



**Barcelona
Supercomputing
Center**

Centro Nacional de Supercomputación



Grandes proyectos de la RES

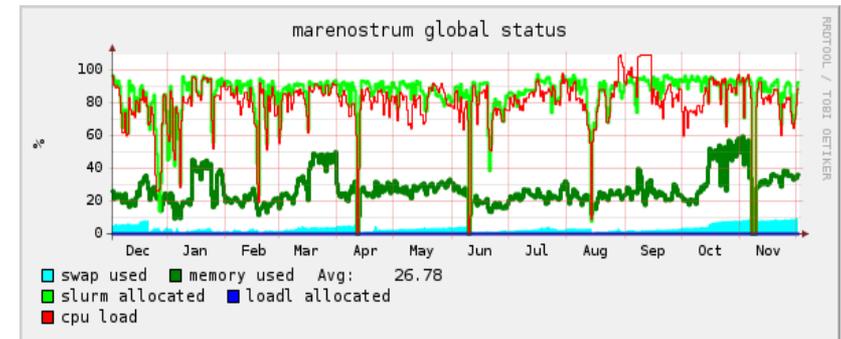
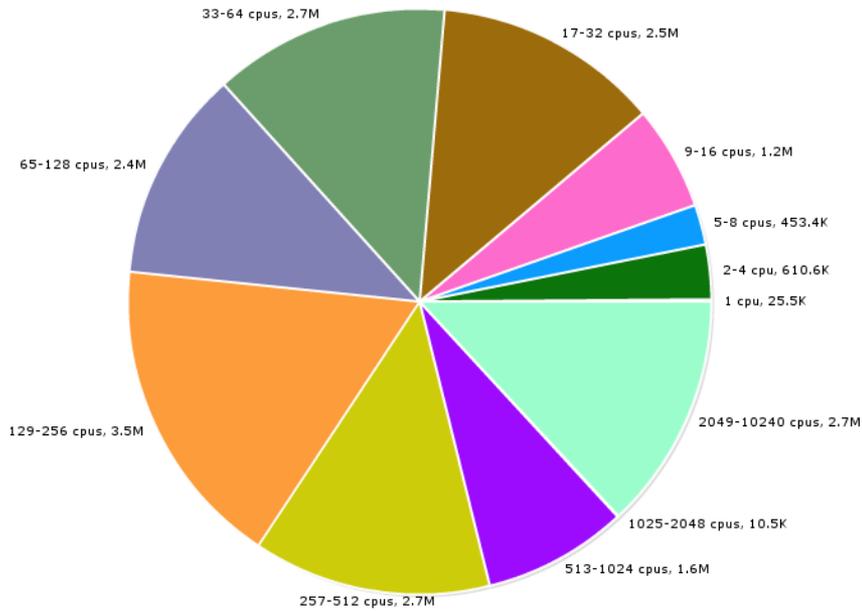
David Vicente
HPC Support group manager

Que es un “grand challenge” ?



MareNostrum Cpu Hours distribution

Start date: 2009-09-01
End date: 2009-12-02





Dinámica de Fluidos
Astrofísica
Genómica computacional
Ciencias de la Tierra



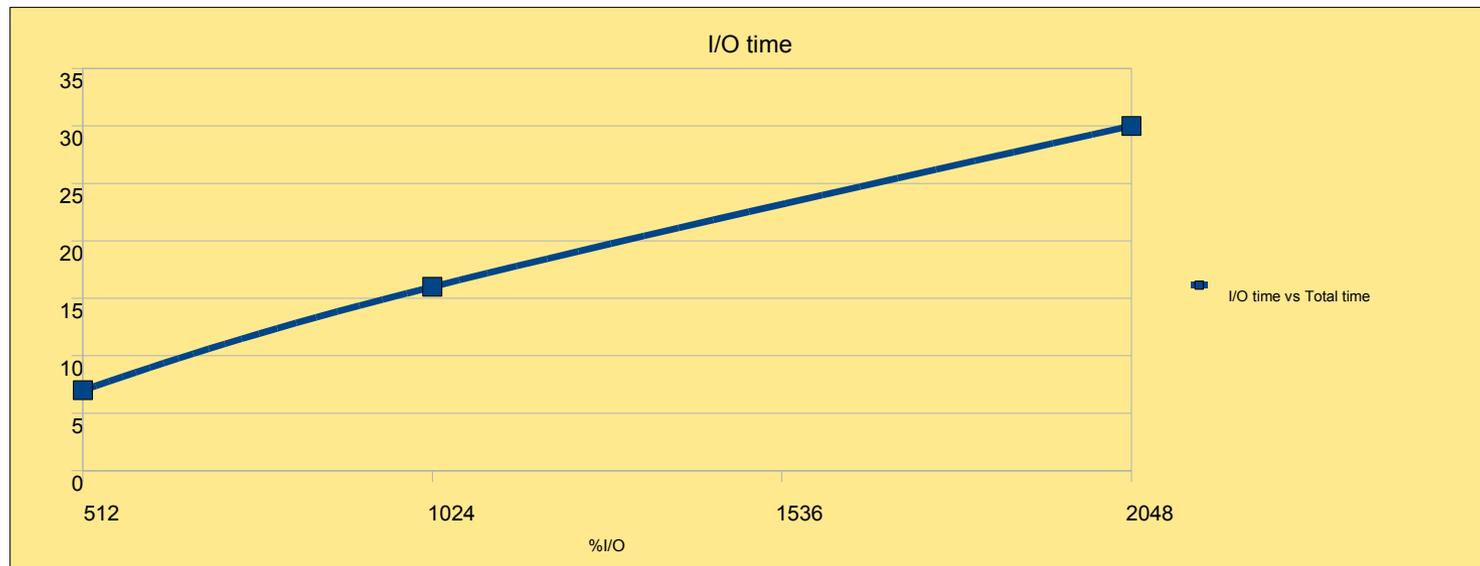
Channel 2000: A Highly Parallel Algorithm for Computing Wall Turbulence



IP : Javier Jimenez Sendin, Escuela de Aeronauticos (UPM)



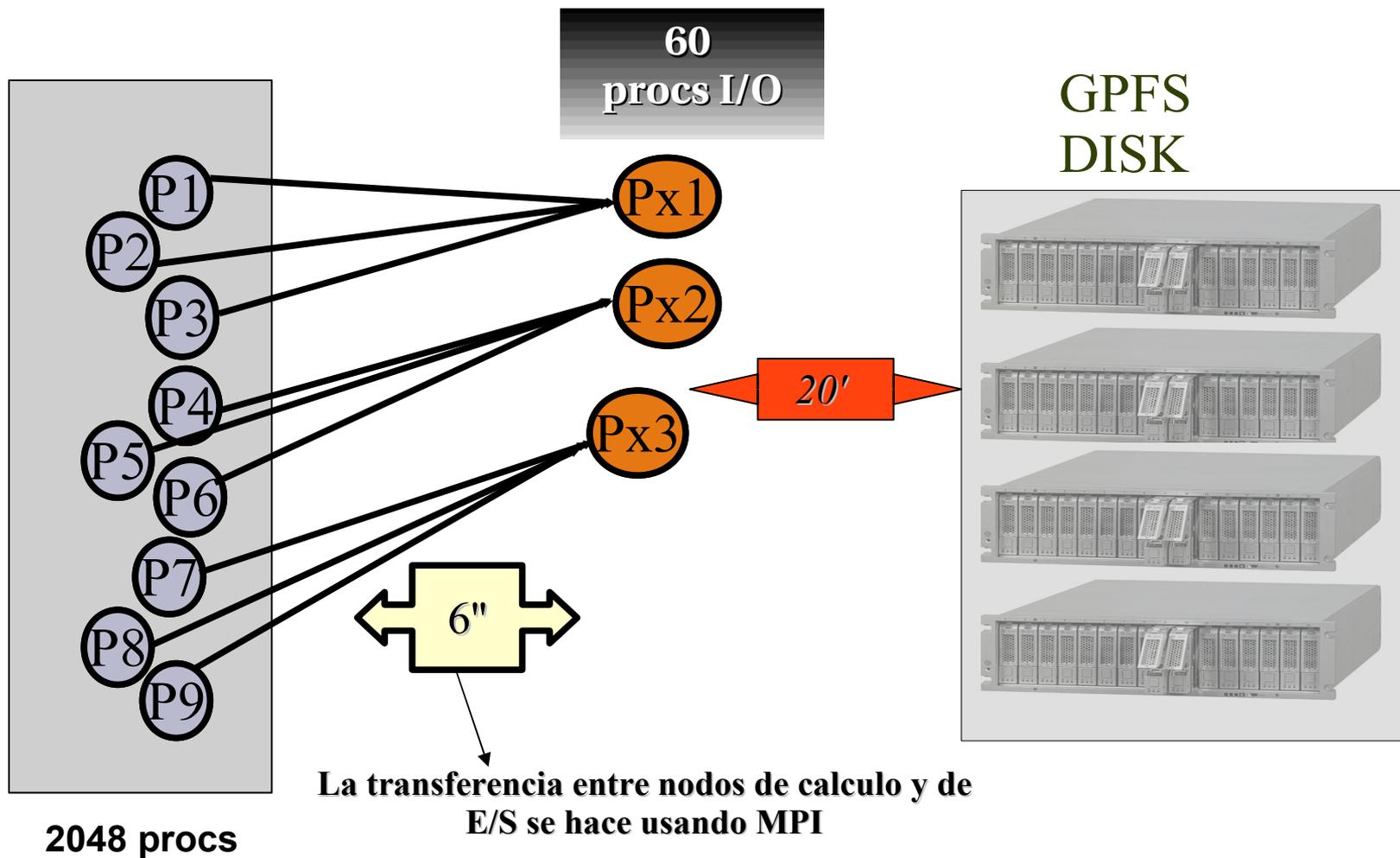
- Escritura de restarts cada 200 steps (unos 60GB)
 - Implementación Original :
 - Escritura gestionada por el nodo master .



7% con 256
>30% con 2048

- Posibles soluciones
 - Escritura en paralelo usando los discos locales
 - Escritura en paralelo usando GPFS (global filesystem)
 - Uso de nodos dedicados a E/S

Channel 2000 : E/S



Channel 2000 : Resumen

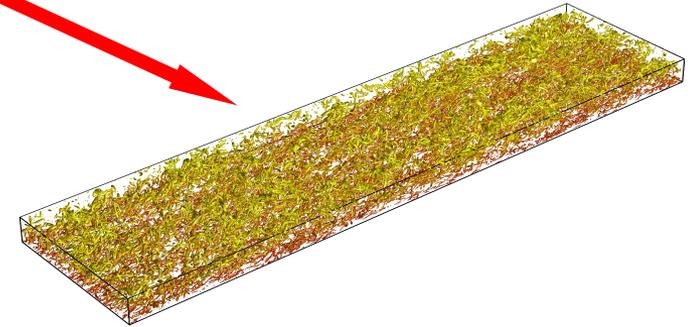
“ This simulation is used to understand the physics of turbulent boundary layers at high Reynolds numbers “



Matriz: $6144 \times 633 \times 46084 \sim 2 \times 10^{10}$

Memoria 400 GB distribuida

Tiempo por paso : 2048 procesadores 39.7s



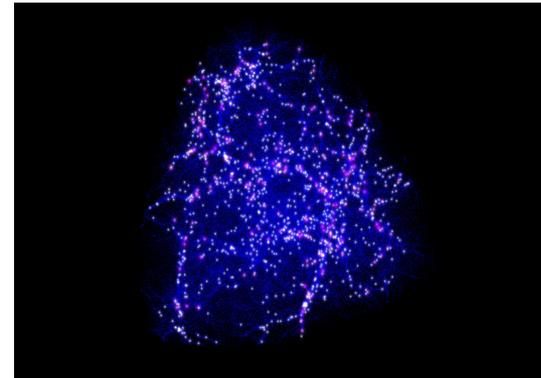
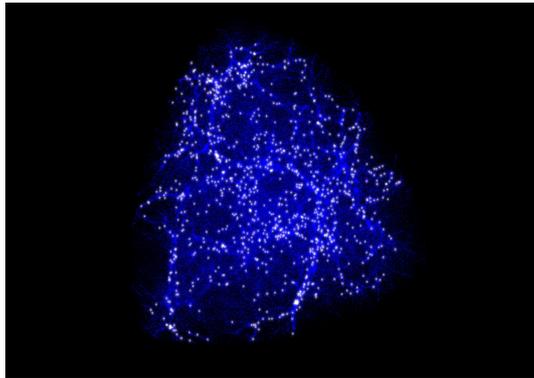
Números globales:

Pasos	125.000
Total horas CPU	6×10^6 (6 millones)
Total horas hombre	4 meses



Este proyecto ha generado la mayor simulación numérica en cosmología hasta la fecha. Esto nos permitirá conocer mejor el origen y evolución del Universo.

- 10+ millones de horas CPU, hasta 4.096 procesadores, 8 TB RAM , +50 TB de resultados
- GADGET N-body simulation con 10^9 - 10^{11} partículas.



IP : Pablo Fosalba, Instituto de Ciencias del Espacio (IEEC-CSIC)

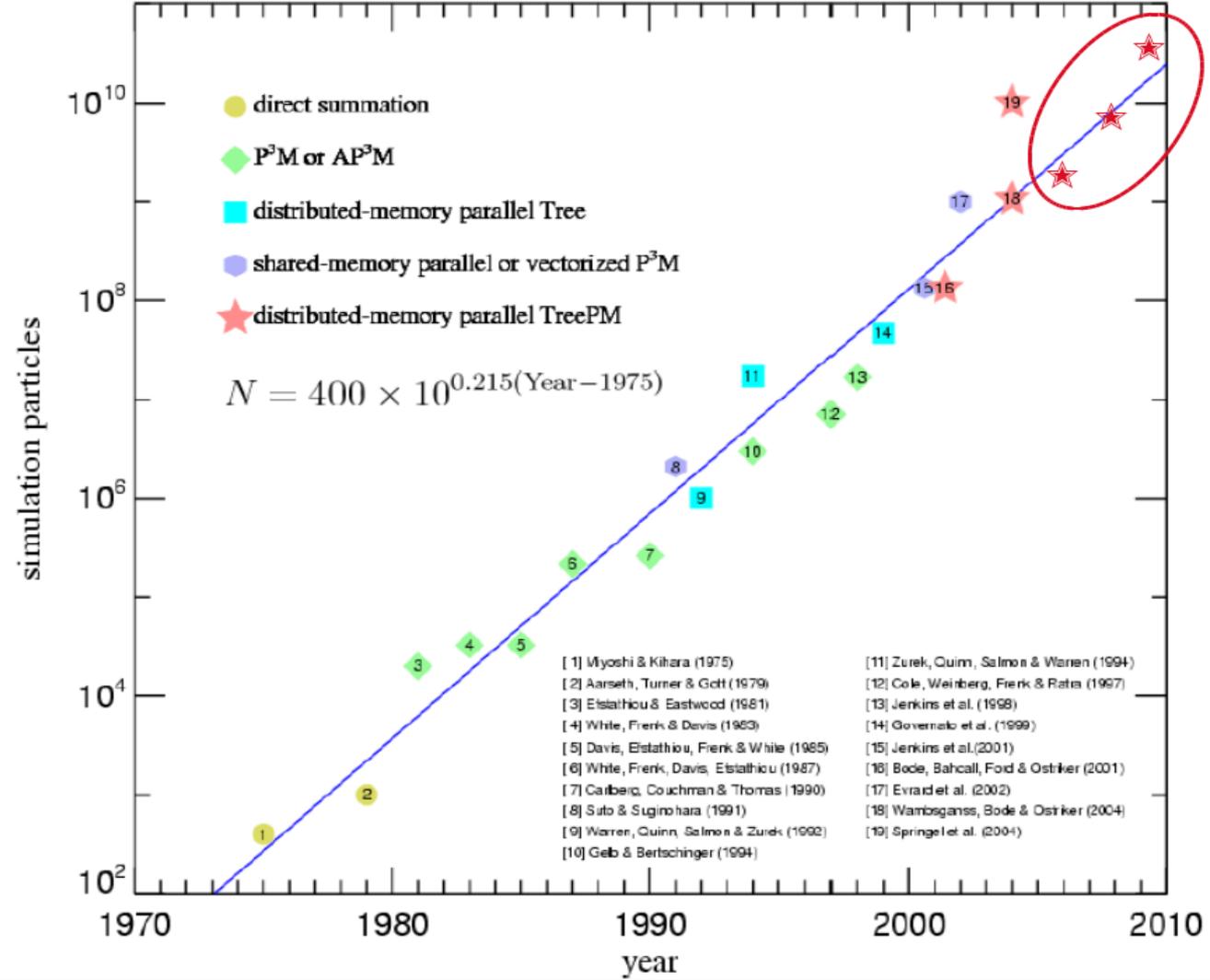
www.ice.cat/mice

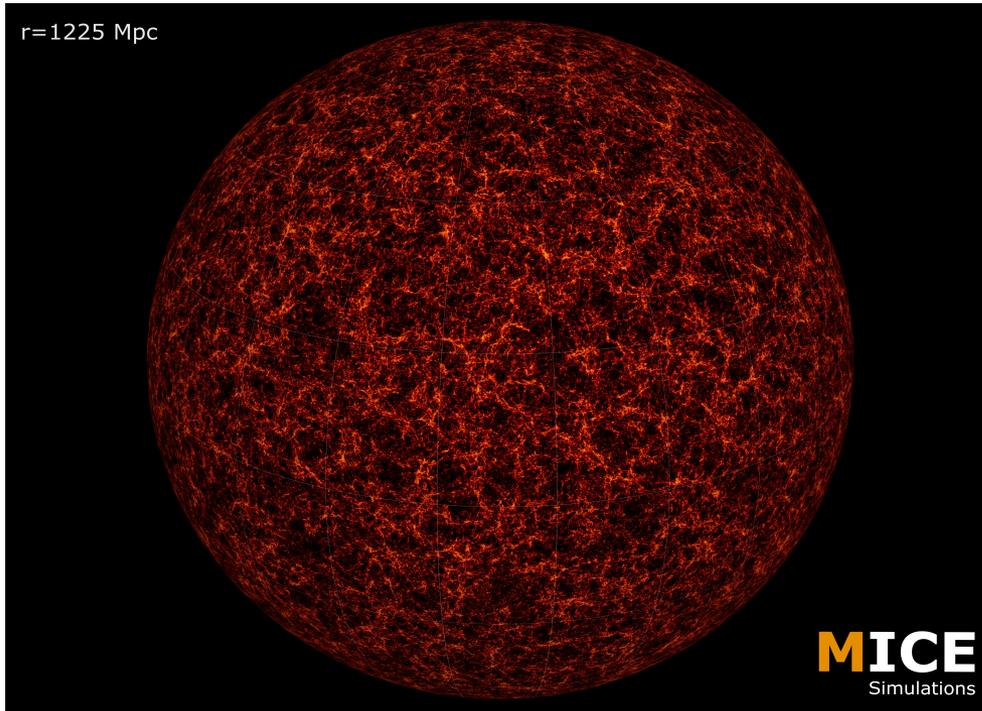
MICE simulation



Cosmological N-body simulations have grown rapidly in size over the

- ▶ Computers double their speed every 18 months (Moore's law)
- ▶ N-body simulations have doubled their size every 16-17 months
- ▶ Recently, growth has accelerated further. The Millennium Run should have become possible in 2010 – we have done it in 2004 !





- Problemas a nivel técnico:

- **Escalabilidad**
- **Desbalanceo de carga**
- **Entrada / Salida**

- Mejoras del código para incrementar la eficiencia de la simulación :

- **Memoria**
- **E/S**
- **MPI (red myrinet)**
- **Balaneo dinámico**



Este proyecto es la mayor comparación de proteínas que se ha realizado hasta el momento. Se integraron las secuencias peptídicas de todo el DNA que se conoce hasta la fecha.

Código : BLAST

procesadores: 4096

Memoria : 5TB (Distribuidos)

Tareas independientes ejecutadas :~ 2 millones (usando GridSs)

Resultados : 8 TB

Duración : 2 semanas con 4096 cpus dedicadas

- IP: Dr Bork (EMBL) & Dr David Torrents (BSC)
- Uso de GridSs para gestionar la simulación
- Uso de los discos locales para mejorar la lectura de las BBDD de BLAST

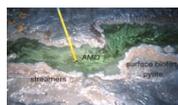
Analysis of protein and function diversity on earth



Organisms

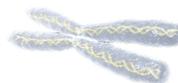


Environmental genomics



We have compared 15 million protein sequences using **BLAST** and *GRID Superscalar*

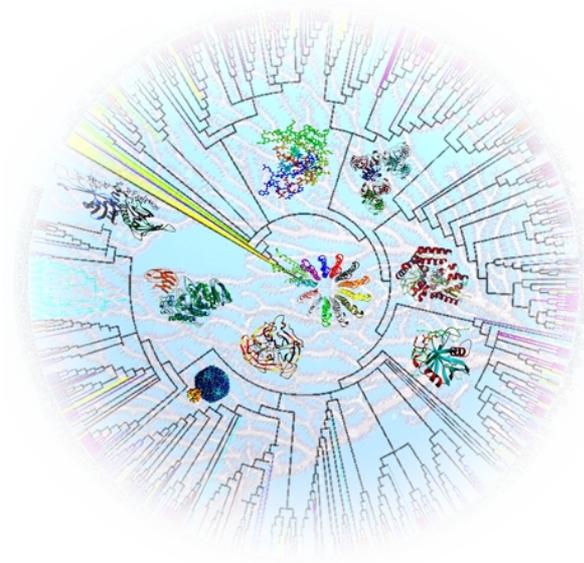
Genomes



15 million Proteins



Query and Database file size: ~ 5 Gb
4000 CPUs (= 1000 exclusive nodes)
Total CPU time: 311,112 hours
5 Tb of results



CALIOPE : sistema Nacional de calidad del aire



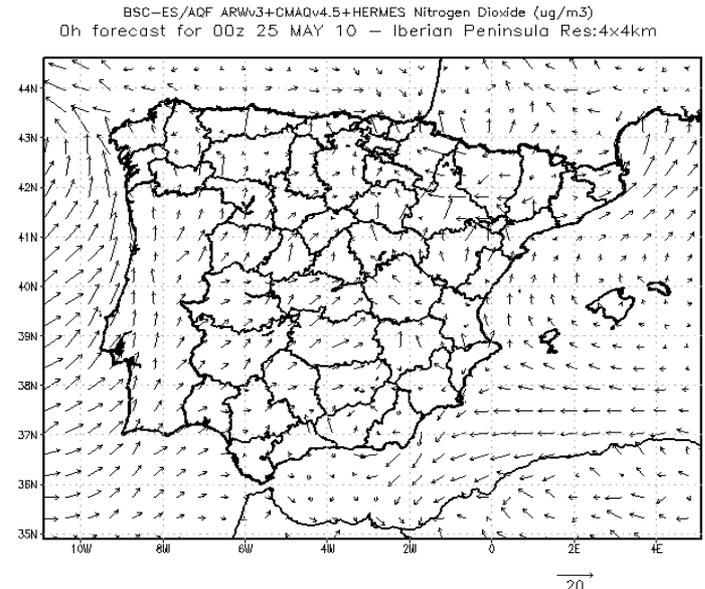
Proyecto : CALIOPE establece un sistema de pronósticos de calidad del aire para España. El objetivo es realizar predicciones detalladas sobre los niveles de contaminantes a los que se ve sometida la población.

Procesadores : 512
Outputs : >1 TB (diarios)

Duración : 10 horas

- Proyecto **CALIOPE** (www.bsc.es/caliope)

- El modelo hace uso de un gran volumen de datos atmosféricos, que son tratados por un conjunto de aplicaciones, entre las que cabe destacar :
 - WRF (Weather Research & Forecasting) : para la predicción meteorológica
 - CMAQ (Community Multiscale Air Quality) : para la simulación del modelo de calidad del aire.

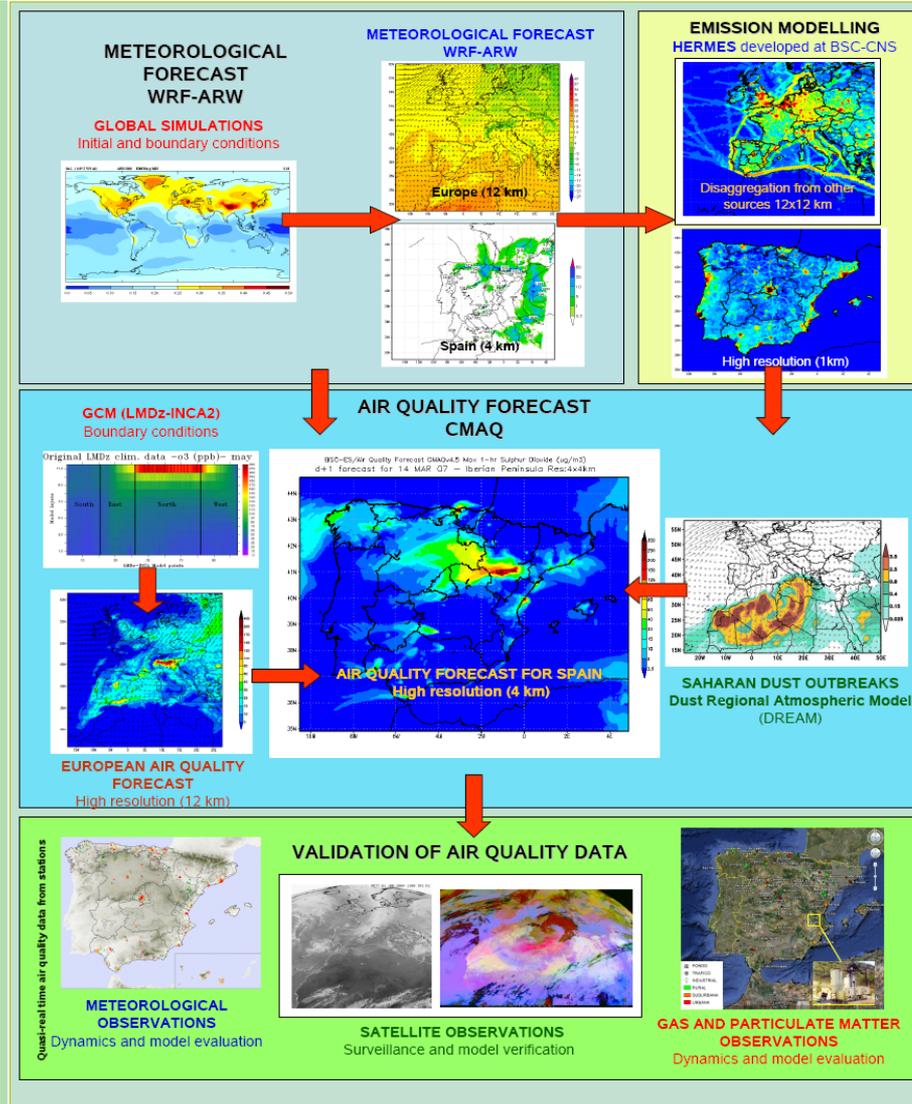


CALIOPE, sistema Nacional de calidad del aire



Workflow diario :

1. Meteorological Forecast.
 - **WRF-ARW** v3.0.1.1 (inputs from global models).
2. Emissions Forecast:
 - **EMEP** (Europe)
 - **HERMES** (Spain).
3. Air Quality Forecast:
 - **CMAQ-CTM** v4.5 (inputs from LMDz-INCA2).
 - Aerosol module (AERO4).
4. Desert dust from Africa Forecast
 - **BSC-DREAM8b**.
5. Results
 - Hourly concentration: **O₃**, **NO₂**, **SO₂**, and **PM₁₀**
6. Forecast Evaluation
 - Spanish and European networks
 - Satellite data (OMI)
 - Ozone sondes.





Gracias !
<http://www.bsc.es>

