

Soluciones para entornos HPC

Dr. Abel Francisco Paz Gallardo.
IT Manager / Project Leader @ CETA-Ciemat
abelfrancisco.paz@ciemat.es

V Jornadas de Supercomputación y Avances en Tecnología



GOBIERNO
DE ESPAÑA

MINISTERIO
DE ECONOMÍA
Y COMPETITIVIDAD



CENTRO EXTREMEÑO DE
TECNOLOGÍAS AVANZADAS

CETA Ciemat



FEDER

Fondo Europeo de Desarrollo Regional

Una manera de hacer Europa



- 1 HPC... ¿Qué? ¿Cómo?
- 2 Computación (GPGPU, UMA/NUMA, etc.)
- 3 Almacenamiento en HPC
- 4 Gestión de recursos

Soluciones para entornos HPC



GOBIERNO
DE ESPAÑA

MINISTERIO
DE ECONOMÍA
Y COMPETITIVIDAD



CENTRO EXTREMEÑO DE
TECNOLOGÍAS AVANZADAS

CETA Cimat



FEDER

Fondo Europeo de Desarrollo Regional

Una manera de hacer Europa

1 HPC... ¿Qué? ¿Cómo?

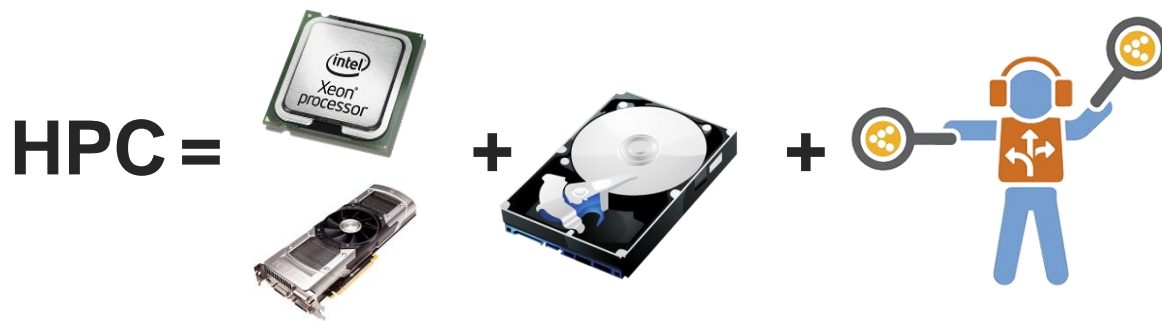
¿Qué?

- **HPC – High Performance Computing** (Computación de alto rendimiento)
- **Objetivo principal:** Resolución de problemas complejos

¿Cómo?

Tres pilares fundamentales:

- Procesamiento = Cómputo
- Almacenamiento
- Gestión de recursos

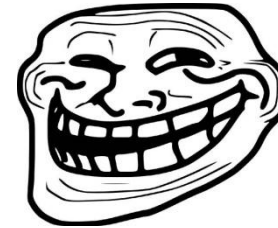


2 Computación (GPGPU, NUMA, etc.)

¿Qué es una GPU?

Primera búsqueda en 2006:

- Gas Particulate Unit → Unidad de partículas de gas ¿?



GPU = Graphics Processing Unit (Unidad de procesamiento gráfico).



FFFFFFF
FFFFFFF
FFFFFFF
FFFUU
UUUU
UUUU
UUUU
UUUU
UUUU



2 Computación (GPGPU, NUMA, etc.)

La cuestión es...

Si una GPU en un videojuego procesa miles de polígonos, texturas y sombras en tiempo real...



¿Por qué no utilizar esta tecnología para procesamiento de datos?

CHALLENGE ACCEPTED



GOBIERNO DE ESPAÑA

MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD



CENTRO EXTREMEÑO DE TECNOLOGÍAS AVANZADAS

CETA Ciemat

Una manera de hacer Europa

CETA-Ciemat/ Noviembre 2012

Soluciones para entornos HPC

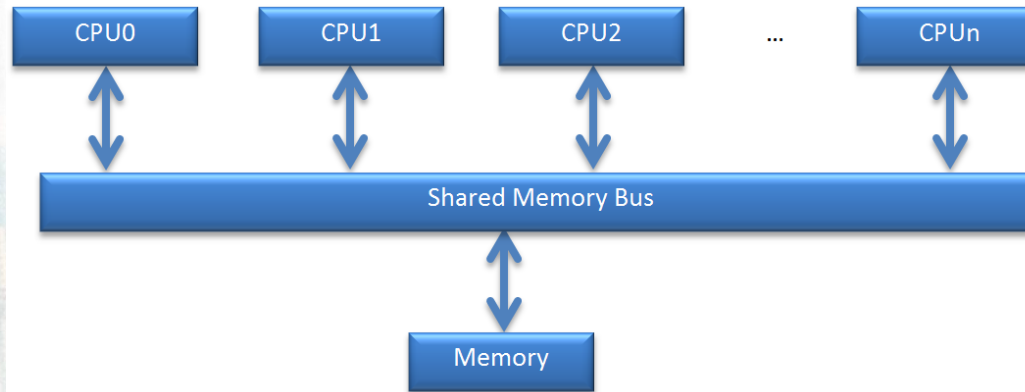
Abel Francisco Paz Gallardo

EXTREMADURA RESEARCH CENTER FOR ADVANCED TECHNOLOGIES

2 Memoria compartida

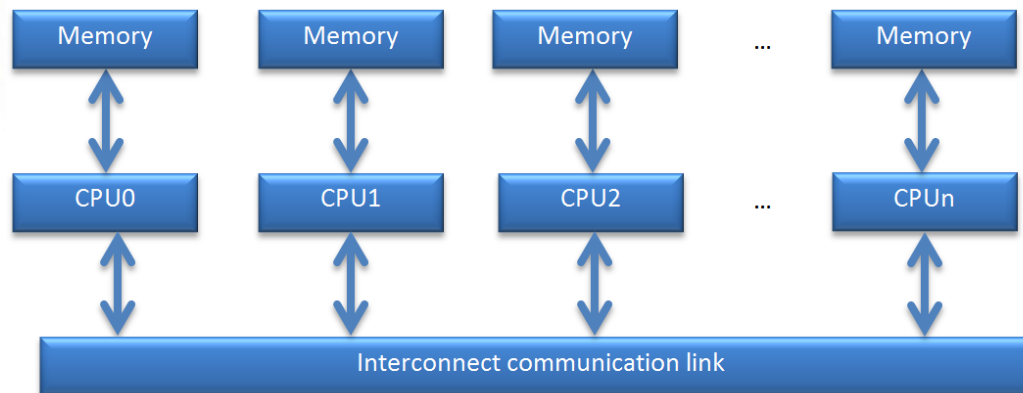
Memoria compartida (Multiprocesadores)

- UMA – Uniform memory access



- Memoria central compartida de forma uniforme
- Límite de escalabilidad según bus

- NUMA – Non Uniform Memory Access



- Memoria distribuida
- Fácil de escalar
- Cada CPU tiene su memoria, caché y E/S

Imgs by **Iakovos Panourgias** - NUMA effects on multicore, multi-socket systems



GOBIERNO DE ESPAÑA

MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD



CENTRO EXTREMEÑO DE TECNOLOGÍAS AVANZADAS

CETA Ciemat

Una manera de hacer Europa

CETA-Ciemat/ Noviembre 2012

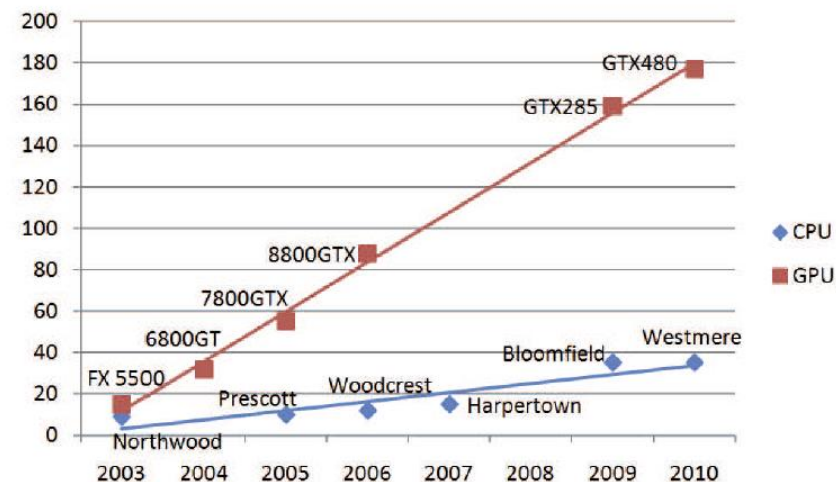
Soluciones para entornos HPC

Abel Francisco Paz Gallardo

2 Computación (GPGPU, NUMA, etc.)

¿Pero entonces no podemos usar sólo la CPU o GPU??

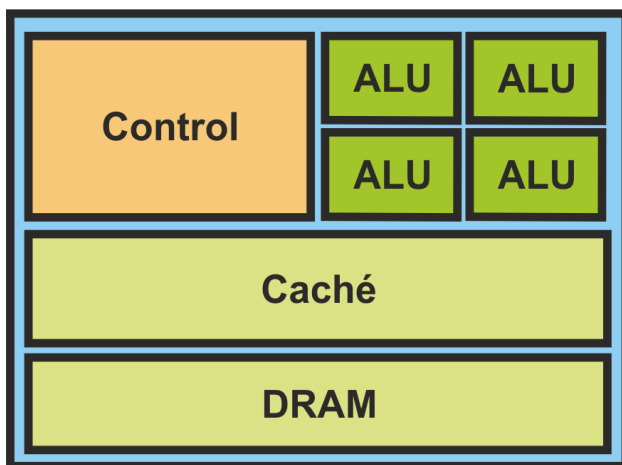
- ◆ Comparativas constantes, gráficas, etc.
#FAIL! pero EPIC!
- ◆ Programación diferente, pero no tanto
Ej. CUDA frente a OpenMP
- ◆ Tenemos información de ambas arquitecturas



La clave está en sacar el máximo partido de ambas, ¡son complementarias!

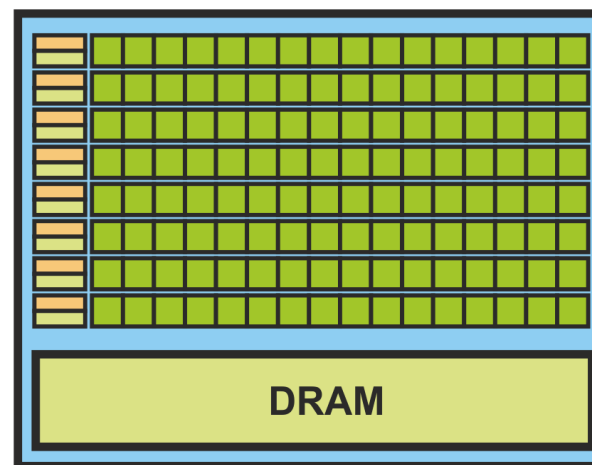
- + Cache
- + Control
- Cores

CPU

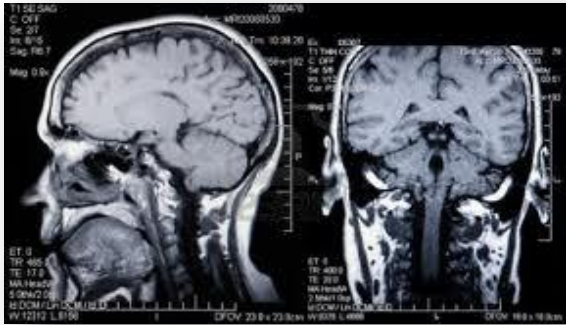


- + Cores
- Cache
- Control

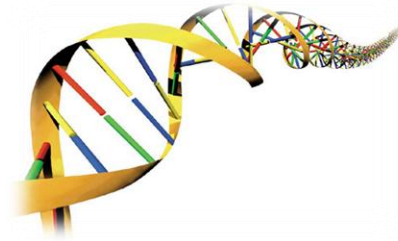
GPU



Procesamiento de Imágenes



Bioinformática



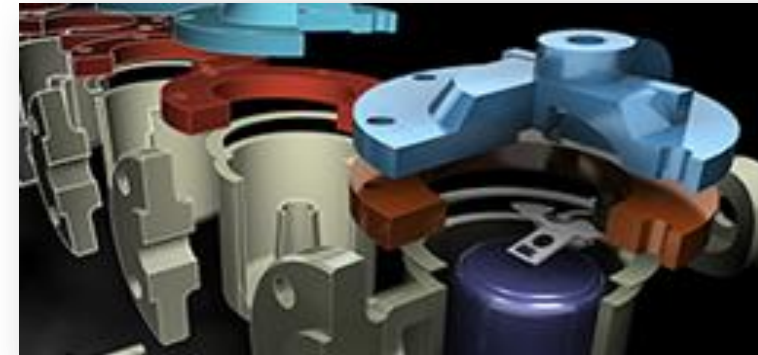
Cálculo financiero



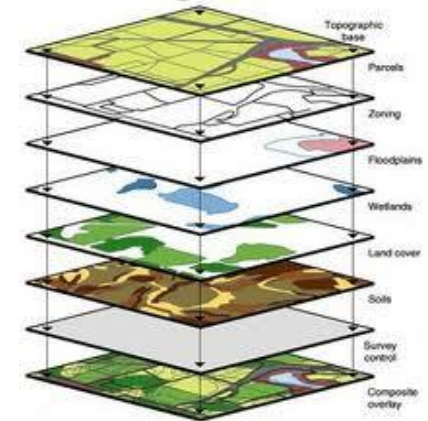
Animación



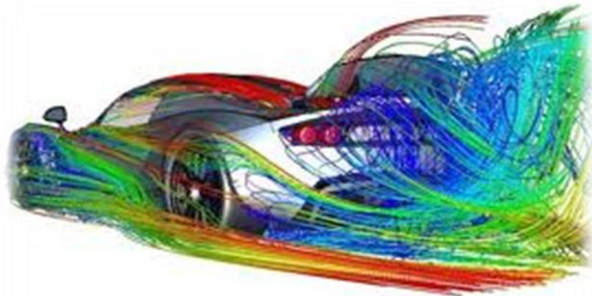
Diseño/Confeción/Manufacturación Asistida por ordenador



Sistemas de Información Geográfica



Dinámica de Fluidos



GOBIERNO DE ESPAÑA

MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD



CENTRO EXTREMEÑO DE TECNOLOGÍAS AVANZADAS

CETA Ciemat

CETA-Ciemat/ Noviembre 2012

Soluciones para entornos HPC
Abel Francisco Paz Gallardo

EXTREMADURA RESEARCH CENTER FOR ADVANCED TECHNOLOGIES

- La secuenciación y acoplamiento de proteínas tienen un alto coste computacional.



Se mejoran los tiempos en el análisis de genoma.



GOBIERNO DE ESPAÑA

MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD



CENTRO EXTREMEÑO DE TECNOLOGÍAS AVANZADAS

CETA Ciemat

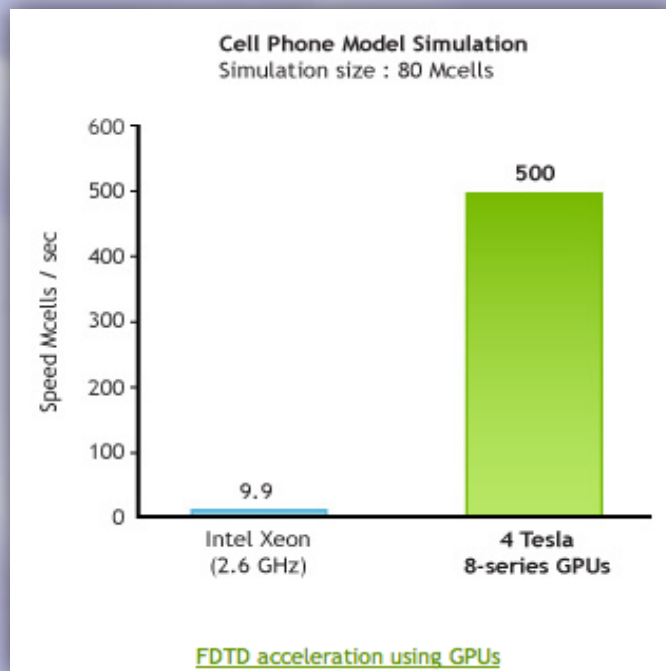
Una manera de hacer Europa

CETA-Ciemat/ Noviembre 2012

Soluciones para entornos HPC
Abel Francisco Paz Gallardo

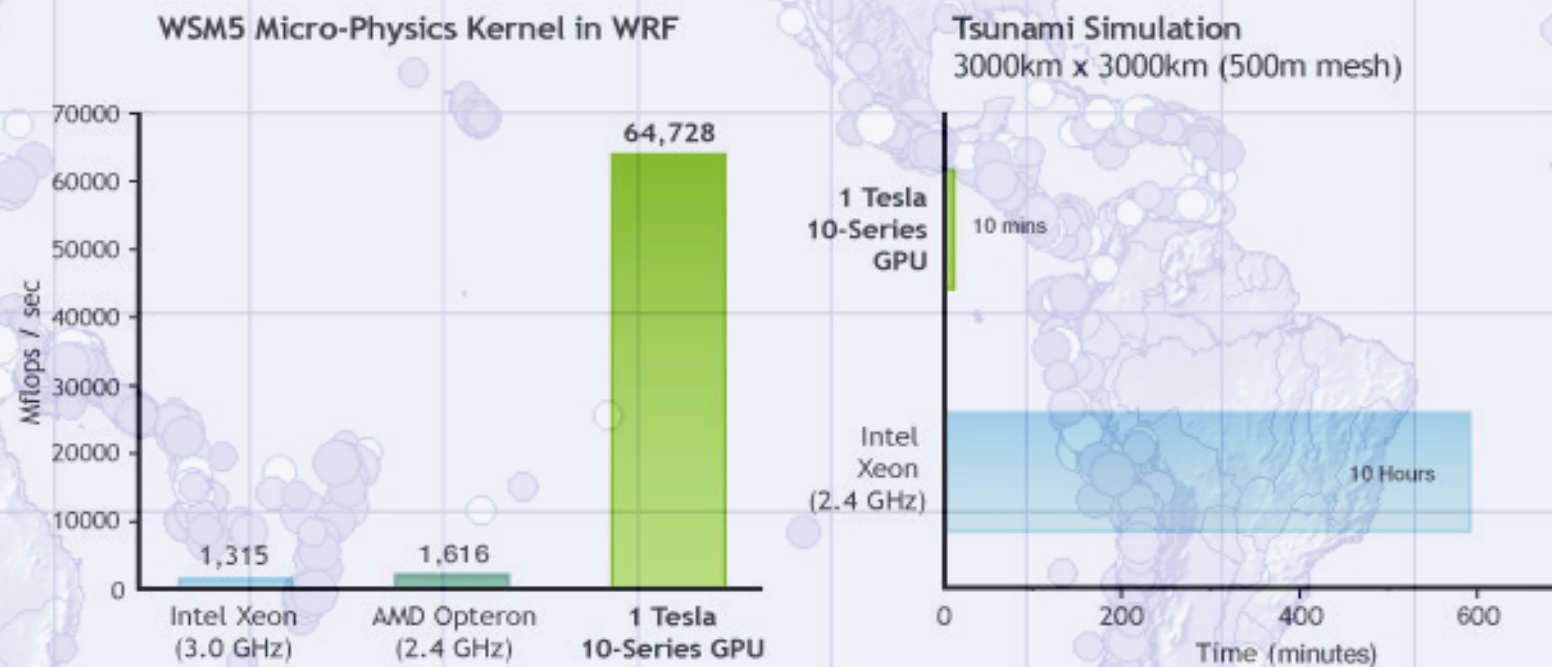
EXTREMADURA RESEARCH CENTER FOR ADVANCED TECHNOLOGIES

- El método de las Diferencias Finitas en el Dominio del Tiempo se utiliza para resolver **problemas electromagnéticos** transitorios utilizando diferencias finitas.
- Utilizado para averiguar **cómo tratar las ondas electromagnéticas y su interacción con los materiales.**



2 Soluciones GPGPU – Dinámica de fluidos

- Las aplicaciones de **modelado de cambio climático** y de **simulación de tsunamis** obtienen aceleraciones, llegando a obtener resultados en tiempo real.



GOBIERNO DE ESPAÑA

MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD



CENTRO EXTREMEÑO DE TECNOLOGÍAS AVANZADAS

CETA Ciemat

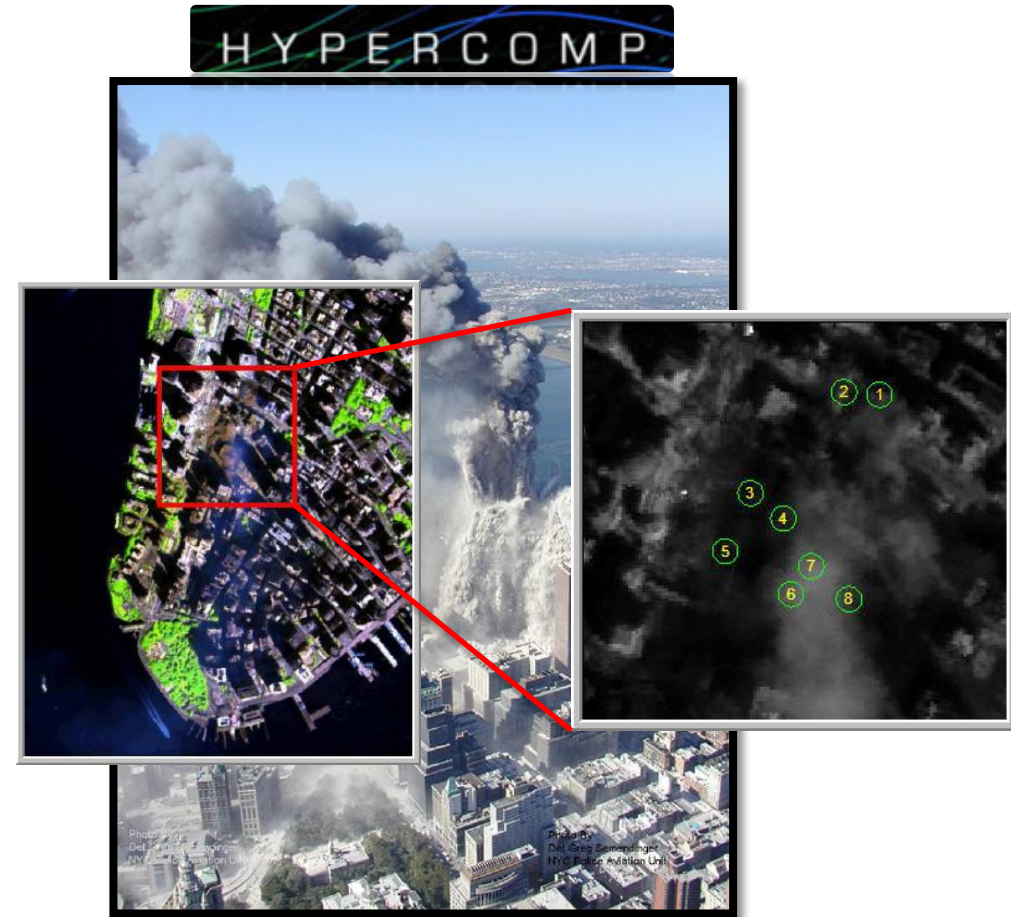
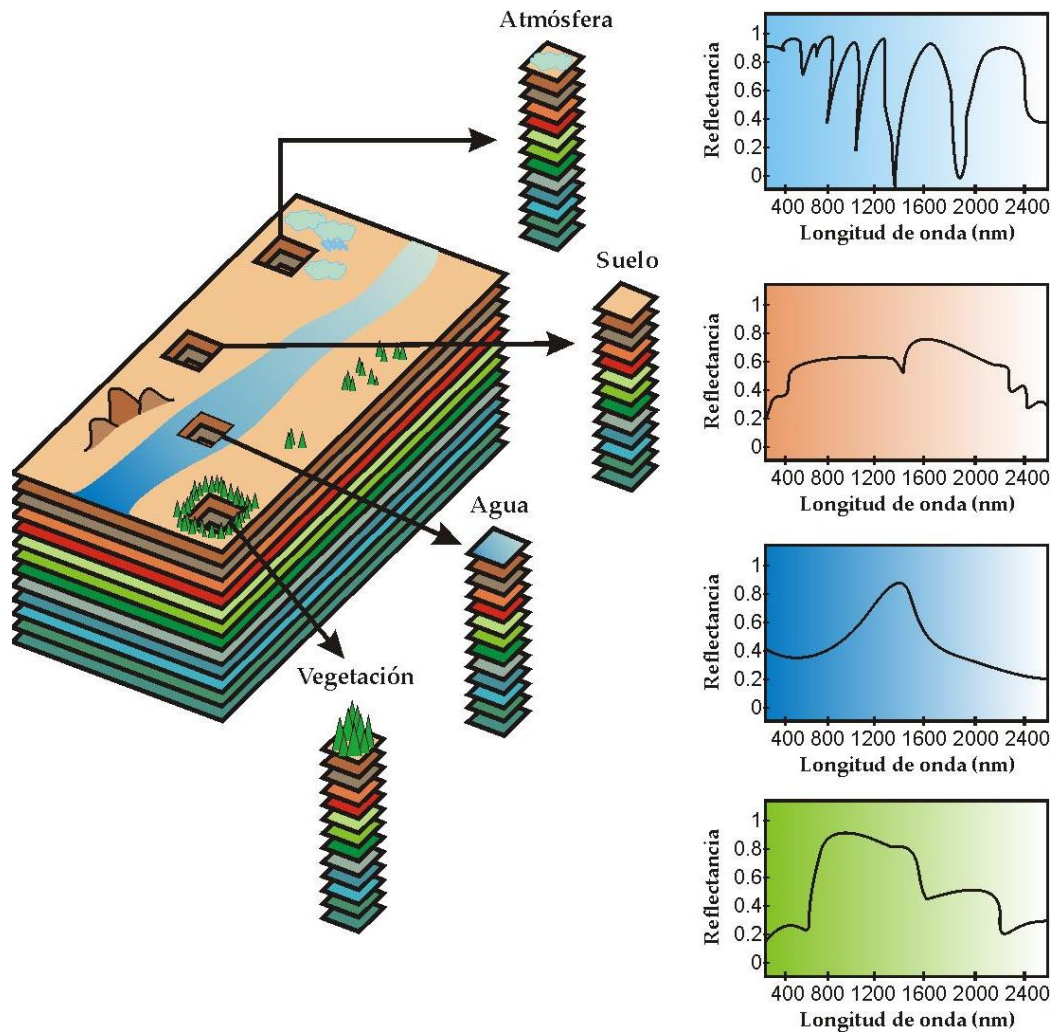
Una manera de hacer Europa

CETA-Ciemat/ Noviembre 2012

Soluciones para entornos HPC
Abel Francisco Paz Gallardo

EXTREMADURA RESEARCH CENTER FOR ADVANCED TECHNOLOGIES

Tesis: DISEÑO E IMPLEMENTACIÓN DE ALGORITMOS PARALELOS DE ANÁLISIS DE IMÁGENES HIPERESPECTRALES EN TARJETAS GRÁFICAS PROGRAMABLES (GPUS)



CPU: 78 segundos

GPU: 2 segundos (40x, Tiempo Real)

Nodo de memoria compartida

- 2 x Bullx S6030



+



=

**64 cores (128 con HT)
1 TB RAM
(16 GB/core)**

2 x 8 octocore Intel(R) Xeon(R) CPU X7550 @ 2.00GHz

◆ Ideal para:

- Procesamiento de **grandes volúmenes de datos**
- Probablemente la mejor arquitectura para procesar datos con **alta dependencia de datos** entre sí

Otros centros:

Acceso uniforme (UMA)

LUSITANIA (CénitS): 256 CPU cores, 2 TB RAM
CESMU: 128 CPU cores, 1,5 TB RAM
CESGA: 128 CPU cores, 0,4 TB RAM

Acceso no uniforme (NUMA)

Finisterrae (CESGA) 144 CPU cores, 2 TB RAM
FC_SCL: 256 CPU cores, 3 TB RAM
CETA-Ciemat: 64 CPU cores, 1 TB RAM



GOBIERNO DE ESPAÑA

MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD



CENTRO EXTREMEÑO DE TECNOLOGÍAS AVANZADAS

CETA Ciemat

Una manera de hacer Europa

CETA-Ciemat/ Noviembre 2012

Soluciones para entornos HPC
Abel Francisco Paz Gallardo

EXTREMADURA RESEARCH CENTER FOR ADVANCED TECHNOLOGIES

CETA-Ciemat GPU HPC Cluster

- Funcionando desde Enero 2010
- **113 TFLOPs** en simple precisión y **37 TFLOPs** en doble precisión
- **700TB** almacenamiento
- **Más de 30 grupos de investigación y universidades de todo el mundo lo utilizan**
- **Problemas científicos de alta complejidad**
 - **Dinámica molecular**
 - **Procesamiento de imágenes**
 - **Visión por computador**
 - **Análisis numérico**
 - **Estudios Imagen médica**
 - **Genómica**
 - ...



GOBIERNO DE ESPAÑA

MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD



CENTRO EXTREMEÑO DE TECNOLOGÍAS AVANZADAS

CETA Ciemat

Una manera de hacer Europa

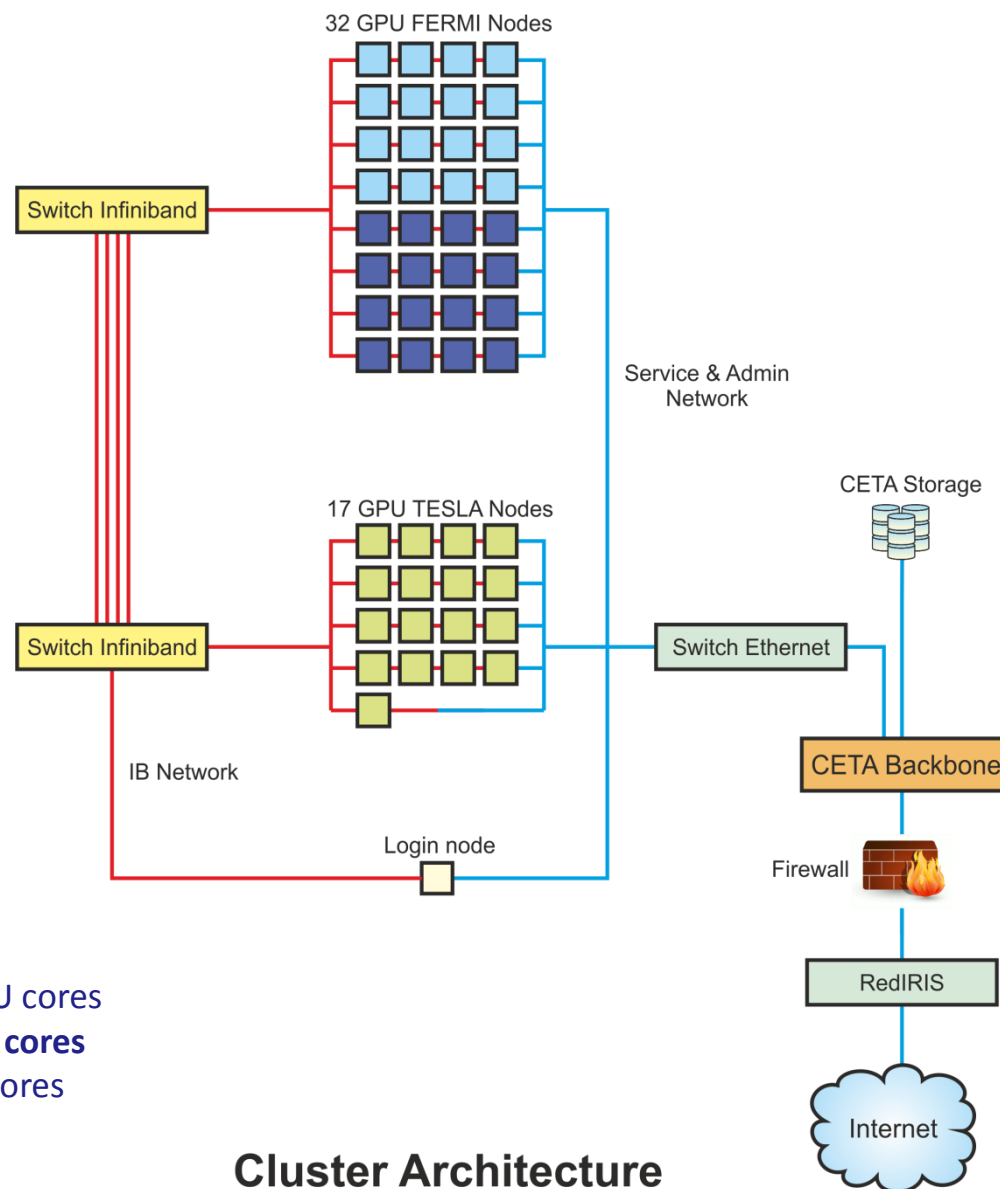
CETA-Ciemat/ Noviembre 2012

Soluciones para entornos HPC
Abel Francisco Paz Gallardo

EXTREMADURA RESEARCH CENTER FOR ADVANCED TECHNOLOGIES

Arquitectura del cluster

- **Red Infiniband:**
 - **Baja latencia**
 - **Comunicación entre nodos**
- **Red de Administración y servicios:**
 - **Administración**
 - **Servicios**
 - **Conectividad con backbone**
- **Uso del cluster:**
 - **Acceso mediante solicitud**
 - **Soporte en la paralelización**



Cluster Architecture

Otros centros:

BSC:	131.072 GPU cores
CETA-Ciemat:	39.712 GPU cores
FCSCCL:	2.880 GPU cores

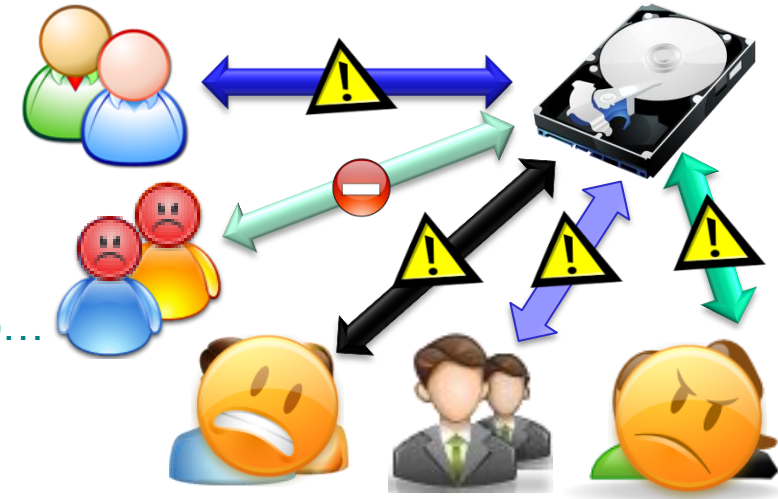
¿Almacenamiento en HPC?

- **Objetivo: Alto rendimiento en operaciones de E/S en disco**

- **Problema con discos normales**

- Discos separados (poco espacio)
- Cuellos de botella si el acceso es compartido

- **Solución 1: Comprar “LA” cabina de almacenamiento...**



... y su software y mantenimiento anual!



- **Solución 2: Distributed Parallel File System (Sistema de ficheros paralelo distribuido)**

Lustre

- **Top500:** 15 de los 30 primeros superordenadores más potentes del mundo utilizan Lustre.
 - Titan [1], el número 1 del Top500 del mundo: **240GB/s de transferencia** (SATA 3 sólo ofrece 6Gb/s!)
- **OpenSource (GNU GPL)**
- **Altamente escalable**
 - Decenas de miles de clientes simultáneos
 - Decenas de Petabytes, incluso Exabytes (K computer, FEFS – Fujitsu Exabyte File System [2])

Lustre en el CETA-Ciemat

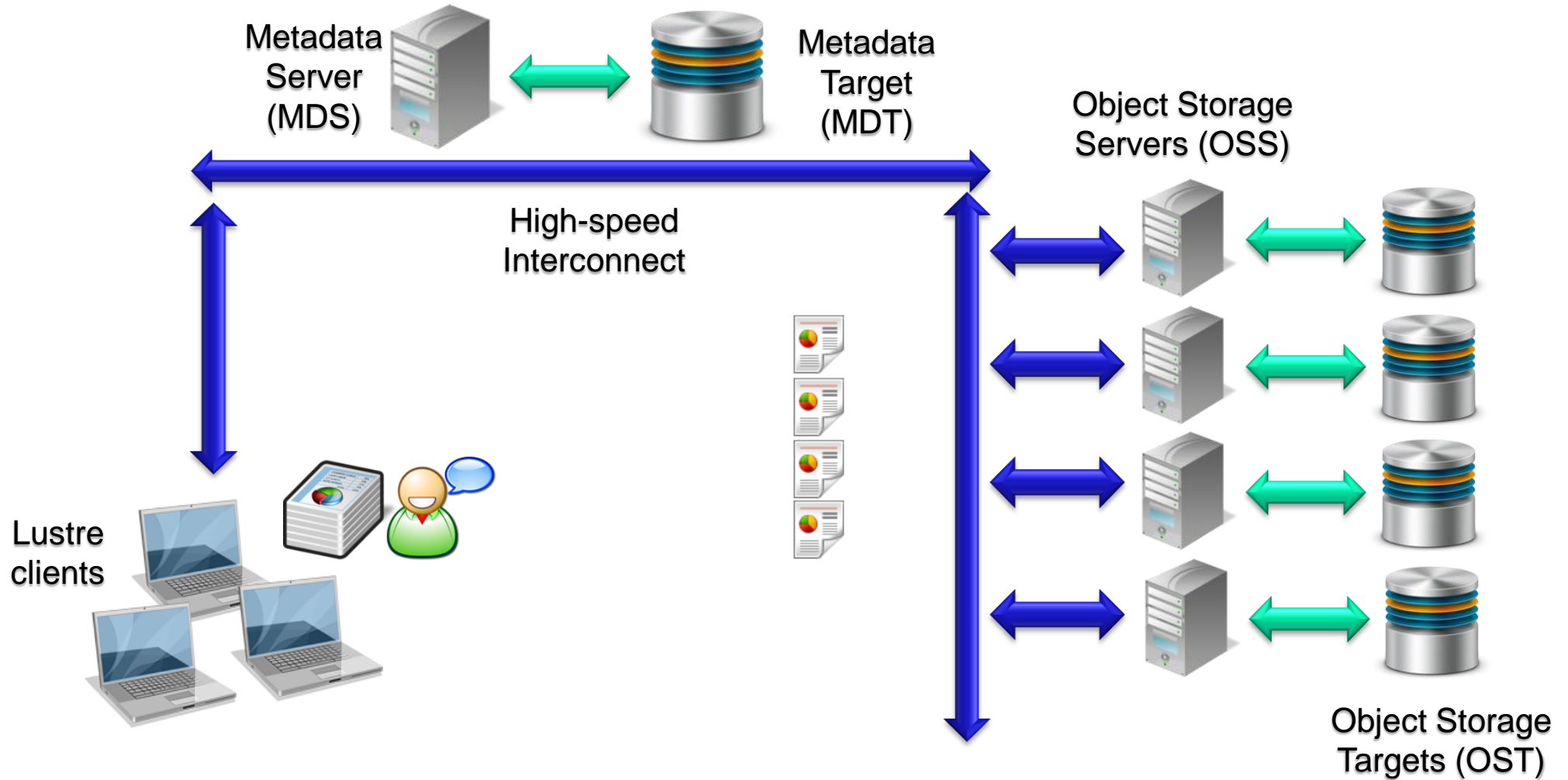
- **Mejorado: con tolerancia a fallos (Distributed parallel fault-tolerant file system)**
 - **Alta disponibilidad en toda su arquitectura**
 - **Robusto: Seguridad de los datos a nivel físico (RAID) + Nivel lógico**
 - **Recientemente actualizado a versión 2.2**
 - **Red GigaEthernet – 20Gbps**

[1] Titan - [http://en.wikipedia.org/wiki/Titan_\(supercomputer\)](http://en.wikipedia.org/wiki/Titan_(supercomputer))

[2] FEFS - <http://phys.org/news/2011-10-fujitsu-high-performance.html>

¿Cómo funciona Lustre?

- Esquema jerárquico



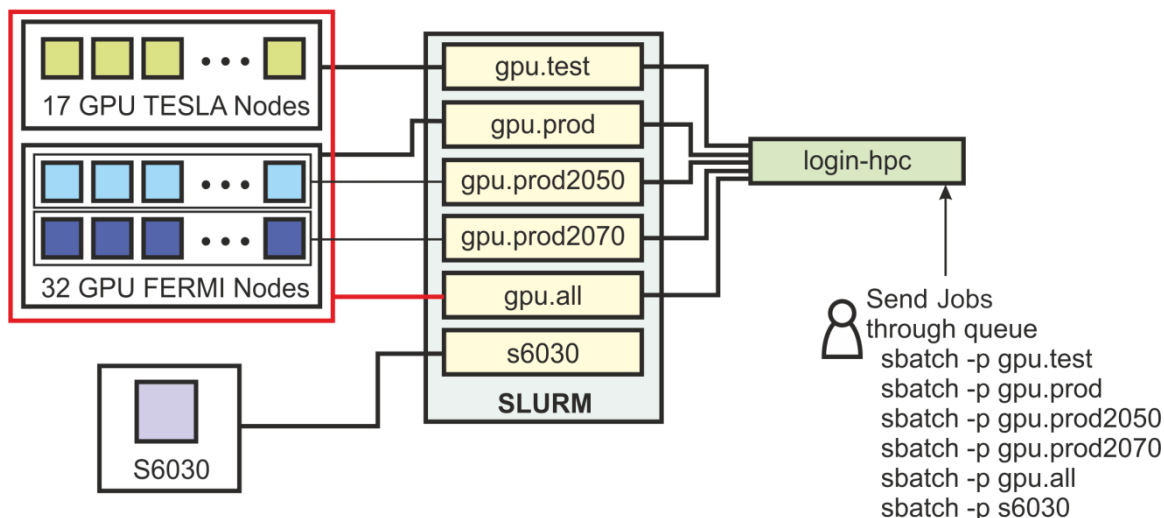
¿Cómo controlar tantos recursos?

- **Gestor de recursos**, también conocido como un gestor de colas, *scheduler*, planificador...

En el CETA-Ciemat

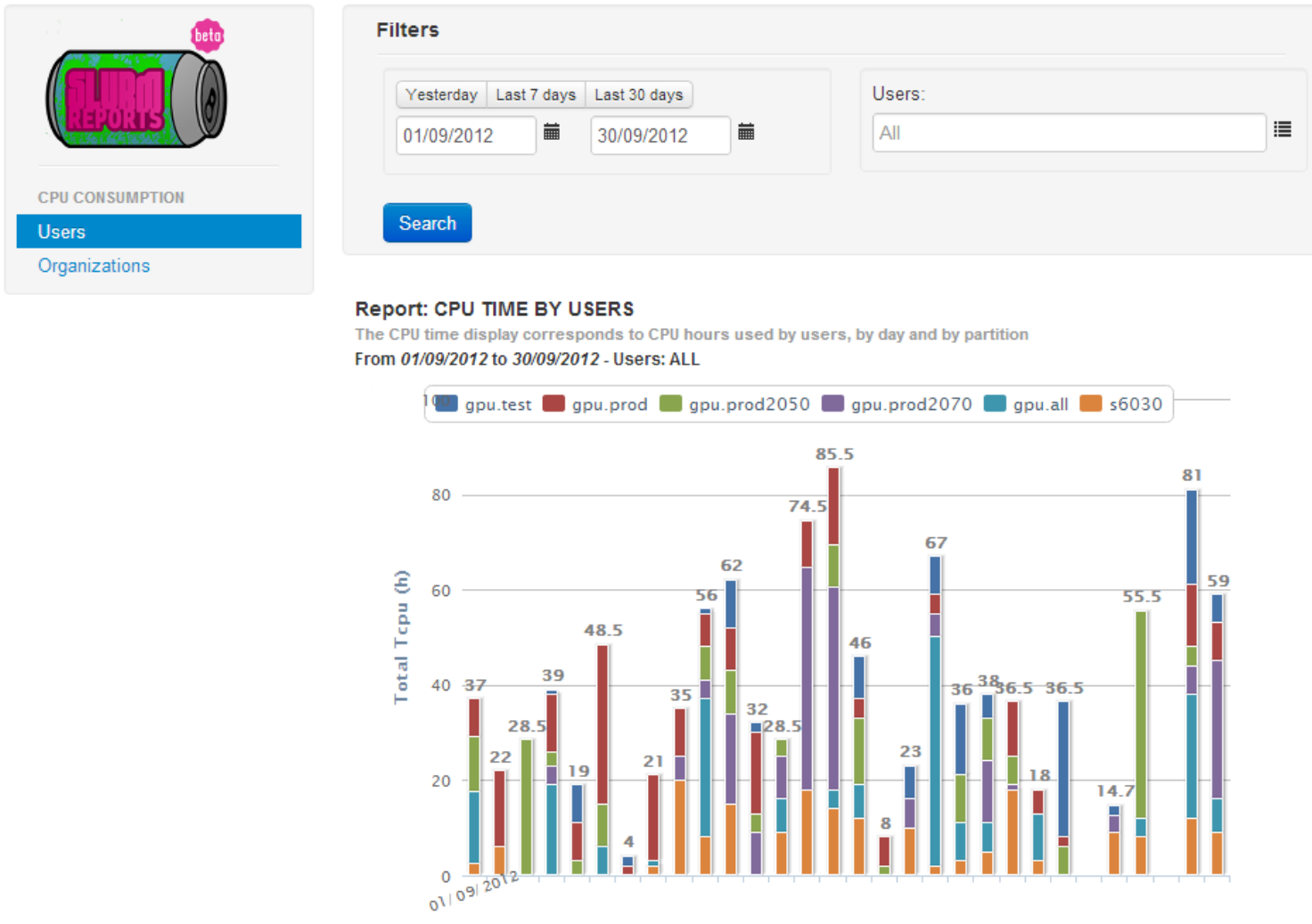
- **Inicialmente, y aún para Grid: SGE** (Sun Grid Engine)
- **HPC: SLURM** (Simple Linux Utility for Resource Management)
 - Altamente escalable: hasta 65.536 nodos, cientos de miles de procesadores...
 - OpenSource
 - **Top500**: 40% del top lo utiliza, entre ellos el número 1.
 - “Green IT”

Particiones de recursos:



Ventajas de ser Open Source

- Facilidad para interconectarlo con otras tecnologías y adaptarlo a tus necesidades



Informes
personalizados de
utilización de
recursos

CETA-Ciemat agradece la aportación del Fondo Europeo de Desarrollo Regional



GOBIERNO DE ESPAÑA

MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD



CENTRO EXTREMEÑO DE TECNOLOGÍAS AVANZADAS

CETA Ciemat

Una manera de hacer Europa

CETA-Ciemat/ Noviembre 2012

Soluciones para entornos HPC

Abel Francisco Paz Gallardo

EXTREMADURA RESEARCH CENTER FOR ADVANCED TECHNOLOGIES

Thanks!



Conventual de San Francisco, Sola 1, 10200 Trujillo
Phone: 927 65 93 17 Fax: 927 32 32 37
www.ceta-ciemat.es



GOBIERNO
DE ESPAÑA

MINISTERIO
DE ECONOMÍA
Y COMPETITIVIDAD



CENTRO EXTREMEÑO DE
TECNOLOGÍAS AVANZADAS

CETA Ciemat



FEDER

Fondo Europeo de Desarrollo Regional

Una manera de hacer Europa