SyeC
Supercomputación y eCiencia

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Resultados del proyecto Supercomputación y e-Ciencia SyeC

Consolider

WP1 – Optimización de aplicaciones

- Optimización de aplicaciones
  - Optimización secuencial
  - Mejora de la escalabilidad con MPI
  - Paralelización híbrida MPI/OpenMP
  - Colaboraciones grupos SyeC:
    - Gromacs: UB + BSC-CS
    - T-coffee: CRG + UAB
    - GEM: CRG + UNIZAR
    - WRF, ALYA: BSC-ES + BSC- CASE
    - NCAR-WACCM: Dos grupos de la UCM
    - SPEV: UV + UM
    - MRGENESIS: UV + BSC-CS
    - RATPENAT: UV + BSC-CASE
    - pDEVA: UAM + BSC-CASE
    - SIESTA: CSIC–ICN2+ BSC-CASE
  - Publicaciones conjuntas (2010)
    - BSC-CASE/CS y ES (5)
    - BSC-CASE y UV (1)
    - BSC-CASE y ICMAB/CIN2 (1)

WP1 – Optimización de aplicaciones

- Uso de modelos de programación para la expresión de workflows: StarSs y MapReduce
  - Colaboraciones grupos SyeC:
    - GASS, GOG: UB + BSC-CS
    - Life Sciences: UB + BSC-CS
    - MSA: CRG + UAB
    - Earth Sciences: BSC-ES + BSC-CS
  - Publicaciones conjuntas (2010)
    - UAB y CRG (1)
    - BSC-CS y UB (1)
- Aplicaciones de biomedicina y procesado de video/imagen en GPUs
  - 3D Fast Wavelet Transform, Generalized Hough Transform, … (UM)
  - Image segmentation (UNIZAR)

WP2-3 – Tecnologías de supercomputación

- WP2 – Software de sistema para futuros supercomputadores
  - Modelos de programación: StarSs, CUDA/OpenCL, MPI
  - Planificación en multicores homogéneo y heterogeneos (aceleradores, asimétricos, …) y técnicas de work stealing
  - Entornos para la especificación y ejecución (gestión de recursos) de aplicaciones workflow: StarSs y MapReduce

- Colaboraciones:
  - BSC-CS, UCM, UM
  - BSC-CS, UAB

- WP3 – Arquitectura para los futuros supercomputadores
  - Jerarquía de memoria para uniprocesadores y mecanismos de prebúsqueda adaptativos
  - Sistemas multiprocesador en chip: interconexión y memoria
  - Redes de interconexión para grandes sistemas paralelos, routing, …
  - Soporte a nivel de arquitectura a los modelos de programación (memoria transaccional, sincronizacion, tareas, …)

- Colaboraciones:
  - UNIZAR, UC, UPC
  - UC, BSC-CS

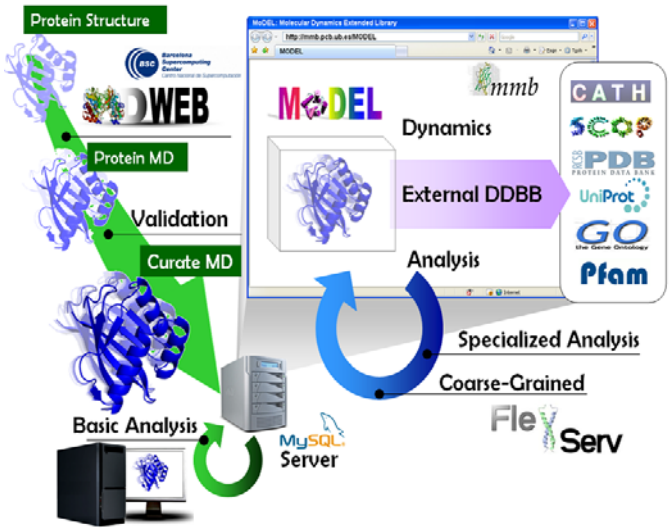
WP2-3 – Tecnologías de supercomputación

- Algunas publicaciones destacadas (2010):
  - IEEE Transactions on Parallel and Distributed Systems (2)
  - IEEE/ACM International Symposium on Microarchitecture (1)
  - IEEE International Symposium on Information Theory (1)
  - ACM International Conference on Supercomputing (3)

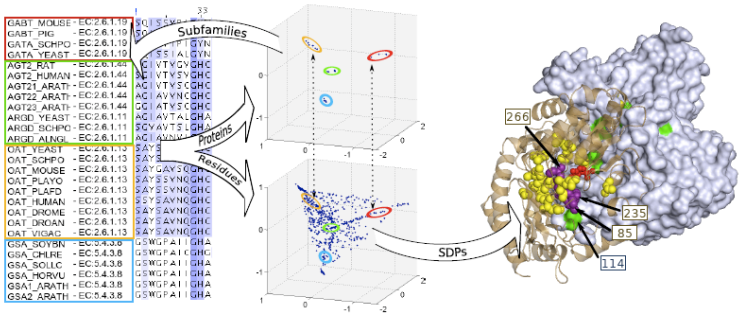
- Internacionalización
  - Inicio de nuevos proyectos FP7: TERAFLUX, TEXT, EnCORE, ScalaLIFE, …
  - Colaboraciones con laboratorios de investigación: IBM, NVIDIA, Intel, …
  - Participación en iniciativas Exascale: PRACE, EESI e IESP
    - Proyecto EU Mont-blanc

WP4 – Simulación de estructuras macromoleculares

(1/3)



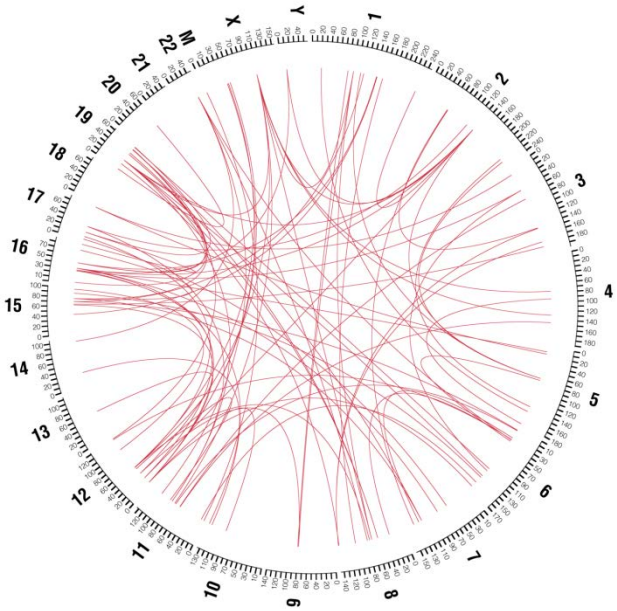
MoDEL infrastructure



Identification of sequence groups and associated significant residues

- Impacto de SyeC
  - Generación de la mayor base de datos de trayectorias de DM de proteínas
  - Optimización de NAMD y GROMACS para HPC
  - Capacidad de despliegue de la infraestructura de computación adecuada
  - Definición de la siguiente generación de simuladores para la determinación del “human protein interactome”
- Algunas publicaciones destacadas (2010)
  - Rausell, Juan, Pazos and Valencia. Protein interactions and ligand binding: from protein subfamilies to functional specificity. *Proc. Natl. Acad. Sciences*
  - Portella and Orozco. Multiple routes characterize the folding of a small DNA hairpin. *Angew. Chem. Int.*
  - Pérez and Orozco. Real time atomistic description of DNA unfolding. *Angew. Chem. Int. Ed. Eng.*
- Colaboración entre (IRB, CNIO) y BSC-CS
- Internacionalización
  - Participación el proyecto ABC
  - Participación en el proyecto EU-ScalaLIFE
  - ERC Victor Guallart

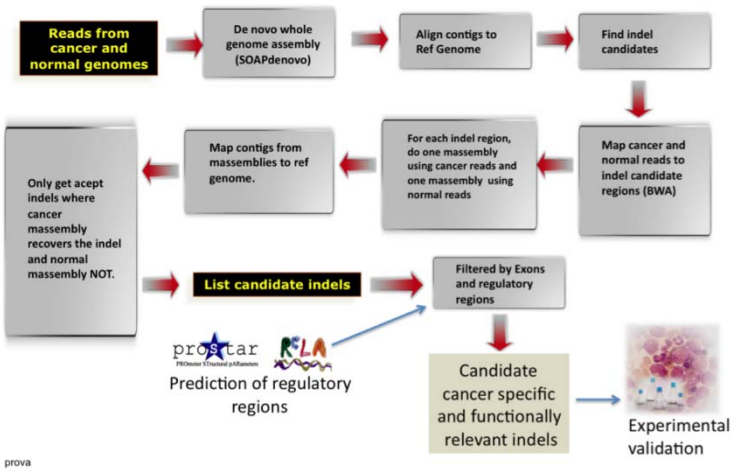
WP4 – “Splicing” alternativos y epigenética (2/3)



Eventos reportados de
inter-chromosomal trans-splicing

- Impacto de SyeC
  - Optimización de códigos genómicos para acelerar el análisis de splicing alternativo en RNA
  - Optimización de los códigos genómicos permite análisis masivos imposibles hasta ahora, con aceleraciones de 2.5x a 3x en las versiones secuenciales por aumento de localidad
  - Podemos llegar a determinar y cuantificar el transplicing en la especie humana
- Algunas publicaciones destacadas (2010)
  - Montgomery et al. Transcriptome genetics using second generation sequencing in a Caucasian population. *Nature*
  - Buschbeck et al. The histone variant macroH2A is an epigenetic regulator of key developmental genes. *Nat. Struct. Mol. Biol.*
- Colaboración entre (IRB, CRG, CNIO) y (UNIZAR, UAB): GEM, T-coffee
- Internacionalización
  - Participación en el proyecto ENCODE

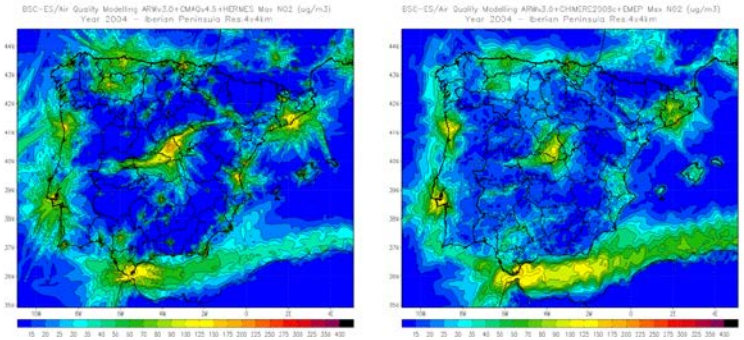
WP4 – Descrifrando las claves de la leucemia (3/3)



BSC-Workflow para detección y anotación de InDel

- Impacto de SyeC
  - Análisis de las interacciones entre genes que producen leucemia. Impacto potencial enorme en terapia para humanos
  - Desarrollo en un tiempo breve de toda la infraestructura necesaria
- Publicación destacada (2010)
  - X.Puente et al., *Nature* (2nd revision)
- Colaboración entre (IRB, CRG, CNIO) y (BSC-CS): workflows
- Internacionalización
  - Participación en el International Cancer Genome Consortium (primera vez que participa un centro de supercomputación)

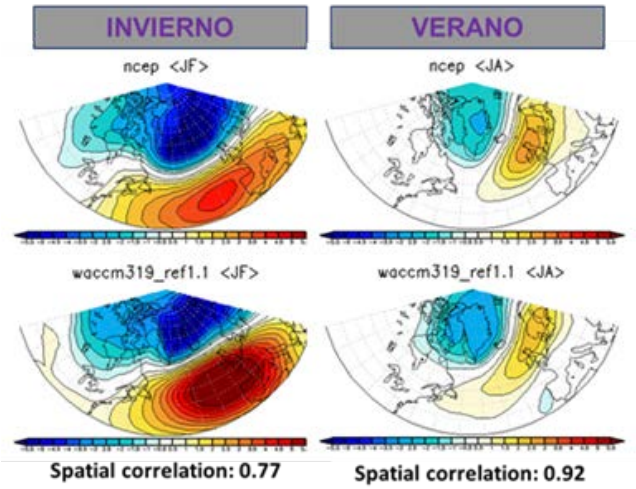
WP5 – Escalabilidad de códigos atmosféricos (1/2)



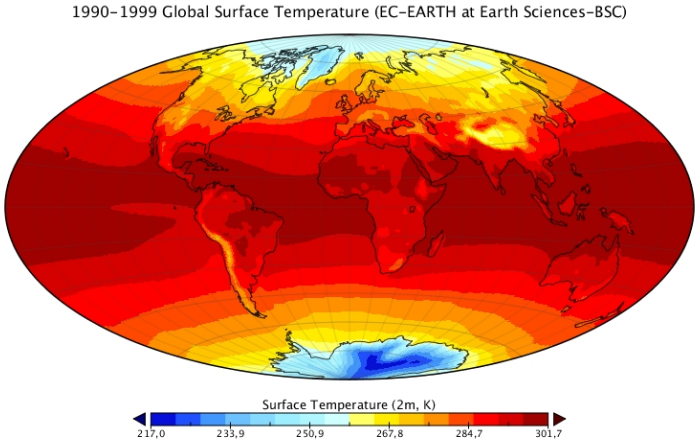
Concentracion de NO_2 en la peninsula Iberica

- Impacto de SyeC
  - Escalabilidad de HERMES (emisiones) para mallas 1x1
  - Comparación WRF-CHIMERE-EMEP con WRF-CMAQ-HERMES y WRF-CHIMERE-HERMES
  - Producción operativa y optimización de CALIOPE
  - Nuevos esquemas numéricos en malla no estructurada con escalabilidad a 10^4 procesadores
- Algunas publicaciones destacadas (2010)
  - M.T. Pay et al. A full year evaluation of the CALIOPE-EU air quality modeling system over Europe for 2004. *Atmospheric Environment*
  - R. Aubry et al. An unstructured CFD approach to numerical weather prediction, *AIAA Conference*
- Colaboración entre (BSC-ES, CIEMAT) y BSC-CASE
- Internacionalización
  - INERIS – France, IPSL – France (código CHIMERE)
  - EPFL–Switzertland, NCAR–USA (nuevas parametrizaciones urbanas)
  - Naval Research Laboratory (Naval Postgraduate School). (USA). Nuevos esquemas númericos

WP5 – Modelos de circulación global (2/2)

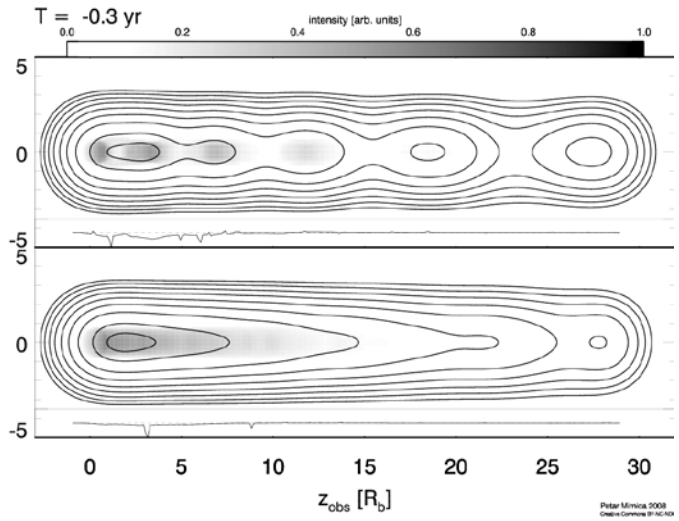


Simulación de la oscilación del Atlantico Norte en invierno y en verano con WACCM



- Impacto de SyeC
  - Puesta en marcha del modelo EC-EARTH (NEMO+IFS+OASIS)
  - Uso de WACCM para simular la evolución de la tropopausa y los efectos del ciclo solar en el clima terrestre
  - Mejora de la escalabilidad de WACCM de 40 a 180 procesadores con una reducción del tiempo de ejecución de 16 a 6 horas
- Algunas publicaciones destacadas (2010)
  - G. Chiodo et al. The solar cycle response in WACCM-3.5 transient simulations, *SCOSTEP Symp.*
  - G. Chiodo et al. Interannual changes in mass consistent energy budgets from ERA-Interim and satellite data, *J. Geophys. Res.*
- Colaboración entre (BSC-ES, UB, UCM) y (BSC-CS, UCM)
- Internacionalización
  - Colaboración estrecha con el NCAR para WACCM
  - Colaboración en el consorcio europeo EC-Earth (ECMWF, KMNI, etc.)

WP6 – Flujos relativistas

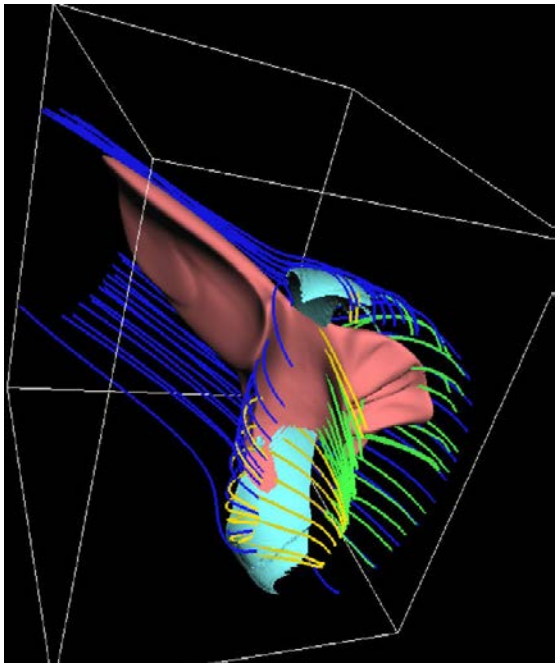


MRGENESIS/SPEV 4D simulation of the radio emission of pc-scale jets

- Impacto de SyeC
  - Validación de simulaciones de flujo térmico relativista con observaciones
  - Escalado de MRGENESIS/SPEV a 2000 cores con paralelización híbrida MPI/OpenMP: reducción del tiempo de cálculo por debajo de 1.5 Millones horas-cpu
  - Paralización hibrida de otros códigos (RATPENAT,…) con buen escalado
- Publicación destacada (2010)
  - Mimica, Giannios and Aloy. Multiwavelength afterglow light curves from magnetized gamma-ray burst flows. MNRAS
- Colaboración entre UV y (UM, BSC-CS, BSC-CASE)
- Internacionalización
  - An ERC Starting Independent Researcher Grant (ERC StG 2010) awarded to Dr. M.A. Aloy

WP6 – Flujos magnetizados y erupciones

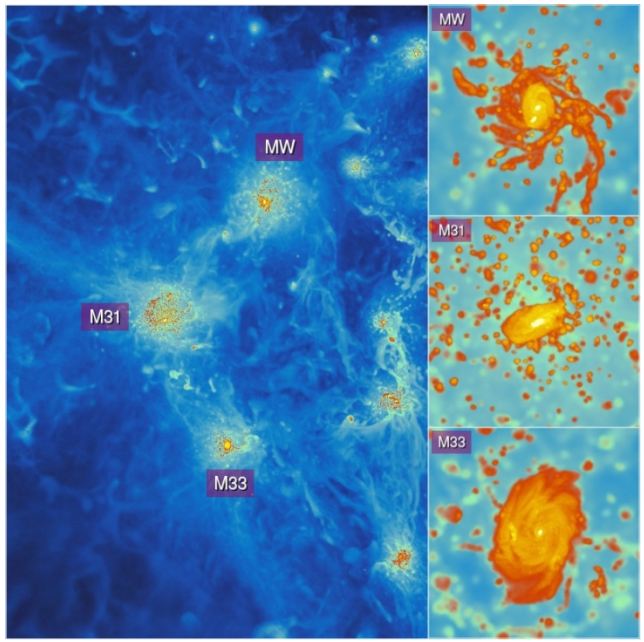
(2/4)



3D numerical model of
X-Ray jet in the solar corona

- Impacto de SyeC
  - First 3D model of observed X-Ray jets in solar coronal holes.
  - First 3D model of subsurface helioseismic waves in sunspots
  - Creación de nuevos códigos MPI para 3D MHD, escalabiliad hasta 1000 cores
  - CPH Stagger Code aplicado al problema de cosmic abundance, incluyendo campos magnético
- Publicación destacada (2010)
  - Fabbian, Khomenko, Moreno-Insertis, Nordlund. Solar abundance corrections derived through three-dimensional magnetoconvection simulations. *ApJ*
- Colaboración entre IAC y BSC-CS
- Internacionalización:
  - Proyecto DEISA: Modelling solar flares and active regions (SunFlare). 2.4 M standard hours

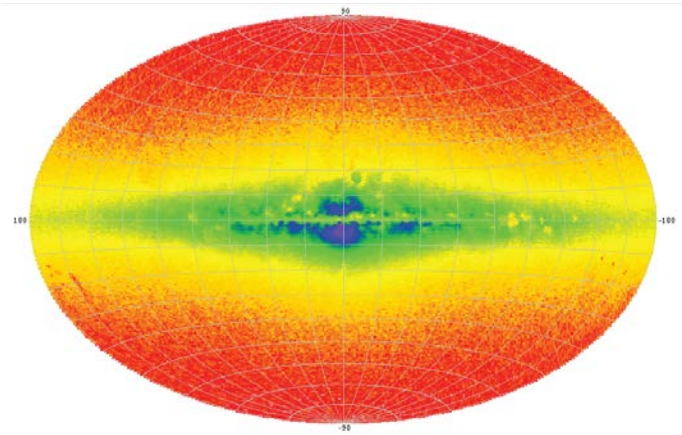
WP6 – Cosmología



Simulación de la formación de nuestro grupo local de galaxias (CLUES)

- Impacto de SyeC
  - Simulaciones hidrodinámicas de la formación del universo a diferentes escalas
  - Nuevo código para generación de condiciones iniciales (GINNUGAGAP) usado en CLUES (Constrained Local Universe Simulation), colaboración CONSOLIDER MultiDark
  - Nuevo código paralelo (MPI) para buscar halos (AHF)
  - Inicio de la paralelización híbrida de pDEVA
- Algunas publicaciones destacadas (2010)
  - Forero-Romero et al. Bullet Clusters in the MARENOSTRUM Universe, *The Astrophysical Journal*
  - Ragone-Figueroa et al. The relation between halo shape, velocity dispersion and formation time, *MNRAA*
- Colaboración entre UAM y BSC-CASE: pDEVA
- Internacionalización
  - Proyecto DEISA: CLUES to Galaxy formation (CLUESGAL). 480 K standard hours

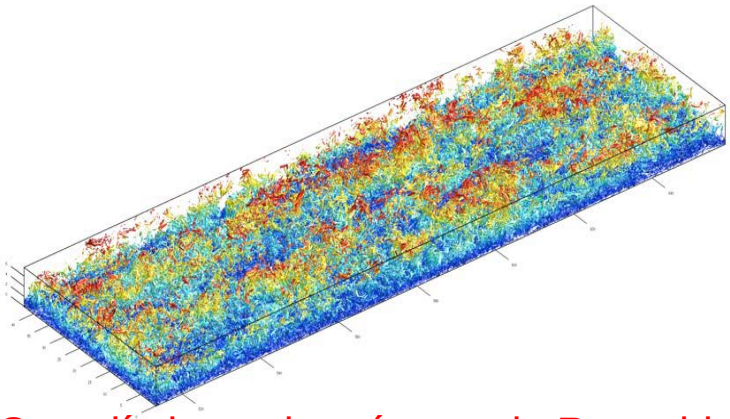
WP6 – GAIA



Gaia Simulations: Density of stars in the Milky Way (30 million objects simulated)

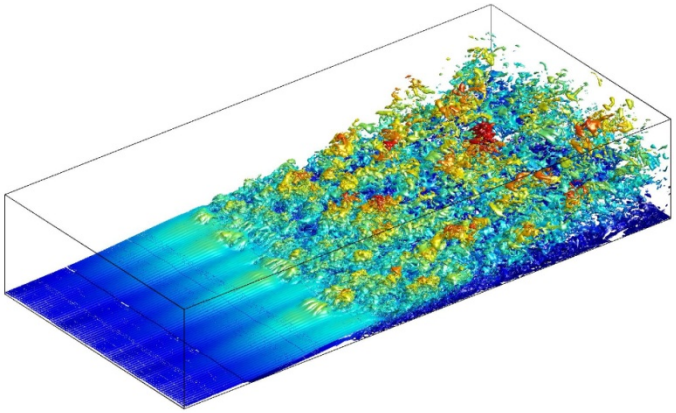
- Impacto de SyeC
  - Simulaciones de universo realista e Instrumentos
  - Producción de 12TB (comprimidos) de datos realistas siguiendo el ciclo de producción marcado por la ESA
  - Definición del reprocesamiento de datos en bruto y sistema de calibración (IDU) acordes con los requerimientos de la misión
  - Despliegue del simulador GASS usando GridSs con 3700 cores
  - Testeo exitoso de la transferencia de alta velocidad (250Mbps) entre BSC y ESAC (Madrid)
  - Uso de HDF5 para los datos en bruto de Gaia
- Publicación destacada (2010)
  - C. Fabricius et al. Reduction of the Gaia astrometric data. *SEA Conference*
- Colaboración entre UB y BSC-CS
- Internacionalización
  - GAIA es una misión internacional de la ESA

WP7 – Simulación DNS de turbulencia



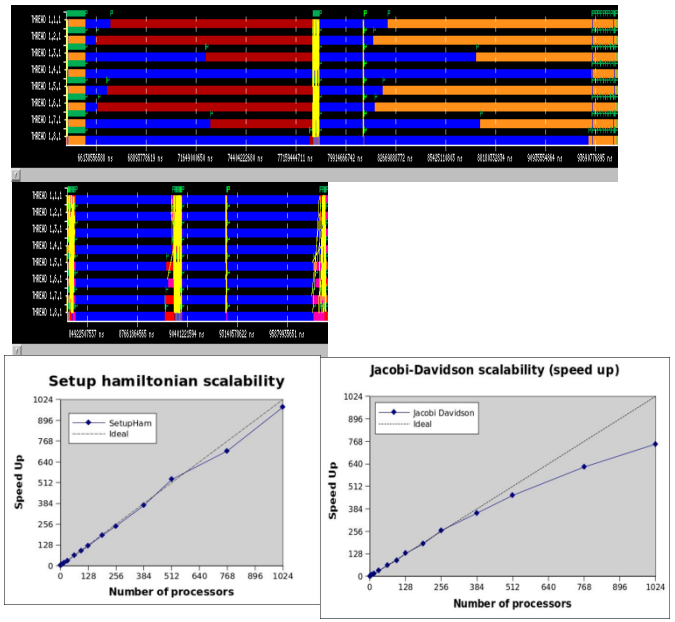
Capa límite a alto número de Reynolds

- Re_θ= 3500-6300 (ant. 4000, Estocolmo)
- 55 x 10^9puntos
- 32 Kcores
- Hybrid (MPI/OpenMP)



- Impacto de SyeC
  - Comparación entre canales y capas
  - Simulación con gradientes adversos de presión
  - Análisis de la capa limite para alto número de Reynols con paredes lisas y rugosas
  - Simulaciones de canal DNS en O(10^4) cores
  - Organización de grandes bases de datos (300 TB)
- Algunas publicaciones destacadas (2010)
  - J. Jiménez et al. Turbulent boundary layers and channels at moderate Reynolds numbers, *J. Fluid Mech.*
  - A. Pinelli et al. Immersed-boundary methods for general finite-difference and finite-volume navier-stokes solvers, *J. Comput. Phys.*
- Futura colaboración UPM y BSC-CS en el acceso a grandes bases de datos.
- Internacionalización
  - Proyecto PRACE: Entrainment effects in rough-wall boundary layers. JSC Jugene: 40 M core hours (2011)
  - Incite (BlueGene/P- Argonne, UT Austin. 75 M hours (2010-11)
  - ERC Advanced Idea (2011-2015)

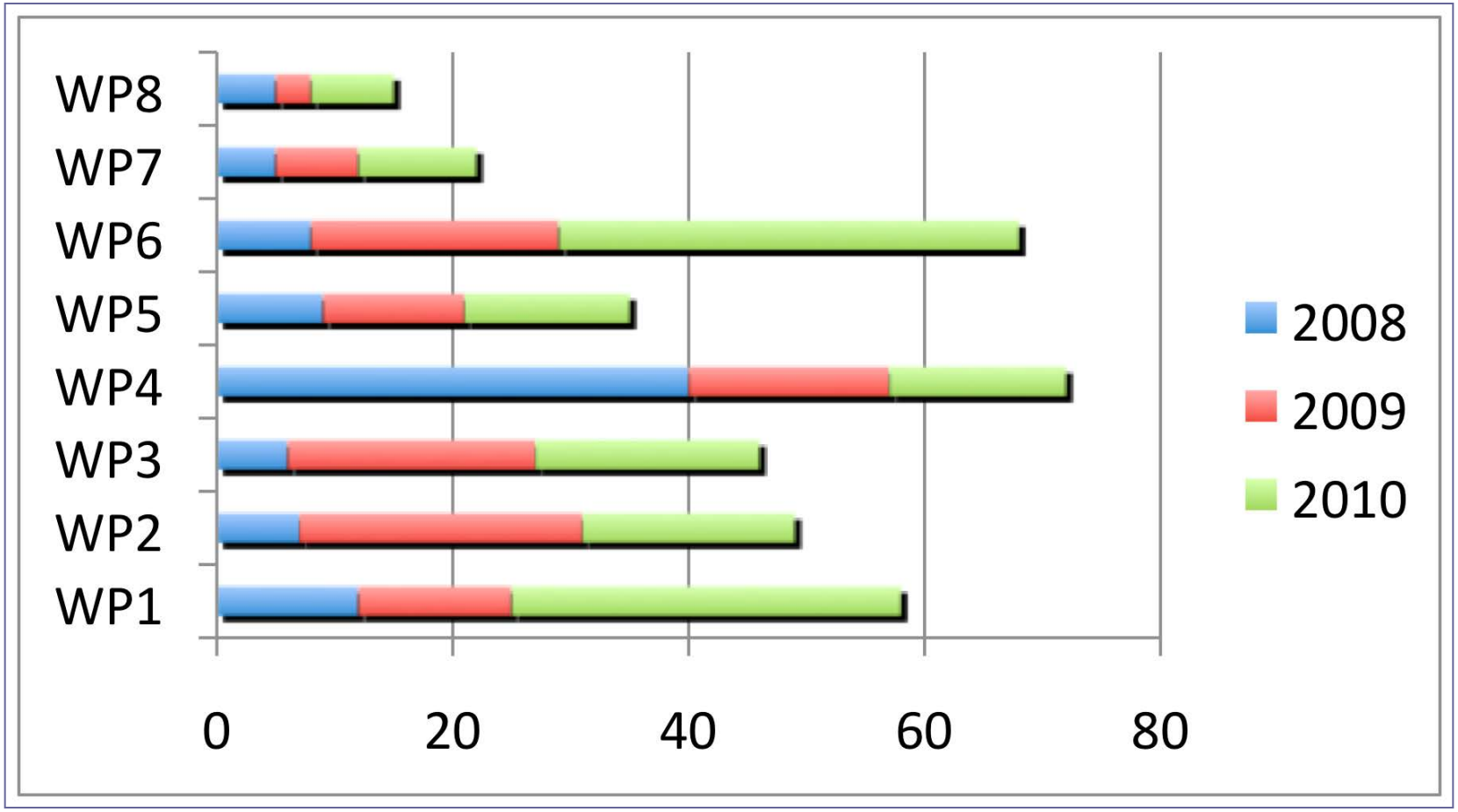
WP8 – Ciencias de los materiales



Mejora de la escalabilidad de SIESTA hasta mas de 1000 cpus

- Impacto de SyeC
  - Escalabilidad de SIESTA hasta mas de 1000 cores
  - Implementación de un esquema paralelo eficiente para calculo de interacciones Van der Waals mediante DFT
  - Desarrollos de nuevos Eigensolvers
- Algunas publicaciones destacadas (2010)
  - Carlos et al., An efficient implementation of a QM–MM method in SIESTA. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling*
  - Roman et al. Stability, Adsorption, and Diffusion of CH4, CO2, and H2 in Clathrate Hydrates. *Phys. Rev. Letters*
- Colaboracion entre (ICMAB/CIN2, UAM) y BSC-CASE
- Internacionalización
  - Organización de 4 Escuelas Internacionales en Ciencias de los Materiales
  - Participación en varios proyectos de la EU

Resultados del proyecto (publicaciones y tesis)



- Tesis doctorales: 20

Resultados del proyecto (difusión y formación)

- RES Users' Training
  - "Programming Models (StarSs)". Barcelona, Septiembre 2010. SyeC participants: BSC-CS
  - "Parallel Computing". Santander, Noviembre 2010. SyeC participants: University of Cantabria, BSC
- PRACE-related events
  - PRACE Training Week. "Tools to analyze the performance of parallel applications (Paraver and Dimemas)". Barcelona, Septiembre 2010. SyeC participants: BSC-CS
  - PRACE Autumn School. "Programming models, performance analysis and debugging tools for Parallel Programmers". Barcelona, Octubre 2010. SyeC participants: BSC-CS, University of Malaga
- Other events
  - "PUMPS Summer School". Barcelona, Julio 2010. SyeC participants: BSC-CASE, BSC-CS.
  - "HiPEAC Computing Systems Week". Barcelona, Octubre 2010. SyeC participants: BSC-CS.
  - "2nd Barcelona Multicore Workshop (BMW-2010)". Octubre 2010. SyeC participants: BSC-CASE, BSC-CS.
  - "Exascale Challenges in Computational Biology", Barcelona, Decembre 2010. SyeC participants: UB, CRG and BSC
  - "Third Coconut meeting", Valencia, September 2010. SyeC participants: UV
  - "Primera Reunión Científica de la Red Española de Gaia (REG)", San Fernando, Junio 2010. SyeC participants: UB
  - Halos Going MAD, Am International Workshop on Halo Finders. Madrid, May 2010. SyeC participants: UAM
  - Annual Meeting of the CLUES Project. Madrid, May-June 2010. SyeC participants: UAM
  - Metal Enrichment from Hydrodynamical Simulations. El Escorial, Madrid. September 2010. SyeC participants: UAM

Resumen resultados del proyecto

- Científicos
  - Conferencias y revistas: 365
  - Tesis: 20
  - Rendimiento/escalabilidad de las aplicaciones: visto en las transparencias
- Colaboración
  - Publicaciones conjuntas entre CS y e-Science: 9
  - Reunión anual plenaria
  - Estancias: 16
  - Organización de eventos: más de 20
- Proyección internacional
  - Propuestas aprobadas: DEISA (2), PRACE (1), FP7 (6), participación en consorcios internacionales: IESP/EESI y 8 más vistas en las transparencias
- Transferencia de tecnología
  - Patentes/spin-offs: 1/1
  - Convenios con empresas: IBM, NVIDIA, Intel
- Visibilidad
  - Joint master: dentro del master tecnológico de la FIB en la UPC
  - Otras actividades vistas anteriormente

Actividades futuras

- Las tareas se desarrollan de acuerdo al plan de trabajo
  - Investigadores: 136 (2007) → 193 (2010)
  - Doctores: 127 (2007) → 151 (2010)
  - Nuevos doctorandos (18) y técnicos titulados superiores (15)
- Continuar las colaboraciones entre grupos de Ciencias de los Computadores y e-Ciencia
  - Potenciar publicaciones conjuntas
  - Participación conjunta en proyectos EU (p.e. ScalaLIFE)
- Nuevas colaboraciones e ideas después de tres años
  - Modelo de programación StarSs (pDEVA, RATPENAT, …)
  - Posibles paralelizaciones híbridas CPU/GPU
  - PRACE: escalabilidad usuarios españoles
  - Proyecto Exascale Mont-blanc

¡Gracias por su atención!