

# Tecnologías GRID

Grids en el mundo científico



**Área de Arquitectura y Tecnología de Computadores**  
*Departamento de Informática de la Universidad de Oviedo*

La tecnología Grid se ha creado y continua desarrollándose mayoritariamente en el **mundo científico**

- La tecnología Grid es un instrumento fundamental de la **e-Ciencia**
- Definición de **e-Ciencia** (“enhanced Science” en inglés)

Conjunto de actividades científicas desarrolladas mediante el uso de recursos distribuidos accesibles a través de Internet

Fuente: Libro blanco de la e-Ciencia en España [FECYT 2004]

Para promover la e-Ciencia en España se ha creado . . .

**Red Española de e-Ciencia** <http://www.e-ciencia.es/>

Grid y Supercomputación al Servicio de la Comunidad Científica Española

# Evolución del Grid hacia la e-Ciencia (1)

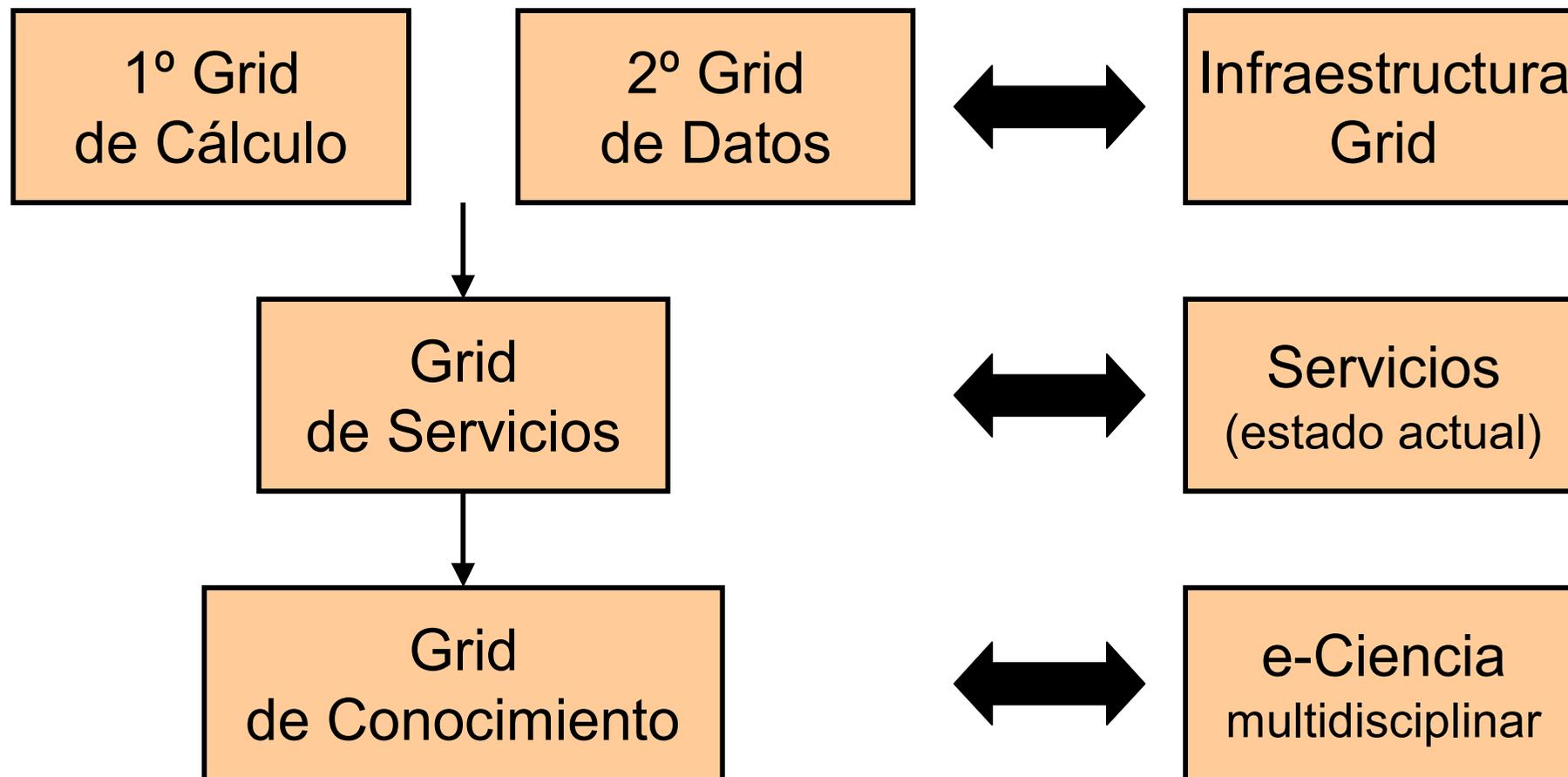


Principales recursos utilizados  
y compartidos mediante la red

Cálculo  
Almacenamiento  
Información

- La e-Ciencia utiliza la tecnología Grid para simplificar el acceso a los recursos y utilizarlos mas eficazmente
- Otras tecnologías también permiten el acceso a recursos distantes (telnet, ftp, P2P, etc)

## Evolución del Grid hacia la e-Ciencia (2)

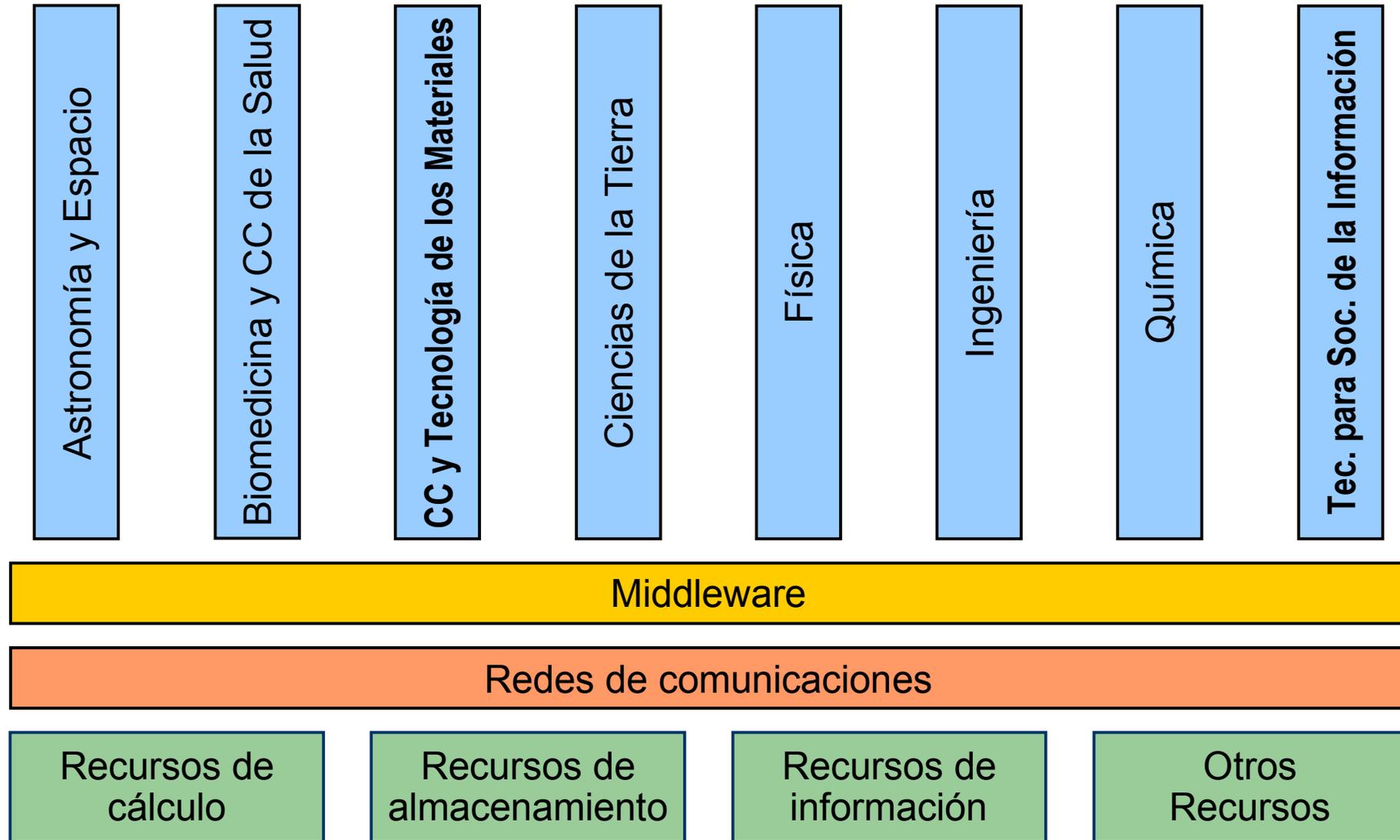


## ¿Qué proporciona el Grid a los científicos?

La infraestructura de e-Ciencia proporciona al investigador  
Recursos (capacidad de cálculo, información, servicios, ...)  
Desde una interfaz única con las siguientes características:

- ▶ **Sencillez:** Uso intuitivo como el de un navegador Web
- ▶ **Transparencia:** No es necesario conocer donde ni como se proporciona el servicio
- ▶ **Rapidez:** El tiempo de respuesta debe ser aceptable para el servicio solicitado
- ▶ **Seguridad:** Permite controlar el acceso a los recursos y a la información
- ▶ **Fiabilidad:** Debe garantizar los resultados obtenidos
- ▶ **Disponibilidad:** Debe estar disponible siempre que se necesite (como la luz, el agua)
- ▶ **Economía:** Coste menor que poseer infraestructura propia

# Taxonomía de la e-Ciencia: Aplicaciones



# Astronomía y Espacio (1)

## Áreas de aplicación:

Astrofísica teórica y cosmología, Astronomía y astrofísica observacional  
Exploración espacial del sistema solar

## Comunidad:

### **Universidades + Centros de excelencia:**

Instituto de Astrofísica de Canarias (IAC) <http://www.iac.es/>

Instituto de Astrofísica de Andalucía (IAA) <http://www.iaa.es/>

Instituto de Ciencias del Espacio (ICE) <http://www.ice.csic.es/>

Instituto de Física de Cantabria (IFCA) <http://www.ifca.unican.es/>

Instituto Nacional de Técnica Aeroespacial (INTA) <http://www.inta.es/>

Laboratorio de Astrofísica Espacial y Física Fundamental (LAEFF) <http://www.laeff.inta.es/>

Centro de Astrobiología <http://cab.inta.es/>

### **Instalaciones de seguimiento espacial**

Centro Espacial de Canarias (Maspalomas)

Estación Espacial de Robledo (INTA – NASA) <http://www.mdsc.org/>

Estación de Seguimiento de Villafranca (INTA – ESA)

# Astronomía y Espacio (2)

El uso de Internet tiene tradición en Astronomía → FITS (*Flexible Image Transport System*)

Formato de datos estándar usado en Astronomía

## Necesidades:

<http://fits.gsfc.nasa.gov/>

Operación de recursos observacionales

Acceso a archivos de datos obtenidos por observatorios

Desarrollo de modelos y su simulación

## Retos Tecnológicos:

Manejo de (múltiples) telescopios robotizados de forma coordinada

Acceso unificado a servicios de observatorios y archivos de datos → **Observatorio Virtual**

OV = Integra datos de múltiples repositorios en un hipercubo; el usuario extrae cortes

OV = Define el concepto de observación como un servicio Grid

Elaboración de modelos y simulación de escenarios → Evolución de Galaxias o el Universo

La nueva astronomía → Análisis masivo de datos + Integración de modelos y datos



# Biomedicina y Ciencias de la Salud (1)

## Áreas de aplicación:

Todas las disciplinas que estudian: Procesos biológicos, Medicina clínica, Salud pública  
Hay una realimentación continua entre clínica y laboratorio

## Comunidad:

Instituto de Salud Carlos III <http://www.isciii.es/>

CSIC ( 26 Centros de Investigación ) <http://www.csic.es/centros.do>

Instituto Nacional de Bioinformática ( 9 Centros ) <http://www.inab.org/>

## Múltiples aplicaciones ya se benefician de la e-Ciencia

Celware: Modelado y simulación de interacciones celulares

<http://www.bii.a-star.edu.sg/achievements/applications/cellware/index.php>

NeuroGrid: Análisis de imágenes (del cerebro)

<http://www.gridbus.org/neurogrid/>

Planificación y soporte en tiempo real de procesos quirúrgicos (GEMSS)

<http://www.it.neclab.eu/gemss/>

<http://www.eu-crossgrid.org/application.htm>

# Biomedicina y Ciencias de la Salud (2)

Investigación sobre modelos animales de enfermedades humanas (BIRN)

<http://www.nbirn.net/> (Biomedical Informatics Research Network)

Estudios a gran escala de patologías complejas integrando:  
Datos genéticos, imágenes, historias clínicas y factores ambientales

CLEF: Clinical e-Science Framework <http://www.itri.brighton.ac.uk/projects/clef/index.html>

INFOGENMED: Integrating Information Genetic & Medical <http://infoгенmed.ieeta.pt/>

## Retos Científicos y Tecnológicos:

Abordar grandes desafíos → {  
Plegamiento de proteínas  
Modelos de células y enfermedades  
Evolución de epidemias

Promover el uso de Grids en la comunidad científica biomédica  
(Favorecer la interacción entre comunidades tradicionalmente aisladas entre sí)

Seguridad (Preservación de datos personales, Conformidad con LOPD, ... )

Gestión de metadatos y desarrollo de ontologías comunes en biomedicina

Hacer entornos más productivos para almacenar, procesar y visualizar datos biomédicos

# Ciencia y Tecnología de Materiales (1)

## Áreas de aplicación:

Caracterización de materiales y cristalografía, Diseño y síntesis de nuevos materiales

## Comunidad: **CSIC / Universidades**

Instituto de Ciencia de Materiales de Madrid (ICMM) <http://www.icmm.csic.es/>

Instituto de Ciencia de Materiales de Barcelona (ICMAB) <http://www.icmab.es/>

Instituto de Ciencia de Materiales de Aragón (ICMA) <http://icma.csic.unizar.es/>

Instituto de Ciencia de Materiales de Sevilla (ICMSE) <http://www.icmse.cartuja.csic.es/>

Instituto Tecnológico de Materiales de Asturias (ITMA) <http://www.itma.es/>

Centro Nacional de Investigaciones Metalúrgicas (CENIM) <http://www.cenim.csic.es/>

Instituto de Cerámica y Vidrio (ICV) <http://www.icv.csic.es/>

Instituto Eduardo Torroja de Ciencias de la Construcción (IETCC) <http://www.ietcc.csic.es/>

Instituto de Ciencia y Tecnología de Polímeros (ICTP) <http://www.ictp.csic.es/>

Centro de física de materiales <http://cfm.ehu.es/>

---

Centro Nacional de Aceleradores de Sevilla <http://centro.us.es/cna/>

Sincrotrón del Vallés (Futuro inmediato)



# Ciencia y Tecnología de Materiales (2)

## Necesidades:

Simulación del comportamiento de materiales

{ Nivel macroscópico (estructural, térmico)  
Nivel microscópico (mecánica estadística  
Dinámica molecular)

## Retos Científicos:

Análisis de la reactividad química en superficies: catálisis

Diseño de materiales con propiedades predefinidas

Materiales para la degradación de residuos y productos tóxicos

Materiales para la nanoelectrónica

Materiales híbridos - Biomateriales

Almacenamiento de energía: pilas de combustible

## Retos Tecnológicos:

Interoperabilidad de programas de simulación

Compatibilidad de información entre programas de simulación

Definición de estándares para el intercambio de datos (metadatos)

Implementación de utilidades de visualización remota de datos de simulación

Seguridad (autenticación, acceso a datos sensibles, etc.)

# Ciencias de la Tierra (1)

## Áreas de Aplicación:

Meteorología: Predicción del tiempo a corto plazo, predicciones estacionales, ...

El sistema tierra: Simulación de la biosfera, Simulación de cambios climáticos

Oceanografía y Modelado de aguas superficiales y subterráneas

Estudios medioambientales → Ej.: Simulación de difusión de contaminantes

## Comunidad de Usuarios:

### **Universidades + Muchos organismos públicos de investigación**

CEDEX: Centro de Estudios y Experimentación de Obras Públicas <http://www.cedex.es/>

CSIC: Consejo Superior de Investigaciones Científicas <http://www.csic.es/>

CIEMAT: Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas <http://www.ciemat.es/>

IEO: Instituto Español de Oceanografía <http://www.ieo.es/>

AEMET: Agencia Estatal de Meteorología <http://www.aemet.es/>

INTA: Instituto Nacional de Técnica Aeroespacial <http://www.inta.es/>

IJA: Instituto Jaime Almera de Ciencias de la Tierra <http://www.ija.csic.es/>

## Ciencias de la Tierra (2)

### Necesidades:

Operación y soporte de grandes recursos observacionales

Almacenamiento de resultados de experimentos globales, regionales y locales

Desarrollo y adaptación de modelos de predicción y de simulación de escenarios futuros

Visualización de datos

### Retos Científicos y Tecnológicos:

Modelización de procesos e interacciones del sistema climático  
(con escalas de alta resolución [ $<1$  Km] )

Modelización y simulación cada vez mas detallada de los entornos terrestres

Desarrollo de un entorno colaborativo (internacional) para el desarrollo y la comparación de resultados de simulación

Desarrollo consolidación y acceso a bases de datos experimentales virtuales (distribuidas)

# Física (1)

## Áreas de Aplicación:

Física de altas energías, Fusión nuclear

Otras áreas: Sistemas complejos, Física estadística, Astrofísica, Física-Química, Fluidos

## Comunidad:

### **Física de altas energías (ó de partículas)**

Universidades + CSIC + CIEMAT + Institutos <http://www.fpa.csic.es/>

Instituto de Estructura de la Materia (IEM CSIC) <http://www.iem.csic.es/>

Institut de Física d'Altes Energies (IFAE) <http://www.ifae.es/>

Instituto Galego de Física de Altas Enerxías (IGFAE) <http://igfae.usc.es/>

Instituto de Física de Cantabria (IFCA) CSIC <http://www.ifca.unican.es/>

**Fusión nuclear** Laboratorio Nacional de Fusión (CIEMAT) <http://www-fusion.ciemat.es/>

### **Otras áreas**

Instituto de Física Interdisciplinar y Sistemas Complejos (IFISC) <http://ifisc.uib.es/>

Centro de Física Miguel Antonio Catalán (CFMAC) <http://www.cfmac.csic.es/>

Instituto de Biocomputación y Física de Sistemas Complejos (BIFI) <http://bifi.unizar.es/>



# Física (2)

## Comunidad Internacional:

CERN (Conseil Européen pour la Recherche Nucléaire) European Organization for Nuclear Research

Desarrolla el LHC (Large Hadron Collider)

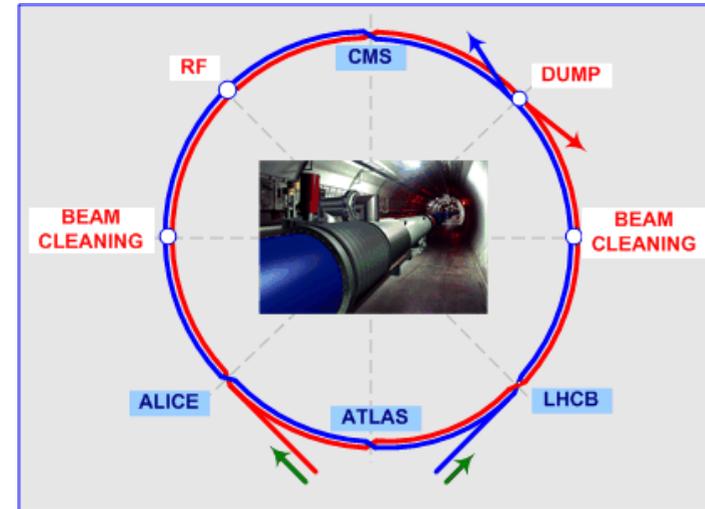
<http://lhc.web.cern.ch/> Hadrón = Partícula subatómica

Es mayor acelerador de partículas del mundo  
(circunferencia de 27 Km enterrada de 50 a 175 m)

Genera una cantidad de datos enorme:  
27 TBytes/día + 10 TB de resumen

Procesamiento de  
datos → Grid

- Fila 0: En el propio CERN
- Fila 1: Conjunto de 11 instituciones académicas en todo el mundo
- Fila 2: Otras 150 instituciones

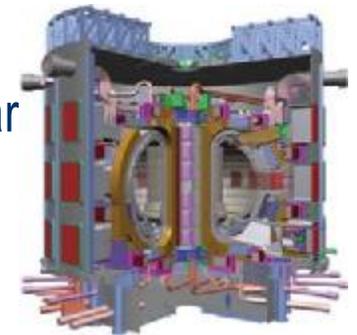


ITER International Thermonuclear Experimental Reactor <http://www.iter.org/>

Reactor demuestra la viabilidad científica y tecnológica de la fusión nuclear

Ubicado en Cadarache (Sur Francia) Duración: 2007-2016

Participantes: UE (Euratom), USA, Japón, Rusia, China, India, Corea



## Necesidades:

Adquisición, almacenamiento y acceso a grandes volúmenes de datos

Física de partículas → Detectores del LHC      Energía → Experimentos con el ITER

Elevadas necesidades de simulación

Simulación de colisiones de partículas      Modelado de confinamiento magnético

Análisis, calibración y visualización de datos

Algoritmos complejos sobre terabytes de información → Búsqueda del Bosón de Higgs  
(Seleccionar 100 colisiones de entre 1 billón)

## Retos Científicos y Tecnológicos:

Gestionar recursos distribuidos

Para satisfacer las necesidades de colaboración → Organizar la interoperabilidad con seguridad

Acceder y procesar elevados volúmenes de datos

Desarrollar la Virtualización sobre commodities      Agregación de recursos nacionales

Desarrollar potentes algoritmos de análisis con respuesta rápida

Simulación de procesos en tiempos limitados → Optimizar infraestructura y Acceso concurrente

Lograr entornos de colaboración y participación remota

Procedimientos comunes y coordinados de instalación y operación de infraestructuras

# Ingeniería (1)

## Áreas de aplicación:

Procesos físicos con aplicación industrial, Simulación de procesos industriales  
Diseño de productos y procesos

## Comunidad:

### Universidades + Centros de investigación públicos

Instituto de Robótica i Informática Industrial (IRI) <http://www-iri.upc.es/>

Laboratorio de Investigación en Tecnologías de la Combustión (LITEC) <http://www.litec.csic.es/>

Instituto de Automática Industrial (IAI) <http://www.iai.csic.es/>

## Características Específicas:

Fenómenos complejos (no se pueden simplificar demasiado)

Simulaciones de muchos millones de puntos y fuertemente acopladas → Mejor Supercomputador que Grid  
Resultados individuales grandes y con relativamente pocos ficheros a compartir

Fuerte integración internacional Grandes industrias en sectores concretos → OrgV establecidas

Cooperación ciencia/industria con fuerte confidencialidad de algunos datos

# Ingeniería (2)

## Retos y Necesidades:

Los grandes retos con comunes en todo el mundo

Ej.: Procesos industriales básicos de turbulencia y combustión

Problema	Puntos de malla	Año previsto
Turbulencia (problema físico)	$10^{12}$	2020
Turbulencia gran escala (problema industrial)	$10^{13}$	2025
Combustión (problema físico)	$10^{15}$	2030

## Lograr una cooperación internacional mas fácil

Mas difícil en el campo industrial por la confidencialidad de las investigaciones

Las redes abiertas de cooperación son mas escasas

## Acceso necesario a grandes instalaciones

Ej: Jet Propulsion Laboratory <http://www.jpl.nasa.gov/>

Los bancos de datos existentes suponen una inversión de años → Uso mediante discos físicos

# Química (1)

## Áreas de aplicación:

Estudio de la reactividad, cinética y dinámica de reacciones, Química atmosférica  
Investigación en nanopartículas, cristales moleculares, superconductores, etc.

## Comunidad:

### Universidades + Institutos públicos (CSIC)

Centro de Química Orgánica Manuel Lora Tamayo <http://www.cenquior.csic.es/>

Instituto de Ciencia y Tecnología de Polímeros (ICTP) <http://www.ictp.csic.es/>

Instituto de Fermentaciones Industriales (IFI) <http://www.ifi.csic.es/>

Instituto de Química Médica (IQM) <http://www.iqm.csic.es/>

Instituto de Química Orgánica General (IQOG) <http://www.iqog.csic.es/>

Instituto de Carboquímica (ICB) <http://www.icb.csic.es/>

Instituto de Catálisis y Petroleoquímica (ICP) <http://www.icp.csic.es/>

Instituto de Tecnología Química (ITQ) <http://itq.webs.upv.es/>

Instituto Nacional del Carbón (INCAR) <http://www.incar.csic.es/>

Instituto de Investigaciones Químicas (IIQ) <http://www.iiq.cartuja.csic.es/>

## Necesidades y Retos Científicos:

- Reducción de los subproductos generados en los procesos químicos industriales
- Mejora de la calidad del aire urbano
- Desarrollo de nuevos materiales de alto valor estratégico y tecnológico
- Motores moleculares (Proteínas que transforman energía química en mecánica)
- Reconocimiento molecular (Interacciones moleculares y supra-moleculares)

## Retos Tecnológicos:

- Acceso a software estándar de química computacional
- Optimización y paralelización masiva de los códigos mas usuales
- Acceso rápido a bases de datos químicos (cristalográficos y de otros tipos)
- Elaboración de modelos de sistemas químicos complejos
- Simulaciones basadas en un gran número de datos
- Herramientas de visualización y análisis de resultados

# Tecnologías para la Sociedad de la Información (1)

Para las TSI la e-Ciencia es más un fin que un medio

La e-Ciencia intenta resolver problemas complejos con gran impacto económico y social

Para las TSI el objetivo básico es mejorar la infraestructura de e-Ciencia

## Comunidad:

Mas de 15 grupos de investigación en áreas de

{  
Sistemas Distribuidos  
Programación Concurrente  
Supercomputación, ...

## Áreas y Retos a Afrontar:

**Arquitecturas Grid** Extender la arquitectura OGSA y su implementación  
Mejorar el soporte para Grid en los SOs

**Middleware** Se necesita un marco general para una autoridad de registro / autenticación  
Mejorar la calidad del middleware desarrollado en proyectos de investigación

**Gestión de datos** Manejar con mas efectividad grandes volúmenes de datos  
Soportar mejor el manejo de datos heterogéneos

# Tecnologías para la Sociedad de la Información (2)

**Desarrollo de software** Nuevos modelos, lenguajes y entornos para Grids  
Paralelización y distribución automática de tareas

**Gestión del Conocimiento** WorkFlow y colaboración basados en el conocimiento  
Soporte para organización virtuales  
Web y Grid Semánticos

## Desafíos Científico-Tecnológicos:

Hacer el Grid mas invisible (transparente)

Pasar de: Grid Computacional → Grid del Conocimiento

Hacer compatible privacidad con transparencia

Grids para sistemas móviles y empotrados

Pasar del “plug & play” al “connect & share”

Incremento del ancho de banda de la red y optimizar su gestión