Barcelona Supercomputing Center. The results presented were achieved in the 1400th generation, after 52 hours using 512 PowerPC 970FX processors. The best blade obtained for the conditions stated above was achieved in the 1370th generation and its efficiency is 79.4%. The generated power by this blade is very close to the optimum for the considered case. So, an automatic design environment and an aerodynamic simulation are combined to obtain an improved windmill blade. This procedure allows to solve very complicated and costly non linear models.

Thanks to the increasing supercomputing capabilities, such as MareNostrum, an evolutionary algorithm based system, can use a realistic simulator to evaluate each individual while keeping a reasonable population size.
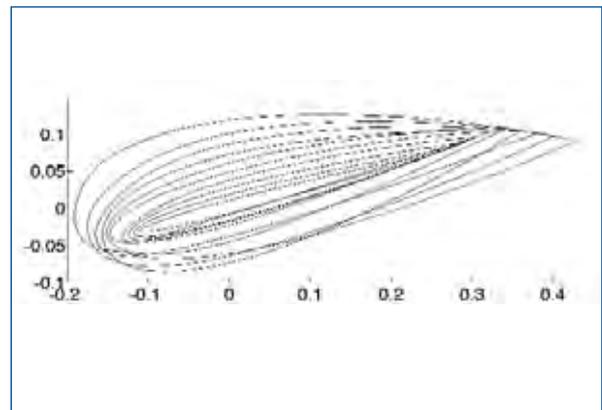
Jordi Faraudo
Universitat Autónoma de Barcelona

Molecular Dynamics Simulations of Biomimetic Membranes

The project identifies new mechanisms of electrostatic charge regulation in biological or artificial phospholipid membranes, using all atom molecular dynamic simulations. Using these charge regulation mechanisms, biomimetic membranes can tune electric interactions with biomolecules into repulsive or attractive interactions as desired. The simulations describe in detail the phospholipids (typically 100-200 phosphoglycerides with >40 groups), solvent (around 10000 water molecules) and divalent counterions and salt (between 100-200 ions). These very hard, high demanding simulations can be performed only at the world-leader Supercomputing centers.

In MareNostrum, the team requires 128-256 processors running about 72 hours to obtain data from equilibrium states (at the nanosecond scale). In this project the team has identified the mechanisms used by model phospholipid systems (resembling biological membranes) in order to regulate their surface charge in presence of minute concentrations of specific electrolytes. In particular, the team has shown that phospholipids which have the appropiate structure and design of the headgroup are able to capture and dehidrate divalent ions in solutions (such as Ca2+) even in the minute concentrations (10E-7 M) typical of biological systems by developing transverse phospholipid-ion correlations. In this way, some membranes with phospholipids with the appropiate design develop the intriging functionality of charge regulation. They are even capable of invert the sign of the charge of the membrane, with profound consequences for the physical properties of the membrane. These new mechanisms allow the team to explain the puzzling experimental results obtained using X-ray techniques. Future applications of these exciting results include the design of new nanoestructured, biologically inspired functional materials which take advantage of charge regulation in the same way that biological membranes do. Prior to the team's work, the mechanism employed by living systems to perform such functionality was unknown so it was impossible to design new biologically inspired materials with these functionalities. The team envisages a large range of possible applications of these new materials, from pharmaceuticals to nanomachines. Future perspectives include the simulation of basic cell processes which make extensive use of the mechanisms analyzed in the project such as signalling based on divalent ions, a process of great importance for cell life which still has important unresolved physical aspects.

The Molecular Dynamics Simulations allow the team to obtain a detailed description of atomistic mechanisms in the interaction between cations and membranes which cannot be discriminated with current experimental techniques.

Also, it is clear that these large scale computer simulations developed in this research project are only possible in top Supercomputing facilities. The team's experience in MareNostrum shows that a typical equilibration run for a realistic phospholipid layer (previous to a production run) requires 128 processors during 1-2 days. A typical production run for this system (1-2 nanoseconds in physical time) require 2-4 days with 256-128 processors. The team's experience also shows that currently MareNostrum is the only supercomputing facility in Europe able to perform this kind of simulations. Other supercomputing facilities, employed in previous research projects showed a significantly smaller performance in the calculations, and the team were restricted to simulations of smaller systems (128 amphiphillic molecules, their counterions and 103 water molecules in 1 nsec production run as compared with 512 amphiphillic molecules, counterions, added salt and 10 4 water molecules in 2-4 nsec production runs).

Taking into account the sluggish dynamics of these systems, the gain obtained in MareNostrum significantly improves the results and allows the team to study realistic systems inaccessible in other supercomputing facilities.
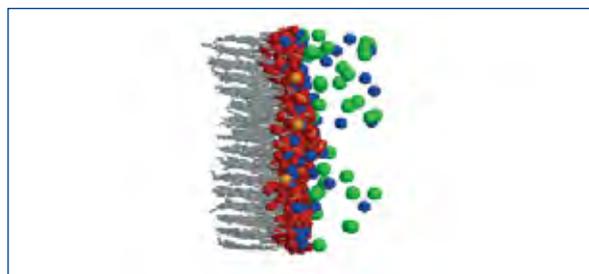
A fragment of a Biomimetic membrane with negatively charged phospholipids (DMPA2-) containing an excess of adsorbed divalent ions. Note how negative Cl- ions closely approach the membrane, indicating that membrane charge has been reversed by divalent ion adsorption

Publications:
J. Faraudo and A. Travesset, "Charge inversion induced by Transverse correlations in Phospholipid systems", Submitted to Phys. Rev. Lett. (May 2006).

J. Faraudo and A. Travesset, "Interaction between Phosphatidic Acid domains in Membranes and divalent ions" submitted to Biophysical Journal (June 2006).

Kafui Tay and Fernando Bresme "Hydrogen bond structure and vibrational spectrum of water at a passivated metal nanoparticle", Journal of Materials Chemistry (2006)

Domingo Giménez Cánovas
Universidad de Murcia

Parallel routines optimization and applications

MareNostrum has been used to study two problems with high computational capacities:

1 - The parallelization of algorithms for the solution of Inverse Eigenvalue Problems has been studied. The performance of the system has been characterized by studying the value of parameters that appear in the modelling of algorithms. The model of the execution time has been used to determine optimal conditions.

2 - Algorithms for the solution of Simultaneous Equation Models have been developed. SEM are used in the simulation of economical models, but the algorithms have so far been used only to solve simulated systems.

The use of MareNostrum has enabled the study of larger problems than those previously studied, together with a deeper study of the scalability of the developed algorithms.

Publications:
Georgina Flores Becerra, Algoritmos Secuenciales y Paralelos para la Resolución del Problema Inverso de Valores Singulares, Ph. D. Thesis, Universidad Politécnica de Valencia, May 2006.

José-Juan López-Espín, Domingo Giménez, Solution of Simultaneous Equations Models in high performance systems, In XXIX Congreso Nacional de Estadística e Investigación Operativa, 15-19 May 2006, Tenerife, Spain.

José-Juan López-Espín, Domingo Giménez, Solution of Simultaneous Equations Models in high performance systems, in PARA06, WORKSHOP ON STATE-OF-THE-ART IN SCIENTIFIC AND PARALLEL COMPUTING, Umeà, Sweden, June 18-21, 2006.

Vicent Gimenez Gómez
Universidad de Valencia

Monte Carlo numerical computations of the properties of hadrons

All atoms are made of particles. Some of them, like electrons, are belived to be elementary, but other, like protons and pions, collectively known as hadrons, are composed by strongly interacting quarks. The force between quarks is similar to an elastic string: small at short distances but strong at large separations. The team numerically simulates the dynamics of confined quarks inside hadrons to study and predict their properties.

Quantum Chromodynamics (QCD) is the present best model of the strong interaction between quarks in the hadrons. However, it is far from being demonstrated that it indeed describes all aspects of the strong forces. In particular, non-perturbative phenomena such as the confinement of quarks inside hadrons, the symmetry breaking pattern of the interchange of massless left- and right-handed quarks, the violation of time-reversal symmetry and surprisingly large differences in the cross-section of very similar decay processes of the K-meson are completely un-resolved problems. To address such questions one has to resort to non-perturbative methods, in particular to large scale numerical simulations in a discretized version of the continuum QCD theory: the so-called lattice QCD.

The aim of the project is to perform first realistic simulations of lattice QCD. This means to use dynamical fermions taking the strange quark mass into account in the simulation and to run at light quark masses, as close as possible to the "physical point", where the pion mass assumes its experimentally measured value. In addition, the volume of the lattice box, to which the team approximates the continuous infinite space-time, should be sufficiently large, corresponding to a box length of at least 2 fm. This demanding setup can in principle be realized by using a new formulation of lattice QCD: the so-called maximally twisted mass fermions as the preparatory work strongly suggests.

To study the physics problems the team is interested in (hadron spectrum, decay constants, matrix elements relevant to phenomenology, and so on), the gluon field has to be simulated. On MareNostrum, the team is using a highly optimized hybrid Monte Carlo to compute the gluon field inside an hadron including light quark polarization effects. Due to the fact that the team's method is stocastic in nature, the group has to generate several hundreds of preliminary gluon configurations before getting stabilization or equilibration of the simulation parameters. The group calls this process thermalization. The team has generated 400 configurations

but unfortunately equilibration is still quite far. Therefore, for this project it is needed more computer time to finish the thermalization. Due to factors like the large volume of the lattice box ($32^3 \times 64$) and the small masses of the sea quarks, our simulations are very demanding. So demanding that they can only be performed on massively parallel supercomputers such as MareNostrum.
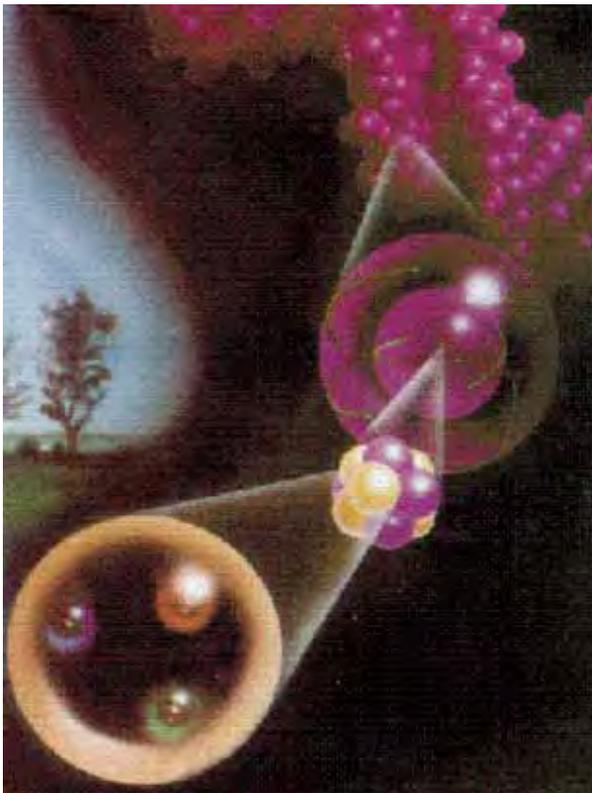
All matter is made up of atoms. Atoms contain electrons surrounding in orbits a small nucleus. The atom's nucleus is made of protons and neutrons very tightly bounded by the nuclear force. Both, protons and neutrons, contain strongly interacting quarks.

Antonio González-Arroyo
Universidad Autónoma de Madrid

Particle Production and Field Generation at Preheating after Inflation

This scientific program involves the study of the evolution of Quantum Fields in the early universe, more precisely in the preheating era. There are some properties of the present day Universe which might have its origin in this period. Baryogenesis is certainly one of the most popular, but recently the team has put forward the possibility that primodial magnetic fields also originate in this period.

The team has used MareNostrum to make test runs of the main codes that should be used in the study both of Baryogenesis and Magnetic Field Generation. Results are encouraging. In the first topic the team was able to study the standard model coupled to a singlet inflaton and evaluate the necessary resources for a full scale computation. The necessary ingredient of a CP violating phase has not yet been tested. Parts of the results obtained with MareNostrum were used in the work presented by Alfonso Sastre for his DEA at Autonoma University. Concerning primodial magnetic fields the team also tested the codes and obtained some configurations to analyse the spatial distribution of the field. Scientific Computing is a crucial tool to investigate an increasing number of areas of fundamental and applied research. Availability of these resources is a measure of the degree of intellectual and technological development of a society. Reversely these resources provide the ground to train new generations in the use of this fundamental tool. It is unconceivable that a particular country could have a strong technological development in information technology without a similar development in frontier areas of computing, like those attainable with MareNostrum. It is also important to emphasize that these big scale computers are not in contradiction with smaller scale resources available to research groups and Universities. It is only with this network of different scale resources that one can take full advantage of them. In the team's opinion the great supercomputing centers are useful for collaborations with mature and well-developed codes. The team emphasizes that these centers should also favour run-time parallel computing, as opposed to PC farms. It would be a waste of resources to treat MareNostrum as an array of processors processing independent codes.

Magnetic field intensity in 3-space soon after the end of inflation

Pilar Hernández
Universidad de Valencia

Non-perturbative aspects of QCD in flavour physics

Quantum Chromodynamics (QCD) is the theory that describes the strong forces that keep quarks confined into protons and neutrons. This theory is extremely predictive. All the rich hadronic phenomenology observed in Nature can in principle be derived from a few fundamental parameters: the strength of the strong force and the masses of the quarks. These predictions are very challenging in practice, because the theory cannot be solved by analytical methods and supercomputing resources are needed to solve it numerically, through a discretization of the system in a space-time lattice.
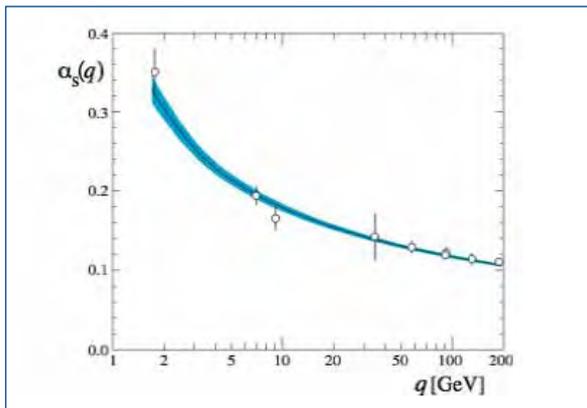
The experimental measurements of the force between two quarks as a function of the 1/distance together with the prediction of QCD (blue band).

Publication:
L. Giusti, P. Hernandez, M. Laine, C. Pena, J. Wennekers and H. "On the K->pipi amplitudes with a light charm quark'' Wittig arXiv: hep-ph/0607220, Physical Review Letters. Submitted on 20th July 2006.

Javier Jiménez Sendín
Universidad Politécnica de Madrid

DNS in turbulent channels

The efficiency of turbines can be increased by increasing the curvature of the blades. Curvature is related to adverse pressure gradients which decelerates the flow. However higher curvature can have a negative effect on the development of the flow, including separation, vibration and the loss of the desired efficiency. This type of flow is highly unsteady, but most research is done using steady approximations of the equations that govern the flow. This engineering approach has led to great advances in blade design. However the next step in optimising turbines should involve unsteady calculations. Furthermore high quality simulations of the most complicated and unsteady parts of the flow are necessary. The latter is what has been done in this study using a code that was developed in-house ([MPS]).

The simulations have been done to try to imitate an experiment done over a flat plate. In this experiment the effects of curvature are modelled using a curved upper wall ([HHXFZ]). The simulation has been performed on a flat plate. The curved wall is modelled with a certain velocity profile at the upper wall of the numerical domain. The velocity profile is chosen as to obtain the same starting point of the separation as in the experiment. The Reynolds number of the simulation and the experiment are chosen equal.

The principle objective of this study is the influence of wakes on the flow development on the flat plate. Therefore wakes have been modelled and are periodically imposed at the inlet of the numerical domain. The angle with the freestream has been matched with experiment.

Due to the simulations the team has obtained information about the flow development under severe adverse pressure gradients. One of the important results is the necessary height of the numerical domain necessary to contain the flow field. This is important for future simulations, which include turbulent boundary layers under the influence of a less severe adverse pressure gradient.

For industry and in particular ITP S.A. it is very important to see the influence of the wakes on the flow and the proof that the team's code has the capability to simulate their effect.

In the future more simulations with incoming wakes will be performed to obtain information, about how they can contribute to more efficient turbines ([MPSJJ]).

The last important contribution is that the results of the unsteady simulations and results could be used in the models approximating the Navier-Stokes equations.

The benefits of MareNostrum for the project has been the possibility to do simulations of a model turbine blade flow with typical Reynolds numbers and pressure gradients including the unsteady effects of wakes.
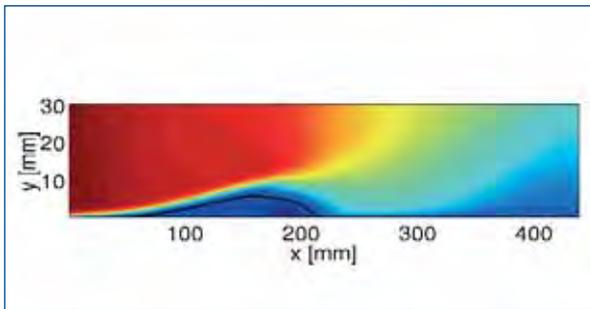
The average velocity. The thick black line indicates the separation bubble.
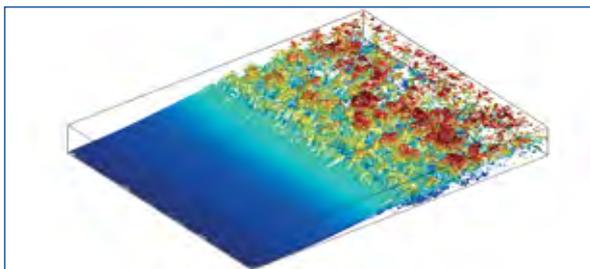
Three dimensional vorticity field. Color indicates height of the structures. Red is high and blue is low.

Publications:
Sergio Hoyas and Javier Jiménez, School of Aeronautics, Universidad Politécnica de Madrid, 28040 Madrid, Spain. "Scaling of the velocity fluctuations in turbulent channels up to Reτ = 2003", Received 25 October 2005, accepted 1 December 2005, published online 11 January 2006. 2006 American Institute of Physics. DOI: 10.1063/1.2162185.

Hoyas et al (2006) "Reynolds number effects in the Reynolds stress statistics and balances", J. of Fluid Mech, in preparation.

M.P.Simens and J.Jimenéz "Alternatives to Kelvin-Helmholtz instabilities to control separation bubbles", ASME paper GT2006-90670. Proceedings of ASME Turbo Expo 2006.

Mark Phil Simens "Separation bubbles and their control" PhD thesis 2006. School of Aeronautics, Madrid, Spain, in preparation.

Bruno Julià Díaz
Universitat de Barcelona

Dynamical Coupled-channel Analysis of Excited Baryons

The inner structure of baryons, e.g. protons, Deltas, has one direct consequence which is the existence of resonances. These are excitations of the quarks and gluons building the baryons and have been extensively studied during the last three decades. Among these studies experiments with electromagnetic probes have proved to be extremely useful to explore the characteristics of such resonant states. In the project the team extracts and interprets their properties by making use of an elaborated model together with a sound reaction theory. The project will build the bridge between experiments and microscopic theories of baryon structure. When baryon structure is explored at small distances, that is at large Q2 in the plot, they are correctly understood as being composed of current quarks and gluons, whose dynamics are well described by Quantum ChromoDynamics. The situation becomes more involved when we explore larger distances, or equivalently when lower Q2 are involved (left part of the picture), then hadronic degrees of freedom start to prevail over quarks and gluons eventually giving way to the usual nuclear physics description of matter. The transition between the two regimes is what interests us.

Making use of the current phenomenological understanding of strong interaction physics in the low momentum transfer regime the team studies the properties of baryonic resonances (which are a clear trace of the existence of structure in nucleons) making use of a dynamical model which incorporates the effect of multi step processes, e.g. Intermediate pion-Nucleon creation. The group uses it to analyse the existing extensive electroproduction database to study the properties of excited states. In the three months allocation at MareNostrum the team has mainly been involved in building a competitive model of the interaction between the pion and the nucleon up to center of mass energies of around 2 GeV. This is an involved project already and has, up to now, been completed partially. The group has managed to get a reasonable model of the S11, P11 and P33 partial waves. These are already of great importance as they host some of the lowest energy resonances, including the well known Delta (1232), the interesting Roper resonance and the first negative partiy excitation of the nucleon, N (1535).

The nature of these resonances, analyse in detail within the model making use of electro production data, is the next near term goal. First, the plan is to extend the pion-nucleon model to other partial waves, this should be finalized in the near future. At the same time the team has developed another parallel code, CCEBA-ELEMAG, which computes electroproduction observables and which can already be used

to fit electromagnetic form factors of the resonances. This code has been tested at MareNostrum but has not yet been used to perform minimizations with the existing database. Right after that has been finalized the team's next important and essential goal is to include three body channels explicitly in the formalism. This has already been done in a serial code, but needs to be parallelized and further tested on the supercomputer. Once the code is ready for production the team should fine tune the previous analisis to encompass the effect of such channels, which should be very relevant in electroproduction above their corresponding production thresholds. Apart from the above, but closely related to it, the group has started another collaboration which will be a very important by product of the current project at MareNostrum. It involves three researchers at CEA/Saclay, in France, Dr. Bijan Saghai, Dr. Jun He and student J. Durand. The main emphasis of this collaboration is the study of baryonic resonances in the production of the eta-nucleon system. This channel is already in our current project, thus the idea is to make use of the developed models and extend them to the photoproduction of the eta-Nucleon system. This is being pursued at the same time and already has a working parallel version.

The impact of these studies in our society will not be visible in the near future. These are studies which deal with our understanding of the way our world works at very small scales. Current technology cannot profit of our theoretical work as of today, but will undoubtly benefit from it at some point in the future when our understanding of the femtometer scale becomes relevant. The use of MareNostrum is of crucial importance to the project and will become even more important as the team increases the amount of experimental information wanted to understand using a dynamical coupled channel formalism.
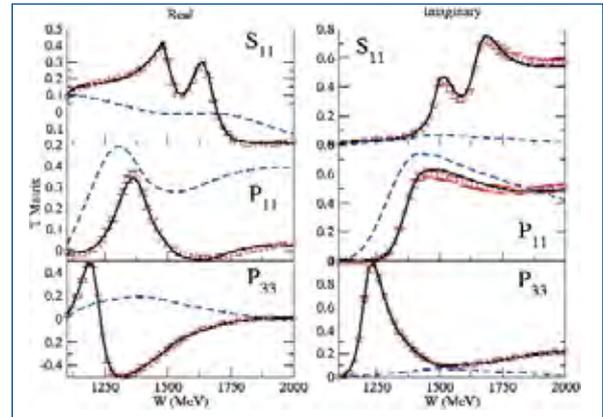
The figure shows part of the current model for pion-nucleon interaction up to 2GeV. The solid black line is the full model, the dashed blue one corresponds to the non-resonant contribution.

Publications:
Extraction and Interpretation of gamma N --> Delta Form Factors within a Dynamical Model. B. Julia-Diaz, Talk given at " N* Analysis Workshop", 4-5 November 2006. Jefferson Laboratory, USA

Simple example of a parallel code and how to run it on a cluster. B. Julia-Diaz, seminar given at the department of Estrutura i Constituents de la Materia, U. Barcelona, Barcelona December 20th, 2006
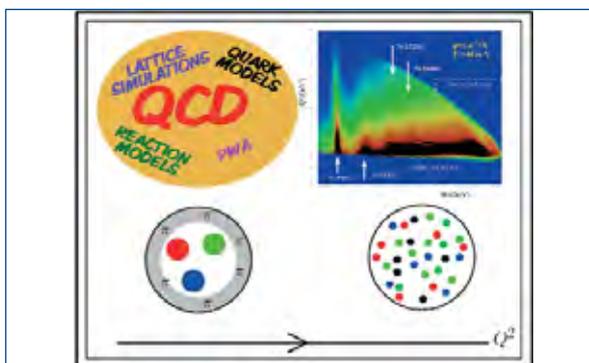
This is a divulgative image, on the upper left part the current theoretical tools to understand baryonic resonances are shown, on the upper right part an experimental image showing how the resonances are seen in photoproduction data, the two images below, show the transition from the confined quark picture with a meson cloud to the free partonic picture.

Thilo Knacke
Technische Universitaet Berlin (Germany)

Numerical simulation of airframe noise

The main objective of this project is the reduction of aircraft noise by suppression of local noise sources. To assess the feasibility of reducing noise through passive and active control of unsteady pressures in the aircraft boundary layer and wake, an understanding of the flow physics is a prerequisite. The flow field simulation of the high-lift system can deliver deep insight into the noise generation mechanisms which predominate under approach conditions.

The reduction of aircraft noise is an important issue for aircraft manufacturers to prevent increasing noise exposure caused by the continuous growth in air traffic. During the landing approach, the flow around deployed high-lift devices is known to be a major contributor to the overall sound pressure levels perceived by observers in the distance. A good understanding of the underlying noise generation mechanisms is essential for the development of noise reduction strategies. Insight can be obtained from numerical simulation, which is well suited to provide detailed flow-field information. As the basis for the airframe noise simulation a typical 3-component high lift configuration has been selected. The 30P30N model, where the numbers refer to the slat and flap deflection angles, is simulated under approach conditions in free-flight. The computational grid, used to represent the unsteady 3-dimensional fluid motion, consists of approx. 30 million volumes. In order to avoid a time-consuming unsteady simulation from scratch, the first computations have been run in a steady mode, resulting in a time-averaged solution of the flow field. Starting from this steady solution, time-accurate simulations are currently in progress. At the present state of the simulations, despite a two-dimensional geometry and boundary conditions, initial 3-dimensional vortical structures are evolving. This change of rotation axis enables vortex stretching mechanisms, which are important features of the turbulent flow and can only develop in a sufficiently resolved 3-dimensional approach. In order to overcome the initial transient behaviour and to achieve stable flow field statistics, the simulations have to be run for a longer time.

Thanks to the supercomputing facilities such as MareNostrum, increasingly complex flows are becoming accessible by means of numerical simulation.

Sound waves generated by the unsteady flow around a wing in high-lift configuration.

Enrique Lomba
Instituto de Química Física Rocasolano, CSIC

Ab initio simulation of amorphous SiC and C

So far the team is in the process of generating trajectories of the project ab initio Molecular Dynamics. A thorough analysis of trajectories will be performed and it is expected to obtain good estimates for the sound velocity in molten SeTe alloys near the semiconductor-metal transition and try to correlate changes on the local microstructure with the changes observed in microscopic properties. The use of MareNostrum makes feasible this type of study that requires reasonably sized samples. The team has simulated the process of elimination of Si from SiC during chlorination by means of a series of Molecular Dynamics runs starting from 216 atoms in which 4 atoms of Si are removed and the system is equilibrated during 5 ps keeping the total volume constant (in accordance with experimental estimates of the final density of amorphous C thus obtained). The team observes the formation of very open clusters with sp^2 hybridization and a large degree of sp components. These results are being contrasted with EELS experiments which determine the amount of sp^2 hybridization present in the real sample.

Preliminary calculations indicate that the degree of sp hybridization is overestimated by the simulation. This can be due to several factors. On one hand seems that the estimates of the final density of amorphous C are incorrect, since they include the mesoporous space which can not be simulated using atomistic MD. A more local measure of the density (as obtained from the plasmon peaks in EEL spectra) is needed in order to perform the simulations. This shortcoming will be addressed in future.

Aside from this, a more worrying problem is connected with the fact that in the experimental conditions, the process of Si elimination gives rise to inhomogeneities in the mesoscale range. This issue can only be addressed using extremely large particle numbers and simulation times with a classical potential. The team plans to make use of the ab initio results already obtained to device such a classical force field (and or test existing ones). A relatively large number of particles have been used so that the final sample after elimination of Si contains 108 particles, a number sufficiently representative to obtain estimates of bulk properties. In a full ab initio calculation like this, the use of supercomputer resources is unavoidable.
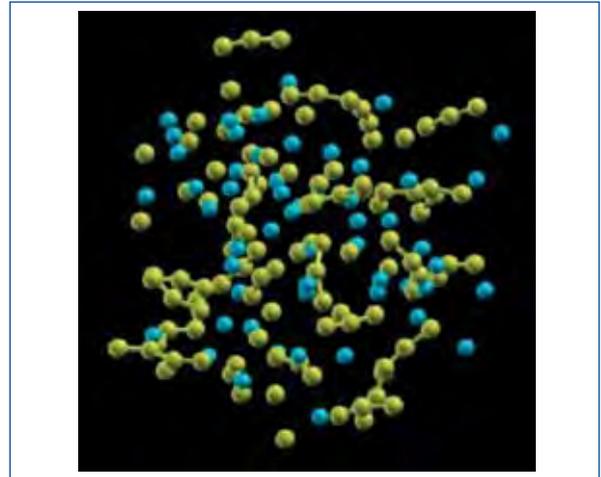
Snapshot of a intermediate stage of the elimination of Si (blue atoms) in a SiC sample. The original zinc blend structure transforms into an amorphous open structure.

Enrique Lomba
Instituto de Química Física Rocasolano, CSIC

Dynamics of Molten SeTe alloys in the semiconductor-metal transition region

The team has performed 30ps long simulations on molten equimolar alloys of SeTe using the VASP ab initio code with 256 atoms in order to analyze the behaviour or time correlation functions, and the dynamic structure factor and thus obtain a microscopic interpretation of inelastic neutron scattering results in the region corresponding to the semiconductor-metal transition.

William Marnane
University College Cork (UK)

Development and Charaterizaton of new LDPC codes, encoders and decoders

Reliable communications are in great demand today. It is, therefore, important to use the transmission systems available as effectively as possible. There are many error correcting codes available one of which is the low-density parity-check (LDPC) code, such codes can achieve a decoding performance just 0.0045dB from the Shannon limit, hence they now possess the unrivaled status as the best known existing codes. The goal of this research is to develop arithmetic and algorithmic innovations in order to improve the existing encoding and decoding hardware architecture and to create superior LDPC codes.

Error correcting codes allow error–free communications to occur over noisy channels by encoding data to be transmitted and using algorithms to detect and correct errors in the received data. The error correction block provides a significant reduction in the cost of operation, as data are not re-sent, also introduces improvements in the system bandwidth and reduce the power consumption of communication systems

A great number of types of error correcting codes exist and are used in telecommunications, BCH, Reed-Solomon (RS) codes are some of the most common codes used from GSM to optical transmissions.

In recent years two classes of codes that exhibit performances near the Shannon limit for noisy channels have been developed. These are Turbo codes and low-density parity check codes (LDPC's). Past research has shown how LDPC's outperform Turbo codes for large block size.

Introduced by Gallager in 1963 Low Density parity check codes had been neglected until the work of MacKay in 1995; these codes can yield high performance on the binary symmetric channel (BSC) as well as on the additive white Gaussian noise (AWGN) channel, and have been shown to outperform the Turbo codes in many applications. The algorithm used for the decoding is called *belief propagation*, known also as the sum-product algorithm. This algorithm uses a graphical representation of the code the *Tanner Graph*.

A full understanding of the state of the art solutions shows how, unfortunately, the hardware implementation of the algorithms used for coding/decoding LDPC codes involves making a number of design choices that have a trade off between the BER performance and the complexity. One of the dominant problems of any implementation of a LDPC decoder depends on the level of messages passing between variable nodes and check nodes. Every edge of the Tanner graph represents a connection that must be implemented in some way, this leads to the classic area versus time design problem.

The only way to ease this problem is to choose a code that has some degree of regularity. One group of codes that possess this quality is the quasi-cyclic codes family. These codes were first introduced by Townsend and have seen recent interest both for the mathematic characteristics and for the possibility of efficient encoding.

Part of the work that has been carried out has been the study and characterization of a new class of quasi-cyclic codes.

A quasi-cyclic code of index t is a linear block code C in which a cyclic shift of any codeword in C by t positions is also a codeword. The generator matrix G for these codes is a matrix where every row is a t circular shift of the previous row. It has been shown that every generator matrix for a quasi-cyclic code can be decomposed in circulant matrices.

Work, that has been carried out, was concerned with the possibility of taking advantage of the structure of quasi cyclic codes for building reduce complexity, smaller and faster hardware implementations of the decoders algorithm. The results have been submitted in a journal paper to IEEE, Transaction of Circuit and System.

The study of quasi cyclic LDPC codes is far from finished. The work outlined above it has been shown, with a particular case, how this class of codes is important both from an information theory point of view and from a hardware realization point of view. The interest is now shifted to the question of the possibility of finding rules that allow building of quasi cyclic LDPC codes with good characteristics. Moreover it is being investigated how these rules could be generalized from the particular case to a much wider family of quasi cyclic LDPC codes. The work being carried out at the moment concerns the answer to this question.

Basically, to test a code and a decoder at a specific noise level, the simulation procedure is as follows. A message vector is randomly created. The message is encoded using the generator matrix to produce a codeword, which is modulated to create the desired signal strength. Noise is generated randomly and added to the codeword to simulate the channel. The corrupted or received codeword is then decoded using an arbitrary decoding algorithm. This procedure is repeated a number of times (called trials) using different codewords for the same conditions—that is, the same code, the same noise level, the same decoder—and the results are recorded and an analysis of the results is

performed. The most significant aspect to realize about this process is that, statistically, the more trials this procedure is run, the more accurate the results. Therefore, the direct simulation process described above is a lengthy and computationally intensive method of testing a code and a decoder at a specific SNR. It is easily understandable the type of benefits of using computer cluster reducing simulation time from days to hours and allowing to study better codes.
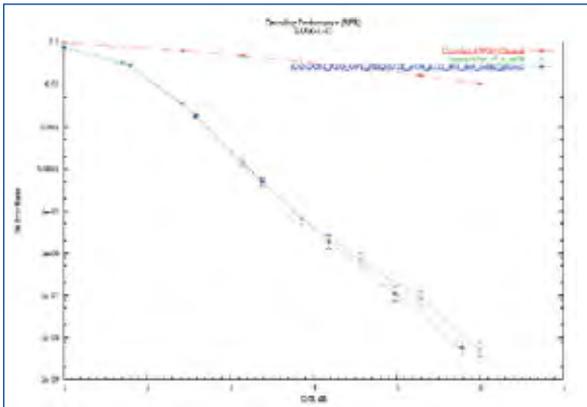
Performance comparison between small size rate 2/3 LDPC codes. In blue a random constructed codes and in green a novel girth ≥ 8 quasi-cyclic LDPC codes.

Publication:
C. Spagnol, A. Byrne, E. Popovici, M.?O'Sullivan, W. Marnane. "Efficient hardware implementation of LDPC decoders", submitted to: IEEE, Circuit and System, Transaction on. (Paper submitted to IEEE: Transactions on Circuits and Systems).

Monty Newborn
McGill University Montreal (Canada)

Parallel Automated Theorem Proving

Programming and experiments were carried out on the parallel automated theorem-proving program called Octopus. Octopus carries out resolution-refutation proofs on theorems presented to it in first-order predicate calculus. It can run on as many processors as available. On MareNostrum, several tests were carried out using as many as 256 processors.

Octopus has been in existence for almost a decade and has participated in several of the annual conferences on Automated Deduction's ATP System Competition. Each of the processors carries out an independent effort to find the proof of a theorem by trying to prove a simpler, or weaker, version of the given theorem. The more processors, the more weakened versions. Some theorems have thousands of weakened versions. The efforts on MareNostrum were aimed to improve Octopus's criteria for selecting weakened versions. The results of this effort will be put to the test soon as discussed above. During the competition, Octopus will run on a network of over 100 processors located in the School of Computer Science at McGill University.
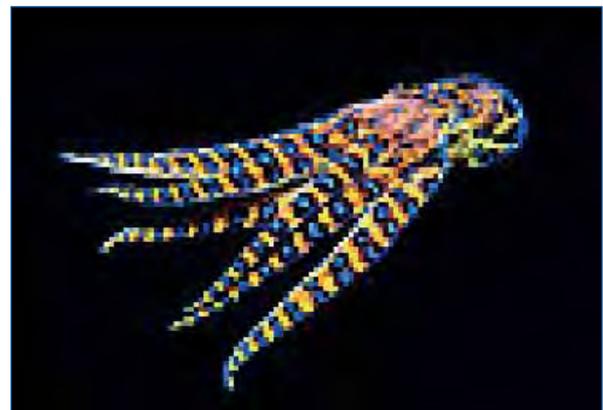
Octopus. A parallel automated theorem-proving program ran on as many as 400 processors while under test on MareNostrum.

Asensi Oliva
Universitat Politècnica de Catalunya

Direct Numerical Simulation of Turbulent Flows

Direct numerical simulation (DNS) is an important area of contemporary fluid dynamics, because it is of interest for improving the understanding of the physics of turbulence and because it is an essential tool for the development of better turbulence models. Recently, relevant improvements to turbulence modelling based on symmetry-preserving regularization models for the convective (non-linear) term have been developed. They basically alter the convective terms to reduce the production of small scales of motion by means of vortex-stretching, precisely preserving all inviscid invariants. They have been successfully tested for relatively high Rayleigh (Ra) numbers. At this stage, high resolution DNS results at higher Ra-numbers are of extreme importance for further progress.

To do so, the team is currently using the new version of the DNS code developed by CTTC on the MareNostrum supercomputer to carry out a simulation of a differentially heated cavity with Ra=10e11, Pr=0.71 (air) and height aspect ratio 4 using a mesh of 111M points. To the best of the authors' knowledge, this will be by far the largest DNS simulation with two wall-normal directions ever performed. Moreover, due to the complex behaviour exhibit, an accurate turbulence modelling of this configuration remains a great challenge. These new DNS results shall give new insights into the physics of turbulence and will provide indispensable data for future progresses on turbulence modelling.
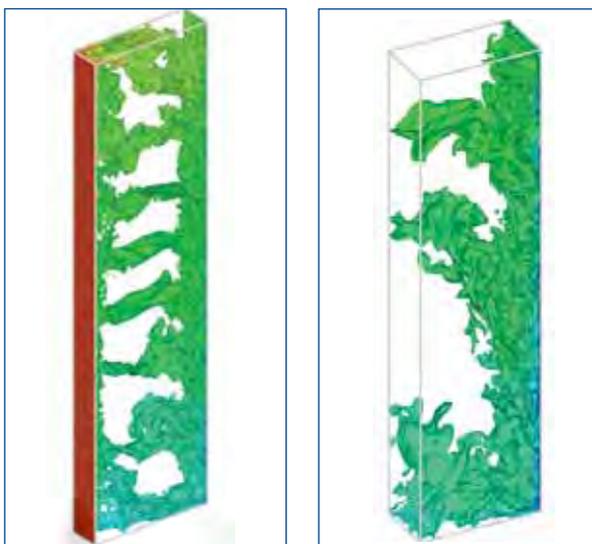
Instantaneous isotherms at Ra=1011. Left: general view of the cavity. Right: zoom around the bottom right corner.

Ignacio Pagonabarraga
Universitat de Barcelona

Lattice Boltzmann simulations of complex fluids

The Project has extended and used an existing code to deal with the fluid dynamics at a mesocopic level. On the one hand, the team has studied suspensions of active particles which move cooperatively through the fluid in which they move. On the other hand, the team has considered the instabilities of fluid/fluid interfaces in confined geometries. The fluid flow generated byself propelling particles induces and effective dissipative interaction between these active particles. As a result, they form aggregates and structures on scales large compared to the propellers' size. The study of these structures and their intrinsic dynamics require the use of intensive computacional means. The team has analyzed the clusters and short and intermediate ordering that particle swarming give rise to. For active particles which can rotate the team has analyzed how their active motion induce translation and give also rise to clusters which can diffuse on a membrana.

These studies open the possibility to continue a more detailed study of the mechanical properties of these materials; in particular their elastic properties and rheological response.

With respect to fluid/fluid interfaces, the team characterized Safmann/Taylor fingers including the effect of meeting and structure in the gap. As a result, the group has quantified the relevance of curvature and stretsses beyond the ususal assumptions of mechanical equilibrium. The group has shown the role that diffusivity plays in the motion of the contact line and used it to control in detail the shape of the finger. By considering finite cross-sections the team has addressed the relation between these finger and other type of instabilities. The studies carried out and results achieved will allow to pursue the study of instable thin liquid films on general solid substrates.

The capabilities of MareNostrum have allowed to reach unprecedented system sizes needed to assess the structures formed by unstable fluid interfaces in extreme regimes for fluid and forcing parameters. Regarding the simulations of propellers, MareNostrum capabilities are required to achieve the long time regime in which structural rearrangements take place and which can not be covered with other computational platforms.

Assumpta Parreño
Universitat de Barcelona

Study of SubAtomic Interactions through Lattice Quantum Chromo Dynamics on Mare Nostrum (SAIL)

Quantum Chromo Dynamics (QCD), the theory governing the interaction between quarks and gluons, the fundamental blocks of matter, has resisted analytical attacks for its solution for decades, so it has been impossible to extract from it all the nuclear phenomenology that it, in principle, contains. The team's approach, lattice QCD (LQCD) provides a method for numerically simulating QCD through a Monte Carlo evaluation of a functional integral defined on a discretized space-time. The team develops the technology and tests the ideas necessary for the LQCD computation of hadronic interactions and is performing a numerical simulation, using QuantumChromoDynamics (QCD), of hyperon-nucleon and hyperon-hyperon interactions. The large uncertainties (and in some cases complete lack of experimental data) in the experimentally extracted low energy parameters of these interactions and the phenomenological interest for the modeling of neutron stars and hypernuclei, have driven to perform a fully-dynamical QCD simulation. The results the team has previously generated in other machines show that the interaction signal is difficult to be discerned from the statistical noise and, consequently, the errors produced are large. Other sources of systematic errors in the calculation are negligible compared to those (discretization errors, for instance, are of the order of a few percent).

Improvements on the results can come about only through a significant boost in the number of gauge configurations (or more propagators per gauge configuration) and/or a different algorithm.

This work at MareNostrum contributes in improving the signal for such interactions, by increasing the statistics of the data. The team is computing more propagators per configuration, as well as exploring new ideas as new types of sources.

Thanks to the increasing supercomputing capabilities such as MareNostrum, it is feasible to make an extensive use of Lattice QCD simulations to study the nonperturbative regime of QCD, and therefore, extract in a realiable way the fundamental quantities of the interactions between hadrons in the low energy regime, in both, the strange and nonstrange sectors.
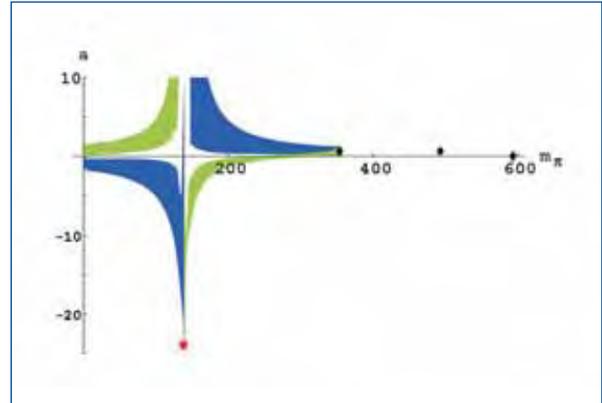
This figure illustrates the allowed regions for the scattering length in the singlet 1S0 nucleon-nucleon channel (in fm) as a function of the pion mass (in MeV). The red point represents the experimental value while the black diamonds are the results of the LQCD simulation

Josep María Porta
McGill University, Montreal

Isolating configuration spaces of polycyclic robots and molecules by extended bound smoothing

The project developed algorithms able to provide complete characterizations of the conformational spaces of molecules. The conformational space of a molecule determines the possible movements of the said molecule. Therefore, this information is basic for finding out the possible interactions between different biological structures (protein-ligand docking, drug-design, virus-cell assembly).

The algorithm will be used for geometric virtual screening. This application is computer intensive and requires significant software development: new algorithms that perform better high-dimensional conformational spaces.

In the context of the project, the team developed algorithms able to provide complete characterizations of the conformational spaces of molecules. The conformational space of a molecule determines the possible movements of this molecule and, thus, this information is basic to find out the possible interactions between different biological structures (protein-ligand docking, drug-design, virus-cell assembly, etc).

Moreover, thanks to the test in MareNostrum the group realized of some scalability issues of the procedures. Right now the team is developing new algorithms that perform better in high dimensional conformational spaces. Those algorithms, though, will be also computational demanding and, thus, it could be necessary to executed them in a supercomputer such as MareNostrum.

Thanks to the supercomputing capabilities of MareNostrum, it is possible to obtain completed characterization of conformational spaces of complex molecules in reasonable times.

The conformational space of cycloctane. Red areas correspond to conformations with higher energy.

José E. Roman
Universitat Politècnica de València

Scalable Eigensolvers in the SLEPc Library

Eigensolvers lie at the core of many computationally intensive applications, in which very large-scale eigenvalue problems have to be solved. The objective of the activity is to enhance the efficiency of these solvers, by proposing modifications of the algorithms that can lead to higher megaflop rates and less overhead due to parallel execution with many processors. The improved algorithms are incorporated into the publicly available library SLEPc.

Research in algorithm scalability aims at improving the efficiency with which scientific codes make use of large computer systems such as MareNostrum. Developers of application codes cannot invest a lot of effort in producing very efficient programs, but efficiency can also be pursued by implementing very specialized versions of the computational kernels that take the largest percentage of the computation. In this case, the research focuses on the solution of large scale eigenvalue problems, by implementing robust and efficient eigensolvers in the SLEPc library. Overall, the result is that a more efficient program implies a faster response time in the computation, a higher throughput in the computing system, and additionally the ability to solve larger or more challenging scientific problems.

The use of MareNostrum gives the team the availability of a large number of processors that would not be available otherwise.

Publication:
V. Hernández, J. E. Román, A. Tomás, "Paralelización de métodos de Krylov para cálculo de valores propios", XVII Jornadas de Paralelismo, Albacete, 2006

Iñigo Toledo
Desafío Español Copa América

Desafío Español Challenge 2007 – America-s Cup Yacht CFD Research

The objective of Desafío Español 2007 America's Cup Yacht CFD Research project is the development of the fastest America's Cup Class yacht for the 2007 America's Cup challenge. Modern America's Cup class yachts are highly refined sailing yachts that represent the product of millions of dollars of aero and hydrodynamic research by experts in their respective fields. The performance of these yachts is defined by an extremely complex balance of aero and hydrodynamic forces, where the difference in performance between the top performing yachts is less than 10 meters per km, or less than approx 1%. Since these performance differences are so small, extremely accurate and consistent performance predictions are required to identify design features which further improve performance. Traditionally, America's Cup yacht hull-form development has principally relied on a combination of potential flow analysis methods and towing tank experiments. However, the accuracy of these traditional techniques is no longer sufficient to reliably advance hull-form design. Over the past 5-10 years Reynolds Averaged Navier Stokes (RANS) codes have evolved to where they can accurately simulate complex free-surface flows, however these codes require significant computational resources. The top teams competing for the America's Cup have had the benefit of well-funded, continuous research and development programmes since the end of 2003.

The resources available at BSC were of immense benefit to the Desafío Español 2007 effort by providing the team access to state-of-the-art CFD capability. The computational resources at BSC gave the team the opportunity to substantially close the R&D technology gap with the top teams and should allow the Desafío Español 2007 sailing team to compete on an equal or better basis with the fastest America's Cup yachts developed by other competing countries.
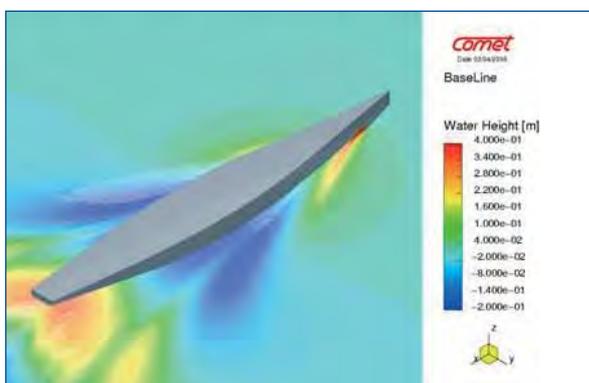
Hull form design applying Computational Fluid Dynamics (CFD).

Daniella Tordella
Politécnico de Turín (Italia)

On the role of turbulent energy and integral scale gradients in turbulent diffusion and intermittency

The project considers the basic properties of the mixing process. Turbulent mixings are very often met in nature and technology. In them the diffusion and transport of pollutants, air-borne particles (e.g. also smells and scents), chemical reactants take place. A shear-free mixing is characterized by the absence of turbulent kinetic energy production and by the presence of a kinetic energy gradient only or by the concurrency of an energy and a macroscale gradient. The team verifies and extends two recent results obtained by means of numerical experiments. The present simulations provide results for mixings with a kinetic energy ratio up to 10^6. Thanks to increasing supercomputing capabilities such as MareNostrum, it is feasible to make an extensive use of turbulence dynamics simulations to study experimentally the finest mechanism of the mixing process, and as certain the response to changes of flow control parameters such as the energy gradient and the correlation distance.
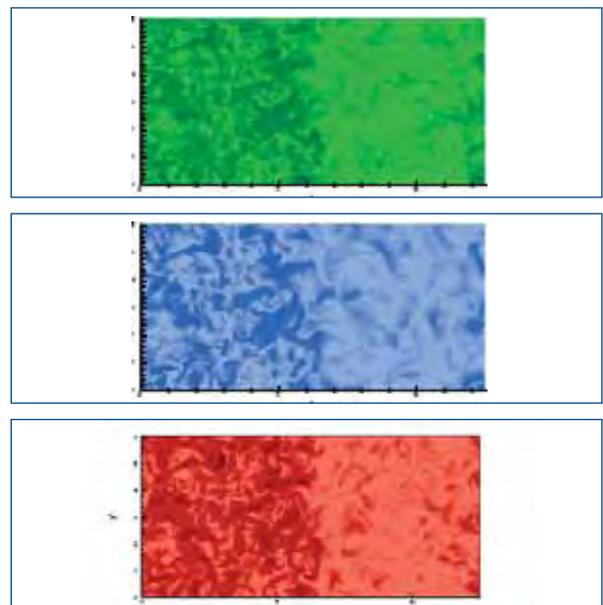
Left: high energy turbulence, right: low energy turbulence. Green: no gradient of integral scale, blue: opposite gradients of kinetic energy and integral scale, red: concurrent gradients of kinetic energy and integral scale.

Markus Uhlmann
CIEMAT

Direct Numerical Simulation of Turbulent Flow with Suspended Solid Particles

The project considers the problem of the motion of a large number (of the order of 10.000) of solid particles suspended in an incompressible fluid. The mixture is driven through a vertical plane channel by a mean pressure gradient which is sufficiently large to generate turbulent motion. In addition, the relative velocity between fluid and particles is high enough to create complicated wake effects, i.e. the particles interact with each other indirectly by means of their wakes.

The specific problem which the project is considering is that of the motion of a large number (~10000) of heavy spherical particles suspended in a fluid. The velocity of the spheres in the range where the wakes behind the particles have important effect on the motion of the particles themselves and upon the fluid statistics. All relevant scales of the fluid motion are resolved. Also, the phase interface is accurately represented, while most previous studies consider the solid phase as point-particles with no spatial extent and therefore no wake effects. The current state-of-the-art in interface-resolved DNS involves only ~1000 spheres, small flow domains and short integration times. This project aims at increasing the system size by an order of magnitude in order to allow for the numerical study of large-scale clustering phenomena and the correct representation of the turbulent scales of motion of the carrier phase. For this purpose the team simulates low-Reynolds-number turbulent flow in a plane channel (which has by now been fairly thoroughly investigated in the single-phase case) with added particles. The channel is oriented vertically and the conditions are adjusted such that the particles are on average at rest. The particle diameter is equivalent to approximately 10 times the smallest flow scales near the wall. The solid volume fraction is approximately 0.5%, i.e. the suspension is dilute and inter-particle collisions are not expected to have a significant effect. The simulations will be run until sufficient Eulerian and Lagrangian statistics have been accumulated. Additionally, a sequence of instantaneous flow fields will be generated for the purpose of visualization and further analysis. No fully-resolved DNS of turbulent particulate flow in a vertical channel has been performed to date. Applications of this type of flow configuration range from important industrial processes such as the combustion of pulverized coal, fluidized bed reactors, pneumatic conveying and coating to geophysical flows and natural sedimentation processes, e.g. in rivers and lakes, and even blood flow.

MareNostrum has allowed the team to scale up the problem to previously unattainable complexity (i.e. number of particles and size of the computational domain). Therefore the team is now capable of investigating the formation of very large scales in the two-phase system.
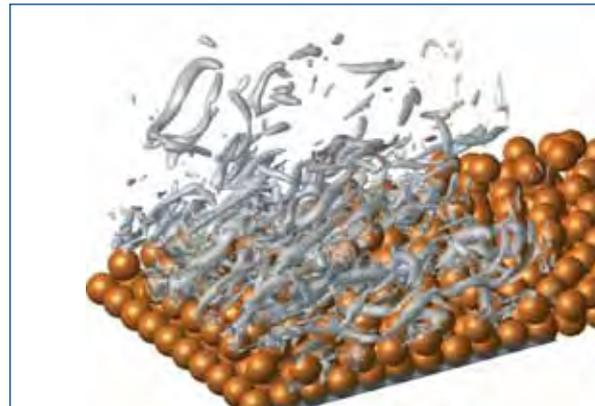
Erosion of sedimented particles in horizontal channel flow.

6.4 Chemistry and Material Science

Carlos Alemán
Universitat Politècnica de Catalunya

Dynamics of an Artificial Muscle based on Conducting Polymers and Calix[4]arene Scaffolds

The electromechanical actuation of an artificial muscle based on redox-active polymers and flexible calix[4]arene scaffolds has been investigated using molecular dynamics (MD) simulations. Initially, short MD simulations were performed to describe each electronic state involved in the actuation mechanism in terms of flexibility/rigidity, as well as to characterize the initial (accordion-like) and final (fully-contracted) conformational states. After this, the accordion-like-to-fully-contracted transition was fully characterized through a 1 μs MD simulation. The results made it possible to describe the dynamics of the molecular actuator from a microscopic point of view and also to identify secondary conformational transitions that may reduce the efficacy of the artificial muscle.
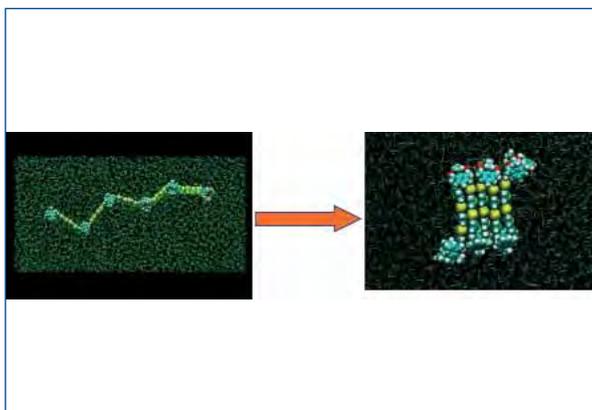
1 μs Atomistic Molecular Dynamics simulation. Molecular actuator +14509 CH2Cl2 molecules PBC and PME

Publications:
D. Zanuy, J. Casanovas, C. Alemán. "Conformational Features of an Actuator Containing Calix[4]arene and Thiophene: A Molecular Dynamics Study". J. Phys. Chem. B 2006, 110, 9876.

D. Zanuy, C. Alemán. "Contraction Process of an Electroactive Actuator Based on a One Microsecond Atomistic Molecular Dynamics Simulation". Chem. Eur. J., March 2006.

Carlos Alemán
Universitat Politècnica de Catalunya

Nanotechnology: De Novo Design using the Self-Assembly of Biomolecules

Molecular dynamics simulations have been used to construct nanofibers based on motifs taken from left-handed beta-helical protein motifs. Thus, the goal was to design nanosystems that show structural stability and can be used as templates for the experimental construction of novel systems. Results showed that a system constructed of four self-assembled replicas of residues 131-165 of galactoside acetyltransferase exhibit remarkable stability under the simulated conditions.
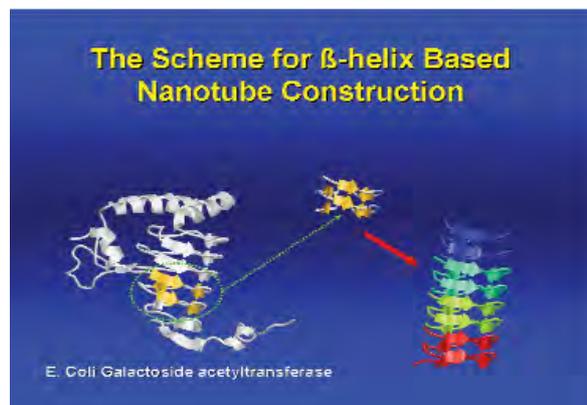
Nanotube constructed by self-assebling four copies of the left-handed, ß-helix building block motif of *E. coli* galactoside acetyltransferase.

Publication:
N. Haspel, D. Zanuy, C. Alemán, H. Wolfsom, R. Nussinov "De Novo Tubular Nanostructure Design based on the Self-Assembly of beta-Helical Protein Motifs", Structure 2006, 114, 1137-1148.

Pedro de Andrés
CSIC

Molecular adsorption on metallic and oxide surfaces

Density Functional Theory was used to study surfaces with or without adsorbed chemical species. These systems are relevant as catalytic models, but they can be important in other fields too, like electronic devices, corrosion, etc. As a way of example, the surface $TiO_2(110)$-$(1x2)$ from a structural and electronic point of view was characterized. The results include finding a potentially active site for catalysis, and the existence of quasi-1D metallic chains. On the methodological side, the project has taken advantage of MareNostrum parallel power to devise a structural multidimensional minimization algorithm where different parameters are simultaneously optimized. This new algorithm shows a Poisson-like distribution, i.e. it performs a most efficient search unlike previous ones showing Gaussian-like distributions. The project has also studied the conductivity of a single H_2 molecule pinched between two Pt tips, and the adsorption of large PTCDA molecules on Ag(110).
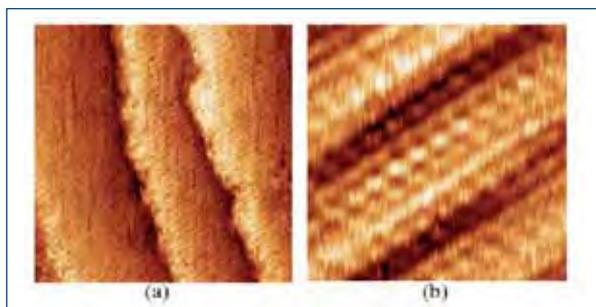
Constant current STM images of 1 x 2 TiO2 (110), V=+1.5V.(a) Monatomic steps and large terraces on a 500 x 500 Å^2 area (1=o.146nA). (b) A high resolution image on a 30 x 30 Å^2 area (1=0.178nA).

Publications:
M. Blanco-Rey, J. Abad, C. Rogero, J. Mendez, M.F. López, J. A. Martin-Gago, and P. L. de Andres, "Structure of Rutile TiO2 (110)-(1x2): Formation of Ti2O3 Quasi-1D Metallic Chains"; Phys. Rev. Lett. 96, 055502 (2006).

M. Blanco-Rey, P. de Andres "Surface Diffraction Structure Determination from Combinatorial Simultaneous Optimization" Surf. Sci. Lett. 600, L91 (2006).

J.A. Verges, P. de Andres; "Electron Transport Through Simple Molecules From Band Structure Formalism"; 3rd Nano Spain Workshop (Pamplona, March 2006).

Andrés Arnau
Universidad del País Vasco

Adsorption of alkanethiols on Au(111) at different coverages up to one monolayer

The project uses density functional theory to study the formation of ordered and disordered structures of water at different coverages on metal and insulating surfaces. At low coverages water-substrate interaction play an important role. However, as coverage increases water-water interactions become more important and its competition with water-substrate interactions determines the final structure.

The study of water adsorption on NaCl(100) surfaces is important to understand chemical reactions occurring at the wetted surface of sea-salt aerosols, which play a key role in tropospheric chemistry, as thin films of water molecules cover most surfaces under ambient conditions. This phenomenon is crucial in biology, material sciences and chemistry. The team find that in the presence of one monolayer of water on the (100) cleavage plane of NaCl, Cl$^-$ ions are displaced at very low energy cost from their crystal lattice positions out towards the plane of water molecules, forming a two-dimensional solvated structure and propose that this is the mechanism behind the catalityc activity of salt at low relative humidity to several reactions involving atmostpheric gases, including ozone depletion.

The study of water adsorption on oxygen covered Ru(0001) surfaces is the most fundamental character. The aim of this study is to understand the formation of wetting layers depending of O coverage on the surface. On clean Ru(0001) surfaces a bilayer structure of water is formed due the matching between the lattice constants of the basal plane of hexagonal ice and Ru(0001). A small amount of oxygen seem to perturb the formation of a perfect wetting layer that, furthermore, contributes to its dissociation as temperature increases or electrons are irradiated towards the surface.

However, at 0.25 ML oxygen coverage an ordered p(2x2) water superstructure seems very likely, according to our DFT calculations. Increasing oxygen coverage on the substrate is expected to influence the formation of water structures due to the fact that hydrogen bonds between the water molecules and oxygen atoms of the surface start playing an increasing role.

Thanks to the increasing supercomputing capabilities such as MareNostrum, it is feasible to make an extensive use of massive computation using large supercells of up to 32 atoms per atomic plane.
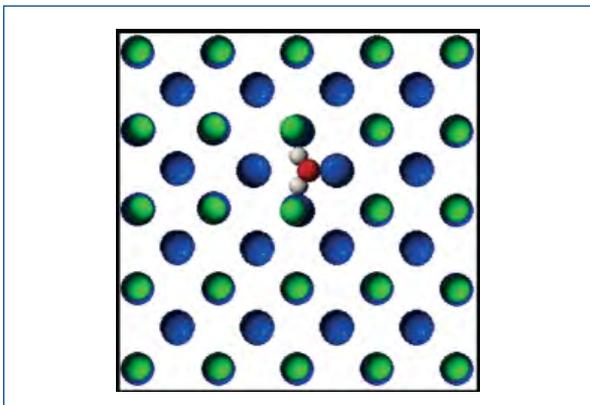
The diffusion of a water molecule on a NaCl(100) surface.

Publications:
N. Gonzalez, N. Lorente, and A. Arnau "Methylthiolate adsorption on Au(111): energetics vibrational modes and STM imaging" Surf. Sci. 600, 4039 a 4043 (2006).

P. Cabrera-Sanfelix, A. Arnau, G. Darling and D. Sánchez Portal "Water Adsorption and Diffusion on NaCl (100)" Journal of Physical Chemistry B (2006).

Carles Bo
ICIQ

Structure of polyoxometalates in solution: from single anions to nano-capsules

POMs constitute an immense class of polynuclear metal–oxygen clusters usually formed by Mo, W or V and mixtures of these elements. POMs molecular size span a broad range, from the small anions to nano-sized giant balls which present big empty cavities and can host other species. POMs have relevant applications in many fields including medicine, catalysis, multifunctional materials, chemical analysis, nanotechnology, etc. In this project, molecular dynamics simulations are being used to investigate the structure of water solutions of POMs and the dynamics of the interactions between POMs and cations.

One of the most significant features of the chemistry of Tungsten and Molybdenum is the formation of molecular clusters known as polyoxometalates (POMs). Nowadays, POMs constitute an immense class of polynuclear metal–oxygen clusters usually formed by Mo, W or V and mixtures of these elements. POMs molecular size span a broad range, from the small and well known Lindqvist anion $[M_6O_{19}]^{n-}$, to nano-sized giant balls such as those recently synthesized by Müller containing more than 100 metal atoms and more than 350 oxygen atoms. Such giant metal oxide molecular clusters present big empty cavities which can host other species. POMs have relevant applications in many fields including medicine, catalysis, multifunctional materials, chemical analysis, nanotechnology, etc.

In the first stage of the project the team were interested in the study of the ion-pairing in Keggin anions with different charge and with different counterions. Since the team already had carried out molecular dynamics simulations for a single POM $Y_n[XW_{12}O_{40}]$ (Y=Li, Na, K; X=Si, P, Al) in water, the team wondered whether the results of those simulation were size-dependent. Therefore, the team studied the dynamics of the same systems but in a simulation box eight times larger than that the team studied previously using in-house computing resources. These kinds out calculation are beyond the computing capabilities at ICIQ. Firstly, the team studied the performace of DLPoly when running in parallel with the systems. The group found that for a system with 3056 atoms, the parallel scaling was very bad beyond 4 CPUs. However, when a larger system (9936 atoms) was considered, the team found:

CPUs	Time (s)	Scaling (%)	Time for 1ns (h)	Time for 1ns (d)
2	156,182	100,000	433,839	18,077
4	76,048	205,373	211,244	8,802
8	41,552	375,871	115,422	4,809
16	21,759	717,781	60,442	2,518
32	14,666	1064,926	40,739	1,697
64	15,235	1025,153	42,319	1,763

Therefore, the team concluded that 32 CPUs was the best choice. In this way, the MareNostrum supercomputer enabled the team to run 1 ns simulation in 2,7 days while the same calculation would cost 49,3 days in 1 CPU. Moreover, the new results obtained indicate that the physical-chemistry properties measured in those systems are size-independent and that the previous conclusions do not change.

In the second stage of the project the team considered larger POMs, such as the wheel shaped compound Na21[Mo57O189V6N6H3(H2O)18] in a box including 9018 external water molecules. In this case, and since the cluster present an interior cavity, the team were interested in the behaviour and ordering of the water molecules inside the cage, and in the effect of the temperature. The results obtained indicated that the 12 water molecules a coordinated to Mo and V atoms and at the same time they also form two six-water rings linked by hydrogen bonds. This water assemblies are stable with temperature, in full agreement with experimental results. In contract with bigger POMs that the team plans to study in the next period, those water molecules do not scape outside the cage.
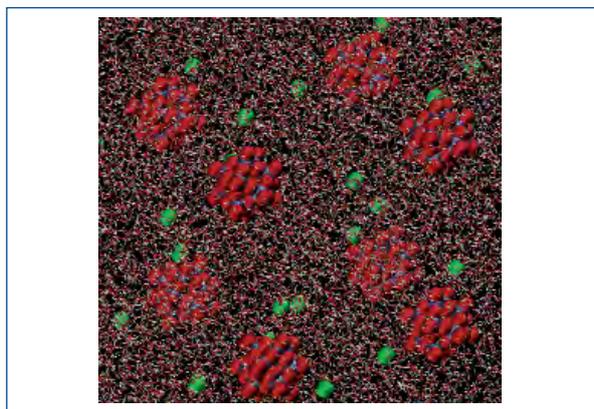
8 $[PW12O40]^{3-}$, 24 Li^+ , 8000 H_2O . 1 ns in 2.7 days.

María José Caturla Terol
Universidad de Alicante

Modeling materials properties at the nanoscale: from defects to nanodevices

Understanding the relationship between microscopic features and macroscopic behavior is crucial to develop new materials with specific properties. The objective of this work is to use molecular dynamics simulations to study the influence of atomic scale defects and nanostructures on the properties of the material. In particular, the team studies the production of defects under irradiation as well as the mechanical properties of materials with nanoscale features such as nanowires and nanocrystals. Molecular dynamics simulations are being performed using the parallel code MDCASK. The aim of these simulations is to understand how defects at a nanoscale can affect materials properties. On one hand defects produced by irradiation are studied, in particular their interaction with dislocations. These calculations are relevant for material properties degradation under irradiation. On the other hand, mechanical properties of new materials, such as nanocrystals, are also studied using this method.

The strength and mechanical properties of metals is controlled by defects called dislocations. Under irradiation these properties can be significantly affected. In order to understand this phenomenon the team used molecular dynamics simulations to study the interaction between defects, produced by irradiation, and dislocations. The team's model system is an f.c.c. lattice (copper). Two edge dislocations are included in the crystalline lattice. These dislocations dissociate into two partial dislocations, separated by a stacking fault. Then, the effect of an energetic particle is modeled by transferring an energy of 20keV to one of the atoms in the lattice. As a result a collision cascade occurs, where many atoms are displaced from their original positions.

Most of these displaced atoms return to their original lattice sites, restoring the perfect lattice. However, a few remain as defects, such as atoms that are occupying positions outside the perfect lattice (interstitials), or atoms that are missing from a lattice site (vacancies). When a dislocation is present the team can also see how the collision cascade affects the nature of this dislocation. On one hand changes in the dislocation structure are observed as a result of the collision cascade. These changes could affect the way the dislocations move under applied stress and therefore could affect the plastic behaviour of this material. On the other hand defects, such as interstials or vacancies, produced by the irradiation interact with the stress field of the dislocation and the group can observe, in the time scale of the simulations, how some of these defects are attracted to the dislocation. Further analysis is needed to fully understand all the mechanisms observed in the simulations.

MareNostrum allows the group to study system sizes that can not be reached with other computational capabilities.
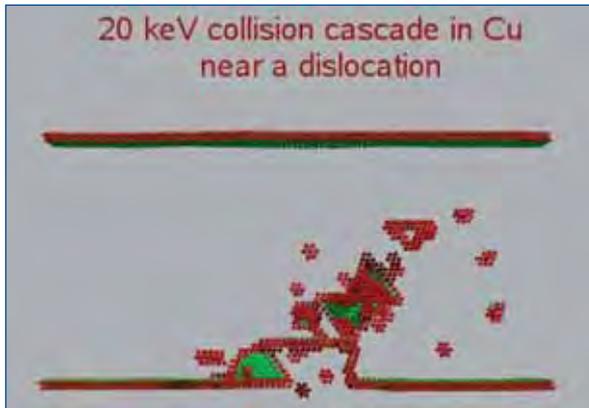
Calculations of collision cascades next to dislocations in Cu.

Publications:
Z. Yao, M. J. Caturla, R. Schäublin, "Study of cascade damage in Ni with MD by different interatomic potentials", accepted in Journal of Nuclear Materials.

C. Ortiz, M. J. Caturla, "Simulation of defect evolution in irradiated materials: the role of intra-cascade clustering and correlated recombinations", submitted.

Albert Cirera
Universitat de Barcelona

Ab initio computation of the effects of vacancies for transparent electronics in In2O3

Tin- indium oxide is an ideal material for flat panel displays because of its high electrical conductivity and high optical transparency, which approach the team to the transparent electronics. Since 2001, Wang et al reported the production of nanobelts of oxides such as SnO2, In2O3 and many others, the potential issues of the dimensionality and space-confined transport phenomena from 0d and 1d towards 2d (the surface of the nanobelt) arise. In the present project, the team simulates by DFT the properties and molecular interactions with the indium oxide surface.

Binary semiconducting oxides, such as SnO2 or In2O3, have distinctive properties and are now widely used as transparent conducting oxide materials and gas sensors. SnO2 nanoparticles are regarded as one of the most important sensor materials for detecting leakage of several inflammable gases owing to their high sensitivity to low gas concentrations. Tin- indium oxide is an ideal material for flat panel displays because of its high electrical conductivity and high optical transparency, which approach the team to the transparent electronics.

In fact, whereas SnO2 crystallizes as cassiterite (P4/mnm) and there is only 2 groups of SnO2 non-equivalent (translation symmetry) per unit cell, In2O3 crystallizes as byxbyite (Ia3) and every unit cell contains 8 units of fluorite-type unit cells, composed by 8 In atoms at the b site, 24 In atom at the d site and 48 oxygen atoms at the e sites according Wyckoff's notation. Therefore, the team should compute unit cells of 80 atoms in In2O3 instead of 6 atoms in SnO2. The practical consequence is the complete electronic structure of In2O3 has been not reported until the end of 90's being necessary the use of the latest supercomputing techniques and facilities. As previously occurred with less complex materials, the deeper theoretical knowledge of the physico–chemistry of In2O3 will allow the technological development of novel and improved applications in a much faster than just with a trial-and-error approach.

In the present work, some extremely demanding calculation of the well known SnO2 surface have been performed as the starting point of the study of the technologically interesting In2O3 and its the surface effects. The study of In2O3 is being performed with first-principles methodology based on the density functional theory (DFT) as implemented in the SIESTA code concerning the following tasks:

1. Preliminary and extremely demanding SnO2 surface calculations.

2. Simulation of the bulk structure. Evaluation of exchange-correlation functional, orbital basis and atomic potential.

3. Bulk electronics and phononic modes.

4. Simulation of low index surface. Determination of lower energy surfaces. Convergence and relaxation of low energy surfaces.

5. Electronic and phononic modes in low energy surfaces.

6. Vacancy description in surface and bulk. Exploitation of results.

7. Molecular modeling –previous results already obtained at CESCA will be used here.

8. Simulation of adsorption of molecules on In2O3 low energy surfaces. Convergence and relaxation.

9. Initials in the design and simulation of nanostructures.

Main results of the first points are:

- Final assessment of previous models of SnO2 and exploitations of results.
- Assessment and test of pseudopotential for indium and oxygen.
- Assessment and Test of LCAO basis
- Construction of unit cell
- Preliminary relaxation
- Test of minimum real space mesh cut-off
- Test of minimum a reciprocal space grid cut-off
- Further relaxation
- Computation of bulk modulus as test of structure relaxation (very expensive due to complex unit cell)
- Computation of local density of electronic states
- Computation of projected density of states
- Construction of surfaces (hkl)={100}, {110}, {111} in 1 to 4 layer systems.
- Relaxation of the corresponding surface systems (extremely expensive due to complex unit cell in multi layer system).
- Computation of the surface energy per number of layer in every case and determination of the most stable surfaces.

The nice description of bulk and surface previously obtained will allow to reliably compute the electronic structure of the surfaces. However, due to some experimental determinations by TEM (personal communication by Dr. Jordi Arbiol –SCT-UB) surface analysis will be extended to (hkl)={110}. Additionally, the team intends to hard focus on the influence

of vacancies (and interstitials if possible). Actually, some works describe its influence on the electronic properties for other materials, like SnO2 where it is suggested as important fact for the transparency. Regarding the main application of In2O3, the team will follow the scheme of work proposed for II-VI semiconductors by Lany et all in a recent paper. Despite the complexity of such computations, some previous computation should be taken into account (for instance) as well as experimental works.

Thanks to the increasing supercomputing capabilities such as MareNostrum, it is feasible to simulate complex systems including broken symmetries: surfaces, molecular adsorption, slabs and defects. MareNostrum offers the possibility to be in touch with real applications which never behave as perfect systems.
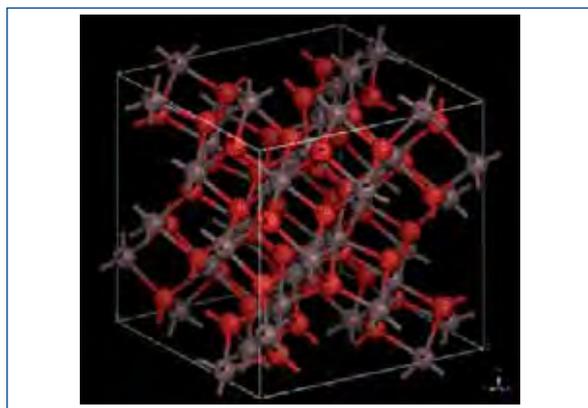
Unit cell model of the In2O3 with 80 atoms. To successfully carry out this huge computational task the latest and bigger supercomputing facilities, like MareNostrum, are required.

Avelino Corma
CSIC

Selective hydrogenation of nitroaromatics catalyzed by gold

The catalytic hydrogenation of nitroaromatics is easily carried out with commercial catalyts. But the selective reduction with H_2 of nitro groups in molecules with other sensitive functional groups (C=O, C=C, C=N) is often difficult to achieve. It has been found that supported gold catalysts, (Au/TiO_2 and Au/Fe_2O_3), are active and highly selective for the reduction of nitro groups with H_2 in substituted nitroaromatics. The project is studying with periodic QM methods the mechanism of gold catalyzed hydrogenation of nitro and other functional groups.

In a first step to understand the mechanism of selective hydrogenation of nitroaromatics catalyzed by supported gold nanoparticles, in this period the team has studied the adsorption of nitrobenzene and 3-nitrobenzostyrene on the Au(111) surface, on a defect containing Au(111) surface, and on three different gold particles, in order to check wether low coordinated gold atoms are able to interact with either the nitro group or the C=C double bond. An almost negligible interaction between these molecules and the perfect metal surface has been found. However, both the nitro group and the C=C double bond can be activated by low coordinated gold atoms located on small nanoparticles (25-38 gold atoms) or on surface defects. The interaction of gold atoms with the C=C bond is stronger than with the nitro group, and it usually involves a deformation of the nanoparticle.

The team shall now introduce the Titanium oxide support in the study in order to analyze how it modifies the properties of the supported gold nanoparticles and to check wether it is involved in the inicial activation of the reactants. The knowledge of the mechanism of selective hydrogenation will help to improve the catalysts actually employed in this important process.

Thanks to the increasing supercomputing capabilities such as MareNostrum, it is possible to investigate the mechanisms of solid catalyzed reactions. The correct modelization of solid catalysts consisting of metal nanoparticles supported on metal oxides involves a large computational effort that can only be achieved in supercomputers such as Marenostrum.
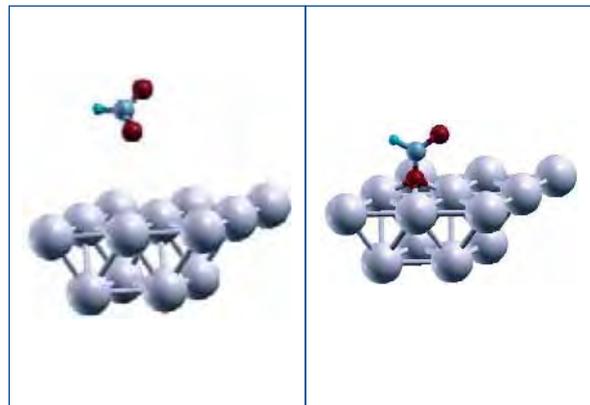
Possible adsorption mode of HNO_2 on Au(111)
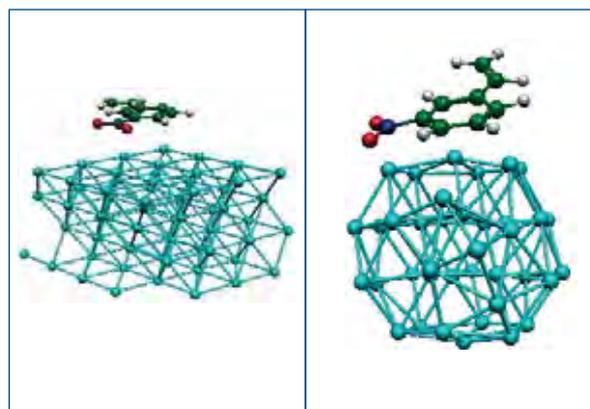
3-nitrobenzostyrene adsorbed on Au(111) surface and on a gold nanoparticle

Jordi Faraudo
Universitat Autònoma de Barcelona

Molecular Dynamics Simulations of Electrostatic Interactions between Macromolecules and/or Macroions in Aqueous Solution

Electrostatic interactions are essential in determining the behaviour and functionality of large molecules or nanoparticles in solution. Applications of interest include nanofunctionalized materials, bio-inspired materials and the design of certain pharmaceutical drugs to mention only a few examples. In the project, the team studies how explicit effects due to the molecular details of the solvent dramatically modify electrostatic interactions at short distances between the interacting molecules or nanoparticles, originating the so-called hydration forces. Simulations of films and monolayers of SDS surfactant adsorbed in water and gold nanoparticles covered by surfactant at water/air interface were performed. In both cases, the molecular detail is accessible only with MareNostrum. The results obtained for large and very large systems in this project on MareNostrum allow the team to demonstrate that the specific behaviour of electrostatic interactions due to the modification of the solvent structure near amphiphilic interfaces is even more important than expected and extends far from the interface. These new effects observed in the simulations are expected to appear at the nanoscale in many chemical systems in aqueous environments. This is particularly important in the design of chemicals and materials tailored on a nanoscale acting in aqueous environments.
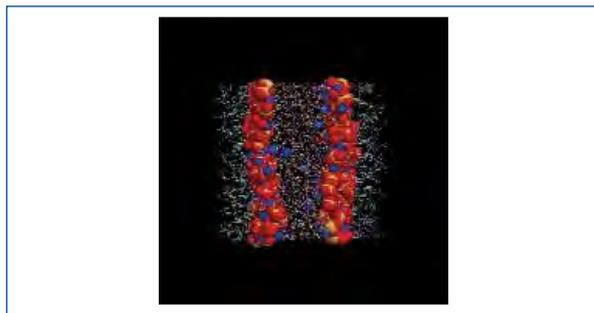
Snapshot of interaction between two surfactant layers.

Publications:
K. Tay and F. Bresme, J. "Hydrogen bond structure and vibrational spectrum of water at a passivated metal nanoparticle", Mater. Chem., 2006, vol 16, page 1956

F. Bresme y J. Faraudo, Molecular Simulations, "Temperature dependence of the structure and electrostatics of Newton Black Films: insights from computer simulations" (2006).

Ernesto García
Universidad del País Vasco

Dynamics of the OH+CO->H+CO2 reaction at low collision energy

The OH + CO reaction is an important reaction in both combustion and atmospheric chemistry. From the theoretical point of view, this reaction is considered as a prototype of complex-forming reactions. The trans- and cis-HOCO complexes associated with deep wells located in the strong interaction region are connected to reactants through small barriers and van der Waals wells and to products through high barriers. Accordingly the dynamics of the reaction can exhibit both an attractive and a repulsive behaviour. In the low collision energy dynamical study reported in this work has been found that the barriers dominate the dynamics of the reaction. Calculations have been performed using a quasiclassical trajectory code parallelized using mpi. The code achieves a very good speed-up even when 1024 processors are used.

The reaction of the hydroxyl radical with carbon monoxide, OH + CO -> H + CO2, is an important reaction in combustion and atmospheric chemistry. It is, in fact, the main source of CO2 in hydrocarbon combustion and the key reaction for the control of the concentration of the OH radicals in the lower atmosphere. This reaction has still to be understood and the related mechanism and the related reaction mechanisms have not yet elucidated. As a matter of fact, for this reaction not only there is no available potential energy surface (PES) of chemical accuracy but also accurate quantum dynamical methods are unfeasible.

Moreover, the OH+CO reaction is particularly difficult to be dealt theoretically due the fact that the PES is rather complex and exhibits several stationary points and multiple reaction paths. More in detail, the PES exhibits two intermediate deep wells corresponding to the trans- and cis-HOCO conformers, which are connected to the reactants through small barriers and van der Waals wells and to the products through high barriers. This is important to the end of elucidating the relative importance of the attractive (wells) and repulsive (barriers) features of the PES on the behaviour of the system at trheshold. To this end, a calculation of the energy dependence of the reactive cross section of OH+CO at low collision energy has been performed. The calculations are based on the quasiclassical trajectory method (the only possibility for calculating full-dimensional detailed cross sections for tetratomic reactions). In this method, Hamilton equations for different initial conditions are integrated. The integration is carried out for a long cpu time because of the complex topology of the PES. Fortunately, the trajectories are independent events and therefore suitable for a natural parallelization. As a matter of fact, this is what the mpi parallel code has led to a very high speed-up even with 1024 processors. Cross section calculations for the OH+CO reaction

with both reactants in the ground vibrotational state have been performed at collision energies as low as 0.02 eV. The excitation function shows indeed a threshold (even if it happens at very low collision energy). After all the shape of the excitation function has a form typical of reactions occurring on a PES having a potential barrier which increases from the threshold with energy up to a constant value. Future work on this reaction will be devoted to the comparison of various more detailed properties (such as the product angular and translational energy distributions) calculated on the available surfaces with experimental results in order to test the accuracy of the proposed PESs.

Thanks to the increasing supercomputing capabilities offered by large scale facilities such as MareNostrum, it has been possible to carry out an extensive computational campaign allowing a detailed study of the dynamics of complex-forming reactions at low collision energy on a very accurate, but computationally demanding, potential energy surface using a high performance parallel code.
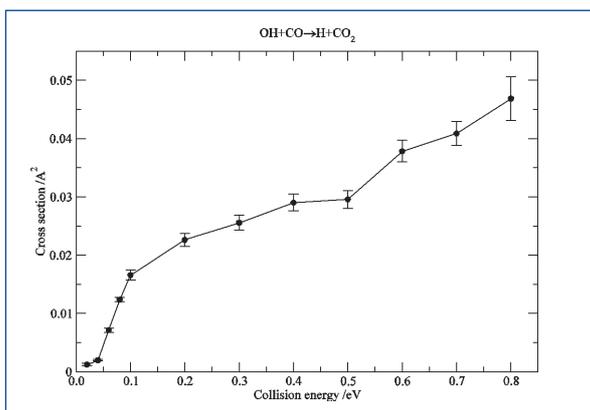
Plot of the excitation function (cross section vs. collision energy) of the OH+CO reaction for both reactants in the ground vibrotational state. The threshold appears is located at very low collision energy.

Gregory Geneste
Ecole Centrale Paris

Ab initio simulation of ionic conductors for solid oxide fuel cells

The team uses highly accurate ab initio DFT calculations to determine the physical properties of various ionic conductors potentially involved in Solid Oxide Fuel Cells. The calculations allow to understand the diffusion paths of oxygen ions and protons in a wide number of crystalline compounds such as BaSnO3 or LaNbO4, so as to determine the relevant parameters that could be optimized in the future to obtain new and functional materials that would make SOFCs suitable for a large-scale industrial application.

The group uses MareNostrum to study within ab initio DFT calculations the properties of ionic conductors for Solid Oxide Fuel Cells. First the group focuses on baryum stannate BaSnO3, a compound for which interesting levels of ionic conduction have been recently reported.

MareNostrum has allowed to progress into two different directions:

1. The structural, dynamical and dielectric properties of $BaSnO_3$ have been widely investigated. The team has shown that this compound remains cubic until very low temperatures, thanks to high-accuracy and high CPU-time consuming phonon bandstructure calculations. This result has been confirmed by the experimentalists of the laboratory, who have carried out X-ray diffraction experiments on $BaSnO_3$ powders down to 7 K. Note that the low-temperature properties of this compound had never been investigated before. The group has also determined the high-pressure behavior of baryum stannate and shown that, interestingly, one optical mode becomes unstable above a critical pressure in the cubic structure.

2. The doping within Lanthanum atoms has been investigated thoroughly. Two doping rates have been tested for the moment: 25% and 50%, which correspond to structures that have been synthesized by the experimentalists of the lab. The team has shown in particular that the La doping atoms have repulsive interactions. The electronic structure has been investigated and shows the existence of a narrow empty band in the gap of the system, consistently with the appearance of oxygen vacancies in the system. The next step is to understand the interactions of water with these vacancies and the formation of the so-called protonic defects that are responsible for the ionic (protonic) conductivity in such systems. MareNostrum allows to perform very consuming CPU-time calculations such as phonon bandstructure computations. Some calculations the team has performed on $BaSnO_3$ in this respect would never have been possible in local workstations. Such calculations are particularly

useful in perovskite systems such as $BaSnO_3$ since they allow to determine whether the cubic structure is stable, and, if not, towards which kind of structure the system is going to transit.
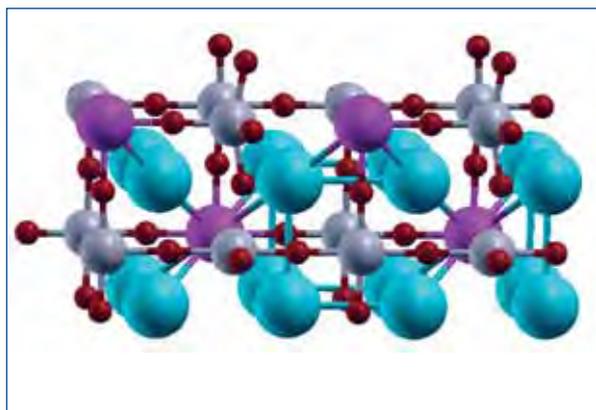
One of the supercell used to simulate a La-doped $BaSnO_3$ ordered system. Blue: Ba atoms, Grey: Sn atoms, Purple: La atoms; Red: Oxygen atoms.One can recognize the perovskite structure ABO3, with B = Sn, substituted by 25% La. In such a system, the insertion of lanthanum must be compensated by oxygen vacancies.

Publications:
E. Bévilllon, G. Geneste, "Unstable polar mode and minimum of teh dielectric constant in cubic BaSnO3 under hydrostatic pressure", accepted in Phys. Rev. B.

E. Bévilllon, G. Geneste, G. Dezanneau, A. Chesnaud, "Reunión Generale du GDR-DFT++", "Etude theorique d'un conducteur protonique", http://lepes.grenoble.cnrs.fr/gdr-dft/

Manuel María González Alemany
Universidad de Santiago de Compostela

First-principles calculations for the electronic structure of semiconductor nanowires

Semiconductor nanowires are one-dimensional nanostructures that constitute ideal building blocks for highly integrated nanoscale electronics and optoelectronics. In this project the team is determining the electronic structure of semiconductor nanowires from first-principles calculations. In particular, the team is using a code developed by them, the PARSEC code (www.ices.utexas.edu/parsec/), which is devised to take advantage of massively parallel environments like MareNostrum.

The historical rates of improvement in productivity experienced by the semiconductor industry during the last decades are based on a continued miniaturization of silicon-based devices. However, existing materials and technologies are approaching their physical limits, which will soon prevent the industry from reaching similar rates of improvement. Finding new materials and technologies adequate for the fabrication of highly integrated devices in the deep nanometer regime would potentially overcome such important limitation. Among the materials investigated on this regard, semiconductor nanowires (NWs) are of special interest. One-dimensional materials like NWs are the smallest dimensional materials that can transport efficiently electrical carriers and optical excitations. Also, NWs with controlled electrical and optical properties can be synthesized for a wide variety of semiconductor species, and are known to work as functional devices like sensors, light emitting diodes, and field effect transistors. As such, NWs represent ideal building blocks for highly integrated nanoscale electronics and optoelectronics.

Despite the technological usefulness of semiconductor NWs being mostly derived from the introduction of electrical carriers in the NWs, little is known about fundamental issues behind doping these nanomaterials. For example, what is the nature of the carriers responsible for electric conductance in the doped NWs (are these states obtained by energy band engineering spatially localized or extended states?), and how size affect their electronic properties (how are the carriers affected by quantum confinement?). Answering such questions is not only important from a fundamental point of view. It is also essential to design future devices and applications based on NWs, as well as to improve characteristics of those devices already existent. Knowing the properties of NWs is a necessary step for determining the optinum working conditions of NW-based electronic devices and opto-electronic devices.

In the activity developed at BSC the team is addressing issues like the ones listed above by performing first-principles calculations on doped semiconductor NWs. In particular, the team is studying NWs constructed from III-V semiconductor

materials, following the enormous attention that these nanomaterials have recently attracted. The results obtained are being very promising. The impurity state introduced by p-type doping the NWs have been already identified, and the effects that the quantum confinement has on this state determined. This is the first time that such a characterization of an impurity state introduced in any nanowire by p- or n-type doping is performed.
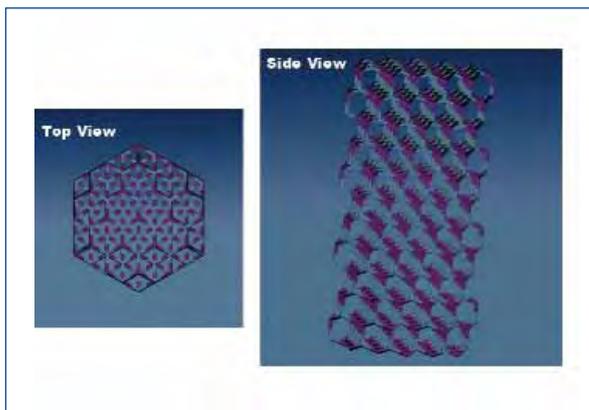
Top and side view of a nanowire constructed from III-V semiconductor materials as studied within our theoretical framework

Eduardo Hernández
ICMAB-CSIC

Electronic structure of materials: thermal effects and ab initio thermodynamics

During this project the team has pursued two main lines of research. On the one hand the team has performed ab initio molecular dynamics simulations directed at understanding the thermal properties and phase behaviour of simple metals Li and Na. The second line of research has focused on the thermodynamics of systems where quantum effects are important, in particular focusing on the case of light impurities in semiconductors, such as H, D, muonium in diamond. This required the performance of Feynman path integral simulations, for which the availability of MareNostrum was crucial.

The team has performed intensive simulations along two different research lines encompassed in the above project, namely: 1) path-integral molecular dynamics studies of pure and hydrogen-doped diamond, and 2) ab initio studies of the thermal behaviour of sodium and lithium. The path-integral simulation technique allows for a correct quantum treatment of the atomic degrees of freedom, which are normally treated approximately by means of classical mechanics. This is particularly important in the case of light atoms such as carbon and hydrogen, where quantum effects are most appreciable.

The thermal behaviour of simple metals Li and Na is currently a topic of great interest in the condensed matter physics community, and the results will be timely and attract large attention. The team has conducted extensive ab initio molecular dynamics simulations which clarify the nature of the re-entrant behaviour observed experimentally in the melting curve of Na, and have performed large two-phase simulations of solid-liquid Li, which have allowed the team to estimate the melting temperature of Li at different pressures. Without a large parallel supercomputing facility such as MareNostrum it would have been impossible for us to conduct the simulations performed. Local computing facilities of the groups involved in these projects could not cope with such large simulations.
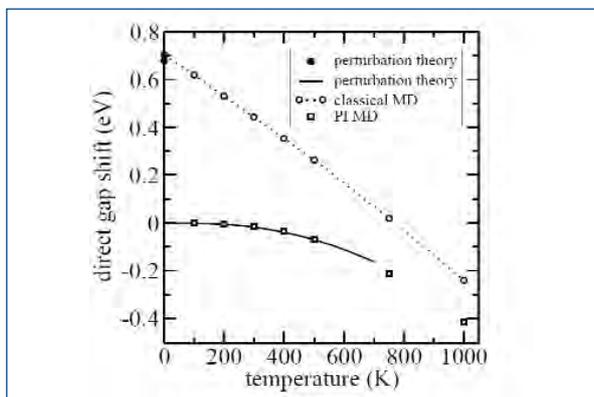
(Coming from page 135) Behaviour of the direct forbidden electronic gap in diamond as a function of temperature. Open circles are the results obtained from classical molecular dynamics simulations of diamond, while the open squares are the results obtained from quantum path-path integral simulations.

Publications:
Rafael Ramirez, Carlos P. Herrero and Eduardo R. Hernandez. "Path-integral molecular dynamics simulations of diamond". Physical Review B, vol. 73, 2452202 (2006).

Carlos P. Herrero, Rafael Ramirez and Eduardo R. Hernandez. "Hydrogen and muonium in diamond: a path-integral molecular dynamics simulation". Physical Review B, vol. 73, 245211 (2006).

Eduardo R. Hernandez and Jorge Iñiguez. "On the nature of the melting line of Sodium". Physical Review Letters (2006).

Francesc Illas
Universitat de Barcelona

Scalability of the VASP code and applications to heterogeneous catalysis

A series of benchmark calculations have been carried out using VASP to check the scalability of this broadly used code on MareNostrum. Good scalability is found but with a maximum of 32 processors. On the other hand, transition state search with the Nudged Elastic Band algorithm exhibits a large scalability since 4 images runs are carried out using 128 processors. Complicated structures may need more images and this makes MareNostrum an ideal architecture for such problems.
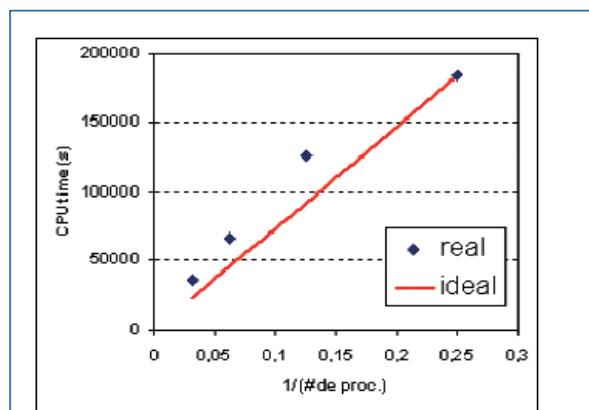
Type of job: geometry optimization of a MgO $6\sqrt{2}$ x $6n\sqrt{2}\sqrt{2}$ x 4 slab, 576 atoms unit cell with 12 Å of vacuum. Relaxation of the two outermost layers.
$E_{cut-off}$ = 415.0 eV and 1 x 1 x 1 k-points

Francesc Illas
Universitat de Barcelona

Towards understanding the molecular mechanisms of heterogeneous catalysis

The atomic structure and thermodynamic stability of the Ag(111)($\sqrt{7}$x$\sqrt{7}$)-R19.1°-CH$_3$S has been studied by means of density functional calculations and atomistic first principles thermodynamics. The unreconstructed and two recently proposed reconstructions have been considered. It is found that, in spite of significant differences in the atomic structure, the different surface models have a very similar surface free energy. It is claimed that the different ordered phases can coexist and that the appearance of one or another depends on the external preparation conditions.

Thanks to increasing supercomputing capabilities such as MareNostrum, it has been possible to study these surface reconstructions involving quite a large unit cell.
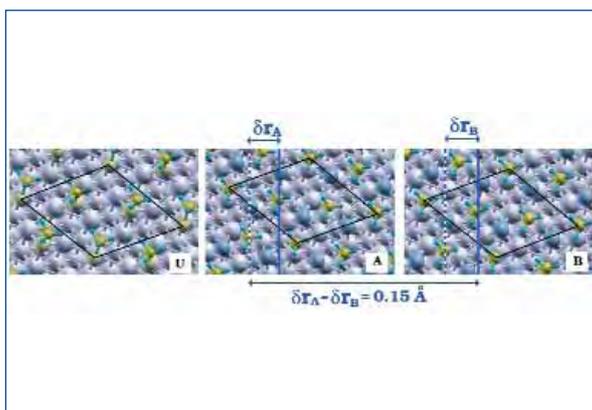
Schematic plan view of the different surface models considered for the U, A and B (top, middle and bottom) atomic models of the ($\sqrt{7}$x$\sqrt{7}$)R19.1°-CH$_3$S phase. The vertical line indicates the displacement of the surface silver atoms of the reconstructed B model relative to A. The Ag atoms of the topmost atomic layer of the A and B models are presented in a darker colour to facilitate the view. The main structural difference between the A and B initial models is highlighted. The optimized geometries are almost identical.

Publication:
D. Torres, P. Carro, R.C. Salvarezza, and F. Illas, "Evidence for the formation of different energetically similar atomic structures in Ag(111) ($\sqrt{7}$x$\sqrt{7}$)-R19.1°-CH$_3$S", Physical Review Letters journal 97, 226103 (2006). (it also has been selected for the December 11, 2006 issue of Virtual Journal of Nanoscale Science & Technology).

Zbigniew Lodziana
Polish Academy of Sciences, Krakow (Poland)

Quantum calculations of electron correlations and the surface properties of magnetite

Surfaces of metal oxides play a prominent role in catalysis and those oxides with closed band gap are very promising materials for novel applications. Unfortunately their properties are poorly known. Magnetite (Fe3O4) is one of such materials. The objective of the present project is to calculate properties of (001)Fe3O4 surface within LDA+U method and possibly resolve the charge-ordering problem on the surface. Magnetic properties of magnetite (Fe3O4) were known to the mankind since the ancient times. Despite enormous research efforts not all mysteries of this material are known today. This especially concerns the hypothesis of charge ordering on the surface at temperatures much higher than metal – insulator transition of bulk material. The team provides explanations of open questions about (001) surface of Fe3O4. The surface is insulating, which is attributed to the particular charge ordering of electrons on d – orbital of the octahedrally coordinated iron.

This material was used as a mariners compass over the centuries. It is almost seventy years since the discovery of metal - insulator transition in Fe3O4. Since that time the researchers are trying to explain unusual properties of this material.

Despite enormous efforts not all mysteries of this material are known today. While an agreement about the bulk properties, nature and origin of metal insulator Verwey phase transition seems to emerge recently; properties of the magnetite surface are still puzzling. This especially concerns the hypothesis of charge ordering on Fe cations at temperatures much higher than TV~125K for the bulk Verwey transition and explanation of observed $\sqrt{2}$x$\sqrt{2}$R45 reconstruction of (001) surface. The surface properties of Fe3O4 are especially important for catalytic applications and novel magnetic or spintronic high temperature devices.

As a result of the present project the team provides explanations of both open questions about (001) surface of Fe3O4. Using quantum mechanical approach of density functional theory the team has shown that the surface of magnetite is insulating, even at temperatures much higher than the metal – insulator transition of the bulk. These properties are attributed to particular charge ordering of electrons on d – orbital of the octahedrally coordinated iron. Such kind of calculations requires extreme accuracy of quantum mechanics and can be obtained only using computational facilities of the supercomputer like MareNostrum.

Calculations of the surface properties of strongly correlated compounds are not possible nowadays on workstations, or even

mid-size machines. Thanks to the supercomputing capabilities of MareNostrum, it was feasible to make extensive calculations of the electronic and structural properties of the surface of magnetite. This material intrigues mankind over the centuries and explanation of the insulating properties of its surface are of crucial importance for future spintronic devices.
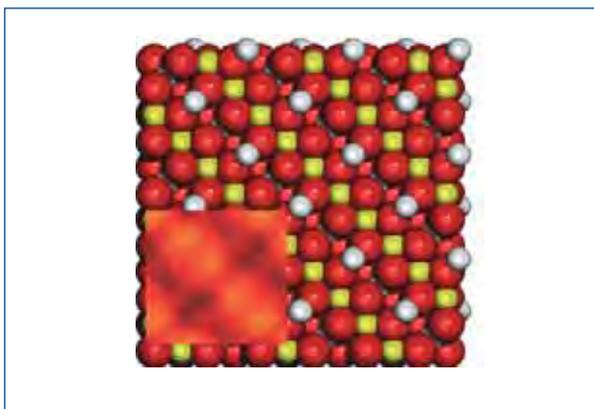
Model of (001) surface of Fe_3O_4, red spheres represent oxygen atoms, yellow spheres are for octahedral Fe and grey are for tetrahedral Fe. Inset in the left lower corner shows calculated STM image of the surface
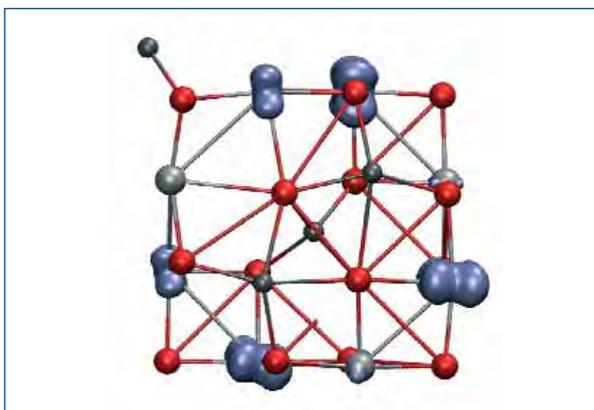
Top view of the octahedrally terminated surface of Fe_3O_4. Red spheres represent oxygen anions, light gray are for octahedrally coordinated Fe, and dark gray represent tetrahedrally coordinated iron. Blue shapes shows the charge ordering of the electronic states below the Fermi level that are responsible for insulating properties of the surface.

Nuria López
ICIQ

Hydrogen production from the degradation of ammonia on Ir nanostructured crystals: structure sensitivity

NH_3 can be stored in the form of a salt and could potentially behave as a energy vector since fulfills most of the DoE requirements for hydrogen content storage. In this case ammonia decomposition to provide the H_2 would be the challenge. Nanostructured Ir has been suggested for low-temperature H_2 extraction from ammonia. The team has employed Density Functional Theory simulations to understand the role of nanostructured surfaces for easy evolution of H_2 from the ammonia and to describe the better active site on several surfaces. On the other hand, selective hydrogenation of different molecules on gold catalysts are being investigated.

The team has obtained Density Functional Theory results relative to the production of hydrogen from ammonia on several Ir surfaces. This is the first time that such a complex study has been considered since the number of reaction steps considered for the dehydrogenation of ammonia plus the formation of nitrogen and hydrogen is large, and the number of surfaces considered is also large, leading to a total of at least 30 different reactions.

With these results it is possible to describe the best reaction site for each step and understand the reasons behind the structure sensitivity observed for the process. Since hydrogenation / dehydrogenation steps are common to a very large group of chemical reactions the team considers that the results are relevant not only for the reaction studied here but also to other that contain such steps.

Therefore the results on the system are scientifically sound in several aspects including their extension to other reaction systems and the new concepts about the best reaction site for such reactions. Moreover, hydrogen production from ammonia storage has been seen as a possible way to overcome the difficulties in hydrogen storage and therefore the search for active catalysts that can provide new ways for hydrogen stripping from ammonia is certainly of interest.

Finally, the study of structure selectivity (this is how much a reaction is dependent on the structure of the catalysts) has never been contrasted with so many different computational experiments for different surfaces as explored now for a complex reaction as ammonia decomposition. This can lead to a better description of the reaction ensemble and to new and more efficient ways of preparing catalysts.

The large number of steps considered in the reactivity described above would have been impossible to run in regular computational resources. Thanks to the increasing supercomputing capabilities such as MareNostrum, the team has been able to shorten the time expended in such simulations and they have provide a much deeper understanding on the trends that govern the structure selectivity of the system.
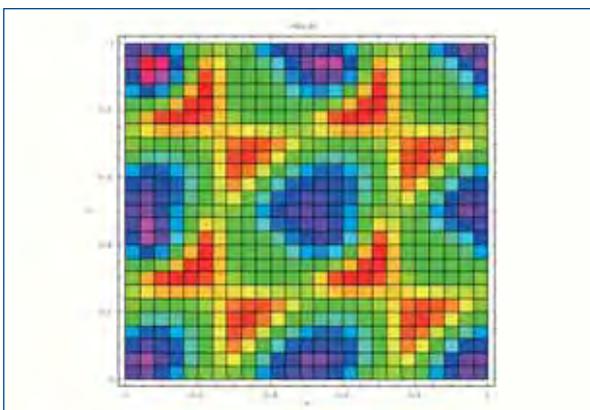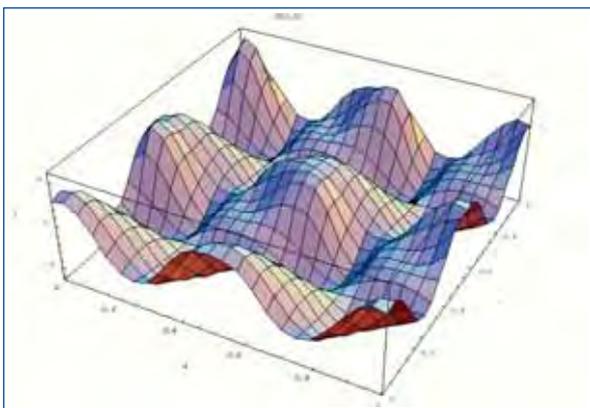
Publication:
N. Lopez, "Structure sensitivity and nanosized effects in the catalytic decomposition of ammonia on Ir", in prep.

Feliu Maseras Cuní
Institut Català d'Investigació Química

A DFT computational approach to supramolecular catalysis

The pumpkin-shaped macrocycle cucurbit[6]uril or CB6 ($C_{36}H_{36}N_{24}O_{12}$) was first synthesized in 1905 but its chemical nature and structure was not revealed until 1981. Although less famous than its structurally similar cousins cyclodextrin and calixarene, it has still found a wide range of applications, such as molecular recognition and catalysis. The synthesis of several new CBn homologues has sparked renewed interest and research into cucurbiturils since the beginning of the new millennium. The project has focused on the catalytic ability of CB6. It has been found to be able to catalyze a cycloaddition reaction, turning a slow and unselective reaction into one that is fast and regiospecific. The reason of interest in this reaction is twofold. The existence of experimental and, most importantly, kinetic data for the reaction is a way of evaluating the computational methods available. This will enable the team to learn what to use in future projects which might not include as much experimental detail, i.e. using computations as more of a predictive tool. Secondly, the computations will gain insight to the catalytic function of CB6 on a molecular level, to answer questions that are beyond today's experimental tools.

In choosing the project's computational method, the team found the use of DFT (density functional theory) to be the best choice for studying the cycloaddition reaction catalyzed by CB6. Cheaper methods, such as hybrid methods (e.g. ONIOM) were not accurate enough. The stationary points were fully optimized and further characterized by frequency calculations, a tremendous task, considering the size of the system (up to 1400 basis functions) that required nearly 60.000 cpu hours. The enzyme-like reaction mechanism of the cucurbituril was elucidated together with energy data for the actual cycloaddition step not accessible through experiments. The calculations showed that the rate-limiting step is the release of product from CB6 (in accordance with experiments), the cylcoaddition step was found to have an activation energy of 20 kcal/mol, a rate-enhancement of 5.5×10^4. Since the results obtained are agree with experiments, the team believes the project's computational approach to be good. In the future, it could therefore be used in a more predictive manner. For example in the rational design of new supramolecular catalysts.
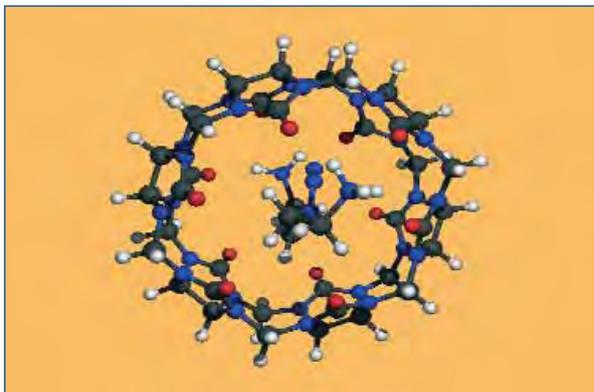
(Coming from page 139) DFT calculations allow the characterization of the process where two small organic molecules undergo a Diels-Alder reaction inside a cucurbituril supramolecule.

Sílvia Picozzi
CNR-INFM (Italy)

Ab-Initio Approach to Multiferroics Based on Perovskites: Exchange Mechanisms and Chemical Trends

In the last few years, there has been an extraordinary interest in materials known as "multiferroics", term coined to highlight their peculiarity in keeping a spontaneous electric and magnetic polarization, even without an applied electric and magnetic field, respectively. These materials might pave the way to unprecedented spintronic applications, exploiting the possibility of storing and manipulating information through their unique magnetic and electric behaviour. However, the comprehension of their peculiarities is still in their infancy. In this respect, first-principles simulations based on density functional theory represent one of the most valid tools in the investigation of novel multifunctional materials. The inherent complexity of these compounds, where a complex crystal structure joins the presence of several different magnetic atoms, constitutes grounds for magnetic frustrations that might be the cause for ferroelectricity. The aim of this research project is therefore to investigate by means of ab-initio methods rare-earth manganese oxides based on perovskites, as a function of the different rare-earth element and structural distortions, in order to deeply understand how magnetism is related to structure.

The team has suggested a new magnetically-driven mechanism for the arising of ferroelectricity in $HoMnO_3$. This is related to the peculiar magnetic ordering in Ho manganite (so-called AFM-E), which shows ferromagnetic zig-zag chains, each antiferromagnetically coupled with its neighbouring chains. In particular, through first-principles calculations based on density functional theory, the team has shown that the peculiar arrangement of the different Mn spins (and in particular the different Mn-O-Mn bond angle - and related Mn-O bond lengths - as a function of the parallel or antiparallel spins) results in a coherent displacement of the O center of mass with respect to the Mn center of mass, therefore giving rise to a permanent dipole i.e. ferroelectricity.

The high complexity of the unit cells in multiferroics renders the simulations quite time-consuming from the computational point of view. Therefore, these simulations become affordable only on supercomputers such as MareNostrum.
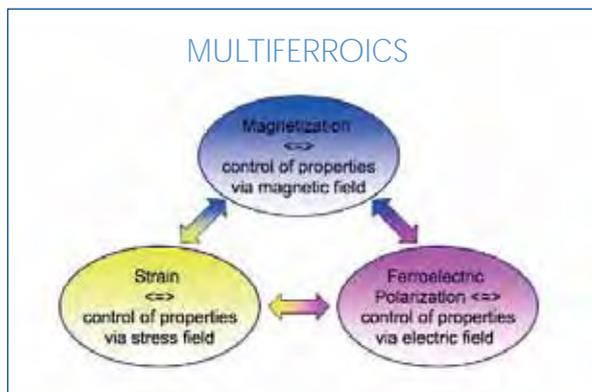
MULTIFERROICS

Interplay between structural, electronic and magnetic degree of freedom in multifunctional multiferroic materials
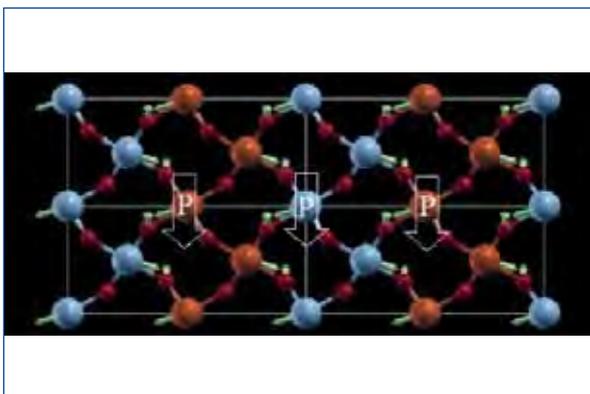

Top-view of the MnO2 plane in the HoMnO3 orthorhombic unit cell. Grey and orange spheres denote Mn with up and down spins, respectively, whereas red spheres denote O atoms. Green arrows denote the displacement of the atoms that lead to ferroelectricity, whereas the big white arrows denote the direction of the resulting ferroelectric polarization.

Publication:
S. Picozzi, K. Yamauchi, B. Sanyal, "Ferroelectricity driven by zig-zag magnetic chains: orthorhombic HoMnO3 as a multiferroic?" American Physical Society, APS March Meeting, March 5-9 (2007).

Josep María Poblet
Universitat Rovira i Virgili

Modelling of Polyoxometalates

The first polyoxometalate (POM), the phosphomolybdate $[PMo12O40]3-$, was reported by Berzelius in 1826. Since then countless structures have been synthesised and characterised. Today, POMs constitute an immense class of polynuclear metal-oxygen clusters usually formed by Mo, W or V and mixtures of these elements. The research activities of the Poblet's group in the field of polyoxometalates are mainly focused on redox properties of polyoxoanions, basicity of mixed species, magnetic and electronic properties of anions with paramagnetic ions, the modeling of large clusters, solvent effects, etc
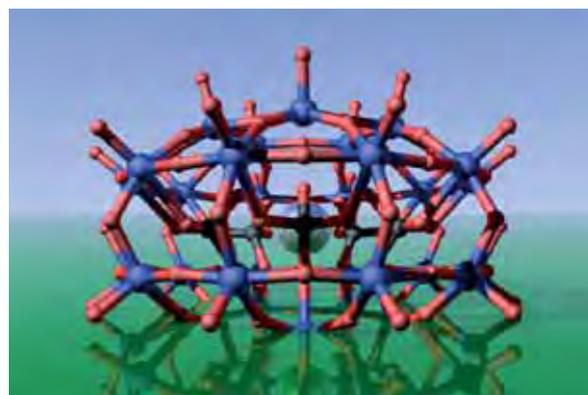

Publication:
Antonio Rodríguez-Fortea, Josep Maria Campanera, Claudia Cardona, Luis Echegoyen i Josep Maria Poblet, "Dancing with the right couple on a fullerene surface: Isomerization of N-ethyl [6,6]-pyrrolidino-Y3N@C80 to the [5,6] regioisomer", Angewandte Chemie, qualified VIP (Very Important Paper), 2006.

Ramón Reigada Sanz
Universitat de Barcelona

Molecular Dynamics study of the effect of Cholesterol in the Cell Membrane

Plasmatic cell membrane is no longer believed as an homogeneous and passive boundary surface. Instead, it is organized into raft domains that move within the lipid bilayer as platforms for the attachment of proteins, determining its functionality beyond a simple inert interface. The affinity of cholesterol for particular lipids of the membrane determines the generation and stability of raft domains. By using Molecular Dynamics, the team studies the interaction and molecular organization of cholesterol-phospholipid bilayer mixtures.
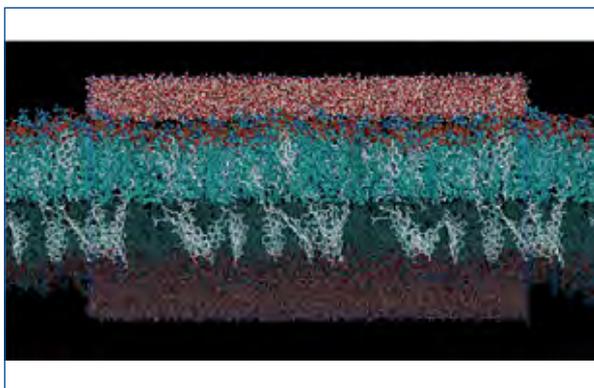
A bilayer composed by cholesterol (20%) and DOPC.

Antonio Rodríguez Fortea
Universitat Rovira i Virgili

First-principles molecular dynamics simulations of the formation mechanisms of polyoxometalates with low nuclearities

The objective of the project is to elucidate the mechanism of formation of the polyoxometalates (POMs) with low nuclearities from the tetrahedral $WO_3(OH)$- building blocks. The team employs first-principles molecular dynamics techniques (Car-Parrinello molecular dynamics, CPMD) including explicitly the solvent molecules (water in this case). Therefore the team is able to model processes involving formation and breaking of bonds (i.e. chemical reactions). In particular, using the metadynamics method, which allows for an acceleration of the dynamics thus making possible the observation of activated processes in reasonable simulation times.

In the simulations the team has observed that the dinuclear structures that show the lowest barriers of formation are not the ones postulated for the first step of the aggregation process. Dinuclear species are found in which the tungsten atoms are four- and five-coordinated, formed by one tetrahedron and one square pyramid that share a vertex (4-5 structure) and species in which both W atoms are five-coordinated, formed by two square pyramids that share an edge (5-5 structure). However, the team does not observe dinuclear species in which the W atoms are four- and six-coordinated, formed by a tetrahedron and an octahedron sharing an edge (4-6 structure) as postulated by Tytko and Glemser. Therefore is concluded that the 4-6 structure presents a higher barrier of formation and lower stability than other dimers. It's also observed that the anions (either the two monomers or the dimer) and the water molecules that solvate them are continuously interacting through hydrogen bonds confirming the requirement of considering the solvent explicitly.

As far as the formation of trinuclear species is concerned, the team has found from the simulations several intermediates that differ in the level of condensation, i.e. in the coordination of W atoms with respect to O atoms. The team observes that the formation of trimers with low coordination numbers (5-4-5, 4-5-5) show a low barrier, but also a low stability since the W-O bond formed breaks up easily. Other species with higher W-O coordination numbers (6-6-4, 6-5-5) are not so easily formed, but they show a higher stability than low coordination isomers. No elimination of water molecules has been observed up to this step of aggregation.

The stability of the monomers at different pH conditions are also analyzed. The team has found that the most-stable species at low-acidic conditions (pH ~ 5-6), the hydrogentungstate anion $(WO_3(OH)^-)$, retains the tetrahedral coordination, but at more

acidic conditions (lower pH) the protonated cation presents an octahedral coordination $[WO_2(OH)(OH_2)_3]^+$.

The team is investigating the first step of the formation of heteropolyoxoanions in order to unravel the role of the heteroatom unit XO_4 in the aggregation process: is it a template from the beginning or is it incorporated in further steps of the synthesis? The elucidation of that role is fundamental in the understanding of the formation process of the Keggin anion. Preliminary results show that the PO_4 unit is a template from the beginning. In conclusion, progress in the elucidation of the formation mechanism of POMs has been achieved, but some key points are still unsolved.

The MareNostrum machine at BSC is the only place in which such a computational load could be performed in a reasonable amount of real time. The experience using the CPMD program at MareNostrum in the preceding months confirms the excellent behaviour of this code when executed up to 256 processors.

Moreover, the requirement of CPMD of a Myrinet-type net for reasonable parallel applications is fulfilled and the difference between cpu time and elapsed time is minor even when employing 256 processors.
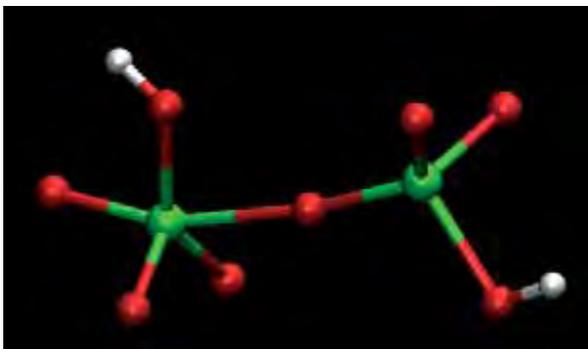
4-5 structure
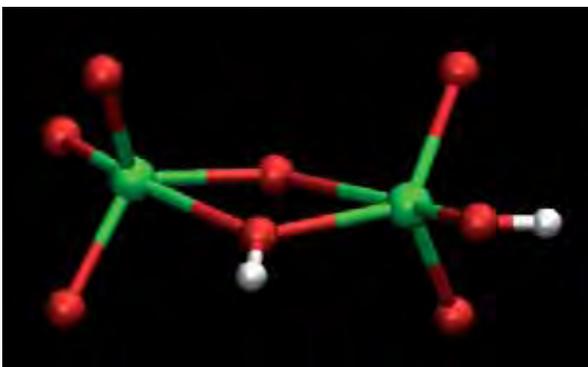
5-5 structure

Ángel Rubio
### Euskal Herriko Unibertsotatea

Spectroscopic properties of biomolecules, nanostructures and extended systems

Graphite possesses highly anisotropic crystal structure, with strong covalent bonding of atoms within and weak van der Waals bonding between the hexagonal symmetry grapheme sheets. Its layered lattice structure translates to a primarily 2D electronic structure, where small pockets of electrons and holes confer semi-metallic properties. The essence of this 2D character is expressed most vividly in the electronic structure of one to few monolayer thick graphene sheets. The recent discovery of massless relativistic behavior of quasiparticles at the Fermi level has aroused great interest in the nature of transport in graphene and other graphitic materials. On account of the nearly linear dispersion of electronic bands at the Fermi level, the quasiparticle mass associated with the charge carrier interaction with the periodic crystalline lattice nearly vanishes leading to extremely high electron mobilities and unusual half-integer quantum Hall effect. Since graphene is the fundamental building block of 0D fullerenes, 1D carbon nanotubes, and quasi-2D graphite, these electronic properties may be expressed in a variety of graphitic materials. The strong electron-phonon (e-p) interaction in graphite is a distinctive characteristic of ineffective screening of the Coulomb interaction in semimetals. The e - p interaction limits the carrier dephasing and high-field transport in solids. In graphite, it is also manifested in the phonon frequency shift upon doping and exceedingly fast electron thermalization.

The team has explored the influence of the nonthermal electron-hole plasma and large amplitude coherent phonon excitation on the femtosecond and picosecond dynamics of the in-plane E2g2 coherent phonon of graphite. The time-dependent frequency of the fast E2g2 coherent phonon is a sensitive probe of the time evolution of the electronic and lattice occupation distributions following the optical excitation. The unusual electronic stiffening of the E2g2 mode, the extremely fast electron-phonon thermalization, and the anharmonic coupling of the E2g2 and E2g1 modes can be attributed to the quasi-2D electronic structure and Kohn anomaly in graphite. The results offer a new paradigm of electron-phonon coupling, where anisotropically excited non-thermal electrons impart exceptional properties to the lattice. Similar interactions are likely to govern the electron-phonon coupling in other graphitic materials, such as carbon nanotubes and graphene, that are of topical interest for high-performance, nanometer scale carbon-based electronic devices.

The group has also started to address the description of electron correlation in the intensity versu voltage characteristic of molecular device constrictions. TWe show that GW describes the Kondo effect and the zero-temperature transport of the Anderson model reasonably well. Combining the GW scheme with density functional theory and a Wannier function basis set,

the team illustrates the impact of correlations by computing the conductance of a hydrogen molecule between Pt chains. Extenisons to more complex situations are under study.
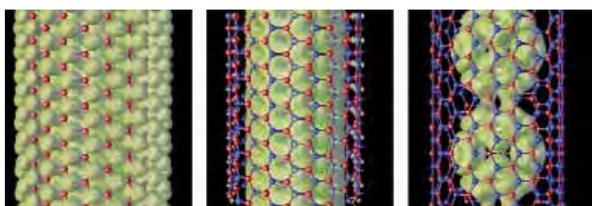
Constant current mode STM image for an applied bias of +/- 1.5eV of a filled Boron-nitride nanotube: either with a C(5,5) tube or with a fullerene inside ("peapod" BN(10,10)@C60). Red, blue and green balls denote B, N and C atoms, respectively. It is seen that for negative bias (mapping of occupied states) the STM images show clear protrusions at the N sites with no clear signature of the filling material. Instead for positive bias we see states within the BN-band gap, those are specific signature of the filling of the tube (with either a carbon nanotube (a) or a fullerenere (b)).

Publications:

M. Dubois, C. Delerue and A. Rubio "Adsorption and electronic excitation of biphenyl on Si(100): A theoretical STM analysis", Physical Review B 75, 041302 (R) (2007)

A. Castro, M.A.L. Marques, H. Appel, M. Oliveira, C.A. Rozzi, X. Andrade, F. Lorenzen, E.K.U Gross and A. Rubio. "Octopus: a tool for the application of time-dependent density functional theory", Physica Statu Solidi (b) 243, 2465 -2488 (2006)

M. Grüning, A. Marini and A. Rubio. "Effect of spatial nonlocality on the density functional band gap ", Physical Review B (Rapid Communications) 74, 161103 -4 (2006)

K.Ishioka, M. Hase, M. Kitajima, L. Wirtz, A. Rubio and H. Petek. "Femtosecond Coherent Quasiparticle correlations of a Noneqyuilibrium plasma in graphite". Physical Review Letters. (2006)

M. Grüning, A. Marini and A. Rubio "Density functionals from many-body perturbation theory: the bandgap for semiconductors and insulators" Journal of Chemical Physics (2006)

D. Varsano, R. Di Felice, M.A.L. Marques and A. Rubio "A TDDFT study of the excited states of DNA bases and their assemblies" Journal of Physical Chemistry B (2006)

L. Wirtz, A. Marini, and A. Rubio "Excitons in boron nitride nanotubes: dimensionality effects" Physical Review Letters {\bf 96} 126104-1,4 (2006)

Eliseo Ruiz
Universitat de Barcelona

Magnetic Properties of large Single Molecule Magnets: Mn$_{25}$ and Mn$_{32}$

The project studies the magnetic properties of different large polynuclear complexes containing manganese atoms. These systems are particularly interesting because some of them are the molecules showing the highest spin known and others exhibit slow relaxation of the magnetization, which means that just a molecule behaves like a nanomagnet. In order to improve the magnetic properties of such molecules is important to increase the total spin of the molecule and their magnetic anisotropy. For such systems, it is impossible to obtain the values and signs of the exchange coupling constants between the paramagnetic centers from the experimental measurements of the magnetic susceptibility due to the large number of states. For instance, for one of the studied compounds, the Mn$_{25}$ complex, the project would need 2.3 10^{13} GB of memory to store the Hamitonian matrix that should be diagonalized several times to perform the fit of the experimental susceptibility. Despite that this huge amount of memory was available, there are several sets of values that will fit perfectly the experimental curve being impossible to extract a proper set of exchange coupling constants. Hence, the use of theoretical methods based on density functional theory is the unique approach to know the magnetic properties at microscopic level. The team can know if the coupling between two paramagnetic centers will be ferromagnetic (positive exchange coupling constant and the same spin in both centers) or antiferromagnetic (negative constant and consequently, opposite spins).

The first studied system was the Mn25 complex, the team has employed the NWChem and Siesta codes to obtain the fourteen exchange coupling constants and with such values, the team has explored between all the possible spin distributions to check which is the ground state. The group obtained a ground state with an S value of 69/2 and 49/2 respectively with NWChem and Siesta, relatively close to the 51/2±1 proposed experimentally. The team found that the inclusion of the only the interactions between first neighbors using the Heisenberg Hamiltonian is not enough to describe properly the magnetic properties. Thus, the possible role of the next-near neighbor interactions is actually being investigated. Likewise, the team has started the study of a similar Mn25 complex synthesized by Prof.
George Christou at the University of Florida where some ligands has been replaced reaching a higher experimental S value of 61/2±1.

Finally, the team studied the magnetic properties of the molecule with highest spin known a Mn19 complex with S

=83/2 and a similar Mn10 complex with S = 22. In the strive for preparing polynuclear complexes with high spin, a common shortcoming is that the local spins of the metal atoms are antiferromagnetically coupled and the total spin in the ground state is not as large as it could be if the coupling was ferromagnetic. However, in these two complexes all the interactions are ferromagnetic and consequently, the total spin is the sum of the local spins of the paramagnetic centers. The team has calculated the eight exchange coupling constants present in the Mn19 complex confirming the ferromagnetic nature of the interaction and also explored the influence of the change of the ligands to improve such magnetic properties especially the increasing of the magnetic anisotropy.

The studies of the Mn19 and Mn10 have been finished. The team is working in the final steps to complete the work devoted to the two Mn25 complexes and has started the study of a Mn32 complex and two complexes Mn12 and Mn22 with wheel shape. Thanks to increasing supercomputing capabilities such as MareNostrum, it is feasible to make an exhaustive study of the magnetic properties of very large polynuclear complexes that requires the calculation of the energy for several spin distributions. Otherwise, the study of such system would be impossible and it would not reach an accurate knowledge of the microscopic exchange interactions that is fundamental to improve the magnetic properties of this kind of molecules.
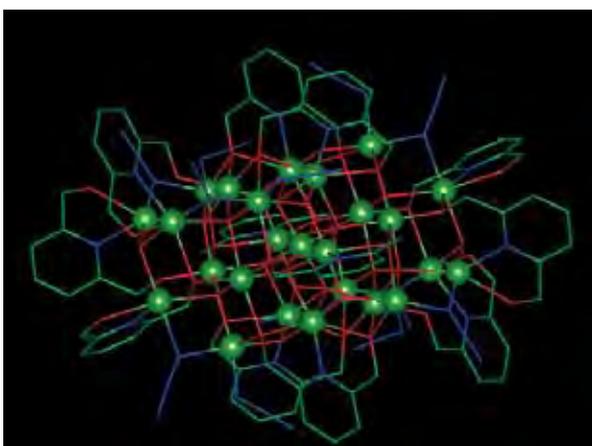
One molecule of the Mn_{25} complex behaves like a nanomagnet.

Enrique Sánchez Marcos
Universidad de Sevilla

Hydration of pharmacologically interesting $Pd2^+$ and $Pt2^+$ complexes

Transition metal ions play a significant role in many biochemical processes, where the usual environments are liquid solution and interfacial regions. The ubiquitous presence of solvent molecules forces the metal-ion-containing molecule to find a delicate equilibrium on its interactions with the other reactants and the solvent molecules. Thus a tipycal reaction process involves a preliminary step where a part of the solvent-molecule shell must be released to allow for the direct interaction with the reactant partner. The knowledge of the solvation structure and its integrity is of paramount importance to get insight into the potential activity of the metal ion compound.

When one is interested in representing a macroscopic system at the molecular level, two key-aspects need to be taken into account: the size of the system and its evolution with time. Regarding system size, the scale of the systems composing our World is in the order of 1024 molecules (that is, one million of milliard of milliards). Dealing with this huge amount of molecules is the price that has to be paid for representing Nature in terms of the smallest chemical entities. Sophisticated (statistical) techniques allow the team to reasonably represent a macroscopic system by less than a hundred molecules. If a deeper insight is desired, quantum mechanical theory must be applied to the submolecular level, describing electronic and nuclear dynamics. Thus a first principles representation of the world can be achieved.

When performing Molecular Dynamics simulations of a system, the computational cost (in terms of real time) for following the system's evolution with time can make the determination of dynamical properties unfeasible. Simulating at a first principles level a picosecond evolution for a system composed of a hundred of atoms using a standard computer can take weeks.

During the activity period at MareNostrum Molecular Dynamics simulations have been performed for two metal cations ($Pt2^+$ and $Pd2^+$) in water solution. To this aim the team has set a system composed of the metal cation plus 70 water molecules inside a cubic box which was repeated along the three directions in space to simulate the macroscopic solution. For each of these two systems the team has run two different statistical trajectories, from whose analysis the team will obtain information on properties as the structure of the solution, the behaviour of water molecules around the cation or the cation mobility.

A series of simulations for the Pt2+ solution were also performed in which the number of water molecules around the metal cation is constrained to different values. This will allow the team to study the change in relative energy with changing structure of the solution.

This work is not only worth for basic science itself, but it is also a mandatory preliminar step in order to study Pt2+ and Pd2+ -containing complexes with biological activity.

The total simulated time performed during this activity period sums up to ca. 150ps. At MareNostrum and using 90 CPUs-parallel computing, simulating 1ps takes ca. 28 hours of real time.

Taking into account that in order to compute dynamical properties such as the self-diffusion coefficient for a species in solution tens or hundreds of nanoseconds must be simulated, having access to supercomputing facilities such as MareNostrum is mandatory to be able to obtain statistical and dynamical information of the system.
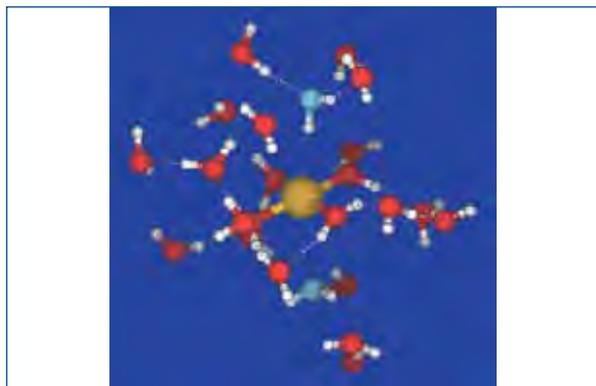

Image of water molecules surrounding a Pt2+ cation, showing strongly bounded water molecules in the plane and cooperatively bonded water molecules in the axial region (meso-shell, in blue).

Miquel Solà Puig
Universitat de Girona

Theoretical Study of the reaction mechanism of human superoxide dismutase

The superoxide dismutase enzyme (SOD) catalyzes the transformation of the radical superoxide, O2- , naturally generated in the living organisms, to molecular oxygen and hydrogen peroxide. The existence of this enzyme is fundamental in the living species to avoid the damage produced by this radical in the organism. However, it is well-known that there is a close relationship between the familial amyotrophic lateral sclerosis (FALS) disease and the CuZnSOD. The project aims to perform a complete study of all reaction paths of the CuZnSOD with the complete enzyme using a Car-Parrinello approach.

The team believes that the knowledge of the SOD reaction mechanism is fundamental to understand the origin of the familial amyotrophic lateral sclerosis disease and, therefore, the team considers that these calculations will shed important light on the advancement of science in this particular issue. The project aims to analyze the reaction mechanism of the CuZnSOD including all atoms of the enzyme (more than 2000) in a QM/MM approach. In this sense, it will be possible to study the effect of each residue close to the reaction center with the final objective of discussing the relation between the FALS and the CuZnSOD.

Thanks to the hours assigned to the project in the MareNostrum, the team had been able to perform the following calculations:

1. Tests in model systems of the active centre to be able to determine the parameters needed to do CPMD QM/MM calculations of the whole enzyme.

2. Tests in the number of atoms included in the QM region.

3. Setting up of the CPMD QM/MM calculations.

4. QM and QM/MM simulations of the first step of the reaction.

5. QM simulations of the second step of the reaction.

6. Study of the role of the solvent in the reaction mechanism.

The group believes that the knowledge of the SOD reaction mechanism is fundamental to understand the origin of the *familial amyotrophic lateral sclerosis* disease and, therefore, the team considers that these calculations will shed important light on the advancement of science in this particular issue.

The complexity and the cost of the ab initio dynamics that are being carried out in the study make this project only possible in a supercomputer like MareNostrum. The program that is being used to perform these molecular dynamics simulation is the CPMD program which is a parallelized plane wave/pseudopotential implementation of Density Functional Theory, particularly designed for ab initio molecular dynamics.
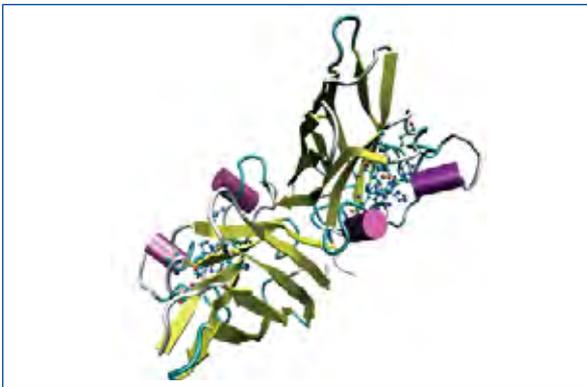
Human superoxide dismutase structructure

César Tablero
Universidad Politécnica de Madrid

Study of the nonradiative recombination

One mechanism of nonradiative recombination is related to a sufficiently large lattice vibration. In a complete nonradiative cycle the levels move across the entire energy gap. First-principles calculations allow the team to explicitly calculate the dependence of the bands on the positions of the atoms without making any additional approximations. This has been the main objective of the work carried out in MareNostrum. Nevertheless, initially it has been necessary to choose and to characterize a material with a deep level that gives place to an intermediate band when increasing the impurity concentration.

Materials with an isolated partially filled intermediate band are of great interest as new materials for high-efficiency solar cells, as up- and down-converters and as mid-infrared lasers with low nonradiative decay because of the properties of this band. Within the photovoltaic field, these types of materials could revolutionize modern solar cells because solar cells based on these materials can reach efficiencies higher than those of single gap solar cells (63.17% versus 40.7%). As well as there is a great interest in and discussion with the effect of intermediate bands versus levels in nonradiative recombination. Levels situated within the semiconductor band gaps promote nonradiative recombination. On the other hand it has been suggested that an increase in the density of impurities will suppress nonradiative recombination.

Initially, in this project, it has been necessary to choose and to characterize a material with a deep level that gives place to an intermediate band when increasing the impurity concentration.
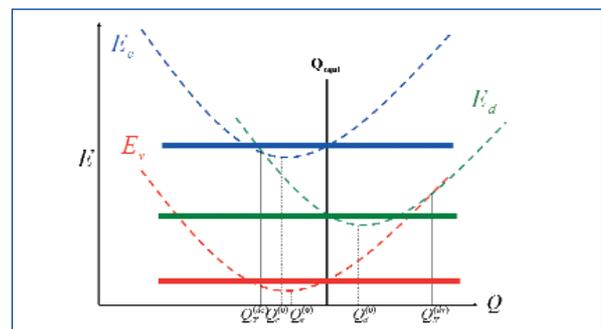
Configuration coordinate diagram for non-radiative recombination

6.5 Success Stories

Astonomy Space and Earth Sciences

"Thanks to this prediction model, in June 2006 the Atmospheric Research of National Technical Aerospacial Institute (INTA) and the Atmospheric Observatory in Izaña of INM took the principal phase of the campaign called "Trompeta" (Tropical Monitoring Phase in the Atmosphere)".

José María Baldasano, *Barcelona Supercomputing Center*

Biomedicine and Health Sciences

"Thanks to increasing supercomputing capabilities such as MareNostrum, it is now feasible to apply maximum-likelihood based classification approaches to the tremendous problem of classification of structurally heterogeneous cryo-electron microscopy data sets. These datasets may contain up to billions of data points, and the proposed approach may take up to several months of CPU to separate a single dataset. Therefore, these developments would not be feasible without the availability of supercomputers such as MareNostrum".

José María Carazo, *CNB-CSIC*

"Thanks to increasing supercomputing capabilities such as MareNostrum, it is feasible to make an extensive use of molecular dynamics simulations to derive properties of protein structures that are not accessible through the analysis of a single structure as contained in the PDB".

Modesto Orozco, *Barcelona Supercomputing Center*

"We require MareNostrum to run the computationally intensive algorithm TBLASTX to find regions of homology between two full genomes (i.e. Mouse vs. Human). Running TBLASTX to find potentially conserved regions between different species is an essential component of the homology-based gene prediction tool SGP which has been developed by our group. Generally, TBLASTX comparison among species with large genomes, such as Human and Mouse, would require 7-10 days on a 20-25 processor grid. However, MareNostrum´s processing speed and especially parallelization potential allows the same work to be performed in approximately 12 hours (on 256 CPUs and excluding queuing time). This is important because it allows us to execute TBLASTX to find homologies among many different genomes at the same time, in a reasonable time frame, which would be impossible to achieve locally or using a smaller computer grid. This, in turn, allows us to develop and improve SGP for large-genome species to a level which would be unattainable by using a smaller grid".

Roderic Guigó, *Centre de Regulació Genòmica*

Physics and Engineering

"Given the extreme difficulty of performing full field, real-time diagnostics on massive dynamic fracture and fragmentation experiments, computer simulations emerge as the only viable approach to these phenomena. The huge scale disparity, from the dimensions of the system to the characteristic length-scales of the fracture processes, together with the long simulation times required, inevitably leads to supercomputing platforms such as MareNostrum. We are currently resolving four orders of magnitude in the scale disparity, which is essential if meaningful conclusions are to be extracted".

Irene Arias, *Universitat Politècnica de Catalunya (UPC)*

"Our Molecular Dynamics Simulations allow us to obtain a detailed description of atomistic mechanisms in the interaction between cations and membranes which cannot be discriminated with current experimental techniques. Also, it is clear that these large scale computer simulations developed in our research project are only possible in top Supercomputing facilities. Our experience in MareNostrum shows that a typical equilibration run for a realistic phospholipid layer (previous to a production run) requires 128 processors during 1-2 days. A typical production run for this system (1-2 nanoseconds in physical time) require 2-4 days with 256-128 processors. Our experience also shows that currently MareNostrum is the only Supercomputing facility in Europe able to perform this kind of simulations. Other Supercomputing facilities, employed in our previous research projects showed a significantly smaller performance in our calculations, and we were restricted to simulations of smaller systems (128 amphiphillic molecules, their counterions and 103 water molecules in 1 nsec production run as compared with 512 amphiphillic molecules, counterions, added salt and 10 4 water molecules in 2-4 nsec production runs). Taking into account the sluggish dynamics of these systems, the gain obtained in MareNostrum significantly improves our results and allows us to study realistic systems inaccessible in other Supercomputing facilities".

Jordi Faraudo, Universitat Autónoma de Barcelona

"MareNostrum has allowed us to scale up the problem to previously unattainable complexity (i.e. number of particles and size of the computational domain). Therefore we are now capable of investigating the formation of very large scales in the two-phase system".

Markus Uhlmann, Universitat Politècnica de Catalunya (UPC)

Chemistry and Science and Technology of Materials
"The MareNostrum machine at BSC is the only place in which such a computational load could be performed in a reasonable amount of real time. Our experience using the CPMD program at MareNostrum in the preceding months confirms the excellent behaviour of this code when executed up to 256 processors. Moreover, the requirement of CPMD of a Myrinet-type net for reasonable parallel applications is fulfilled and the difference between cpu time and elapsed time is minor even when employing 256 processors".

Antonio Rodriguez Fortea, Universitat Rovira i Virgili

"Massively parallel machines such as MareNostrum are essential to advance our understanding of artificial muscles able to mimic the performance of natural muscles via large-scale simulations. The research program of the BSC has allowed explore different microscopic aspects related with the mechanism and efficiency of this engineered devices enabling molecular simulations that are orders of magnitude larger than common parallel computers permit".

Carlos Alemán, Universitat Politècnica de Catalunya (UPC)

"The total simulated time performed during this activity period sums up to ca. 150ps. At MareNostrum and using 90 CPUs-parallel computing, simulating 1ps takes ca. 28 hours of real time. Taking into account that in order to compute dynamical properties such as the self-diffusion coefficient for a species in solution tens or hundreds of nanoseconds must be simulated, having access to supercomputing facilities such as MareNostrum is mandatory to be able to obtain statistical and dynamical information of the system".

Enrique Sanchez Marcos, Universidad de Sevilla