

21 de Marzo de 2003

Las tecnologías de la información y las comunicaciones al servicio de la investigación

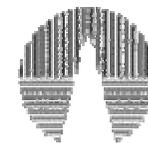
Relevancia y Aplicaciones de las TIC en Biomedicina y Biotecnología

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

UCM



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



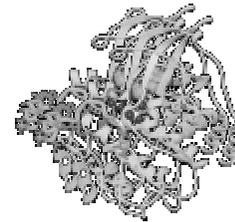
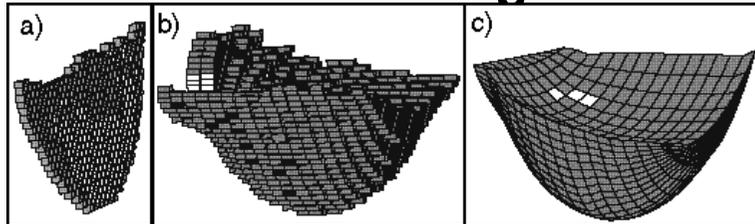
Laboratorio de Computación Avanzada, Simulación y Aplicaciones Telemáticas
Centro de Astrobiología CSIC/INTA
Asociado al NASA Astrobiology Institute

Presentar la tecnología Grid como plataforma para satisfacer las demandas computacionales de las aplicaciones en Biomedicina y Biotecnología

- **Aplicaciones en Biomedicina y Biotecnología**
- **Demandas de las aplicaciones**
- **Limitaciones de la plataforma actual**
- **El Grid como plataforma futura**
- **Componentes del Grid**
- **El papel de la red de comunicación en el Grid**

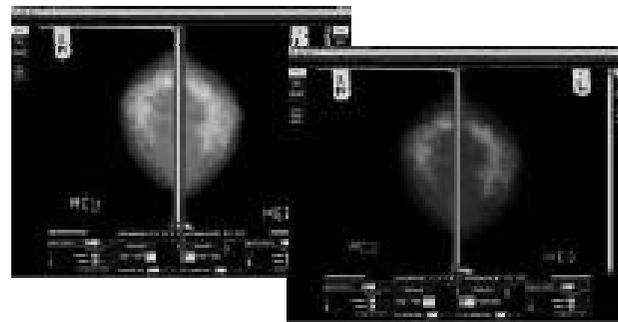
Simulaciones de fármacos contra el SIDA, Anthrax...

Tratamiento de imágenes



Acceso a **bases de datos distribuidas** para realizar el mejor diagnóstico

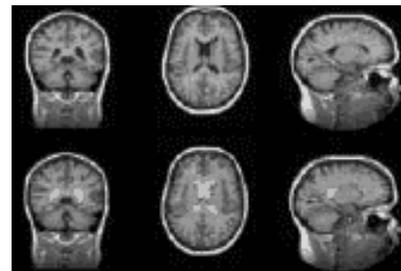
Comparación de secuencias



Visualización

Telemedicina

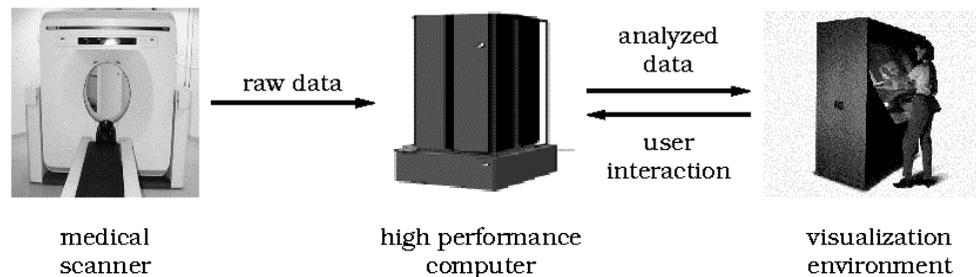
Entornos de colaboración



Acceso a los siguientes servicios y recursos:

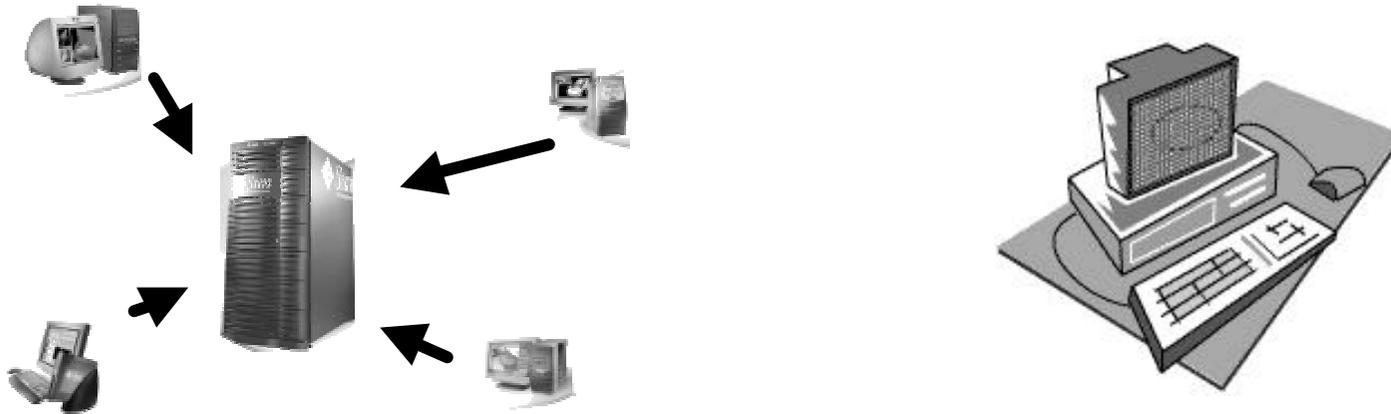
- **Alto Rendimiento**
- **Alta Productividad**
- **Bases de datos distribuidas geográficamente**
- **Entornos de colaboración**
- **Visualización**

Unión de todos los recursos anteriores por medio de una red de comunicación con alto **ancho de banda**, **baja latencia** y **calidad de servicio** y siempre garantizando **confidencialidad, eficiencia y fiabilidad**



Solución clásica

- Computación centralizada basada en servidor

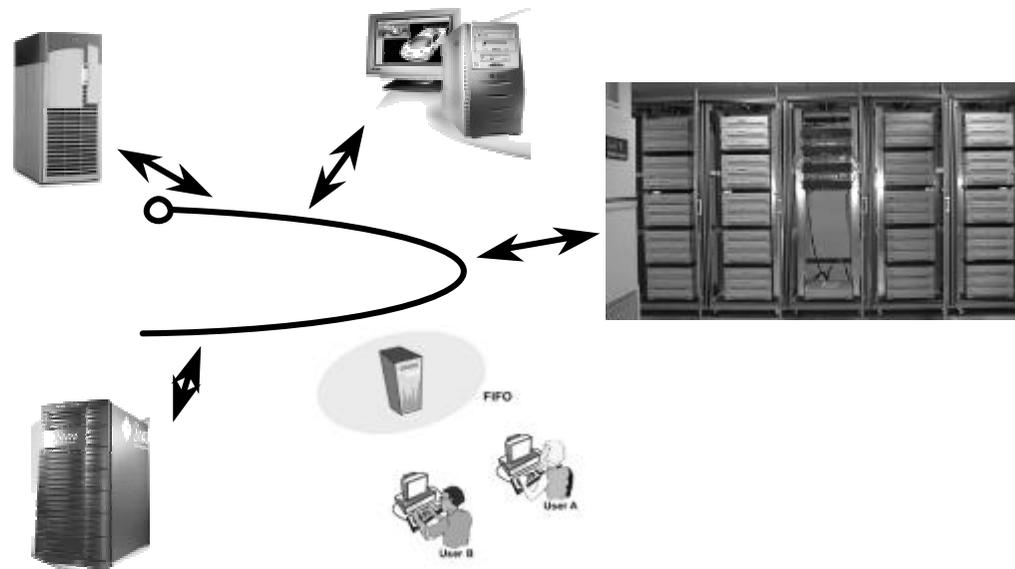


Problemas de la supercomputación basada en servidor

- Falta de escalabilidad
- Equipos muy caros
- Mantenimiento muy caro
- Una vez adquiridos pasan mucho tiempo desaprovechados
- Las demandas de cálculo son puntuales
- Problemas de fiabilidad

Solución distribuida en Intranet

- Utilización de los equipos de una red departamental para ejecutar trabajos por medio de una herramienta de gestión de carga

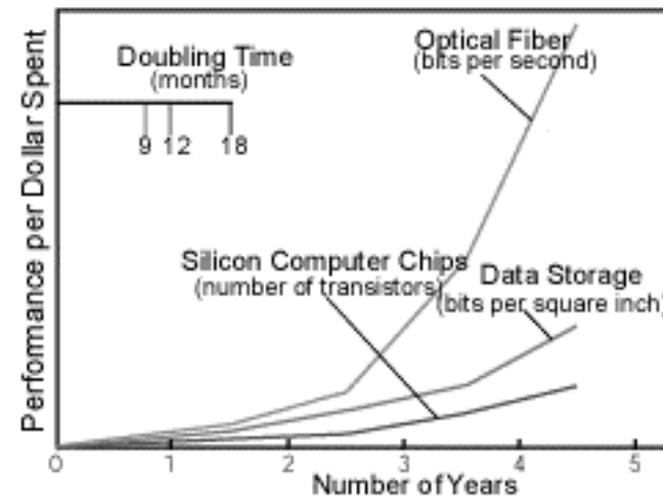


Problemas de la computación distribuida en Intranet

- Escalabilidad limitada a la organización en picos de demanda
- No puedo amortizar mis recursos cuando están desaprovechados
- No puedo compartir recursos con otras organizaciones

Necesidad de más potencia de cálculo

- 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

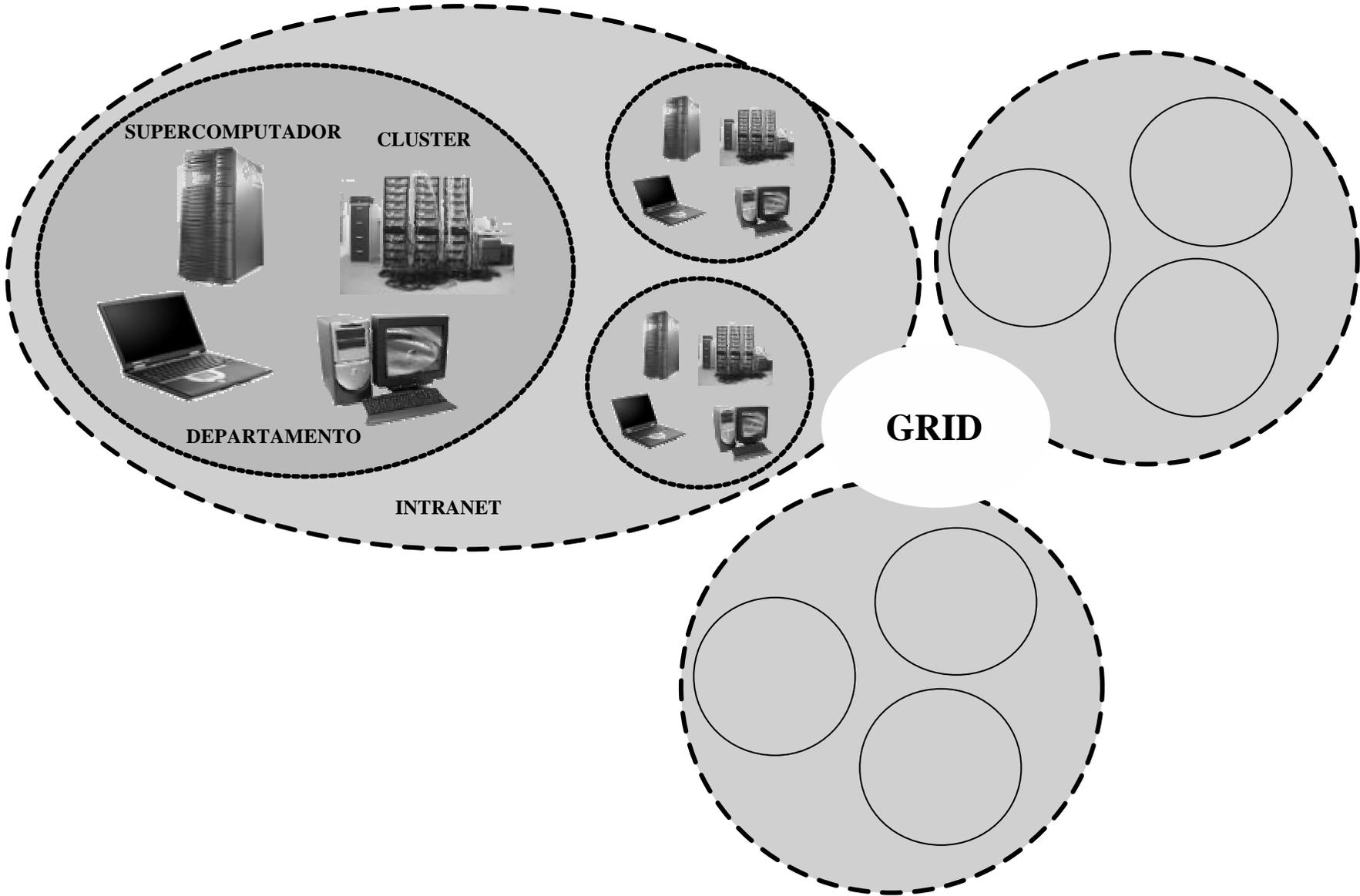
Un único centro no podrá analizar el volumen de información generado

La red permitirá de forma eficiente usar recursos distribuidos

(1 orden de magnitud de diferencia entre procesamiento y red)

*“It’s hard to make
predictions, especially
about the future”*

Yogi Berra



- ✍ Nueva tecnología cuyo objetivo es la **compartición de recursos en Internet** de forma uniforme, transparente, segura, eficiente y fiable

- ✍ Análoga a las **redes de suministro eléctrico**:
 - Ofrecen un único punto de acceso a un conjunto de recursos distribuidos geográficamente **en diferentes dominios de administración** (supercomputadores, clusters, almacenamiento, fuentes de información, instrumentos, personal, bases de datos...)

- ✍ **La tecnología Grid es complementaria a las anteriores**
 - Permite interconectar recursos en diferentes dominios de administración respetando sus políticas internas de seguridad y su software de gestión de recursos en la Intranet
 - **Acceso a grandes bases de datos distribuidas, recursos de cálculo y visualización**

Beneficios

- ✍ Alquiler de recursos
- ✍ Amortización de recursos propios
- ✍ Gran potencia de cálculo a precio bajo sin necesidad de adquirir equipamiento
- ✍ Mayor colaboración y compartición de recursos entre varios centros
- ✍ Creación de organizaciones virtuales
- ✍ Negocios basados en proveer recursos

Relevancia

✍ **Inversión EU FP6:** 300 millones de euros

✍ **Empresas involucradas:** IBM, HP, Sun Microsystems...

Aplicaciones

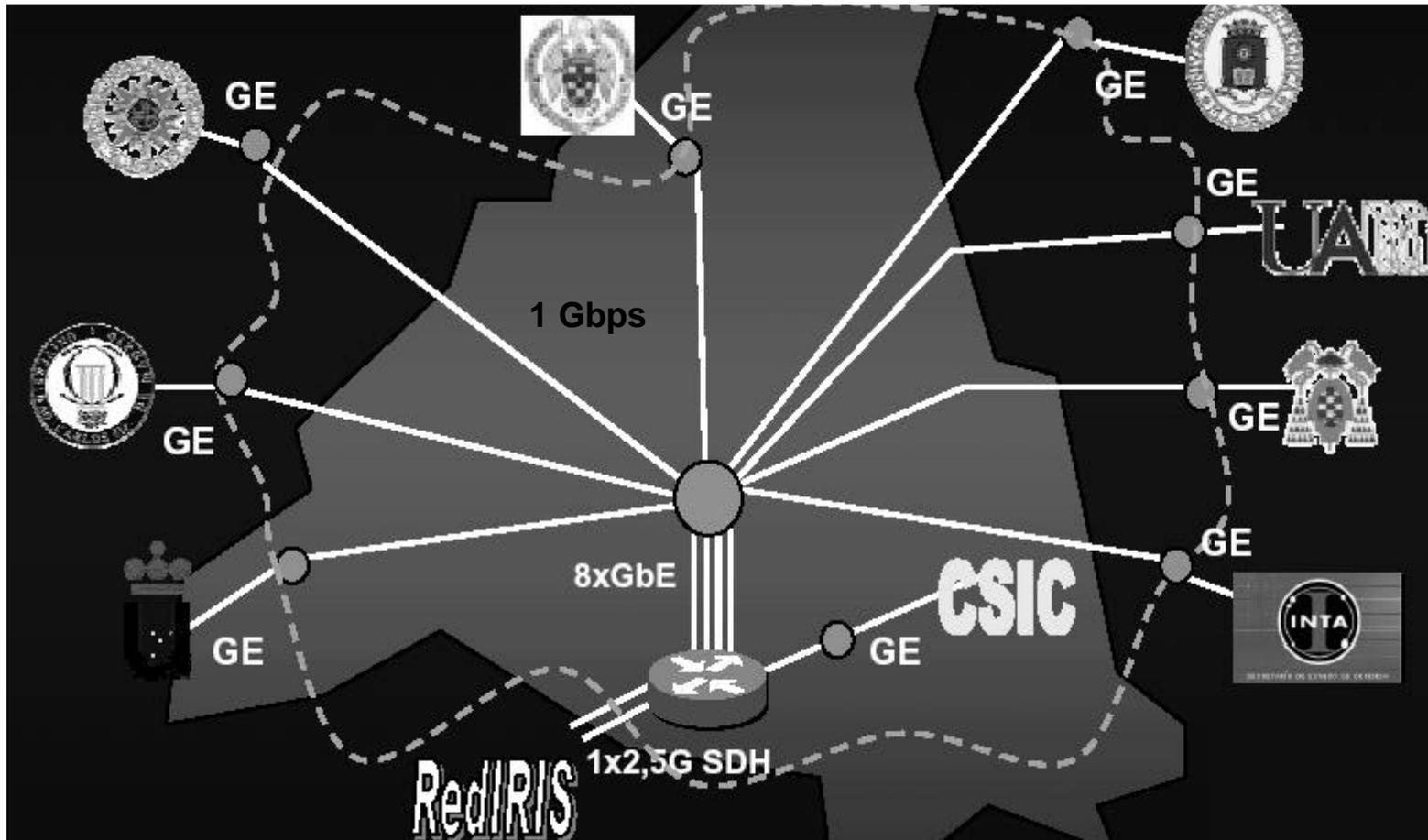
DRM Servicios y Herramientas Grid Status
Cactus MPI globusrun GridWay Nimrod/G Condor-G

GASS Metacomputing Globus GRAM
GridFTP Directory Security Replica I/O
Service Interface Catalog

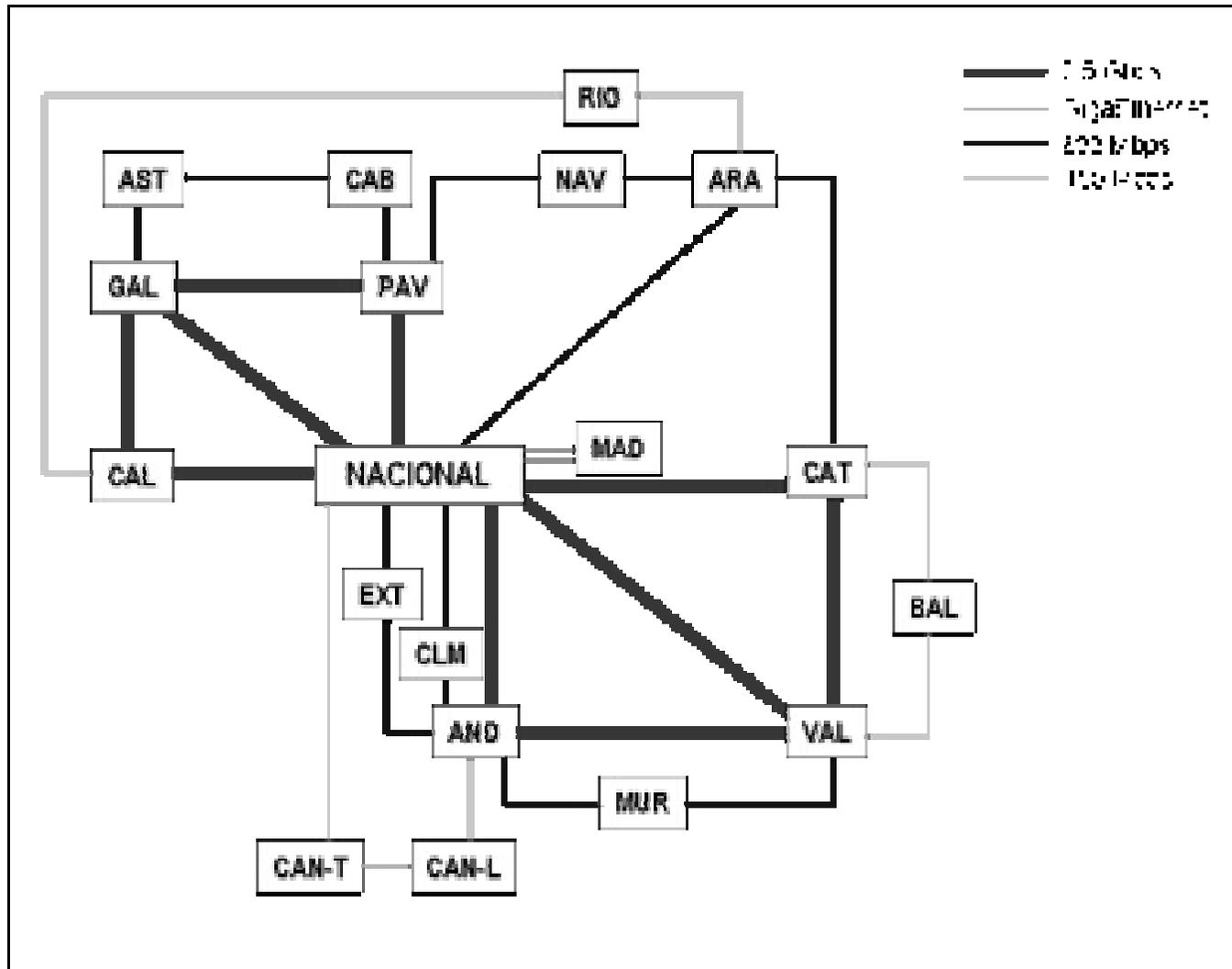
Condor MPI Servicios TCP UDP
LSF PBS locales Linux AIX Solaris
SGE

RED

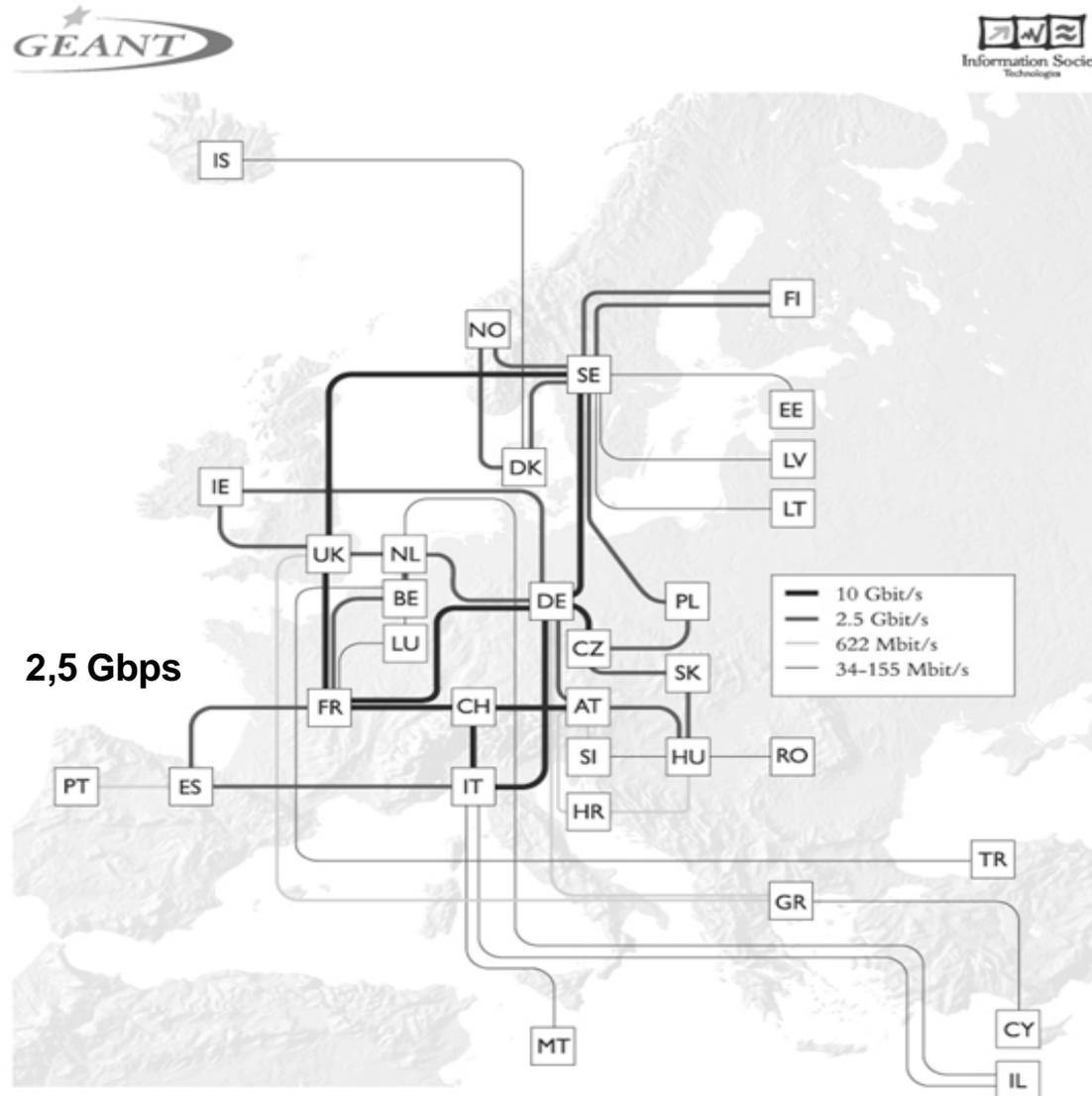
Conexión en la Comunidad de Madrid por medio de la nueva Red Telemática de Investigación



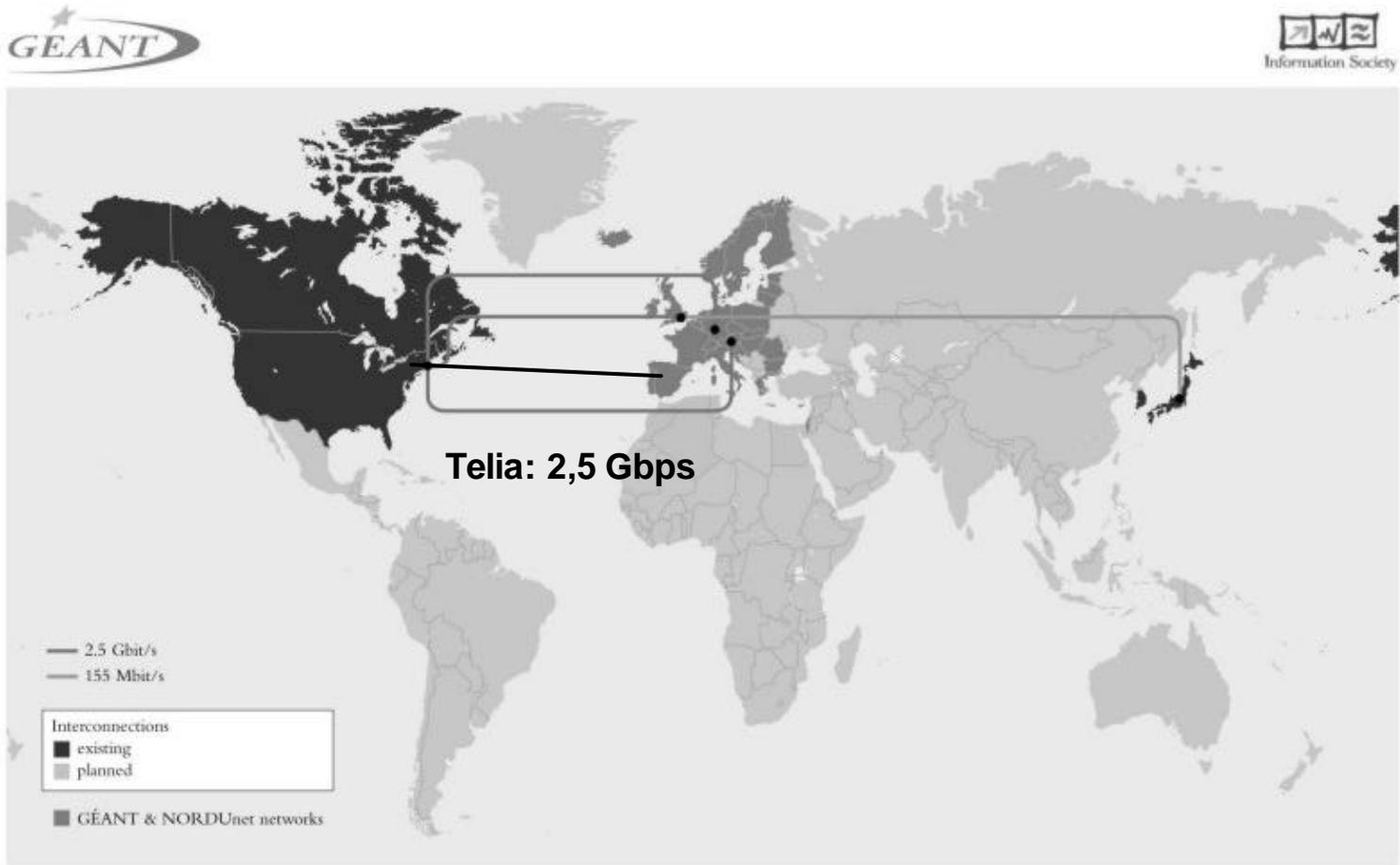
Conexión nacional por medio de Rediris-2



Conexión por medio de la nueva red paneuropea Geant



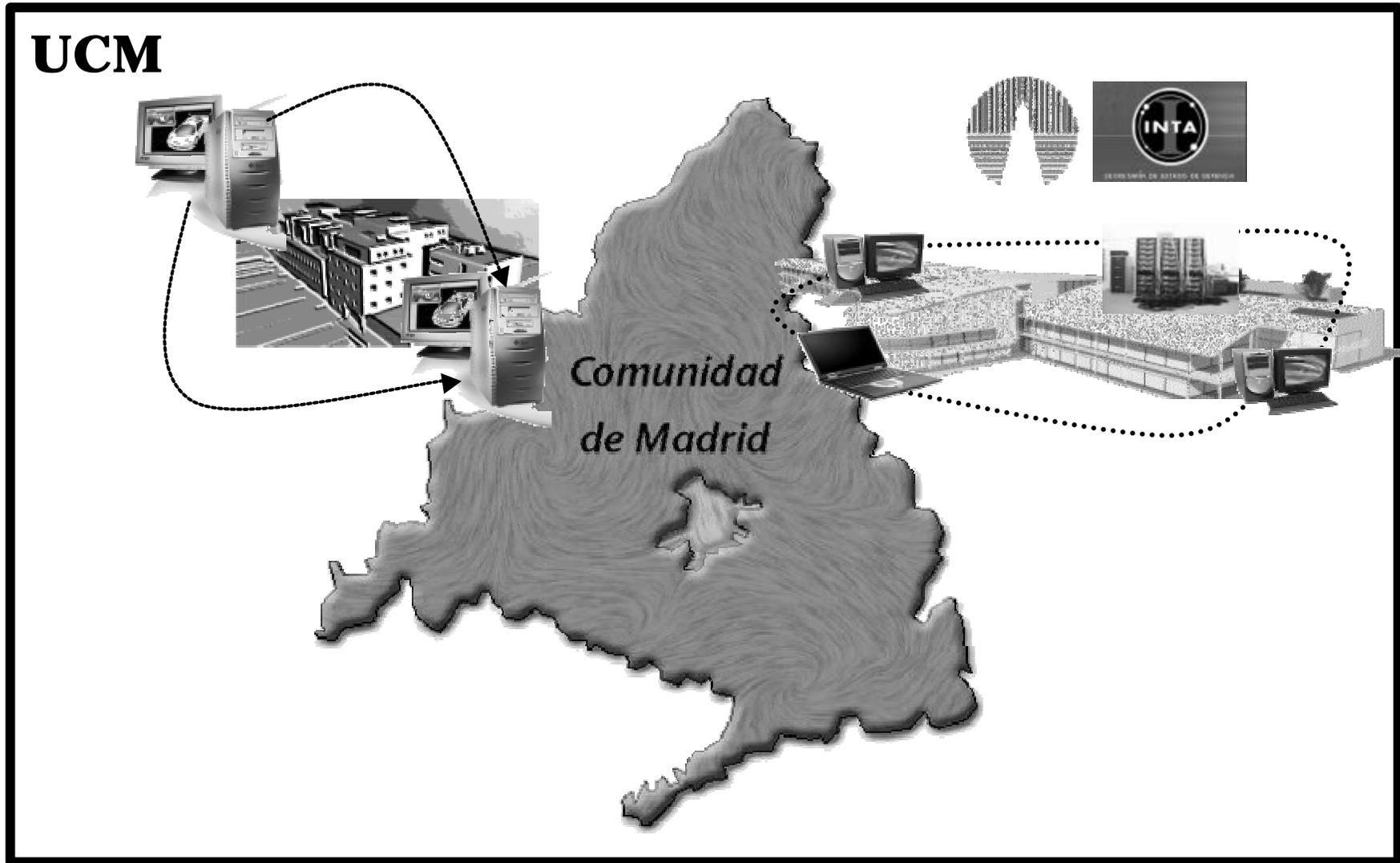
Estados Unidos Geant



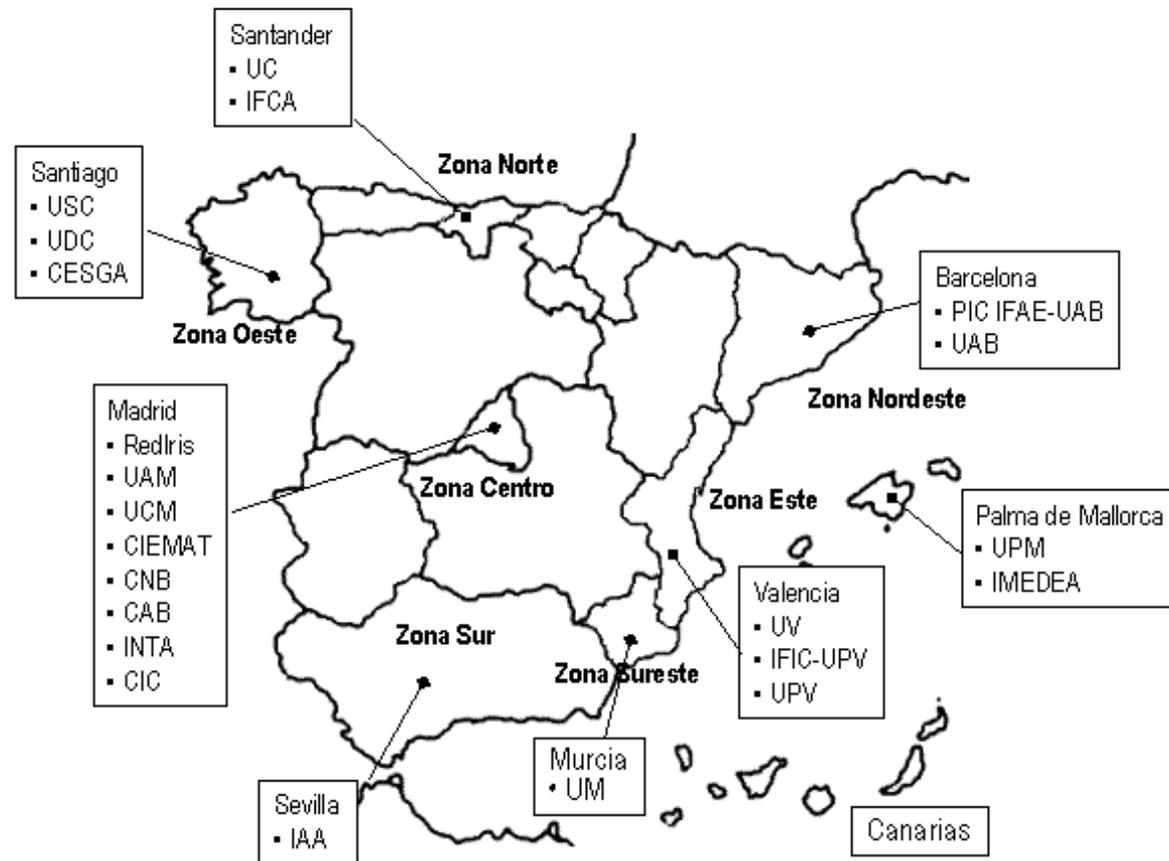
 Multi-Gigabit pan-European Research Network
Global Connectivity August 2002


DANTE
www.dante.net

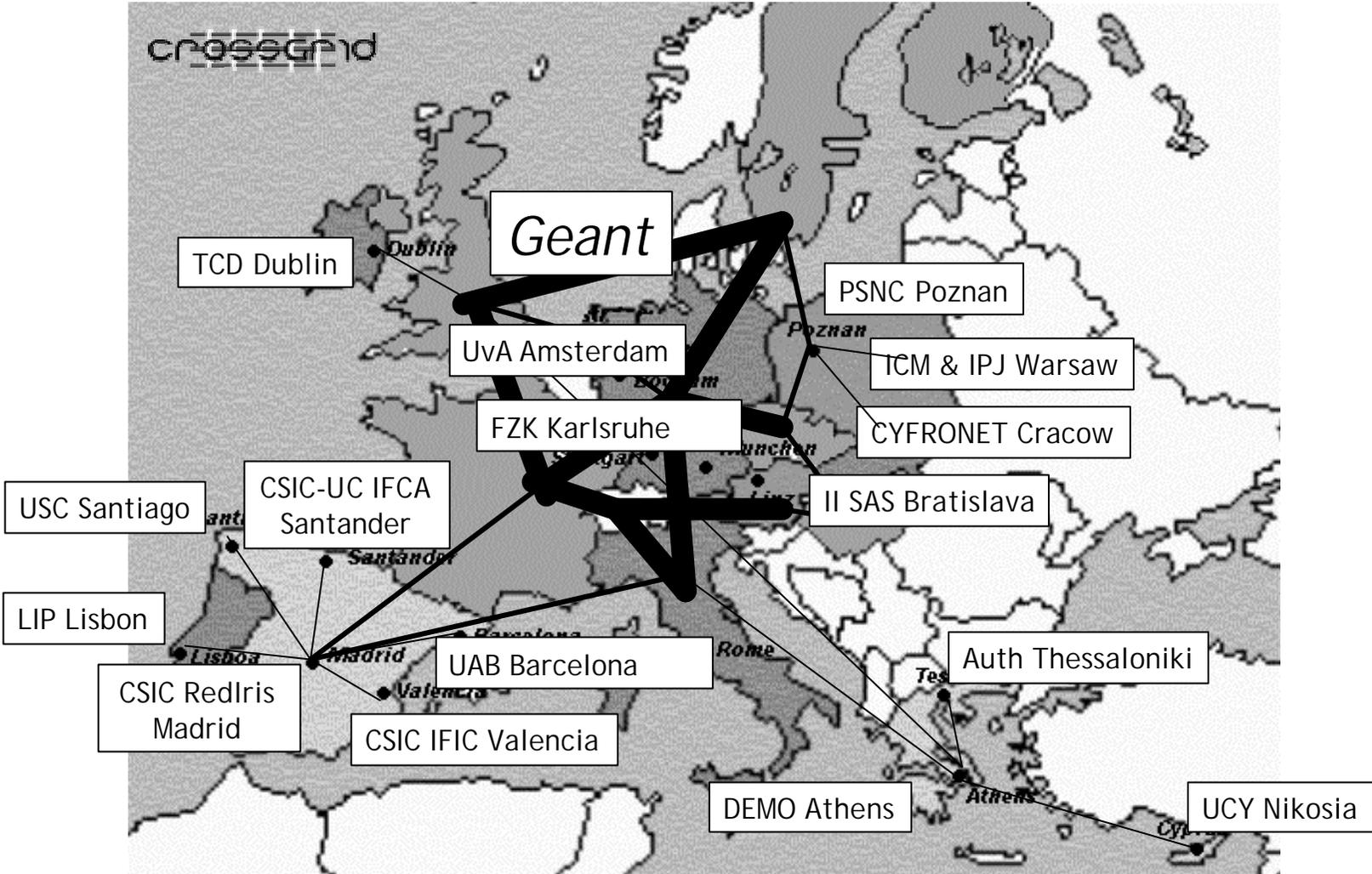
www.cab.inta.es/~CABGrid/



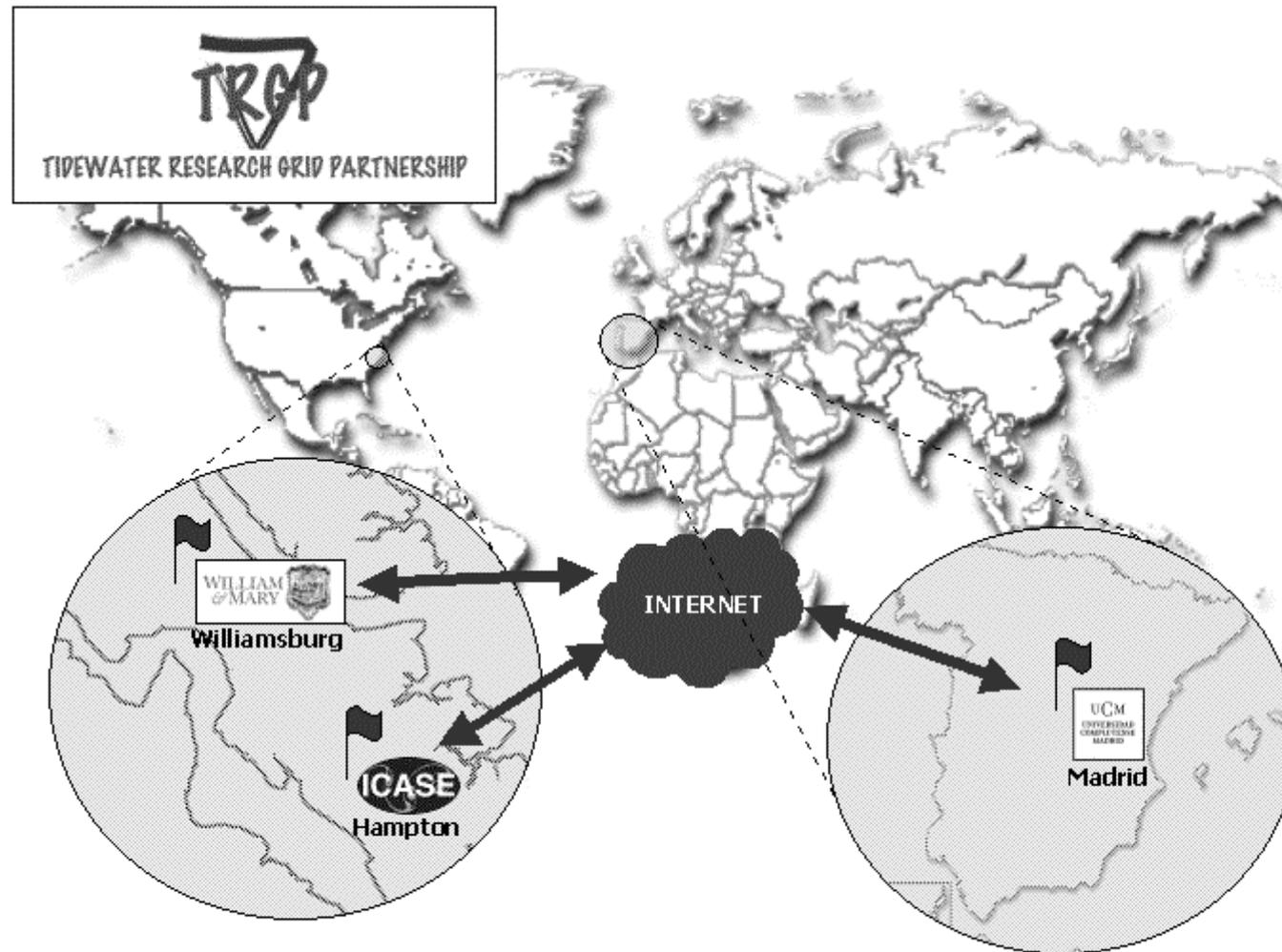
www.rediris.es/gt/iris-grid/



www.crossgrid.org



www.tidewaterrgp.org



“Cuando Internet sea tan rápido como los buses internos de un computador, éste se desintegrará en la red en un conjunto de recursos de propósito específico”

Gilder Technology Report, junio 2000

