UNICORE and the Fastest Supercomputer in Europe
Introduction to BSC

Research

MareNostrum Architecture

UNICORE Access to MareNostrum
Introduction to BSC

Research

MareNostrum Architecture

UNICORE Access to MareNostrum
Consorium composition:
- Spanish Government (MEC)
- Catalan Government
- Universitat Politècnica de Catalunya (UPC)
- Generalitat de Catalunya
- Departament d'Universitats, Recerca i Societat de la Informació
Investigate, develop and manage technology to facilitate the advancement of Science.
• Researching Supercomputing and Computer Architecture

• Collaborating in R&D e-Science projects with prestigious scientific teams

• Managing BSC supercomputers to accelerate relevant contributions
Introduction to BSC

Research

MareNostrum Architecture

UNICORE Access to MareNostrum
• Continuation of CEPB association (European Center for Parallelism, Barcelona)

• Tools for performance analysis

• Programming models

• Operating Systems
Introduction to BSC

Research

MareNostrum Architecture

UNICORE Access to MareNostrum
- 4.812 IBM PowerPC 970 FX processors (dual processors)
- 9.6TB Memory - 4GB ECC 333 DDR memory per node
- 140+93 disk
- 3 network Myrinet
Compute Racks (RC01-RC27)
- BC chassis w/OPM and gigabit ether switch
- JS20+ nodes w/myrinet daughter card

Storage Server Racks (RS01-RS07)
- 15 storage servers 6/rack
- stT 100 3/rack
- XP100 3/rack

Myrinet Racks (RM01-RM04)
- 10 clos 256+256 myrinet switches
- 2 Myrinet spines 1280s

1 Gigabit Network Racks
- 1 Force10 E600 for Gb network
- 4 Cisco 3550 48-port for 10/100 network

Operations Rack (RO1)
- 7316-TF3 display
- 2 p615 mgmt nodes
- 2 HMC model 7315-150
- 3 Remote Async Nos
- 3 Cisco 3550
- 1 BC chassis (BCIO)
JS20 Processor Blade
- 2-way 2.2 GHz Power PC 970 SMP
- 4GB memory (512KB L2 cache)
- Local IDE drive (40 GB)
- 2x1Gb Ethernet on board
- Myrinet daughter card

Blade Center
- 14 blades per chassis (7U)
- 28 processors
- 56GB memory
- Gigabit ethernet switch

6 chassis in a rack (42U)
256 links (1 to each dir.
250MB/s each dir.)
Hardware: Installation
<table>
<thead>
<tr>
<th>State / Country / Year</th>
<th>Computer / Processors Manufacturer</th>
<th>Computer Family Model</th>
<th>Inst. type Installation Area</th>
<th>$R_{\text{max}}$</th>
<th>$R_{\text{peak}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>DOE/NNSA/LLNL United States/2005</td>
<td>BlueGene/L eServer Blue Gene Solution / 65536 IBM</td>
<td>IBM BlueGene/L eServer Blue Gene Solution</td>
<td>Research</td>
<td>136800</td>
<td>183500</td>
</tr>
<tr>
<td>IBM Thomas J. Watson Research Center United States/2005</td>
<td>BGW eServer Blue Gene Solution / 40960 IBM</td>
<td>IBM BlueGene/L eServer Blue Gene Solution</td>
<td>Research</td>
<td>91290</td>
<td>114688</td>
</tr>
<tr>
<td>NASA/Ames Research Center/NAS United States/2004</td>
<td>Columbia SGI Altix 1.5 GHz, Voltaire Infiniband / 10160 SGI</td>
<td>SGI Altix SGI Altix 3700</td>
<td>Research</td>
<td>51870</td>
<td>60960</td>
</tr>
<tr>
<td>The Earth Simulator Center Japan/2002</td>
<td>Earth-Simulator / 5120 NEC</td>
<td>NEC Vector SX6</td>
<td>Research</td>
<td>35860</td>
<td>40960</td>
</tr>
<tr>
<td>Barcelona Supercomputer Center Spain/2005</td>
<td>MareNostrum JS20 Cluster, PPC 970, 2.2 GHz, Myrinet / 4800 IBM</td>
<td>IBM Cluster JS20 Cluster</td>
<td>Academic</td>
<td>27910</td>
<td>42144</td>
</tr>
</tbody>
</table>
Campos Plasencia
University of Zaragoza

- Search of nuclear fusion materials
- w-up of crystal particles

José Caturla
University of Alicante

- of the behavior of materials under conditions of pressure and temperature

Orozco
Institute of Bioinformatics

- Protein dynamics of all native proteins
- Folding simulation

Javier Jiménez Sendin
Technical University of Madrid

- Turbulent channel simulation with Re numbers of friction of 2000

Markus Uhlmann
CIEMAT

- Direct Numerical Simulation of Turbulent Flow with Suspended Solid Particles

Gustavo Yepes Alonso
Autonomous University of Madrid

- Hydrodynamic simulations in Cosmology
- Simulation of a universe volume of 500 Mpc (1,500 million light years)
Introduction to BSC

Research

MareNostrum Architecture

UNICORE Access to MareNostrum
The UNICORE Network Gateway

Job and User Credentials

NJS

Gateway

IDB

Gateway

IDB

Gateway

IDB
The UUDB contains...
• The Target System Interface (TSI) is located in the MareNostrum 'login nodes'.

• Same environment than locally submitted jobs.

• The TSI translates an Abstract Job Object.
A new TSI is being developed.

Targets the SLURM batch system.

It will be needed when SLURM replaces LoadLeveler as a production batch system at BSC.

The job script from
The IDB does not allow an accurate description of MareNostrum resources:

• No support for classes
• No arbitrary constraints
• CPU limits (Unicore) vs. 'wall clock' limits (BSC)
The MareNostrum supercomputer will be used by local and remote users.

UNICORE provides a mechanism to access several sites.

Common to several sites.

Batch system...
Thank you!