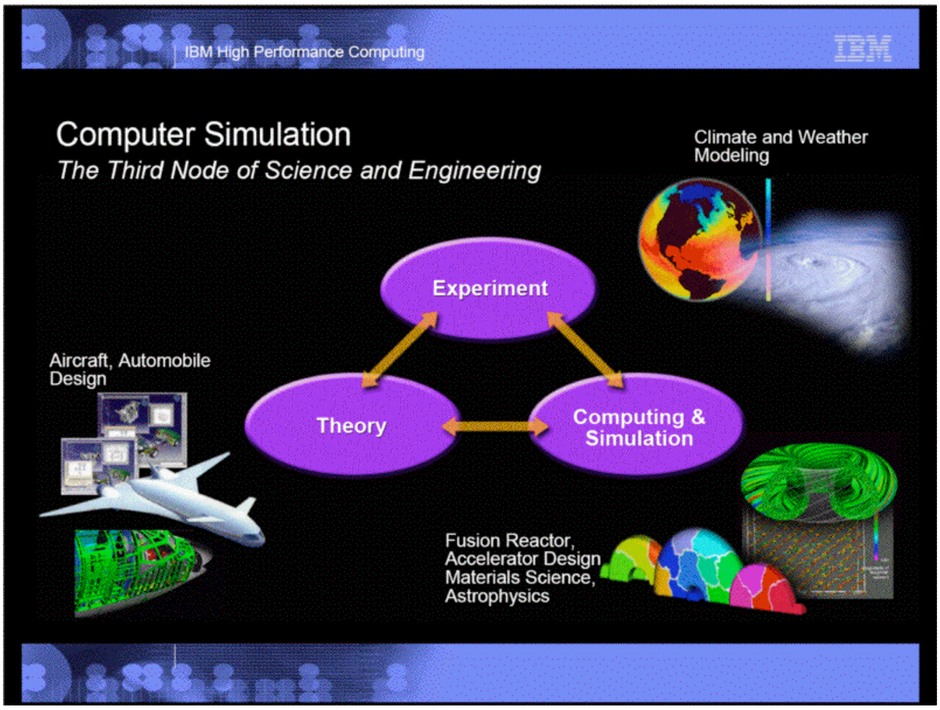
Barcelona
Supercomputer
Center
National Supercomputer Facility

# Centro Nacional de Supercomputación

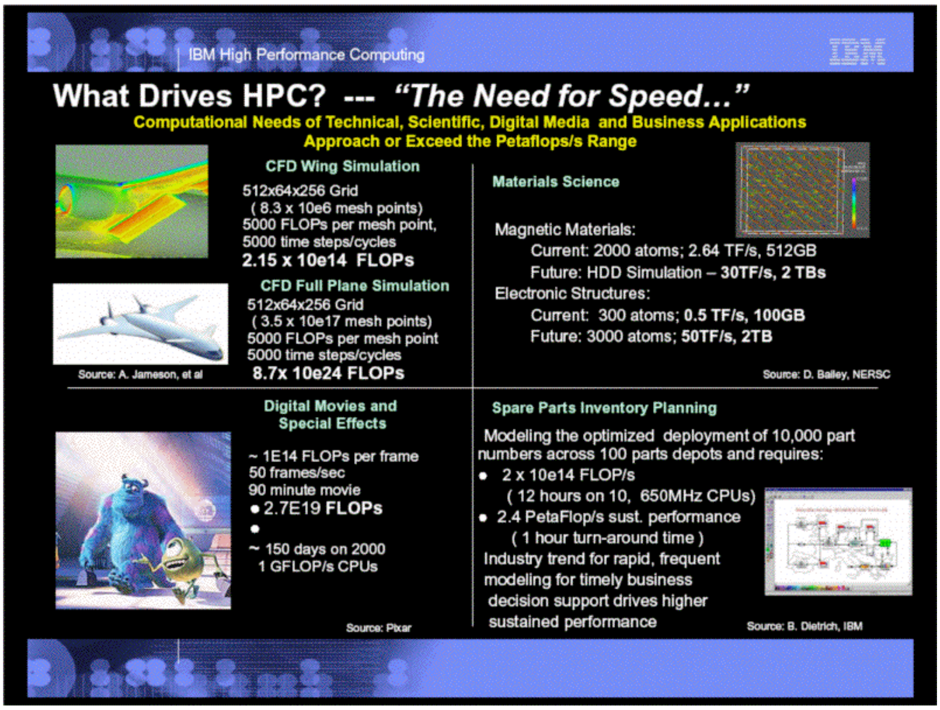
Prof. Mateo Valero
BSC Director







Barcelona Supercomputer Center
National Supercomputing Facility

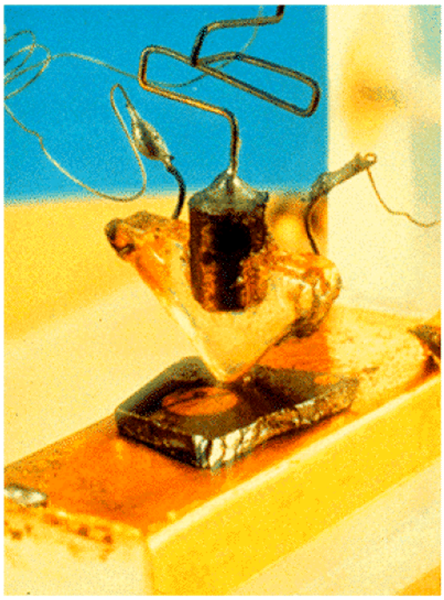




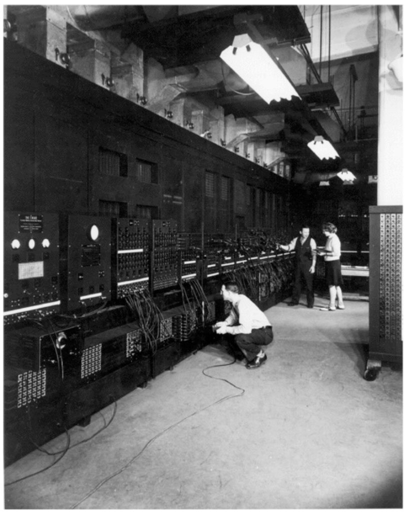

Barcelona Supercomputer Center
National Supercomputing Facility

# Technological Achievements

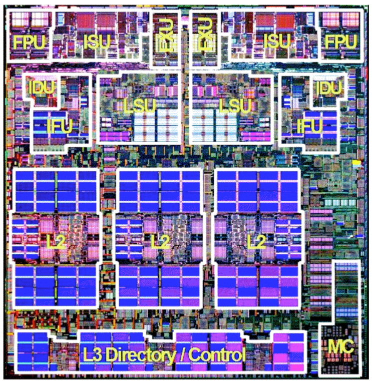
- Transistor (Bell Labs, 1947)
    - DEC PDP-1 (1957)
    - IBM 7090 (1960)
- Integrated circuit (1958)
    - Kilburn, TI/Fairchild
    - IBM System 360 (1965)
    - DEC PDP-8 (1965)
- Microprocessors (1971)
    - Intel 4004



# In 50++ years ...



Eniac, 1946 ...
1900 vacuum tubes




Power5, 2004
First dual core SMT processor
276M transistores
389 $mm^2$
1,8 GHz
1,9 MB L2 cache



# Technology Directions: SIA Roadmap

| Year | 1999 | 2002 | 2005 | 2008 | 2011 | 2014 |
|---|---|---|---|---|---|---|
| Feature size (nm) | 180 | 130 | 100 | 70 | 50 | 35 |
| Logic trans/$cm^2$ | 6.2M | 18M | 39M | 84M | 180M | 390M |
| Cost/trans (mc) | 1.735 | .580 | .255 | .110 | .049 | .022 |
| #pads/chip | 1867 | 2553 | 3492 | 4776 | 6532 | 8935 |
| Clock (MHz) | 1250 | 2100 | 3500 | 6000 | 10000 | 16900 |
| Chip size ($mm^2$) | 340 | 430 | 530 | 620 | 750 | 900 |
| Wiring levels | 6-7 | 7 | 7-8 | 8-9 | 9 | 10 |
| Power supply (V) | 1.8 | 1.5 | 1.2 | 0.9 | 0.6 | 0.5 |
| High-perf pow (W) | 90 | 130 | 160 | 170 | 175 | 183 |
| Battery pow (W) | 1.4 | 2 | 2.4 | 2.8 | 3.2 | 3.7 |

Barcelona Supercomputer Center
National Supercomputing Facility

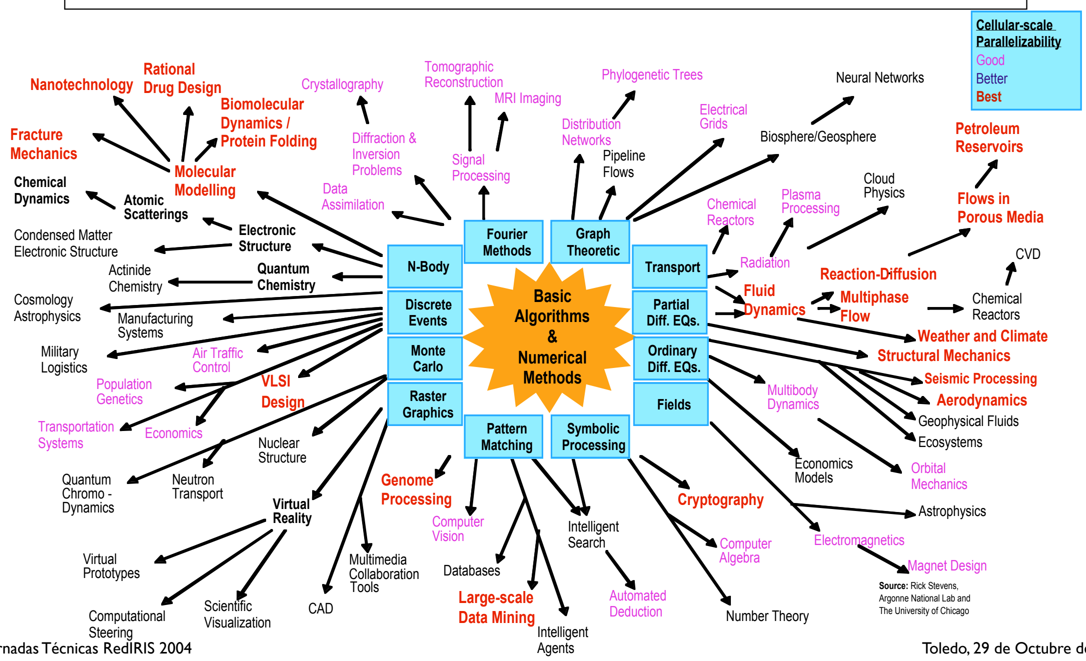
# Applications for Supercomputers

- Aircraft/car simulations
- Life Science (Proteins, Human Organs,…)
- Atmosphere
- Stars
- Nanomaterials
- Drugs
- Regional/Global  Climate/Weather/Pollution
- High Energy Physics
- Combustion
- Image Processing



Barcelona Supercomputer Center
National Supercomputing Facility

# Suitable applications for massively parallel systems



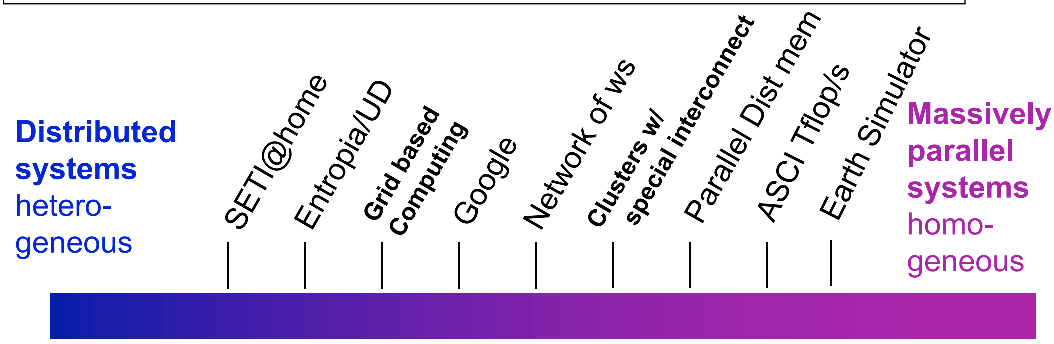
# Throughput vs. Parallel programming

- Throughput
  - Multiple, unrelated, instruction streams (programs) that execute concurrently on multiple processors
  - Multiprogramming n tasks on p processors: each task receives p/n processors
- Parallel Programming
  - Multiple related, interacting instructions (single program) that execute concurrently to increase the speed of a single program
  - 1 task on p processors, each processor receives 1/p of the task: reduce response time





# Distributed and Parallel Systems




- Gather (unused) resources
- Steal cycles
- System SW manages resources
- System SW adds value
- 10% - 20% overhead is OK
- Resources drive applications
- Time to completion is not critical
- Time-shared
- SETI@home
  - ~ 400,000 machines
  - Averaging 40 Tflop/s

- Bounded set of resources
- Apps grow to consume all cycles
- Application manages resources
- System SW gets in the way
- 5% overhead is maximum
- Apps drive purchase of equipment
- Real-time constraints
- Space-shared
- Earth Simulator
  - 5000 processors
  - Averaging 35 Tflop/s

# Google

- Gloogle query attributes
  - 150M queries/day (2000/second)
  - 3B document in the index
- Data centers
  - 15,000 Linux systems in 6 data centers
    - 15 TFlop/s and 1000TB total capability
    - 40-80 1U/2U servers/cabinet
    - 100 MB Ethernet switches with Gigabit Ethernet uplink
  - growth from 4,000 systems (June 2000)
    - 18M queries then
- Performance and operation
  - simple reissue of failed commands to new servers
  - no performance debugging
    - problems are not reproducible

Source: Monika Henzinger, Google





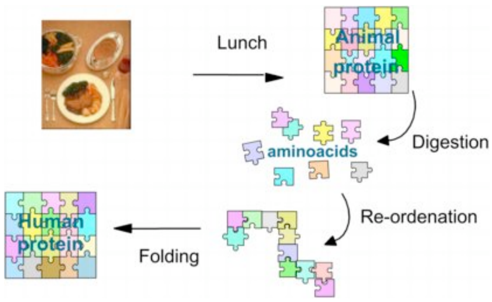
# Life Science future projects

- In silico drug screening
- Clustering of Expressed Sequence Tags
- In silico Structural Genomics
- First ab-initio Calculations of the Electronic Structure of the Smallest Living Organism
- Mixed QM/MM Simulations of the Five Most Important Enzymatic Reactions
- Virtual Cell Project
- Digital human and virtual chirurgy
- Ab-initio Food Science
- Protein Folding Dynamics

Barcelona Supercomputer Center
National Supercomputing Facility

# Better understanding of biological process

- The structure and dynamics of protein folding determine their biological properties
- Some illness related to misfoldings are
  - Alzheimer's Disease
  - Mad Cow Disease
  - Cystic Fibrosis
- Supercomputing will enable the in-silico simulation of protein folding, contributing to novel drug development techniques and new medical treatments





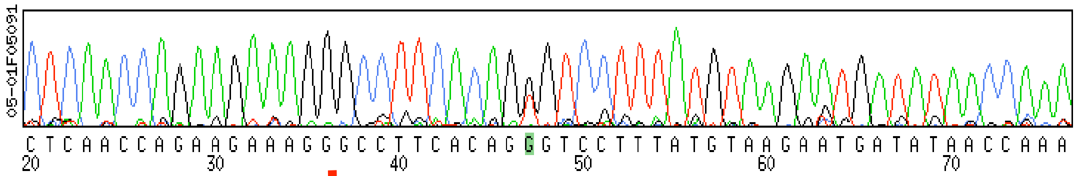

Barcelona Supercomputer Center
National Supercomputing Facility

# Protein Folding Calculation

| Description | Count | Comment |
|---|---|---|
| Atoms | ≈ 32000 | 260 amino acid protein + water |
| Force evaluation / time step | $10^9$ | Pairwise atom-atom interaction |
| FLOPs / force evaluation | 150 | Typical molecular dynamics |
| FLOPSs / time step | $1.5 \times 10^{11}$ | |
| Each time step | $\approx 10^{-15}$ s | 0.5 - 5 femto second |
| Total simulation time single step | $10^{-3}$ s | Protein folds in ≈1 mili second |
| Total time steps | $2 \times 10^{11}$ | |
| FLOPs / simulation | $3 \times 10^{22}$ | Total FLOPs to fold a protein |
| Execution time | $3 \times 10^{7}$ s | 1 year |
| Required FLOPS | $\approx 10^{-15}$ | 1 Petaflop |

# Personalized medicine requires massive supercomputing resources

- Each individual has a different genetic sequence. Its understanding would help to identify specific risks
- Sequencing the DNA of the human population implies the analysis of 1.8 million TB/s (300 TB/person * 6 billion person)
- The comparison of the genetic profiles of an individual against a central database would allow the development of life-long rational healthcare with food and drugs adequate to the individual genetic characteristics

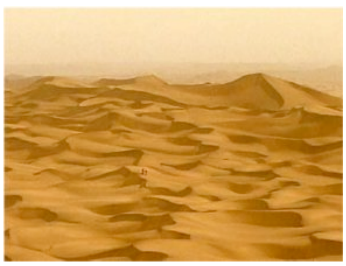







# Protecting and improving our environment

- Supercomputing facilitates accelerated and significant progress in the earth and material sciences
  - Biogeochemical cycles
  - Biodiversity and ecosystems in equilibrium
  - Climatic variations
  - Hydrological predictions
  - Contagious diseases and environmental impact
  - Rational use of resources
  - Reinventing materials and their applications
- Supercomputing will enable to simulate the environmental and social impact of agricultural, industrial and urbanistic policies, as well as the impact of external factors (desertification, catastrophic events) in real time



# CEPBA: European Center for Parallelism of Barcelona

UPC

Dept. → DAC

CER → CEPBA

DAC
- HPC experience

CEPBA
- October 1991
- R&D on parallelism
- Training
- Technology transfer
- European Context

RME, LSI, FEN, FA
- Computing needs

Sponsored by:





# CEPBA: joint ventures

Centre de Computació i Comunicacions de Catalunya
- October 1995
- Coordination between CESCA CEPBA

CEPBA IBM Research Institute
- Starting activities November 2000
- Research Cooperation in Deep Computing
- Support research on other fields of Science

Barcelona Supercomputer Center
National Supercomputing Facility

# CEPBA: Service

| Manuf. | Model | Processor | Memory Disc | Peak Performance | Service |
|---|---|---|---|---|---|
| IBM | SP3+p360 | 128 Power3 + 36 Power4 | 64+18 GB 1,8 TB | 336 Gflop/s | 10/01 12/02 |
| Compaq | Alpha Server GS-160 | 16 Alpha 21264 | 8 GB 108 GB | 23,3 Gflop/s | 9/00 |
| Parsytec | CCi-8D | 16 Pentium II | 1 GB 30 GB | 3,2 Gflop/s | 4/98 |
| Silicon Graphics | Origin 2000 | 64 MIPS R10000 | 8 GB 360 GB | 25 Gflop/s | 1/97 |
| Digital | Alpha Server 8400 | 12 Alpha 21164 | 2 GB 32 GB | 10,5 Gflop/s | 12/96 |
| Silicon Graphics | Power Challenge | 12+12 R8000/R10000 | 4 GB 50 GB | | 7/96 - 7/97 |
| Thinking Machine | CM-2 | 2048 1 bit | 256 MB | 640 Mflops | 4/92 - 2/98 |
| Convex | C3480 | 8 | 1 GB 16 GB | 0,4 Gflop/s | 10/91 - 1/98 |
| Parsys | SN 1000 | 32 T800 | 128 MB 2 GB | 64 Mflops | 5/90 - 5/95 |



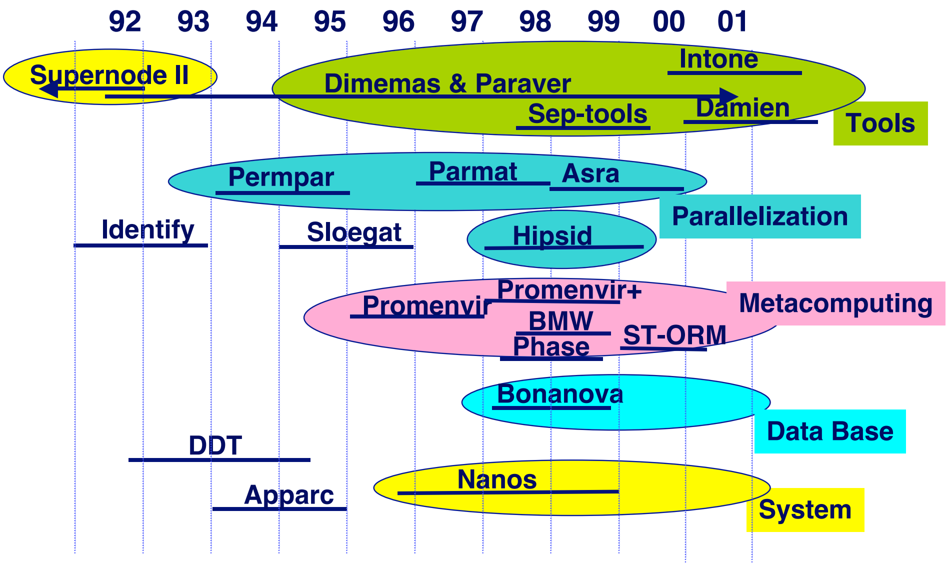
Barcelona Supercomputer Center
National Supercomputing Facility

# CEPBA: European Mobility Programs

- Joint CEPBA - CESCA projects
- Stays and access to resources

| Project | Period | Funding | Visitors |
|---|---|---|---|
| HCM | 1993-1997 | 950 | 112 |
| PECO | 1995-1996 | 160 | 17 |
| TMR | 1996-2000 | 935 | 133 |
| IHP | 2000-2003 | 700 | 142 |
| Total | | 2745 | 404 |

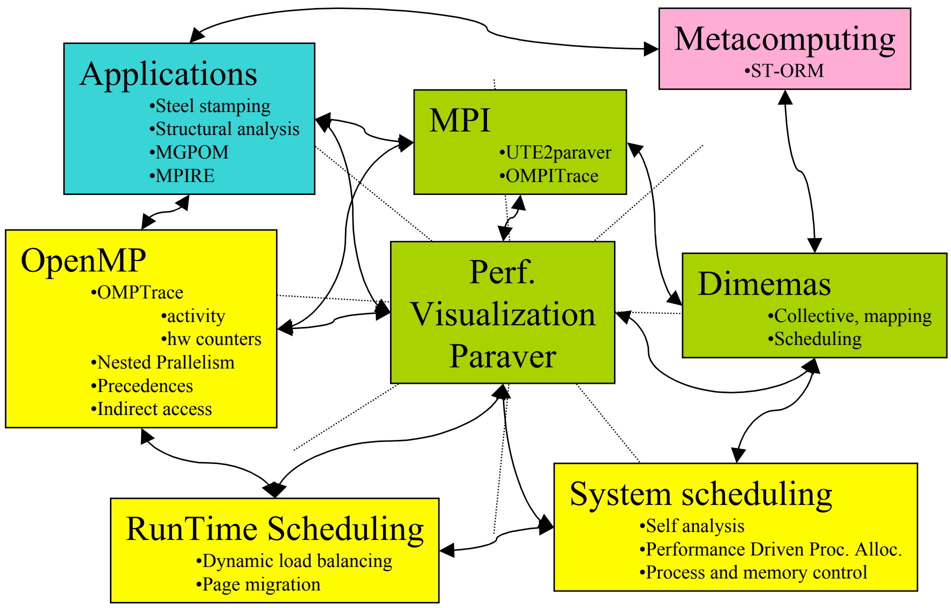
# CEPBA: R&D projects






Barcelona Supercomputer Center
National Supercomputing Facility

# CEPBA: R&D projects




Jornadas Técnicas RedIRIS 2004
Toledo, 29 de Octubre de 2004

# CEPBA: Management of Technology Transfer

1986... 1991... 1994 1995 1996 1997 1998 1999 2000...

R+D projects → 25 Projects

CEPBA →

Technology Transfer Management → 3 cluster projects, 28 Subprojects


- Technical management & Dissemination
- Technological partner & developments

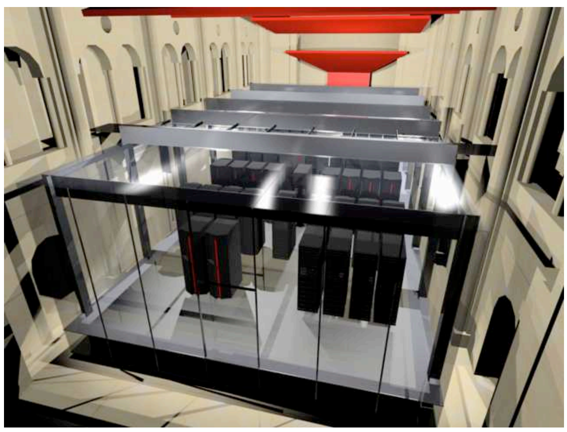


Barcelona Supercomputer Center
National Supercomputing Facility

# The Consortium

- Consortium includes
  - Spanish Government (MEC)
  - Catalonian Government (DURSI)
  - Technical University of Catalunya
- To start operations on January, 2005
- Location

  c/ Jordi Girona 31
  08034 Barcelona



Jornadas Técnicas RedIRIS 2004 Toledo, 29 de Octubre de 2004

Barcelona Supercomputer Center
National Supercomputing Facility

# Motivation

- Significant contribution to advancement of Science in Spain, enabling supercomputing capacity, scientific-technical synergies, and cost rationalization thanks to economies of scale
- Powerful tool to assist research and development centers, public and private, generating impulses for a new technological environment



Barcelona Supercomputer Center
National Supercomputing Facility

# Mission

"Investigate, develop and manage technology to facilitate the advancement of science"