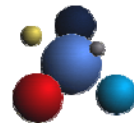


5 de diciembre de 2006
Facultade de Informática
Universidade da Coruña

Una Introducción a la Computación Grid

Ignacio Martín Llorente
asds.dacya.ucm.es/nacho



Grupo de Arquitectura de Sistemas Distribuidos
Departamento de Arquitectura de Computadores y Automática
Universidad Complutense de Madrid

Objetivos de la Presentación

- Describir la tecnología existente para **gestionar recursos distribuidos** dentro de una organización
- Definir **infraestructura grid**
- Describir las **componentes mínimas** para desplegar infraestructuras grid
- Enumerar los **organismos de estandarización** existentes
- Mostrar ejemplos de **infraestructuras grid**
- Describir **modelos de programación** para el grid
- Mostrar **ejemplos de aplicaciones** portadas al grid

Contenidos

1. Plataformas de Computación Paralela y Distribuida
2. Tecnología Grid
3. Infraestructuras Grid
4. Migración de Aplicaciones al Grid

1. Plataformas de Computación Paralela y Distribuida

Contenidos

1. Plataformas de Computación Paralela y Distribuida

1.1. Entornos de Computación

1.2. Sistemas de Gestión de Recursos Distribuidos

2. Tecnología Grid

3. Infraestructuras Grid

4. Migración de Aplicaciones al Grid

1. Plataformas de Computación Paralela y Distribuida

1.1. Entornos de Computación

Plataformas de Computación Paralela y Distribuida

- **Ejecución eficiente** de aplicaciones intensivas en datos o computación

Tipos de Entornos de Computación

Entornos HPC (*High Performance Computing*)

- Su objetivo es reducir el tiempo de ejecución de una única aplicación paralela de memoria compartida o distribuida
- Su rendimiento se mide en número de operaciones en punto flotante por segundo
- Áreas de aplicación: CFD, meteorología...

Entornos HTC (*High Throughput Computing*)

- Su objetivo es aumentar el número de ejecuciones por unidad de tiempo
- Su rendimiento se mide en número de trabajos ejecutados por segundo
- Áreas de aplicación: HEP, bioinformática, finanzas

1. Plataformas de Computación Paralela y Distribuida

1.1. Entornos de Computación

Tipos de Plataformas de Computación

Alternativas Centralizadas



Servidores SMP



Servidores MPP

Alternativas Distribuidas



Clusters dedicados



Clusters no dedicados

Servidores HPC (*High Performance Computing Servers*)

- Arquitecturas de memoria compartida (SMP) o distribuida (MPP)

Perfil de Aplicación

- Ejecución eficiente de aplicaciones HPC y HTC

Ventajas

- Interconexión con ancho de banda alto y latencia baja
- Acceso uniforme al sistema gracias a una única copia del sistema operativo

Inconvenientes

- Baja escalabilidad (para SMPs)
- Modelos complejos de programación (para HPC en MPPs)
- Precio alto



Sistema de Colas Batch
NQE



Clusters Dedicados

- Cluster dedicado y homogéneo de PCs o estaciones interconectados por medio de una red de área de sistema (Giganet, Myrinet...)

Perfil de Aplicación

- Ejecución eficiente de aplicaciones HTC y HPC de grano grueso

Ventajas

- Mejor relación coste/rendimiento para aplicaciones HTC
- Mayor escalabilidad

Inconvenientes

- Requieren modelos de programación de memoria distribuida (librerías de paso de mensajes como MPI) para aplicaciones HPC



Sistema de Gestión de Recursos
PBS



1. Plataformas de Computación Paralela y Distribuida

1.1. Entornos de Computación

Clusters no Dedicados

- Cluster no dedicado y heterogéneo de PCs o estaciones interconectados por medio de una red de área local (Fast ethernet...)

Perfil de Aplicación

- Únicamente ejecuta aplicaciones HTC

Ventajas

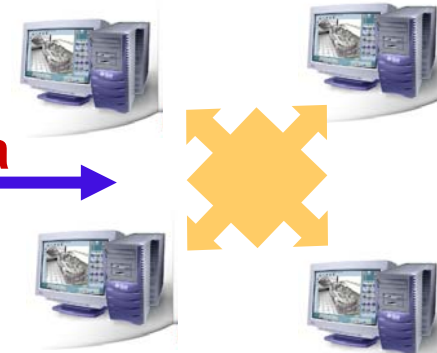
- Mínima relación coste/rendimiento para aplicaciones HTC
- Mayor escalabilidad

Inconvenientes

- Interconexión con ancho de banda bajo y latencia alta
- Requiere capacidades de gestión adaptativa para usar los tiempos ociosos de los recursos dinámicos



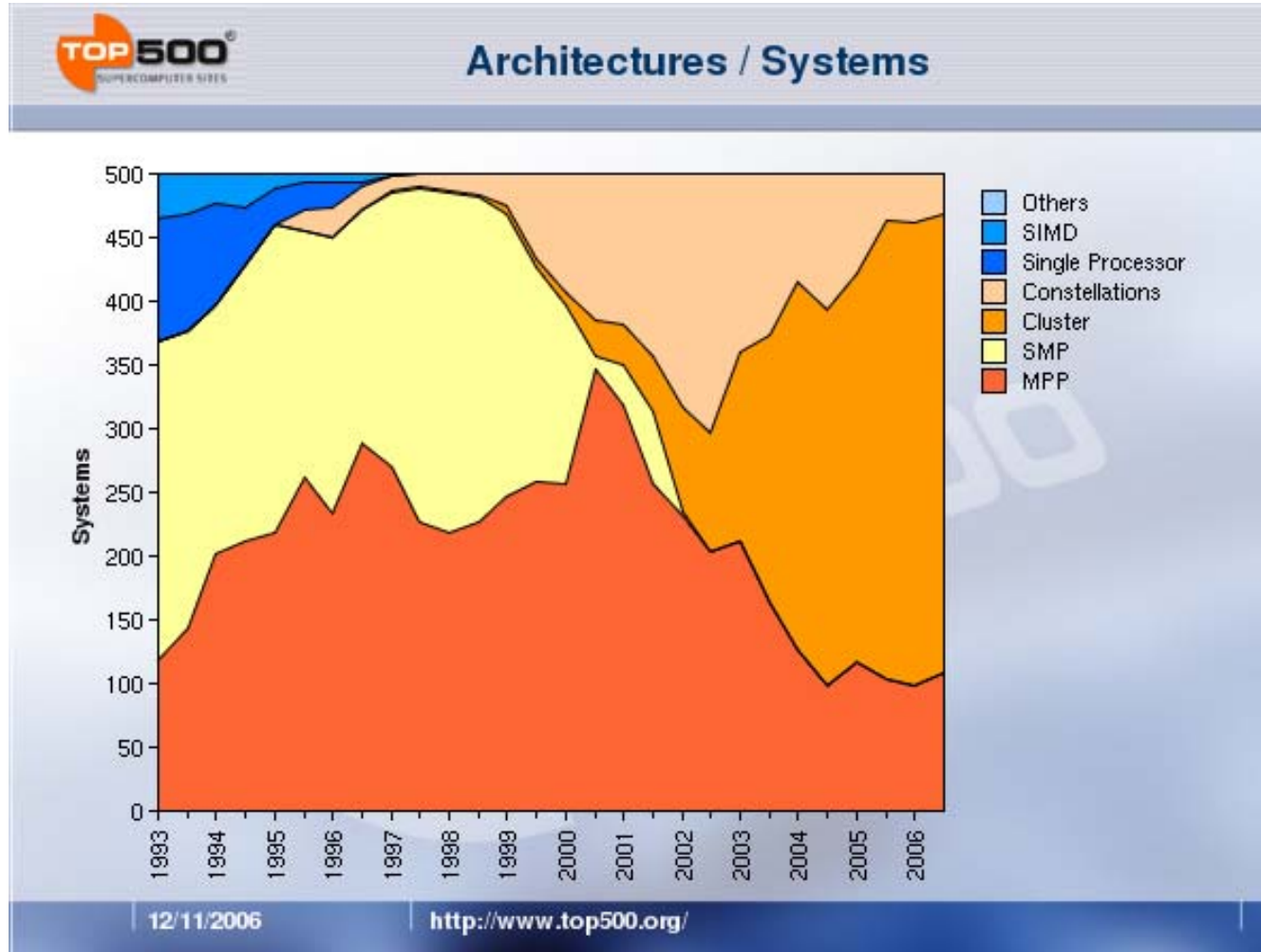
Sistema de Gestión de Carga
Condor



1. Plataformas de Computación Paralela y Distribuida

1.1. Entornos de Computación

Evolución de los Entornos de Computación



1. Plataformas de Computación Paralela y Distribuida

1.1. Entornos de Computación

Evolución de los Entornos de Computación

TOP500 List - November 2006 (1-100)

R_{max} and R_{peak} values are in GFlops. For more details about other fields, check the [TOP500 description](#).

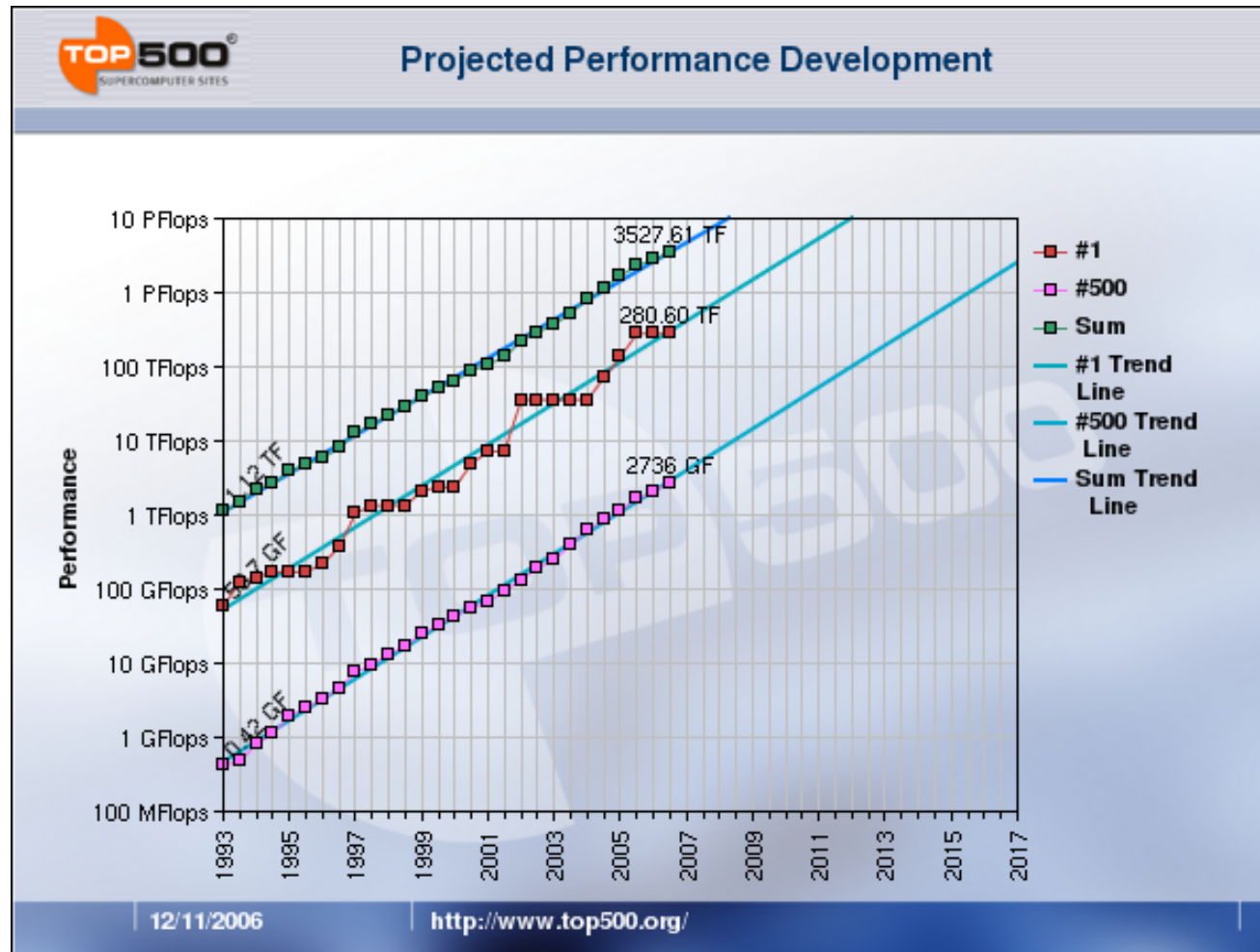
[next](#)

Rank	Site	Computer	Processors	Year	R_{max}	R_{peak}
1	DOE/NNSA/LLNL United States	BlueGene/L - eServer Blue Gene Solution IBM	131072	2005	280600	367000
2	NNSA/Sandia National Laboratories United States	Red Storm - Sandia/ Cray Red Storm, Opteron 2.4 GHz dual core Cray Inc.	26544	2006	101400	127411
3	IBM Thomas J. Watson Research Center United States	BGW - eServer Blue Gene Solution IBM	40960	2005	91290	114688
4	DOE/NNSA/LLNL United States	ASC Purple - eServer pSeries p5 575 1.9 GHz IBM	12208	2006	75760	92781
5	Barcelona Supercomputing Center Spain	MareNostrum - BladeCenter JS21 Cluster, PPC 970, 2.3 GHz, Myrinet IBM	10240	2006	62630	94208
6	NNSA/Sandia National Laboratories United States	Thunderbird - PowerEdge 1850, 3.6 GHz, Infiniband Dell	9024	2006	53000	64972.8
7	Commissariat a l'Energie Atomique (CEA) France	Tera-10 - NovaScale 5160, Itanium2 1.6 GHz, Quadrics Bull SA	9968	2006	52840	63795.2
8	NASA/Ames Research Center/NAS United States	Columbia - SGI Altix 1.5 GHz, Voltaire Infiniband SGI	10160	2004	51870	60960
9	GSIC Center, Tokyo Institute of Technology Japan	TSUBAME Grid Cluster - Sun Fire x4600 Cluster, Opteron 2.4/2.6 GHz and ClearSpeed Accelerator, Infiniband NEC/Sun	11088	2006	47380	82124.8
10	Oak Ridge National Laboratory United States	Jaquar - Cray XT3, 2.6 GHz dual Core Cray Inc.	10424	2006	43480	54204.8

1. Plataformas de Computación Paralela y Distribuida

1.1. Entornos de Computación

Evolución de los Entornos de Computación



1. Plataformas de Computación Paralela y Distribuida

1.2. Sistemas de Gestión de Recursos Distribuidos

Gestión de Plataformas de Computación

Las plataformas de computación se gestionan por medio de diferentes sistemas de gestión de recursos distribuidos (**DRM systems**):

- Sistemas de Colas Batch para servidores HPC
- Sistemas de Gestión de Recursos para clusters dedicados
- Sistemas de Gestión de Carga para clusters no dedicados

Capacidades de los Sistemas DRM

Los sistemas DRM comparten **muchas capacidades**:

- Colas batch
- Planificación de trabajos
- Gestión de recursos

Beneficios de los Sistemas DRM

Sus beneficios en **minimización de costes** y **maximización de rendimiento** se deben fundamentalmente a la mayor utilización de los recursos

1. Plataformas de Computación Paralela y Distribuida

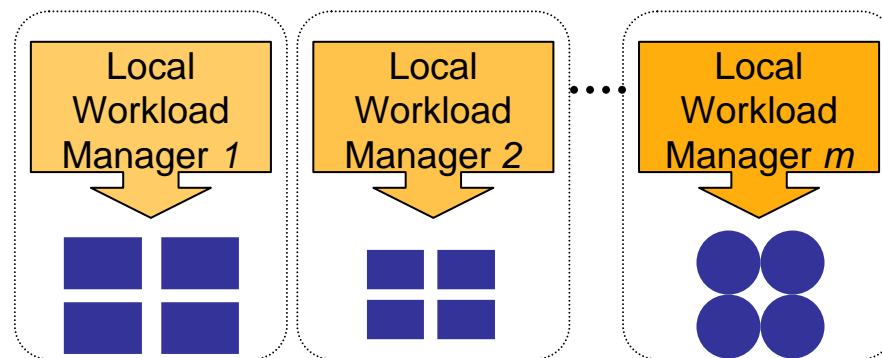
1.2. Sistemas de Gestión de Recursos Distribuidos

Sistemas DRM

<i>Independent Suppliers</i>	<i>Open Source</i>	<i>OEM Proprietary</i>
<i>Platform Computing</i> LSF	<i>Altair</i> Open PBS	<i>IBM</i> Load Leveler
<i>Altair</i> PBS Pro	<i>University of Wisconsin</i> Condor	<i>Cray</i> NQE
	<i>Sun Microsystems</i> SGE	

Silos Verticales de Computación dentro de la Organización

- Los sistemas DRM no proporcionan un interfaz y marco de seguridad comunes y, por tanto, **su integración no es posible**
- La falta de interoperatividad implica la existencia dentro de la misma organización, de **plataformas independientes de computación** (silos verticales) responsables de funciones distintas que:
 - Requieren **habilidades específicas de administración**
 - **Incrementan los costes operacionales**
 - Generan **sobre-provisionamiento** y **desbalanceo global de carga**

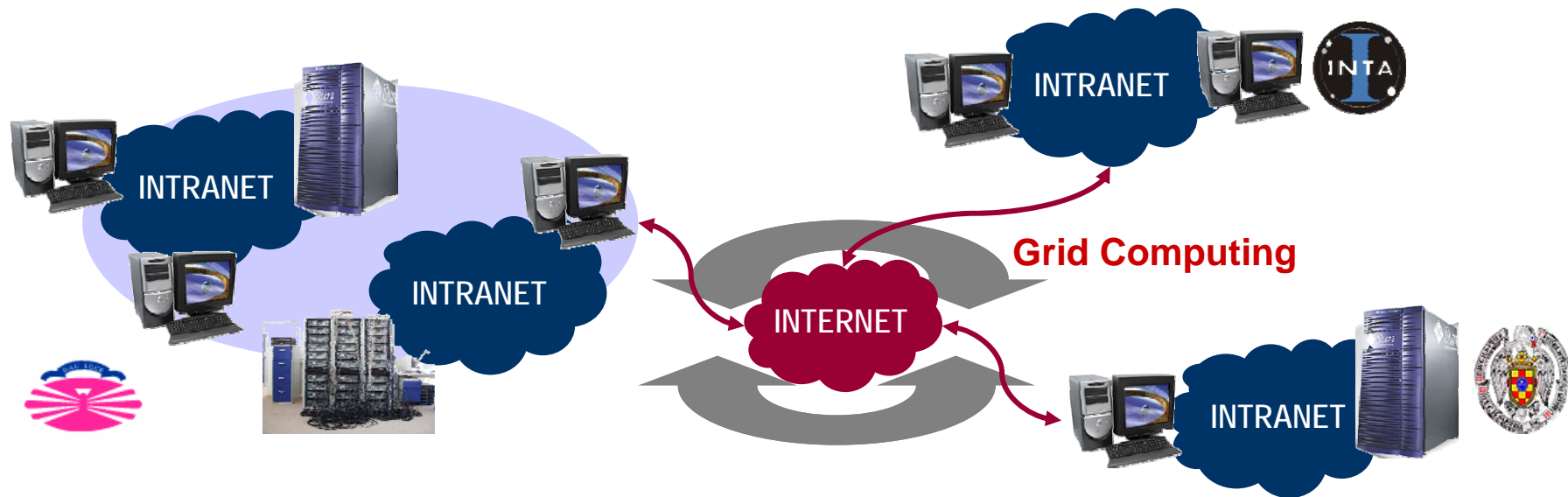


1. Plataformas de Computación Paralela y Distribuida

1.2. Sistemas de Gestión de Recursos Distribuidos

Incapaces de Interconectar Recursos de Diferentes Organizaciones

- Estas tecnologías no son indicadas para el despliegue de infraestructuras computacionales donde los recursos están repartidos diferentes dominios de administración, **cada uno con su propia política de seguridad y sistema DRM**



Contenidos

1. Plataformas de Computación Paralela y Distribuida

2. Tecnología Grid

2.1. Definición y Filosofía

2.2. Viabilidad Tecnológica

2.3. Tecnología Grid

2.4. Organismos de Estandarización

2.5. El Globus Toolkit

3. Infraestructuras Grid

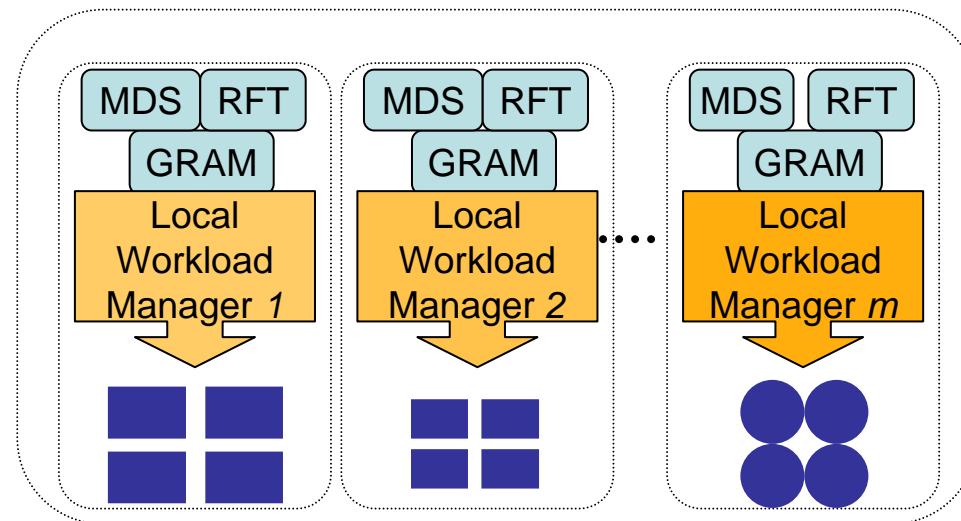
4. Migración de Aplicaciones al Grid

2. Tecnología Grid

2.1. Definición y Filosofía

Definición de Infraestructura Grid

- Una infraestructura grid ofrece una capa común para poder **integrar estas plataformas computacionales no compatibles** (silos verticales) por medio de la definición de un conjunto consistente de interfaces para acceder y gestionar recursos compartidos
- Los **servicios grid** incluyen, entre otros, descubrimiento y monitorización de recursos, asignación y gestión de recursos, infraestructura de seguridad y transferencia de ficheros



2. Tecnología Grid

2.1. Definición y Filosofía

La Filosofía Grid

Un Grid es un sistema que...

- 1) ...coordina **recursos que no están sujetos a un control centralizado...**
- 2) ...usando protocolos e **interfaces estándar**, abiertos y de propósito general...
- 3) ...para proporcionar **calidades de servicio no triviales**.

Ian Foster

What is the Grid? A Three Point Checklist (2002)

La Tecnología Grid es Complementaria a las Anteriores

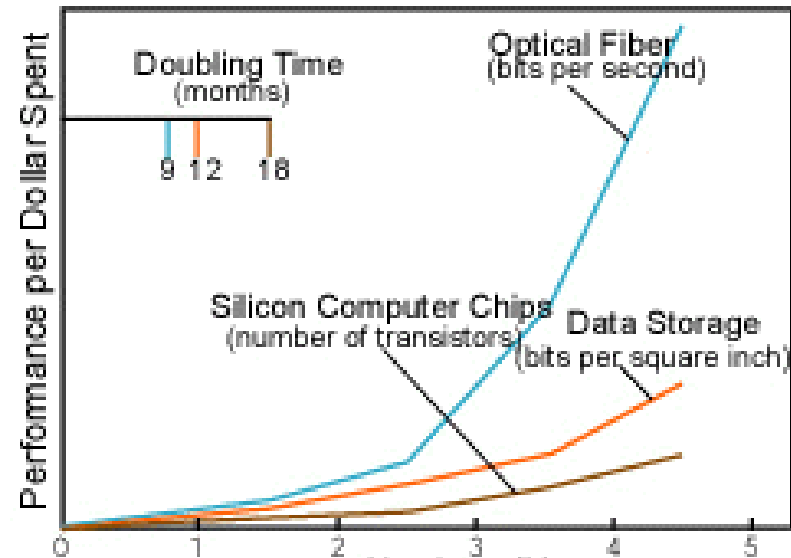
- Interconecta recursos en **diferentes dominios de administración** respetando sus políticas internas de seguridad y su software de gestión de recursos en la Intranet
- Una nueva tecnología dentro del área global de Computación de Altas Prestaciones, para satisfacer las demandas de **determinados perfiles de aplicación**

2. Tecnología Grid

2.2. Viabilidad Tecnológica

*La capacidad de almacenamiento se dobla cada 12 meses
El ancho de banda de red se dobla cada 9 meses
El rendimiento de un procesador se dobla cada 18 meses*

- 1986 to 2000
 - Computers: x 500
 - Networks: x 340,000
- 2001 to 2010
 - Computers: x 60
 - Networks: x 4000



Moore's Law vs. storage improvements vs. optical improvements. Graph from **Scientific American** (Jan-2001) by Cleo Vilett, source Vined Khoslan, Kleiner, Caufield and Perkins.

Conclusiones:

*Un único sistema no será capaz de analizar los datos que almacenen sus discos
Un único centro no podrá analizar el volumen de información generado
La red permitirá de forma eficiente usar recursos distribuidos*

2. Tecnología Grid

2.3. Tecnología Grid

gLite (glite.web.cern.ch/glite/)



UNICORE (www.unicore.org/)



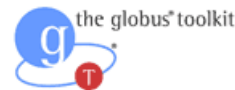
OMII (<http://www.omii.ac.uk/>)



GRIA (www.gria.org)



Globus Toolkit (www.globus.org)



- **Estándar de hecho** para el despliegue de soluciones grid, incorpora gran parte de los estándares actuales
- Usado en la mayoría de las infraestructuras actuales
- Código Open-Source, licencia Apache 2.0
- Comunidad Open-Source dev.globus (basada en la comunidad Apache)
- Existe versión comercial (UNIVA)

2. Tecnología Grid

2.4. Organismos de Estandarización

Nuevos Organismos



European Telecommunications Standards Institute (www.etsi.org)

- Organización independiente de desarrollo de estándares y sin ánimo de lucro
- Se centra fundamentalmente en los aspectos de interoperabilidad requeridos para la siguiente generación de infraestructuras

Organismos Principales



Open Grid Forum (www.ogf.org) = GGF + EGA

- Grupos de trabajo, dirigidos por la comunidad, que están desarrollando una serie de documentos técnicos semejante a los RFCs (*Request for Comments*) de IETF's, que documentan los estándares de Internet
- Servicios OGSA sobre estándares y APIs de OASIS



OASIS (www.oasis-open.org)

- Organización internacional sin ánimo de lucro que promueve estándares industriales para el entorno e-business
- Servicios web necesarios para construir la arquitectura OGSA: WSRF, WSDM, WSS...

2. Tecnología Grid

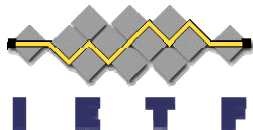
2.4. Organismos de Estandarización

Otros Organismos



World Wide Web Consortium (www.w3.org)

- Organización internacional que tiene como objetivo promover protocolos comunes e interoperables en la Web.
- Infraestructura de bajo nivel: SOAP, WSDL...



Internet Engineering Task Force (www.ietf.org)

- Comunidad internacional dedicada a la evolución de la arquitectura de Internet y a su operación de forma fluida
- Perfil del certificado proxy X.509



Distributed Management Task Force (www.dmtf.org)

- Organización de base que tiene como objetivo desarrollar estándares para la gestión e integración de tecnologías para la empresa.
- *Common Information Model* (CIM) y de *Web-Based Enterprise Management* (WBEM).

Otros Relacionados



2. Tecnología Grid

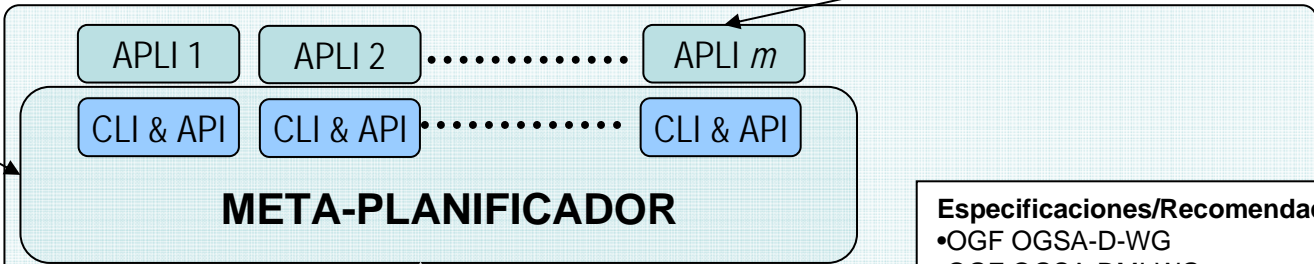
2.4. Organismos de Estandarización

Ejemplo de Estándares



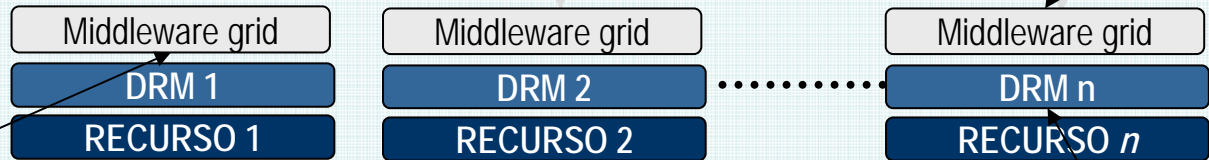
- Especificaciones/Recomendaciones en Aplicaciones**
- OGF DRMAA-WG
 - OGF SAGA-RG
 - OGF JSDL-WG
 - OGF GridRPC-WG

- Recomendaciones en Planificación**
- OGF GSA-RG



- Especificaciones/Recomendaciones en Datos**
- OGF OGSA-D-WG
 - OGF OGSA-DMI-WG
 - OGF GFS-WG

- Especificaciones para Acceso a Recursos**
- OASIS WSRF



- Especificaciones/Recomendaciones en Computación**
- OGF WSRF (W3C..)
 - OGF GRAAP-WG
 - OGF OGSA-BES-WG
 - OGF OGSA-RSS-WG
 - OGF OGSA-HPCP-WG
 - OGF Grid-FTP



- Especificaciones en Acceso DRM**
- OGF DRMAA-WG

El Globus Toolkit, un Estándar de Hecho en Computación Grid

Permite **compartir recursos localizados en diferentes dominios** de administración, con diferentes políticas de seguridad y gestión de recursos

Globus es...

- un middleware software
- un conjunto de librerías, servicios y APIs

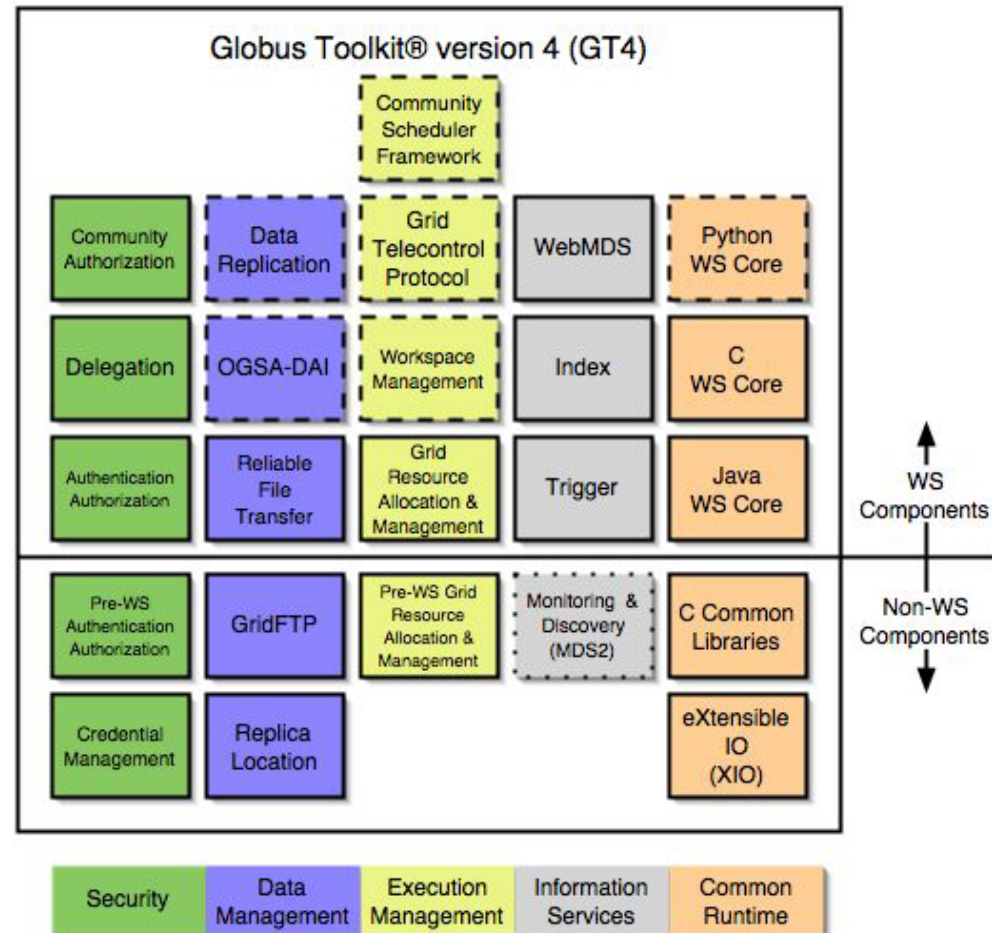
Globus no es...

- una herramienta de usuario o planificador
- una aplicación

2. Tecnología Grid

2.5. El Globus Toolkit

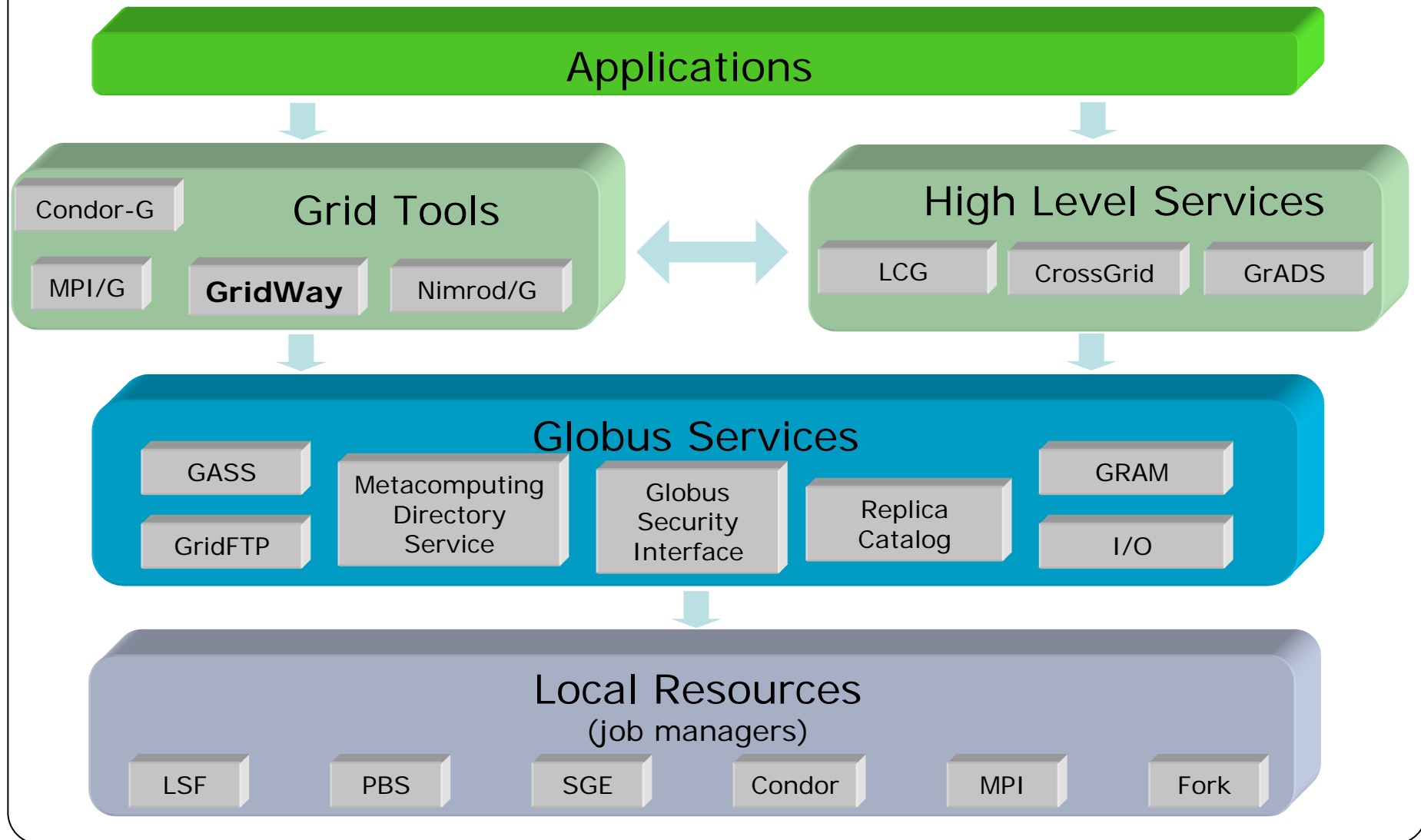
Componentes Principales del Globus Toolkit



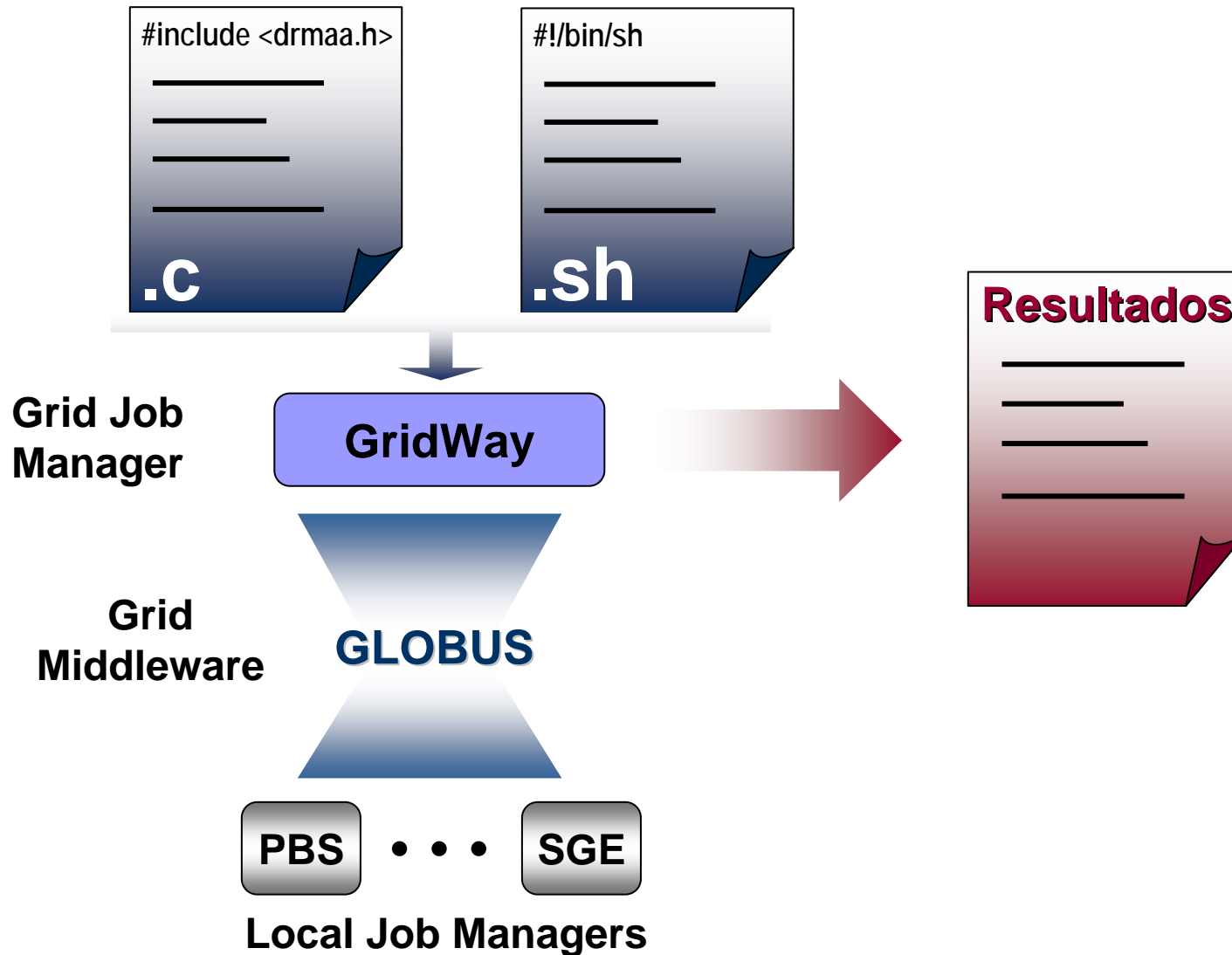
2. Tecnología Grid

2.5. El Globus Toolkit

Estructura en Capas de una Infraestructura Grid Basada en Globus



Modelo de Gestión de Recursos en un Grid



2. Tecnología Grid

2.5. El Globus Toolkit

1.- Pasos en la Planificación de un Trabajo



• ¿Dónde ejecuto mi trabajo?	Resource selection
• ¿Qué necesito?	Resource preparation
• ¿Cómo envío mi trabajo?	Job submission
• ¿Cómo va la ejecución?	Job monitoring
• ¿Existe un recurso mejor?	Job migration
• ¿Cómo recupero la salida?	Job termination

2.- Características de un Grid

- Múltiples dominios de administración y autonomía
- Heterogeneidad
- Escalabilidad
- Variabilidad

3.- Variabilidad en un Grid

Disponibilidad Dinámica

- Adición y exclusión de recursos
- Cancelación de trabajos

Precio Dinámico

- Hora del día
- Carga del recurso

Grid

Carga Dinámica

- Recursos compartidos
- Recursos ociosos pueden saturarse

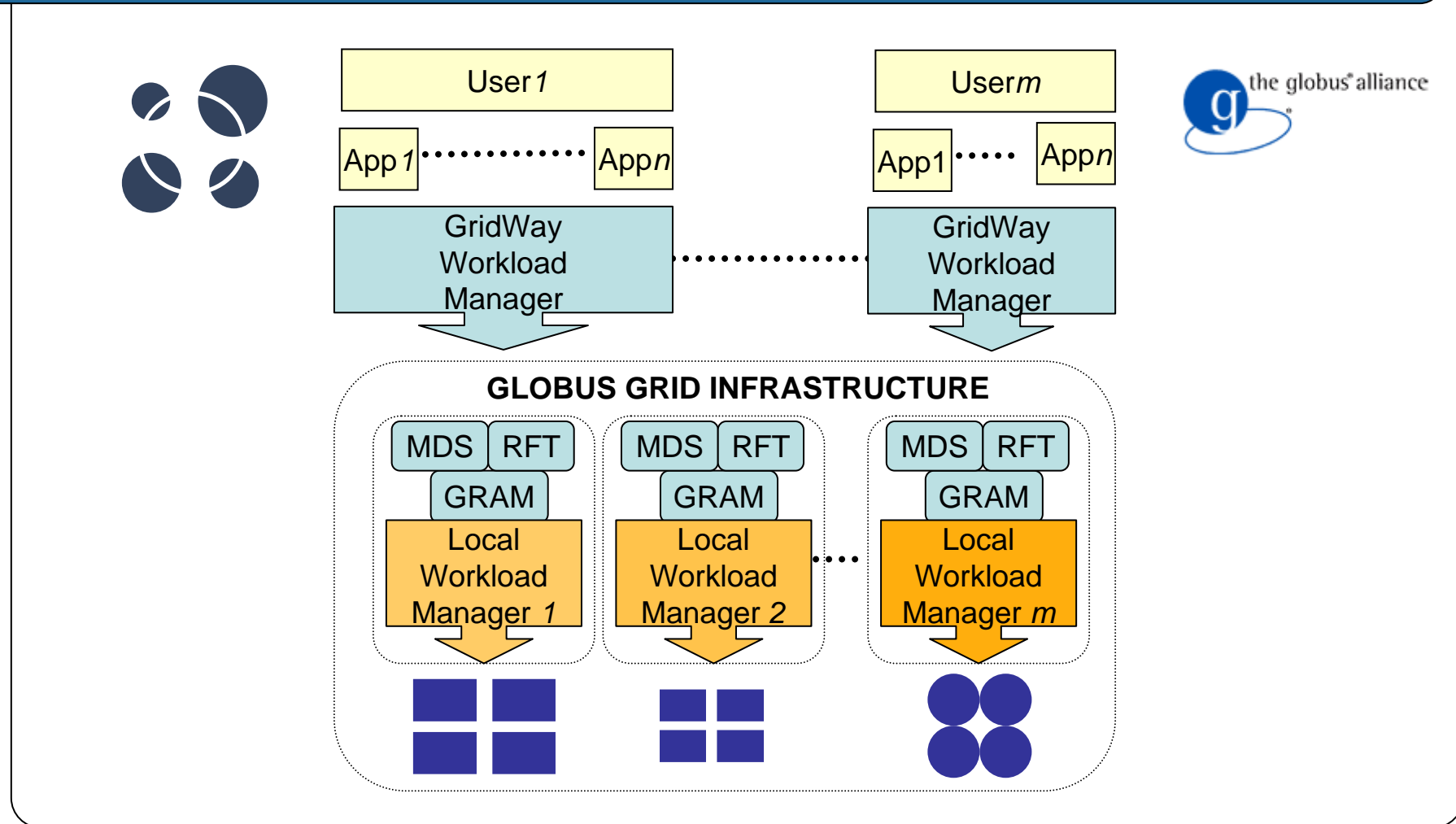
Alta Tasa de Fallos

- Red
- Recurso

2. Tecnología Grid

2.6. Ejemplos de Infraestructuras

GridWay Proporciona Desacoplamiento entre las Aplicaciones y las Plataformas Computacionales Subyacentes



Contenidos

1. Plataformas de Computación Paralela y Distribuida
2. Tecnología Grid
- 3. Infraestructuras Grid**
 - 3.1. Infraestructuras Grid entre Socios
 - 3.2. Evolución de la Computación Grid
 - 3.3. Ejemplos con el Globus Toolkit
 - 3.4. Iniciativas Relacionadas
4. Migración de Aplicaciones al Grid

3. Infraestructuras Grid

3.1. Infraestructuras Grid entre Socios

Tipos de Infraestructura Grid

- **Entornos Grid desacoplados**

✓ Recursos heterogéneos, dinámicos y autónomos interconectados por medio de redes públicas



- **Entornos Grid acoplados**

✓ Recursos homogéneos, dedicados y control centralizado interconectados por medio de redes dedicadas



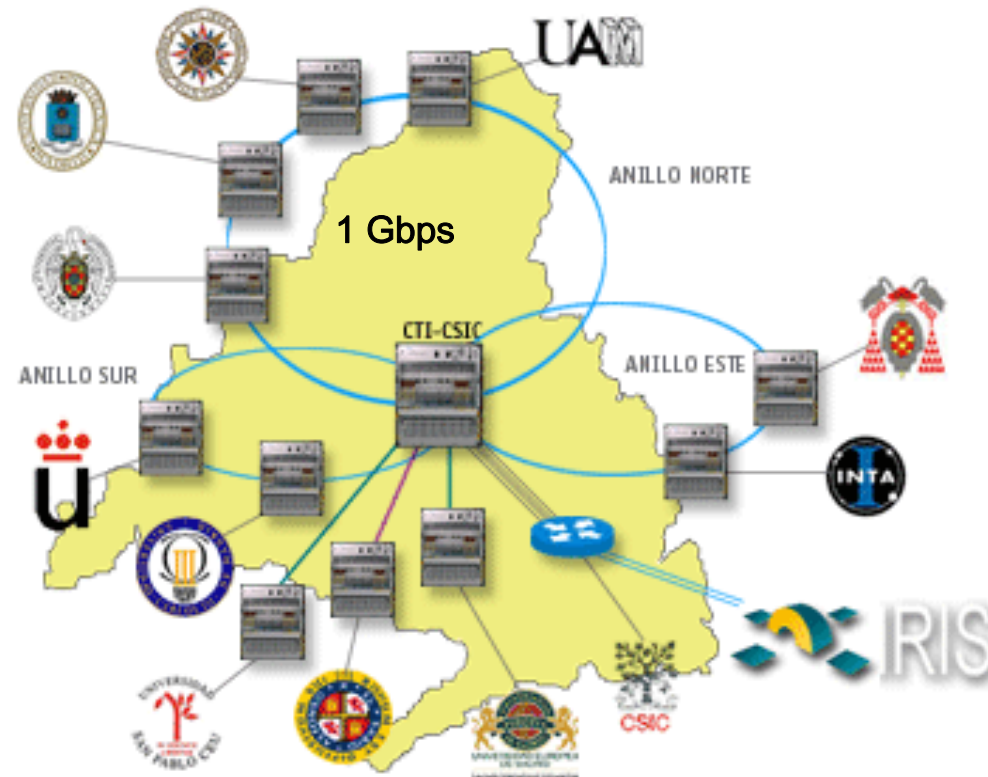
Beneficios

- Acceso puntual a más **recursos**
- Mayor **colaboración y compartición** de recursos entre varios centros
- Creación de **organizaciones virtuales**

3. Infraestructuras Grid

3.1. Infraestructuras Grid entre Socios

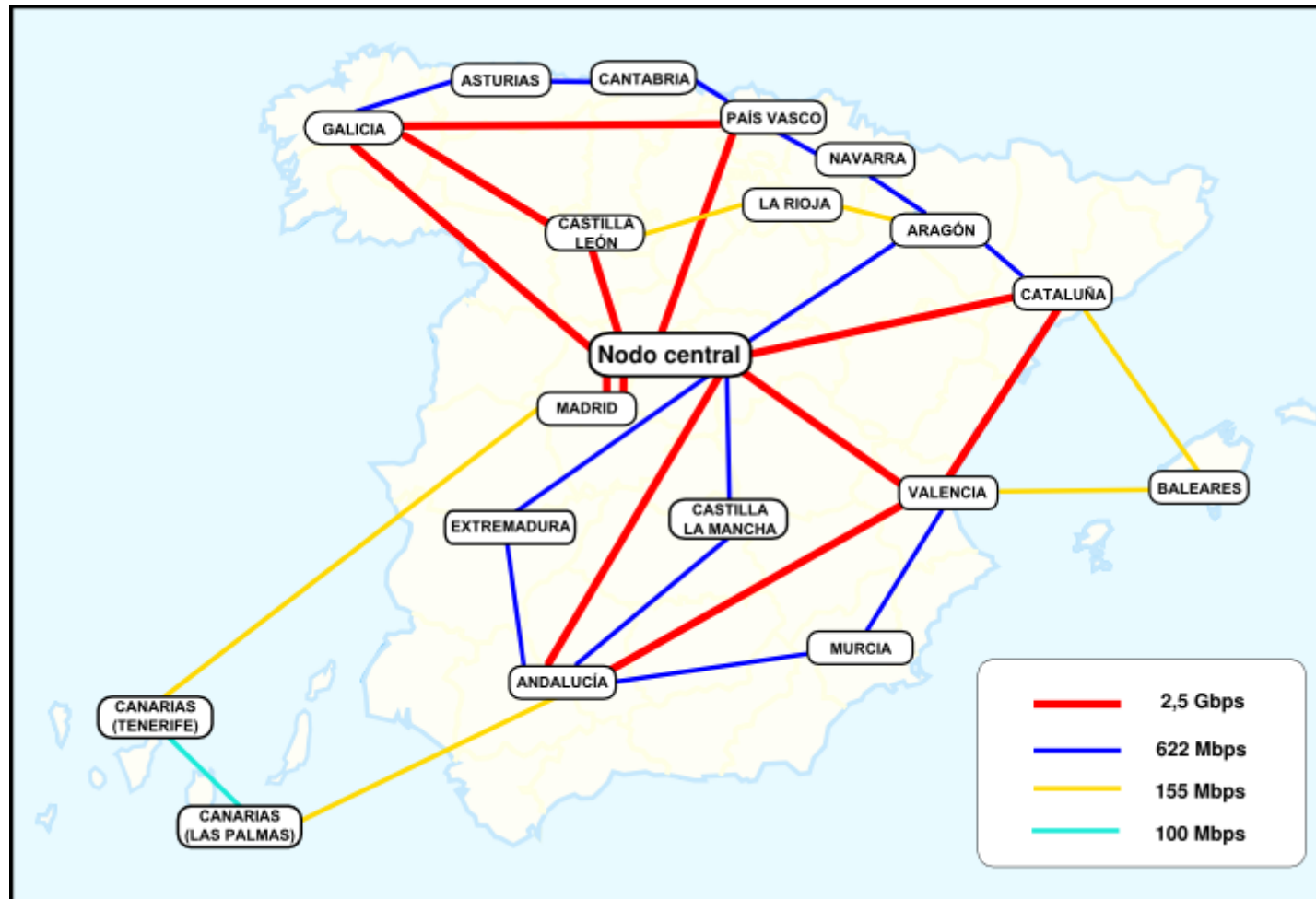
Comunidad de Madrid: **Red Telemática de Investigación**



3. Infraestructuras Grid

3.1. Infraestructuras Grid entre Socios

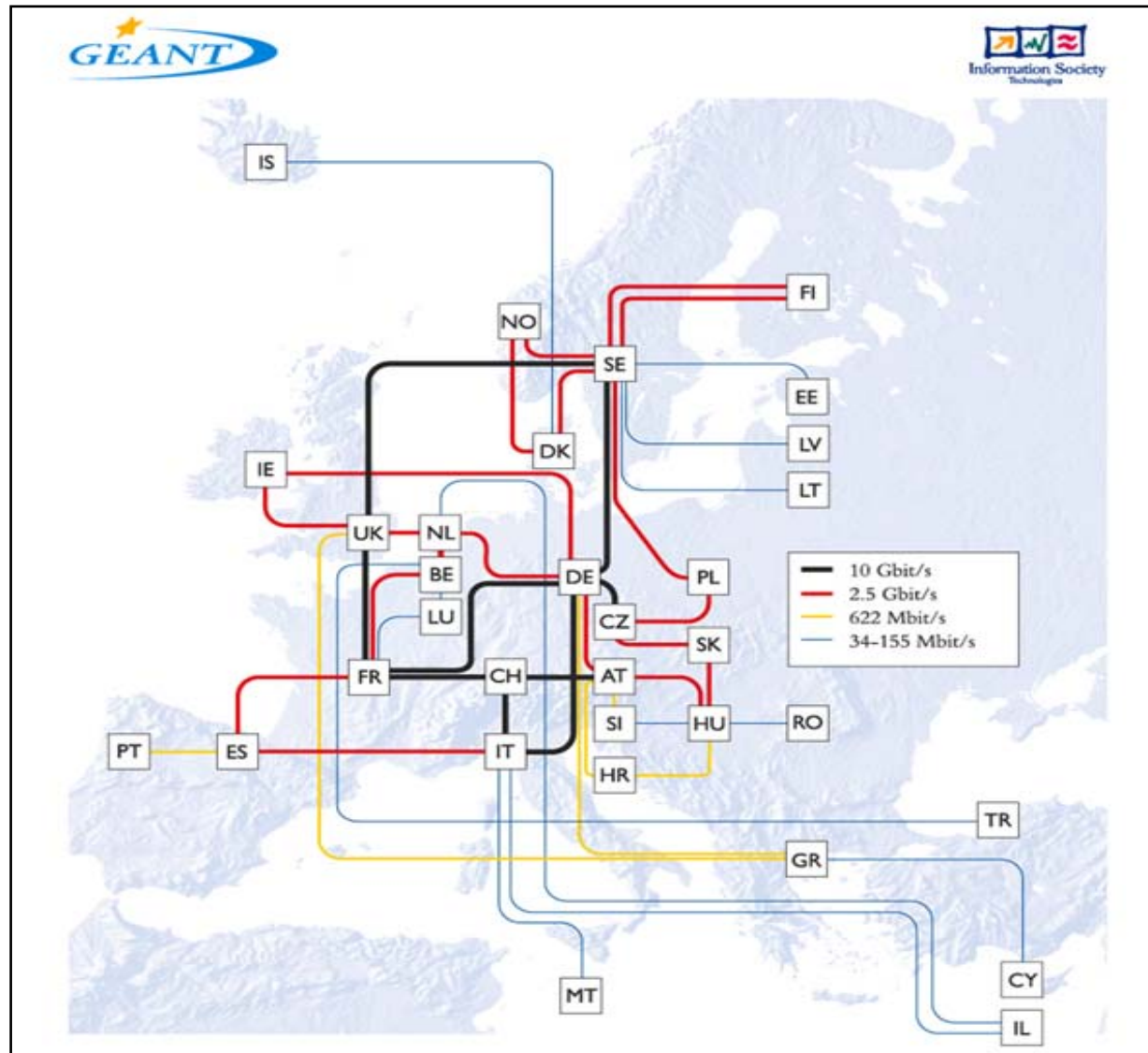
España: **RedIris-2**



3. Infraestructuras Grid

3.1. Infraestructuras Grid entre Socios

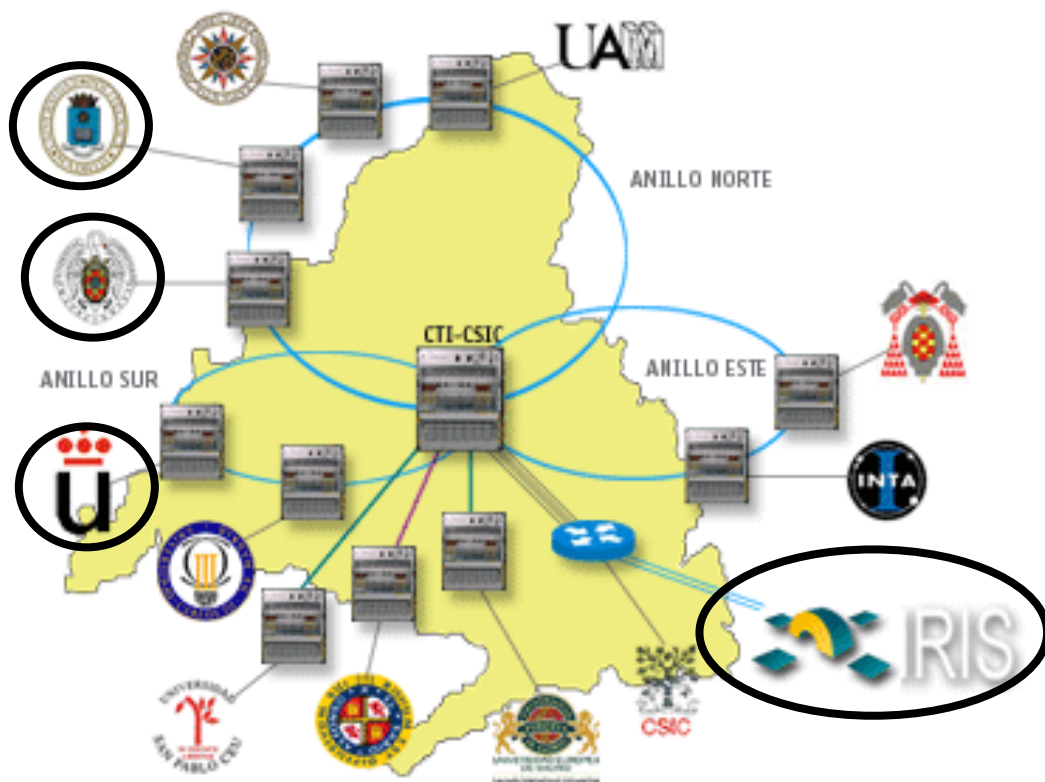
Europa: **Geant**



3. Infraestructuras Grid

3.1. Infraestructuras Grid entre Socios

www.GRIDIMadrid.org

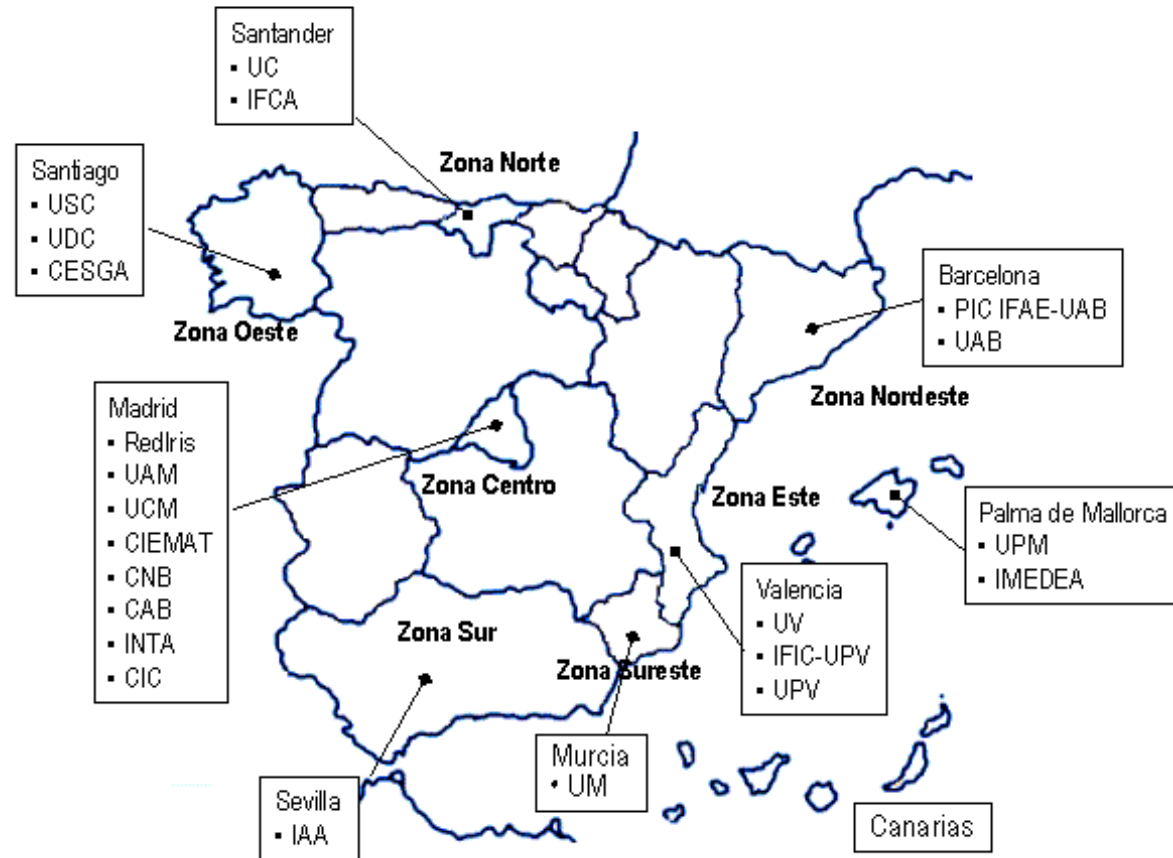


Infraestructura de Investigación

3. Infraestructuras Grid

3.1. Infraestructuras Grid entre Socios

www.Irisgrid.es

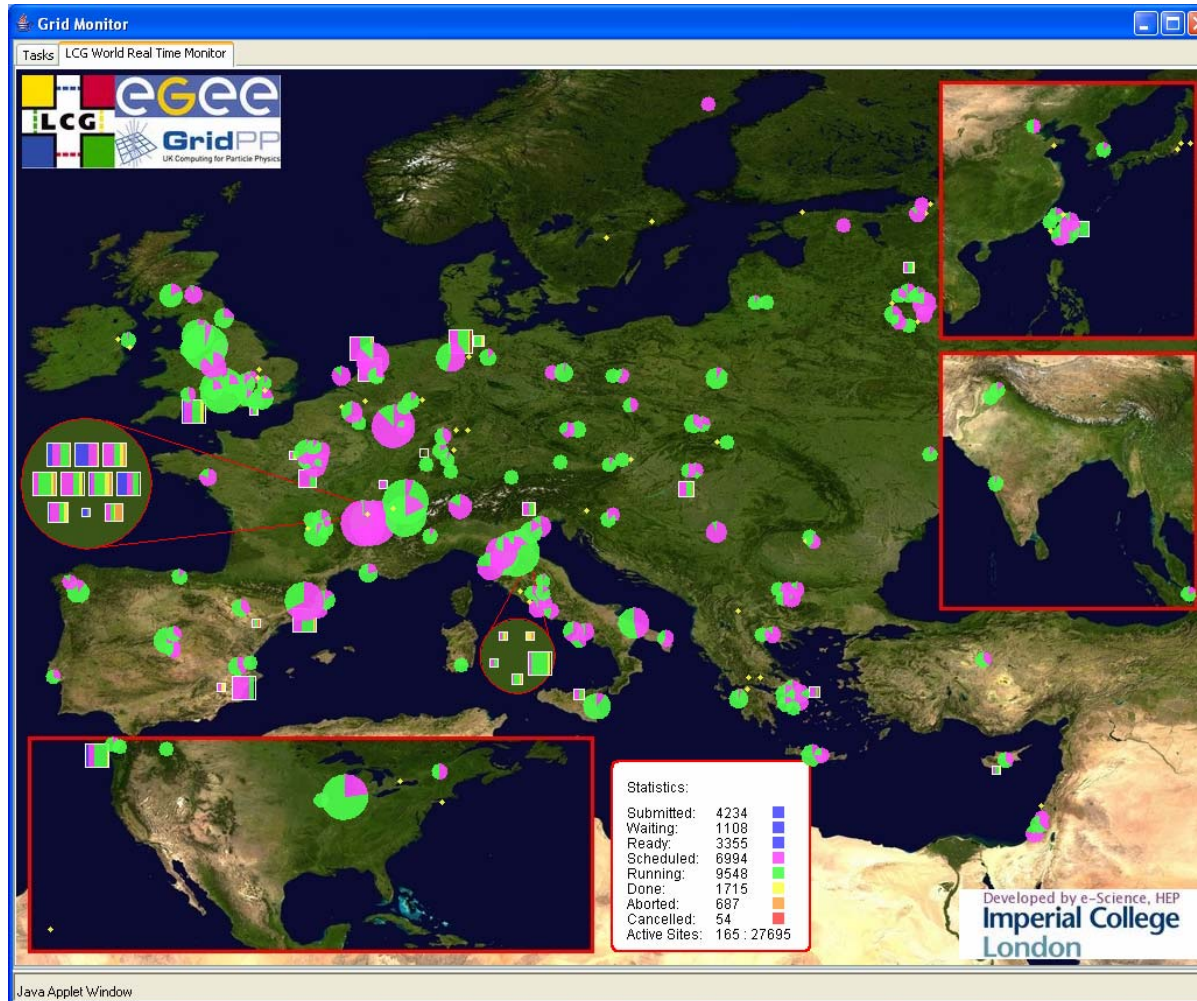


Infraestructura de Investigación

3. Infraestructuras Grid

3.1. Infraestructuras Grid entre Socios

www.eu-egee.org



Infraestructura de Producción

3. Infraestructuras Grid

3.2. Evolución de la Computación Grid

Etapas en la Evolución de la Computación Grid



Fuente: Platform Computing, "The Evolution Of Grid: The Three Stages of Grid Computing".
Available at <http://www.platform.com/grid/evolution.asp>

3. Infraestructuras Grid

3.2. Evolución de la Computación Grid

Descripción de las Etapas

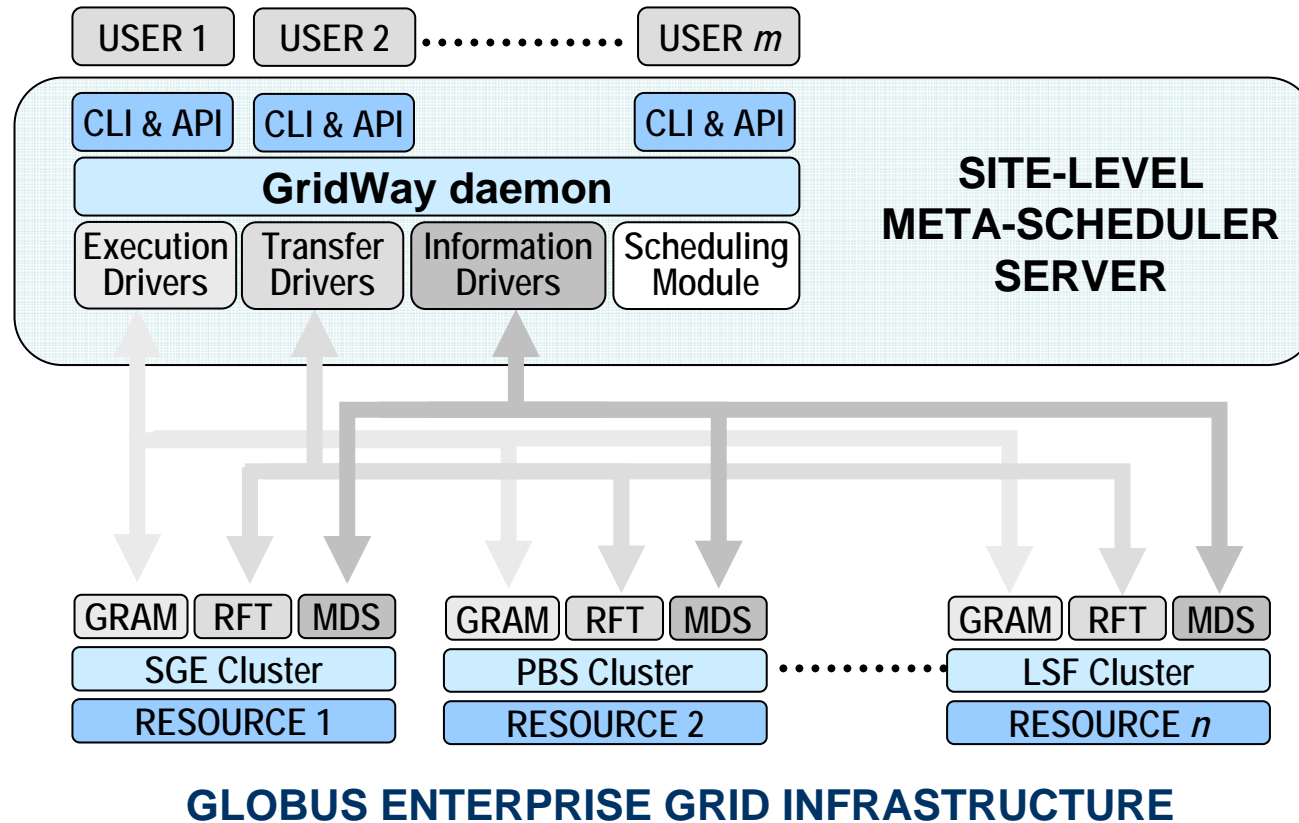
Infraestructuras Grid en la Comunidad Científica

	Enterprise Grid	Partner Grid	Utility Grid
Infraestructura	Recursos internos gestionados por diferentes sistemas DRM que podrían estar distribuidos geográficamente	Recursos distribuidos en diferentes organizaciones o dominios de administración gestionados por diferentes sistemas DRM	Recursos proporcionados por proveedores de servicios externos
Objetivos	Permitir la compartición de diferentes recursos para mejorar la colaboración interna y alcanzar un mayor retorno de la inversión en TIC	Proporcionar compartición fiable y segura de recursos a gran escala entre socios o participantes en la cadena de valor	Proporcionar recursos bajo demanda
Beneficios	<ul style="list-style-type: none">• Minimizar costes• Maximizar prestaciones	<ul style="list-style-type: none">• Acceder a más recursos para satisfacer picos de demanda• Proporcionar soporte para hacer frente a proyectos colaborativos	<ul style="list-style-type: none">• Flexibilidad para ajustar la capacidad• Acceder a capacidad ilimitada• Transformar costes TIC fijos en variables

3. Infraestructuras Grid

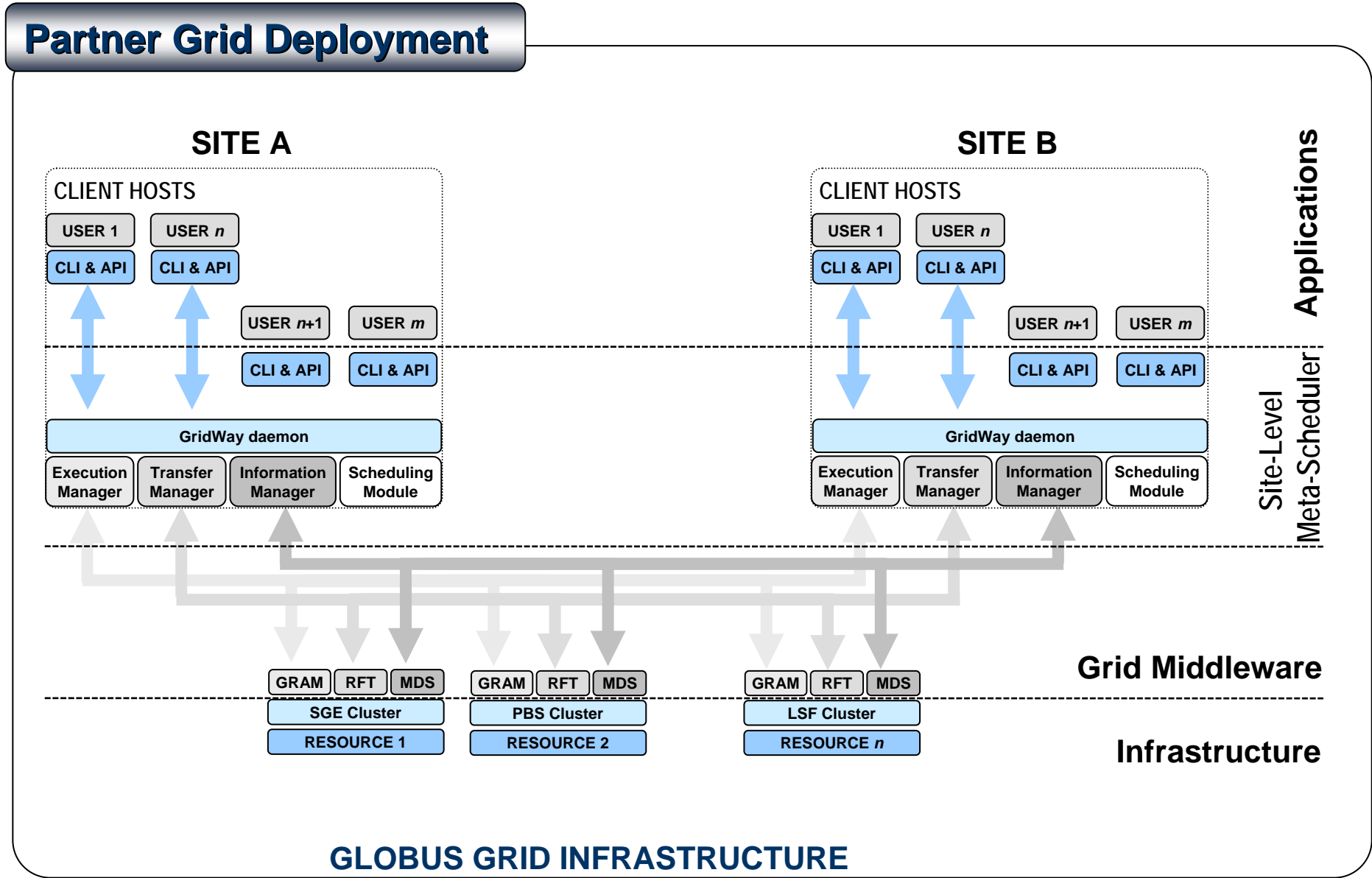
3.3. Ejemplos con el Globus Toolkit

Enterprise Grid Deployment



3. Infraestructuras Grid

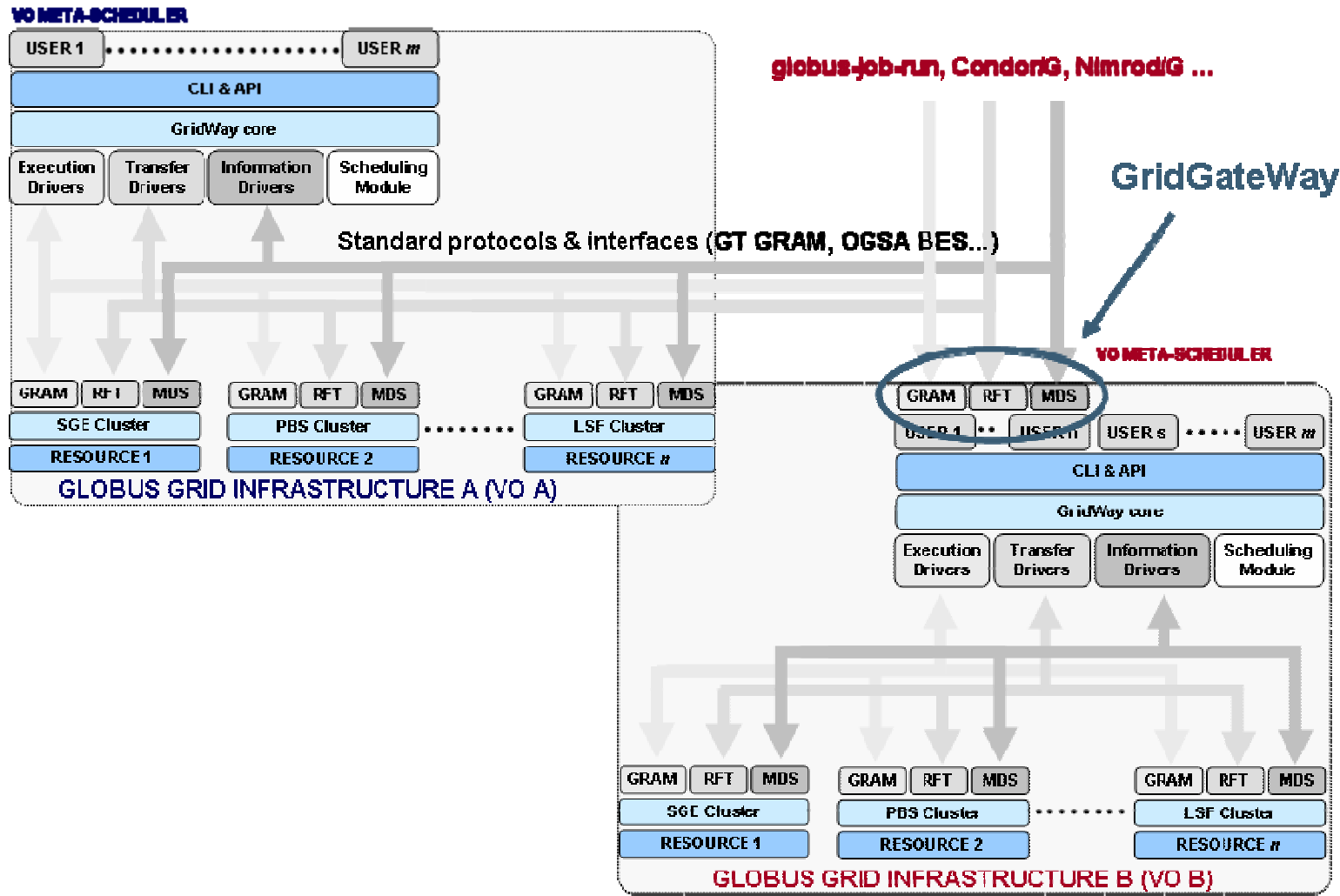
3.3. Ejemplos con el Globus Toolkit



3. Infraestructuras Grid

3.3. Ejemplos con el Globus Toolkit

Outsourced Grid Deployment



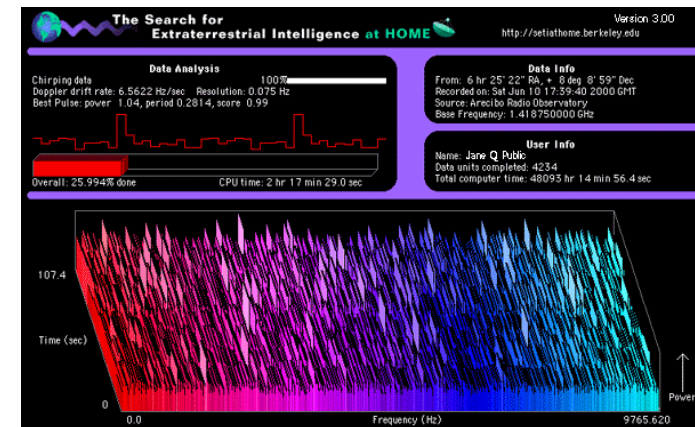
3. Infraestructuras Grid

3.4. Iniciativas Relacionadas

Iniciativas de Computación en Internet

- Plataformas software para usar **recursos cedidos voluntariamente**
- **Tecnología complementaria** a las anteriores que permiten **interconectar recursos individuales** en lugar de servidores y clusters con diferentes DRMs

Objetivo: Análisis de datos de telescopio (Arecibo, Puerto Rico) en búsqueda de señales



Estadísticas Noviembre de 2006

Users: 274.887; Hosts: 436.290; Countries: 223
Average floating point operations per second: 455,297 TeraFLOPS

4. Migración de Aplicaciones al Grid

Contenidos

1. Plataformas de Computación Paralela y Distribuida
2. Tecnología Grid
3. Infraestructuras Grid
- 4. Migración de Aplicaciones al Grid**
 - 4.1. Paradigma HPC
 - 4.2. Paradigma HTC
 - 4.3. Ejemplos de Aplicaciones Portadas al Grid

4. Migración de Aplicaciones al Grid

4.1. Paradigma HPC

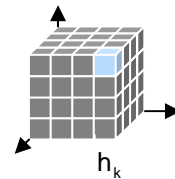
Aplicaciones HPC (*High Performance Computing*)

- Su **objetivo** es reducir el tiempo de ejecución de una única aplicación paralela
- Su **rendimiento** se mide en número de operaciones en punto flotante por segundo
- **Áreas de aplicación:**
 - **Estudio de fenómenos a escala microscópica** (*dinámica de partículas*)
 - Resolución limitada por la potencia de cálculo del computador
 - Cuantos más grados de libertad (puntos), mejor reflejo de la realidad
 - **Estudio de fenómenos a escala macroscópica** (*sistemas descritos por ecuaciones diferenciales fundamentales*)
 - Precisión limitada por la potencia de cálculo del computador
 - Cuantos más puntos, más se acerca la solución discreta a la continua

Ecuación no lineal de Schrödinger
Ecuaciones de Maxwell-Bloch

$$a \frac{\partial u(\vec{x})}{\partial x} + b \frac{\partial u(\vec{x})}{\partial y} = 0 \quad \forall \vec{x} \in \Omega$$

$$u(\vec{x}) = f(\vec{x}) \quad \forall \vec{x} \in \partial\Omega$$



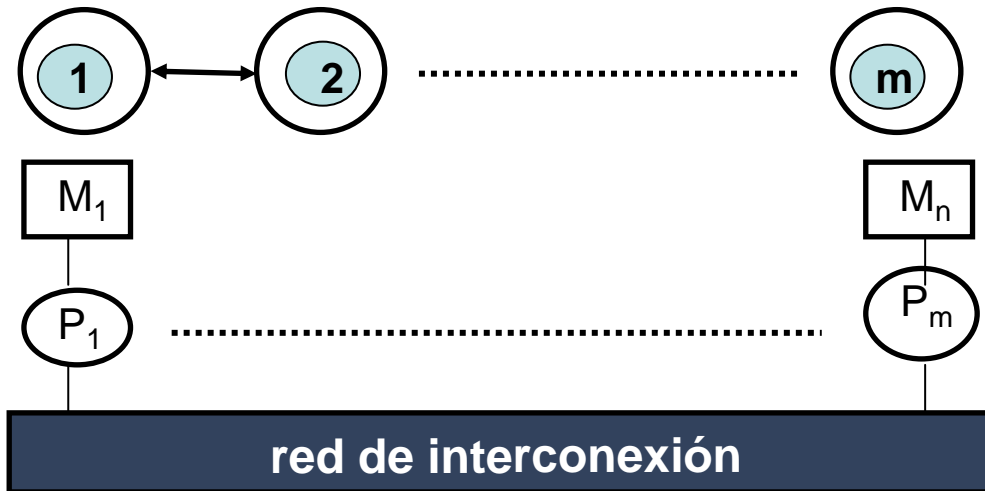
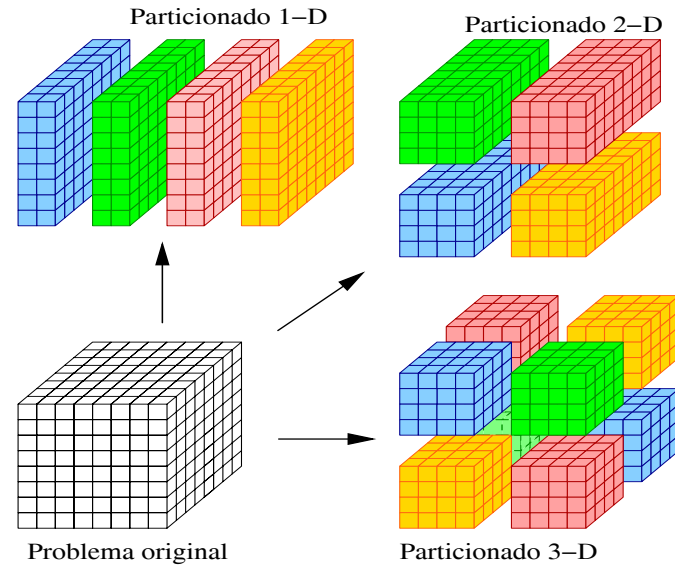
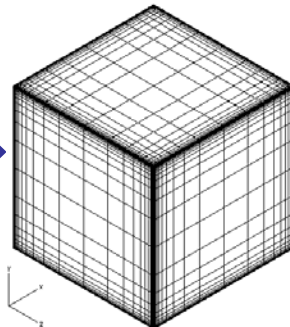
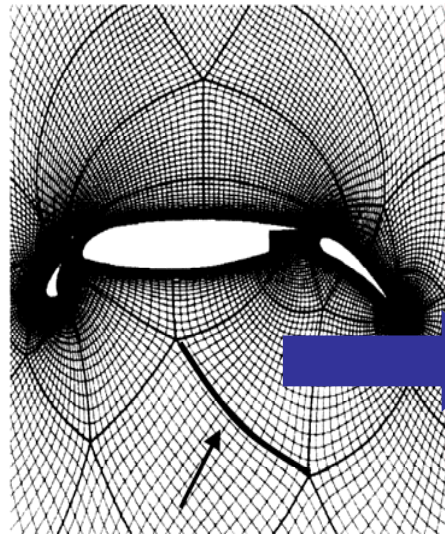
Esquemas numéricos

$$\frac{\partial u}{\partial x} = \frac{u_{i+1} - u_{i-1}}{h}$$
El diagrama muestra una malla 2D de puntos conectados por líneas. Una línea apunta desde el texto de la ecuación a un punto central en la malla.

4. Migración de Aplicaciones al Grid

4.1. Paradigma HPC

Modelo de Programación



4. Migración de Aplicaciones al Grid

4.1. Paradigma HPC

Memoria Distribuida: MPI

```
MPI_Init(&argc, &argv);
MPI_Comm_rank(MPI_COMM_WORLD, &mi_rango);
MPI_Comm_size(MPI_COMM_WORLD, &numero_procesos);
if (mi_rango !=0) {
    sprintf(mensaje,"¡Saludos desde el proceso %d!", mi_rango);
    MPI_Send(mens, strlen(mensaje)+1, MPI_CHAR, destino, etiqueta, MPI_COMM_WORLD);
} else {
    for (origen = 1; origen < numero_procesos; origen++) {
        MPI_Recv(mens, strlen(mensaje)+1, MPI_CHAR, origen, etiqueta, MPI_COMM_WORLD, &est);
        printf("%s\n", mens);
    }
}
MPI_Finalize();
```

Memoria Compartida: OpenMP

```
C$OMP DO PRIVATE(I), REDUCTION(+:BG_SUM), REDUCTION(*:BG_PROD)
C$& REDUCTION(MIN:BG_MIN), REDUCTION(MAX:BG_MAX)
    DO I = 1,N
        BG_SUM = BG_SUM + A(I)
        BG_PROD = BG_PROD * A(I)
        BG_MIN = MIN(BG_MIN, A(I))
        BG_MAX = MAX(BG_MAX, A(I))
    END DO
```


4. Migración de Aplicaciones al Grid

4.1. Paradigma HPC

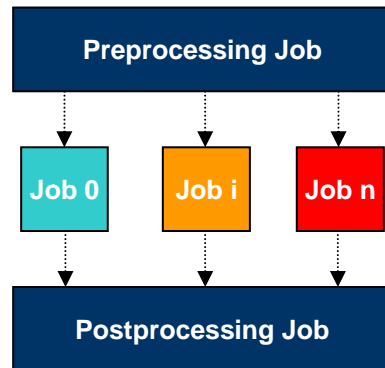
¿Es el Modelo HPC Eficiente en un Entorno Grid ?

4. Migración de Aplicaciones al Grid

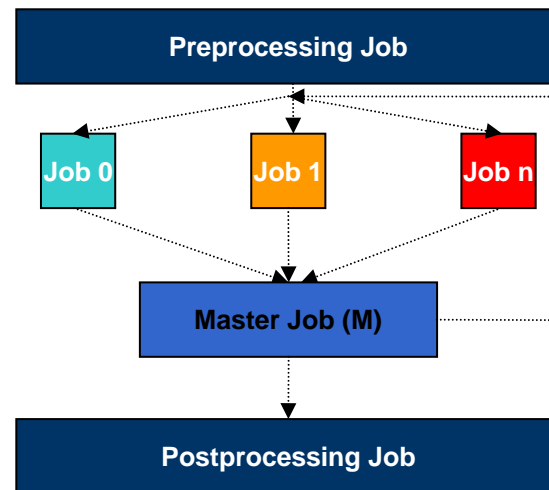
4.2. Paradigma HTC

Aplicaciones HTC (*High Throughput Computing*)

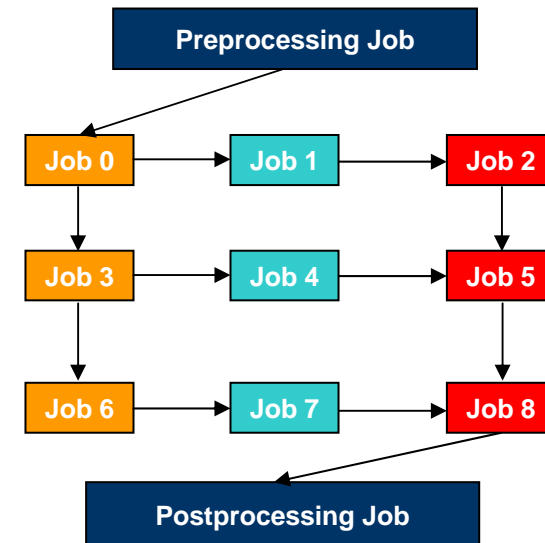
- Su **objetivo** es aumentar el número de ejecuciones por unidad de tiempo
- Su **rendimiento** se mide en número de trabajos ejecutados por segundo
- **Áreas de aplicación:** HEP, bioinformática, finanzas...



**HTC
Síncrono**



Master-slave



Flujos de trabajo

4. Migración de Aplicaciones al Grid

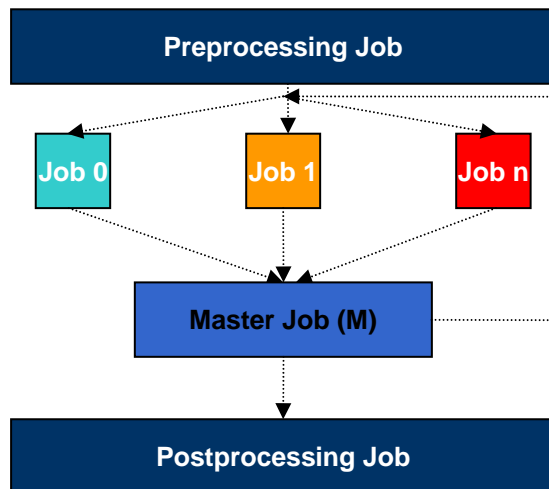
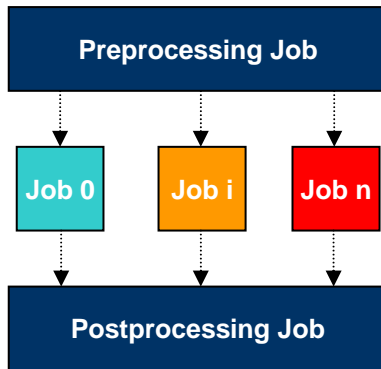
4.2. Paradigma HTC



4. Migración de Aplicaciones al Grid

4.2. Paradigma HTC

DRMAA (*Distributed Resource Management Application API*)



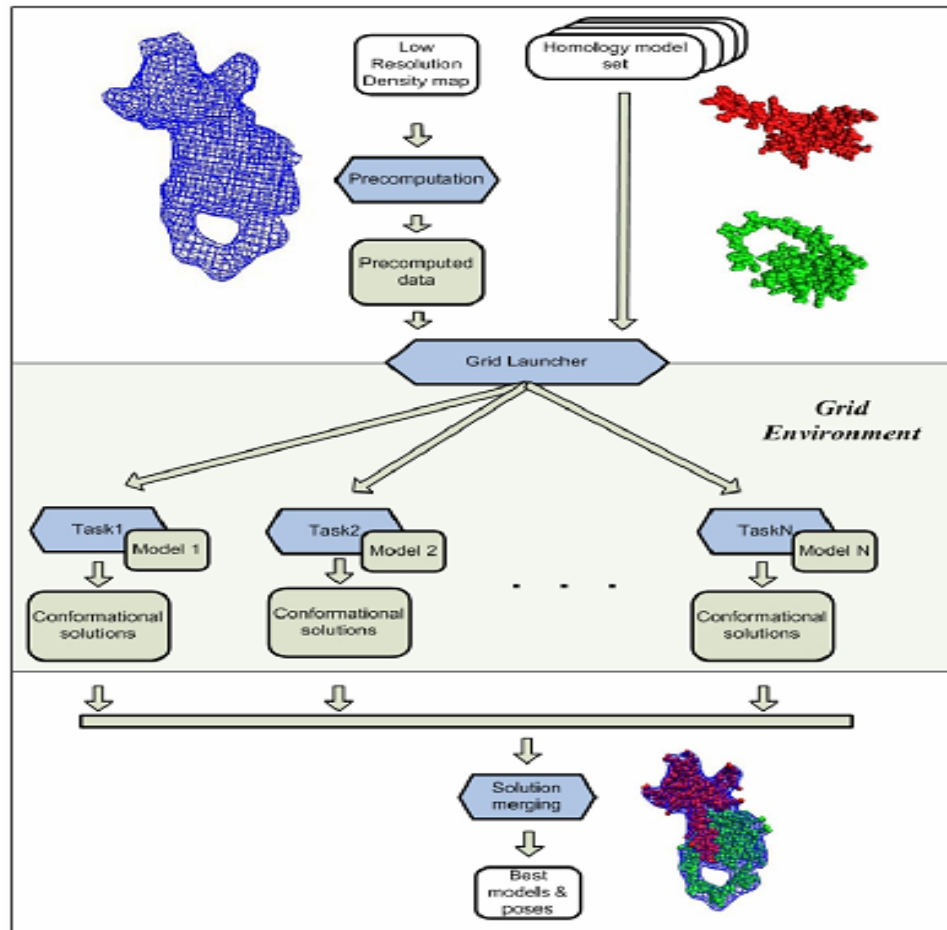
```
rc = drmaa_init(contact, err);  
// Ejecuta el trabajo inicial y espera  
rc = drmaa_run_job(job_id, jt, err);  
rc = drmaa_wait(job_id, &stat, timeout,  
rusage, err);  
// Ejecuta n trabajos y espera  
rc = drmaa_run_bulk_jobs(job_ids, jt, 1,  
JOB_NUM, 1, err);  
rc = drmaa_synchronize(job_ids, timeout, 1, err);  
// Ejecuta el trabajo final y espera  
rc = drmaa_run_job(job_id, jt, err);  
rc = drmaa_wait(job_id, &stat, timeout,  
rusage, err);  
rc = drmaa_exit(err_diag);
```

```
// Ejecuta el trabajo inicial y espera por él  
rc = drmaa_run_job(job_id, jt, err_diag);  
rc = drmaa_wait(job_id, &stat, timeout, rusage, err_diag);  
while (exitstatus != 0) {  
// Ejecuta n Esclavos y espera por ellos  
rc = drmaa_run_bulk_jobs(job_ids, jt, 1,  
JOB_NUM, 1, err_diag);  
rc = drmaa_synchronize(job_ids, timeout, 1,  
err_diag);  
// Ejecuta el Maestro y espera a su finalización  
rc = drmaa_run_job(job_id, jt, err_diag);  
rc = drmaa_wait(job_id, &stat, timeout, rusage, err_diag);  
rc = drmaa_wexitstatus(&exitstatus, stat,  
err_diag);}
```

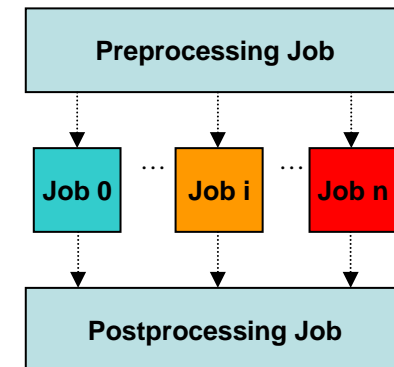
4. Migración de Aplicaciones al Grid

4.3. Ejemplos de Aplicaciones Portadas al Grid

Multi-Resolution Docking (CIB/CSIC)



HIGH THROUGHPUT



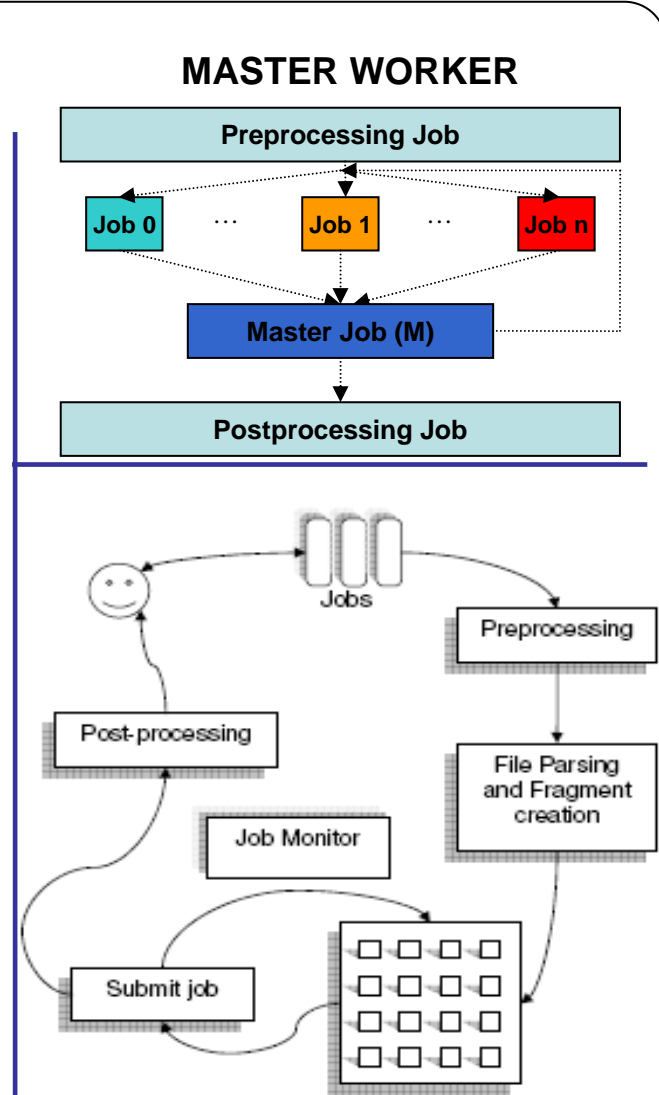
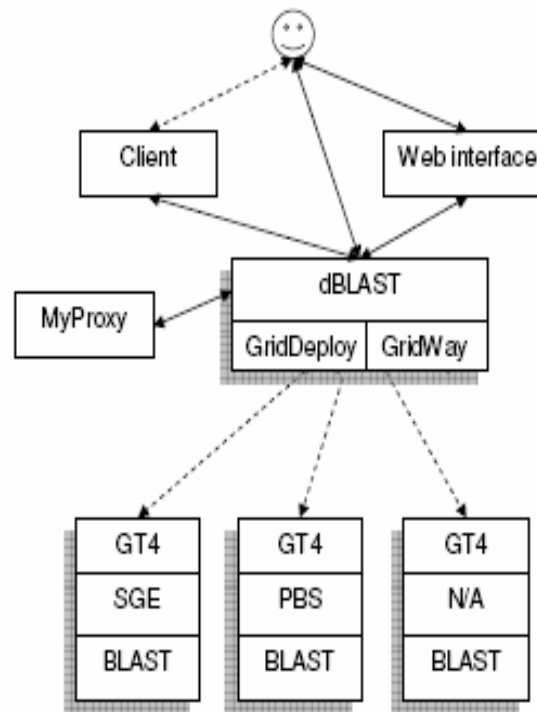
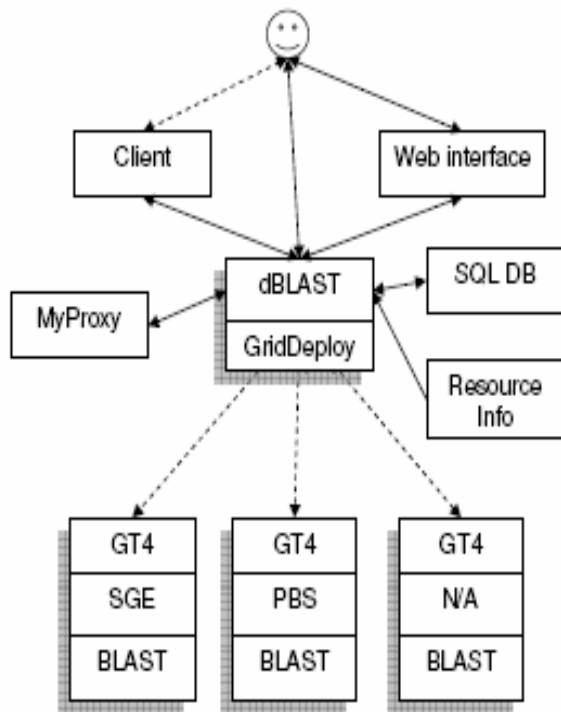
Más información:

http://sbg.cib.csic.es/Research/Docking_EM/index.html

4. Migración de Aplicaciones al Grid

4.3. Ejemplos de Aplicaciones Portadas al Grid

Dynamic BLAST (University of Alabama)



Más información: <http://www.cis.uab.edu/ccl/>

4. Migración de Aplicaciones al Grid

4.3. Ejemplos de Aplicaciones Portadas al Grid

CD-HIT (CNIO- Instituto de Salud Carlos III)

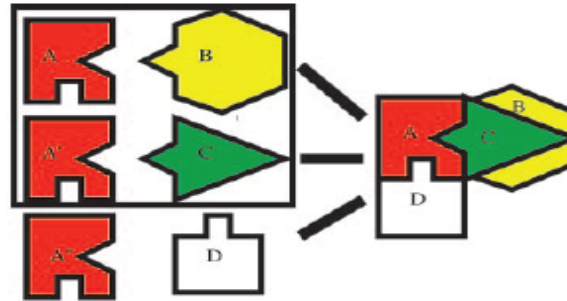
BioGridNet



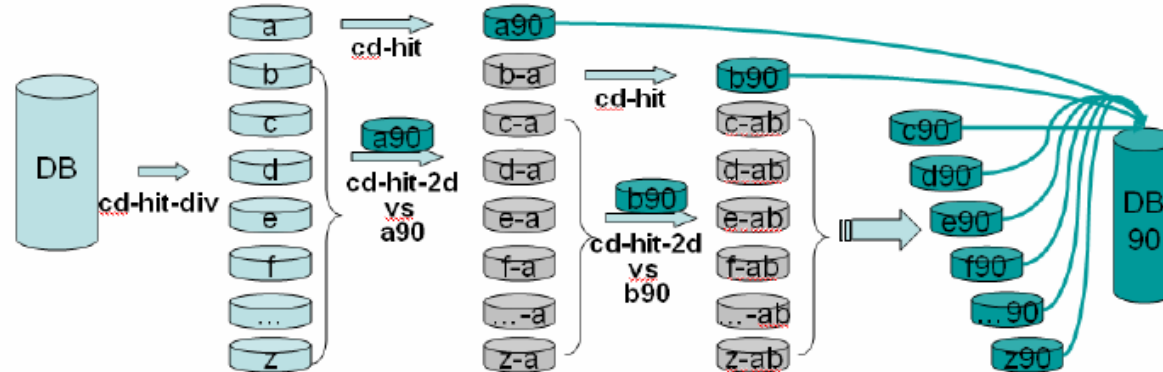
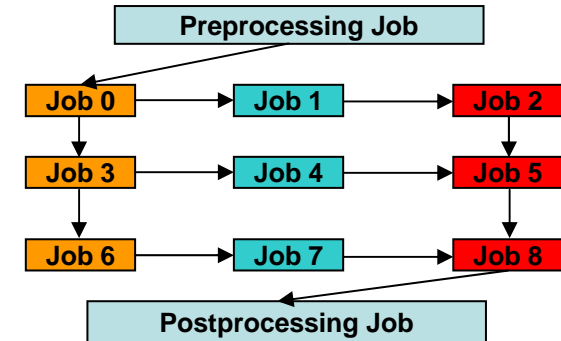
La Suma de Todos

CONSEJERÍA DE EDUCACIÓN
Comunidad de Madrid

www.madrid.org



WORKFLOW



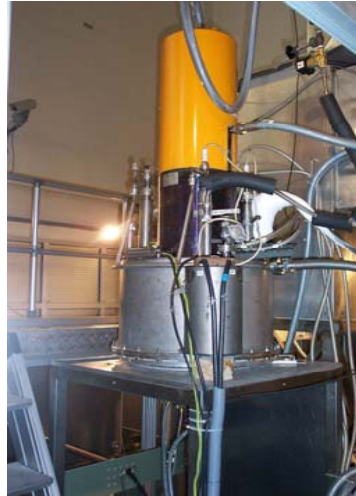
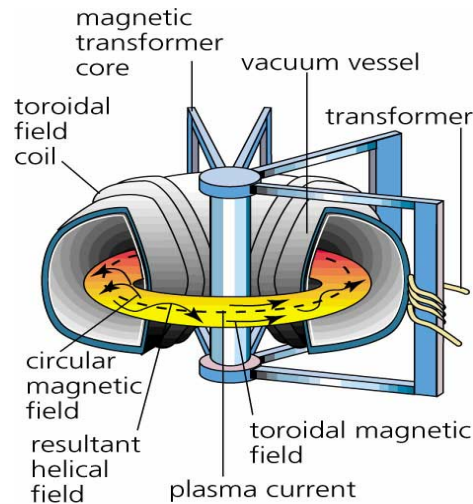
Más información: <http://www.cnio.es/>



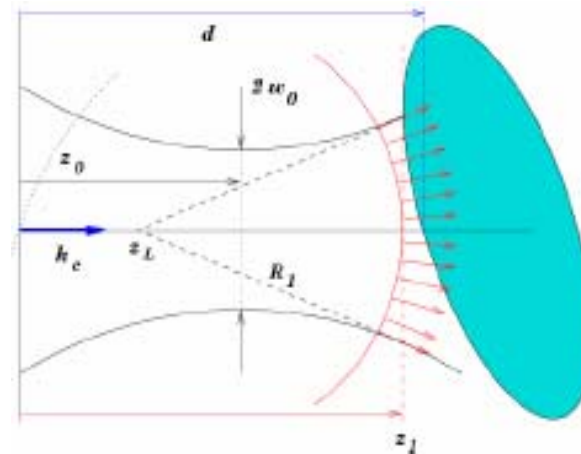
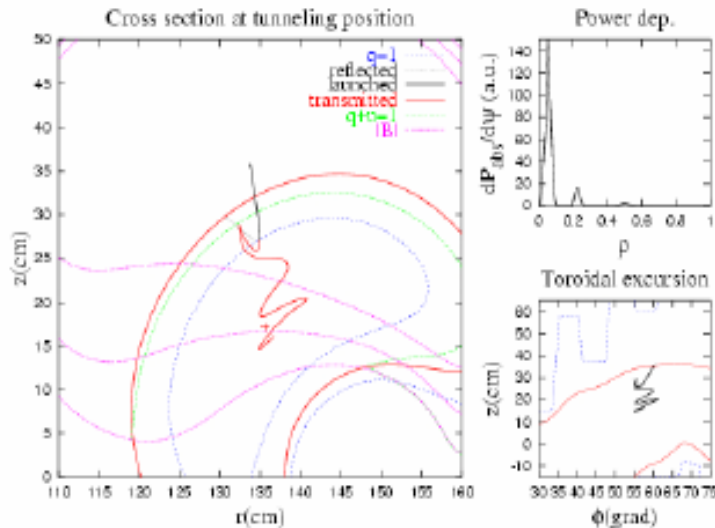
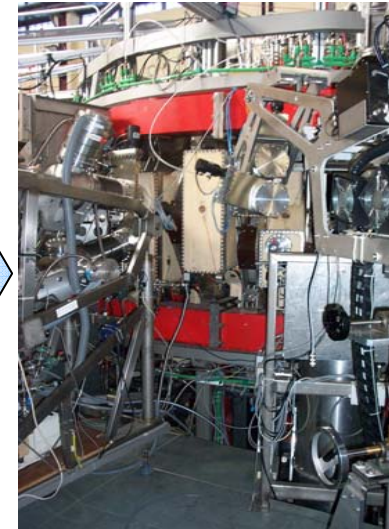
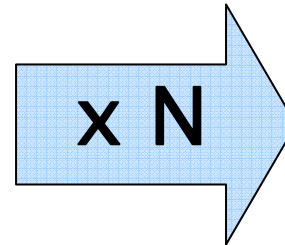
4. Migración de Aplicaciones al Grid

4.3. Ejemplos de Aplicaciones Portadas al Grid

Massive Ray Tracing in Fusion Plasmas



EGEE
Enabling Grids for
E-science in Europe

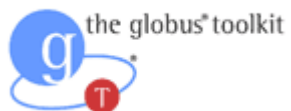


Beam Simulation:

Bunch of rays with beam waist close to the critical layer (100-200 rays) x (100-200 wave numbers) $\sim 10^5$

GridCafé

gridcafe.web.cern.ch/gridcafe/



www.globus.org



www.eu-egee.org



www.ibm.com



www.ogf.org