

Tecnologías Grid

Aplicaciones

Master en Sistemas y Servicios Informáticos para Internet
Área de Arquitectura y Tecnología de Computadores
Universidad de Oviedo



Aplicaciones

Aplicaciones en el mundo científico

Aplicaciones en el mundo científico

- Los cálculos científicos requieren gran cantidad de computación y datos
 - ▣ La tecnología grid es perfecta para esto
- e-Ciencia (*enhanced Science*)

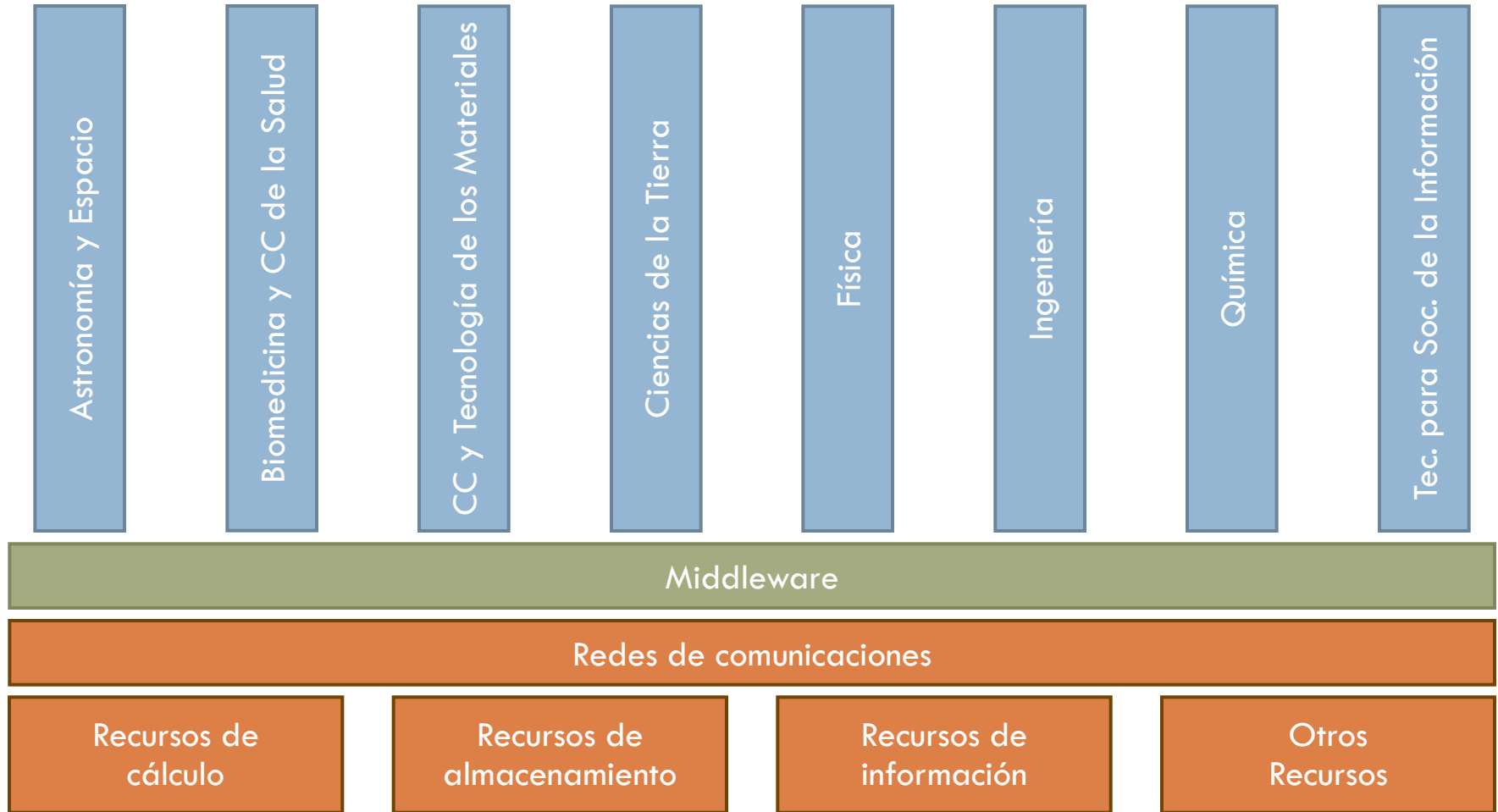
Libro blanco de la e-Ciencia en España

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

- Red Española de e-Ciencia
 - ▣ Grids y supercomputación al servicio de la comunidad científica española
 - ▣ <http://www.e-ciencia.es/>

Aplicaciones en el mundo científico

□ Taxonomía



Aplicaciones en el mundo científico

- Astronomía y espacio (I)
 - ▣ Áreas de aplicación
 - Astrofísica teórica y cosmología
 - Astronomía y astrofísica observacional
 - Exploración espacial del sistema solar
 - ▣ Comunidades en España
 - Instituto de Astrofísica de Canarias (IAC)
 - Instituto de Astrofísica de Andalucía (IAA)
 - Instituto de Ciencias del Espacio (ICE)
 - Instituto de Física de Cantabria (IFCA)
 - Instituto Nacional de Técnica Aeroespacial (INTA)
 - Laboratorio de Astrofísica Espacial y Física Fundamental (LAEFF)
 - Centro de Astrobiología
 - Centro Espacial de Canarias (Maspalomas)
 - Estación Espacial de Robledo (INTA – NASA)
 - Estación de Seguimiento de Villafranca (INTA – ESA)

Aplicaciones en el mundo científico

- Astronomía y espacio (II)
 - ▣ Necesidades
 - Operación de recursos observacionales
 - Acceso a archivos de datos obtenidos por observatorios
 - Desarrollo de modelos y su simulación
 - ▣ Retos científicos y tecnológicos
 - Manejo de múltiples telescopios de forma coordinada
 - Observatorios virtuales
 - Acceso unificado a servicios de observatorios y archivos de datos
 - Elaboración de modelos y simulación de escenarios
 - Evolución de galaxias o el universo
 - Nueva astronomía
 - Análisis masivo de datos + integración de modelos y datos

Aplicaciones en el mundo científico

- **Biomedicina y Ciencias de la Salud (I)**
 - ▣ **Áreas de aplicación**
 - Procesos biológicos
 - Medicina clínica
 - Salud pública
 - ▣ **Comunidades en España**
 - Instituto de Salud Carlos III
 - CSIC (26 Centros de Investigación)
 - Instituto Nacional de Bioinformática (9 Centros)
 - ▣ **Aplicaciones ya en uso**
 - Celware: Modelado y simulación de interacciones celulares
 - NeuroGrid: Análisis de imágenes (del cerebro)
 - Planificación y soporte en tiempo real de procesos quirúrgicos (GEMSS)
 - Investigación sobre modelos animales de enfermedades humanas (BIRN)
 - CLEF: Clinical e-Science Framework
 - INFOGENMED: Integrating Information Genetic & Medical

Aplicaciones en el mundo científico

- Biomedicina y Ciencias de la Salud (II)
 - ▣ 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
 - Elaboración de modelos y simulación de escenarios
 - 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

Aplicaciones en el mundo científico

- Ciencia y Tecnología de Materiales (I)
 - ▣ Áreas de aplicación
 - Caracterización de materiales y cristalografía
 - ▣ Comunidades en España
 - Instituto de Ciencia de Materiales de Madrid (ICMM)
 - Instituto de Ciencia de Materiales de Barcelona (ICMAB)
 - Instituto de Ciencia de Materiales de Aragón (ICMA)
 - Instituto de Ciencia de Materiales de Sevilla (ICMSE)
 - Instituto Tecnológico de Materiales de Asturias (ITMA)
 - Centro Nacional de Investigaciones Metalúrgicas (CENIM)
 - Instituto de Cerámica y Vidrio (ICV)
 - Instituto Eduardo Torroja de Ciencias de la Construcción (IETCC)
 - Instituto de Ciencia y Tecnología de Polímeros (ICTP)
 - Centro de física de materiales
 - Centro Nacional de Aceleradores de Sevilla
 - Sincrotrón del Vallés

Aplicaciones en el mundo científico

- Ciencia y Tecnología de Materiales (II)
 - ▣ Necesidades
 - Simulación del comportamiento de materiales
 - Nivel macroscópico (estructural, térmico)
 - Nivel microscópico (mecánica estadística, dinámica molecular)
 - ▣ Retos científicos y tecnológicos
 - 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
 - Interoperabilidad de programas de simulación
 - Compatibilidad de información entre programas de simulación
 - Definición de estándares para el intercambio de datos (metadatos)

Aplicaciones en el mundo científico

□ Ciencias de la Tierra (I)

▣ 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

▣ Comunidades en España

- CEDEX: Centro de Estudios y Experimentación de Obras Públicas
- CSIC: Consejo Superior de Investigaciones Científicas
- CIEMAT: Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas
- IEO: Instituto Español de Oceanografía
- AEMET: Agencia Estatal de Meteorología
- INTA: Instituto Nacional de Técnica Aeroespacial
- IJA: Instituto Jaume Almera de Ciencias de la Tierra

Aplicaciones en el mundo científico

□ Ciencias de la Tierra (II)

▣ 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
- Modelización y simulación más 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)

Aplicaciones en el mundo científico

- Física (I)
 - ▣ Áreas de aplicación
 - Física de altas energías, Fusión nuclear
 - Sistemas complejos, Física estadística, Astrofísica, Física-Química, Fluidos
 - ▣ Comunidades en España
 - Universidades + CSIC + CIEMAT + Institutos
 - Instituto de Estructura de la Materia (IEM CSIC)
 - Institut de Física d'Altes Energies (IFAE)
 - Instituto Galego de Física de Altas Enerxías (IGFAE)
 - Instituto de Física de Cantabria (IFCA) CSIC
 - Laboratorio Nacional de Fusión (CIEMAT)
 - Instituto de Física Interdisciplinar y Sistemas Complejos (IFISC)
 - Centro de Física Miguel Antonio Catalán (CFMAC)
 - Instituto de Biocomputación y Física de Sistemas Complejos (BIFI)

Aplicaciones en el mundo científico

□ Física (II)

▣ Necesidades

- Adquisición, almacenamiento y acceso a grandes volúmenes de datos
- Elevadas necesidades de simulación
- Análisis, calibración y visualización de datos

▣ 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 potentes algoritmos de análisis con respuesta rápida
- Lograr entornos de colaboración y participación remota
 - Procedimientos comunes y coordinados de instalación y operación de infraestructuras

Aplicaciones en el mundo científico

- Química (I)
 - ▣ Áreas de aplicación
 - Estudio de la reactividad, cinética y dinámica de reacciones
 - Química atmosférica
 - Investigación en nanopartículas, cristales moleculares, superconductores, etc.
 - ▣ Comunidades en España
 - Centro de Química Orgánica Manuel Lora Tamayo
 - Instituto de Carboquímica (ICB)
 - Instituto de Catálisis y Petroleoquímica (ICP)
 - Instituto de Tecnología Química (ITQ)
 - Instituto Nacional del Carbón (INCAR)
 - Instituto de Investigaciones Químicas (IIQ)

Aplicaciones en el mundo científico

□ Química (II)

▣ Necesidades

- 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
- Reconocimiento molecular

▣ Retos científicos y tecnológicos

- Acceso a software estándar de química computacional
- Optimización y paralelización masiva de los códigos más usuales
- Acceso rápido a bases de datos químicos
- Elaboración de modelos de sistemas químicos complejos
- Simulaciones basadas en un gran número de datos

Aplicaciones en el mundo científico

- Ingeniería (I)
 - Áreas de aplicación
 - Procesos físicos con aplicación industrial
 - Simulación de procesos industriales
 - Diseño de productos y procesos
 - Comunidades en España
 - Instituto de Robótica i Informática Industrial (IRI)
 - Laboratorio de Investigación en Tecnologías de la Combustión (LITEC)
 - Instituto de Automática Industrial (IAI)
 - Características específicas
 - Fenómenos complejos (no se pueden simplificar demasiado)
 - Simulaciones de muchos millones de puntos y fuertemente acopladas
 - Resultados individuales grandes y con relativamente pocos ficheros a compartir
 - Fuerte integración internacional: grandes industrias en sectores concretos
 - Cooperación ciencia/industria con fuerte confidencialidad de algunos datos

Aplicaciones en el mundo científico

- Ingeniería (II)
 - ▣ Retos científicos y tecnológicos
 - Lograr una cooperación internacional más fácil
 - Más difícil en el campo industrial por la confidencialidad de las investigaciones
 - Las redes abiertas de cooperación son más escasas
 - Acceso necesario a grandes instalaciones
 - Ej: Jet Propulsion Laboratory

Aplicaciones en el mundo científico

- Tecnologías para la Sociedad de la Información
 - ▣ Áreas de aplicación
 - Más que aplicar, las TSI desarrollan la infraestructura de e-Ciencia
 - ▣ Comunidades en España
 - Más de 15 grupos de investigación en
 - Sistemas distribuidos
 - Programación concurrente
 - Supercomputación
 - ...
 - ▣ Desafíos científicos-tecnológicos
 - Hacer el grid más invisible (transparente)
 - Pasar de grid computacional a grid de conocimiento
 - Compartir no sólo infraestructura, sino también conocimiento y experiencia
 - Hacer compatible privacidad con transparencia
 - Grids para sistemas móviles y empotrados
 - Incremento del ancho de banda de la red y optimizar su gestión

Aplicaciones

Grandes grid científicos en producción

Grandes grid en producción

- Hubo una tendencia a desarrollar grids nacionales
 - National Grid Service, NGS (Reino Unido)
 - D-Grid (Alemania)
 - IRISGrid (España)
 - INFN (Italia)
 - HellasGrid (Grecia)
 - DutchGrid (Holanda)
 - BalticGrid (Estonia, Letonia y Lituania)
- Tendencia en Europa: integrarse en los grids de dimensión europea
 - European Grid Initiative

Grandes grid en producción



- Grids en Europa (I)
 - EGEE (Enabling Grids for E-science)
 - Serie de proyectos financiados por la Unión Europea
 - EGEE-I (marzo 2004 - marzo 2006)
 - EGEE-II (abril 2006 - abril 2008)
 - EGEE-III (mayo 2008 - abril 2010)
 - Generó el software gLite
 - Infraestructura
 - 250 centros de todo el mundo
 - >40.000 nodos de computación, >10 PBytes, >200 organizaciones virtuales

Grandes grid en producción

- Grids en Europa (II)
 - ▣ Large Hadron Collider Computing Grid (LHC-CG o LCG)
 - Diseñado por el CERN para el acelerador LHC
 - 140 centros de computación, 33 países
 - Los computadores usan Scientific Linux
 - Arquitectura en Capas
 - Usa Globus, Condor, gLite
 - Arquitectura en tres capas
 - Tier 0: en el CERN
 - Tier 1: 11 centros otros países
 - Tier 2: 140 centros de investigación



Grandes grid en producción



- Grids en Europa (III)
 - ▣ Nordic DataGrid Facility (NDGF)
 - Comenzado en 2006
 - Participan Dinamarca, Finlandia, Noruega y Suiza
 - Relacionado con NorduGrid
 - Proyecto para experimentos en el CERN
 - Desarrolladores del middleware ARC
 - La mayor parte de los recursos accesibles con ARC
 - Algunos con gLite
 - Es parte del Tier 1 del LHC
 - Utiliza varias versiones de Linux

Grandes grid en producción



□ Grids en USA (I)

▣ Open Science Grid

- Financiado por NSF (National Science Foundation) y el DOE (Department of Energy)
- Iniciado en 2004
- Integra proveedores de recursos y servicios, investigadores de universidades y laboratorios nacionales y centros de computación
- Área principal: física de partículas
 - Creado para facilitar el análisis de los datos del LHC (70% de dedicación)
- Otras: química, biología, astronomía
- >25000 computadores, >70 instituciones
- Linux, Condor, Globus, PBS, LSF, Sun Grid Engine
- Interopera con otros grids: EGEE, LHC Computing Grid, TeraGrid

Grandes grid en producción



□ Grids en USA (II)

□ TeraGrid

- Financiado por NSF (National Science Foundation)
- Iniciado en 2001
- Integra once entidades
 - Laboratorios, universidades, centros de computación
- Aplicaciones
 - Astronomía, ciencias atmosféricas, bioquímica, química, ciencias de la tierra, cosmología, sismología, genética, geofísica, neurociencia, materiales, visualización
- Más de 4000 usuarios de 200 universidades estadounidenses
- Usa Condor-G, Globus, PBS, SRB...

Grandes grid en producción



- Grids en Japón
 - ▣ NAREGI (*National Research Grid Initiative*)
 - Iniciado en 2003
 - Desarrollado en el National Institute of Informatics
 - Ha creado un middleware propio que sigue OGSA
- Grids en España
 - ▣ IRISGrid (Iniciativa Nacional de Grid)
 - Participa en EGEE-III
 - Despliegue en España de GILDA (red de formación del EGEE-III)
 - Gestión de pkIRISGrid
 - ▣ Otras iniciativas
 - GRID-CSIC
 - Infraestructura Grid de investigación Avanzada del CSIC
 - IberGrid: Iberian Grid Infrastructure Conference



Aplicaciones

Aplicaciones en el mundo empresarial

Aplicaciones en el mundo empresarial

- La tecnología grid ha tenido un gran éxito en la e-Ciencia
 - ▣ Mundo académico
 - ▣ Actividades de investigación
- La tecnología grid se va integrando en la empresa
 - ▣ Pero parece que ha tenido mucho más éxito como cloud
 - Ejemplo
 - Empresa GoGrid
 - Eslogan: Instant Windows and Linux Cloud Servers

Aplicaciones en el mundo empresarial

- Tipos de empresas
 - ▣ Proveedoras de middleware grid y servicios
 - Axceleon, Data Synapse, GridSystems, Parabon, Platform Computing, Univa UD, Digipede, PBS GridWorks, ClusterResources
 - ▣ Vendedoras de hardware, software y servicios
 - HP, IBM, SAS, Oracle, GridwiseTech, EmergenceTech
 - ▣ Centros de datos y hosting
 - AIMES, GoGrid