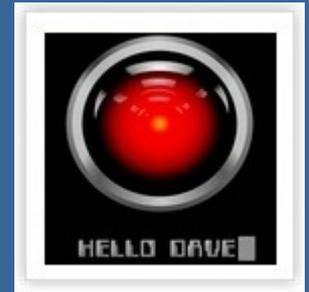


Ordenadores Dedicados y Computación Ciudadana

Alfonso Tarancón, Julio 2011
DFTUZ y BIFI

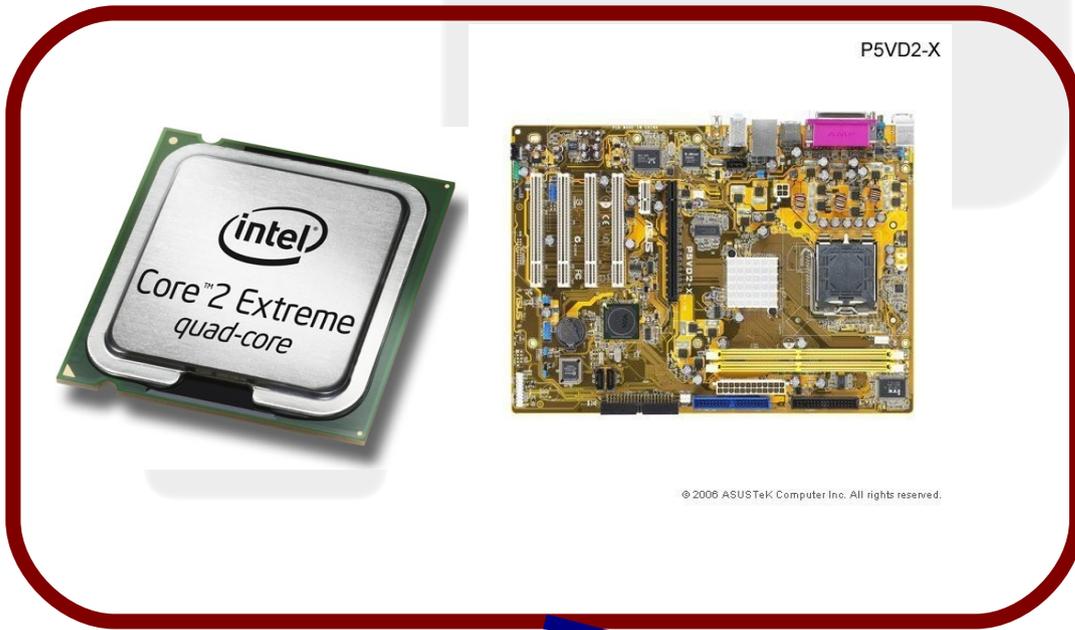


Instituto Universitario de Investigación
**Biocomputación y Física
de Sistemas Complejos**
Universidad Zaragoza



- Aplicaciones y Arquitecturas
- Interludio: Estudio de Aplicación Paralela
- Ordenadores Dedicados: Janus
- Computacion Voluntaria: Boinc e Ibercivis

Aplicaciones y Arquitecturas



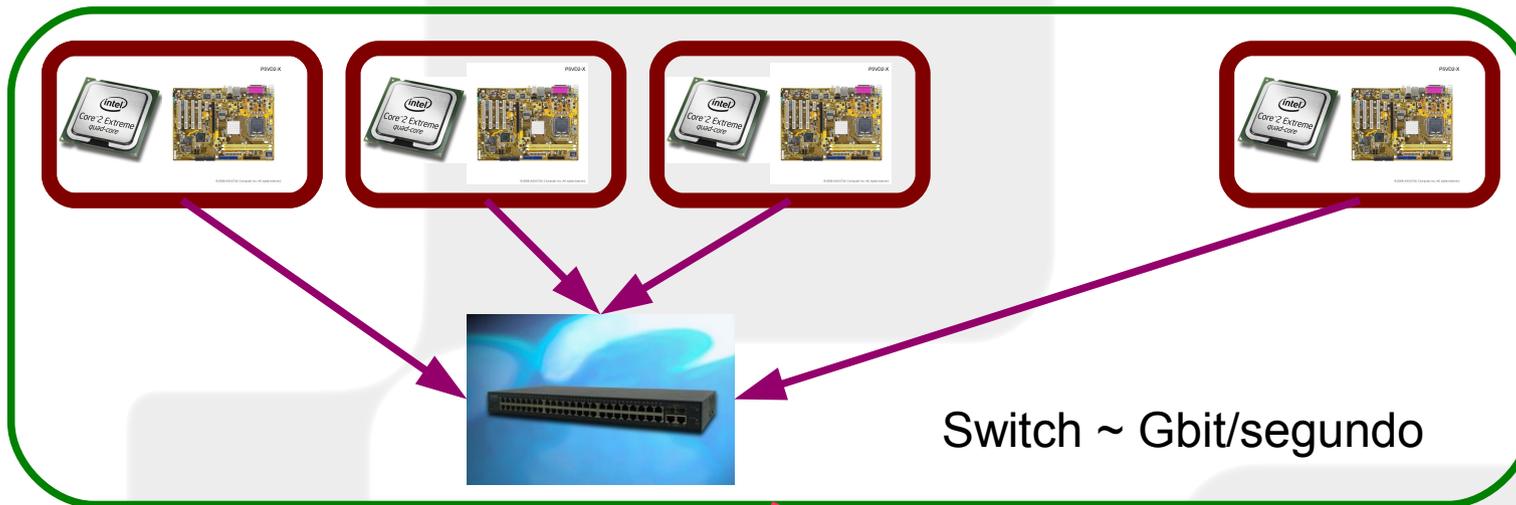
CPU+RAM+Placa Base

Unidad básica de Cálculo

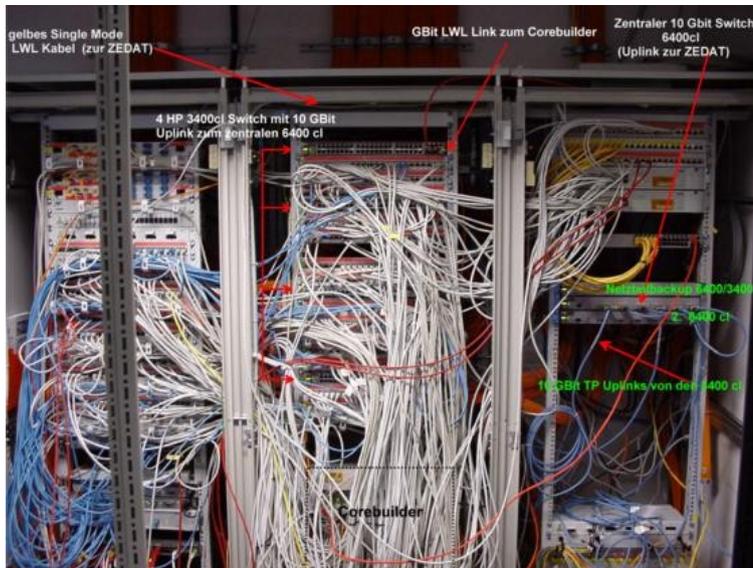
N=1



Clusters



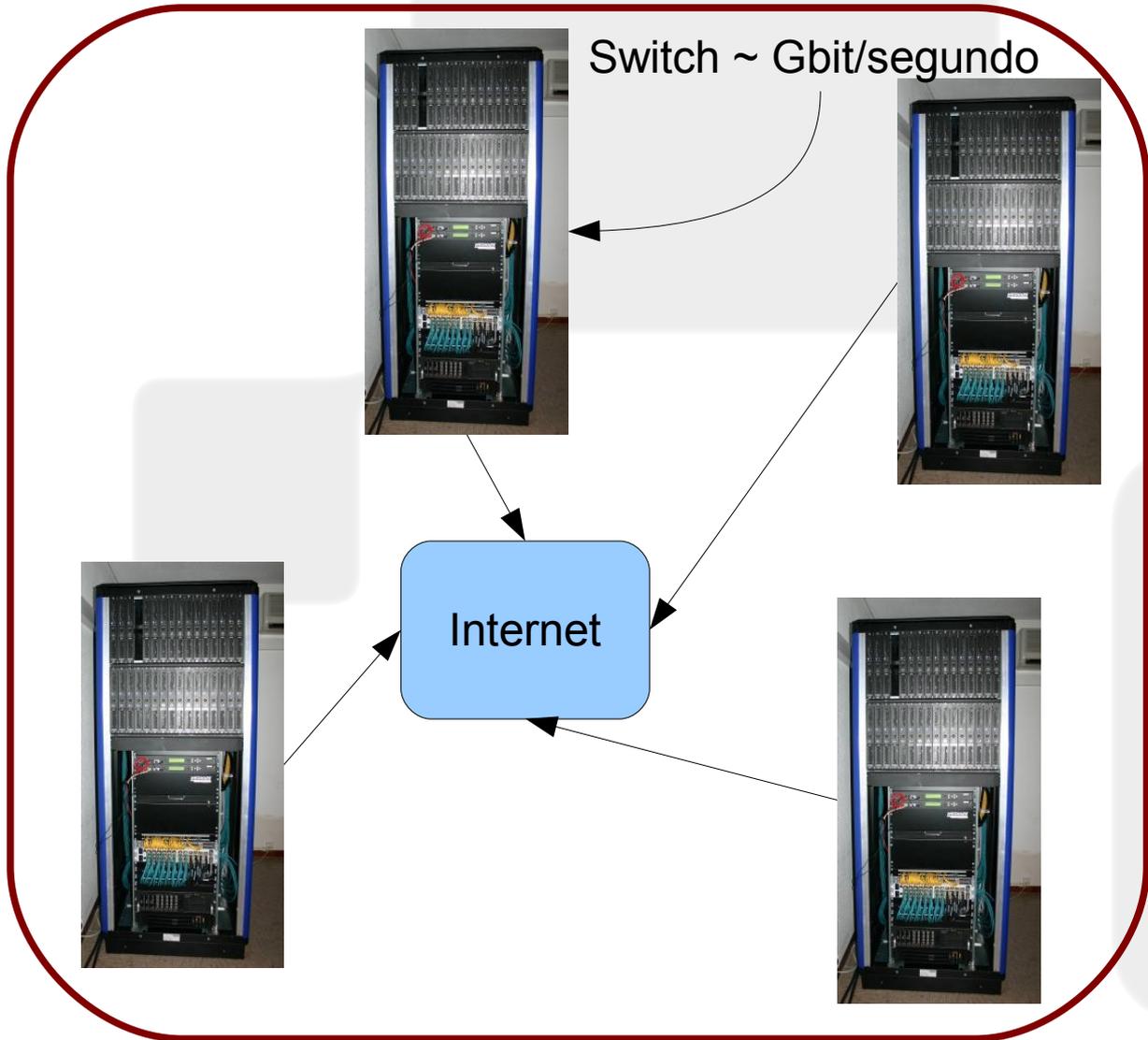
$N \sim 10^2$



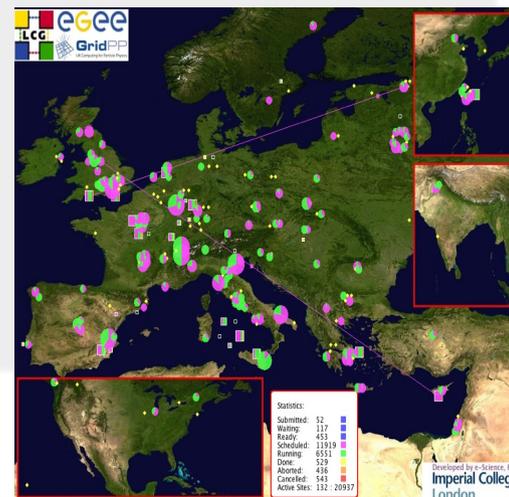
Intel-AMD-PowerPC



GRID



$$N \sim 10^3 - 10^5$$



Intel-AMD-PowerPC

Mare Nostrum (CNS-BSC)



Switch ~ 10 Gbit/segundo

$N \sim 10^4$

Alta Velocidad
Baja Latencia
Casi todos con todos



Copyright 2005. Barcelona Supercomputing Center - BSC

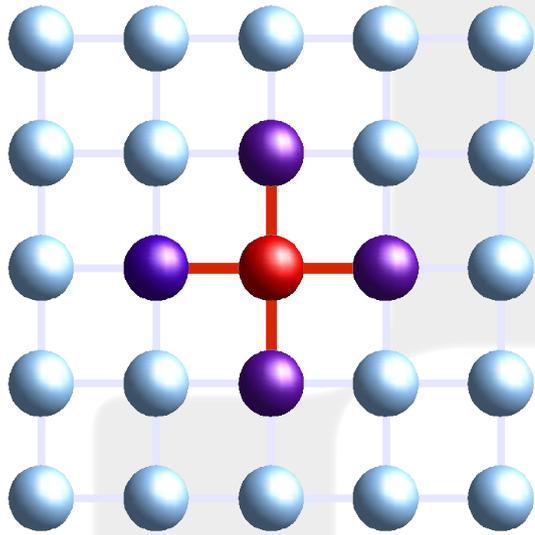


Copyright 2005. Barcelona Supercomputing Center - BSC

La carrera del TOP500

Intel-AMD-PowerPC-SPARC-GPU

Ordenadores Dedicados



$$H = \sum_{n,\mu} \sigma_n \sigma_{n+\mu} + h \sum_n \sigma_n$$

$$Z = \sum_{\text{Configs}} e^{\beta H}, \beta = 1/kT$$



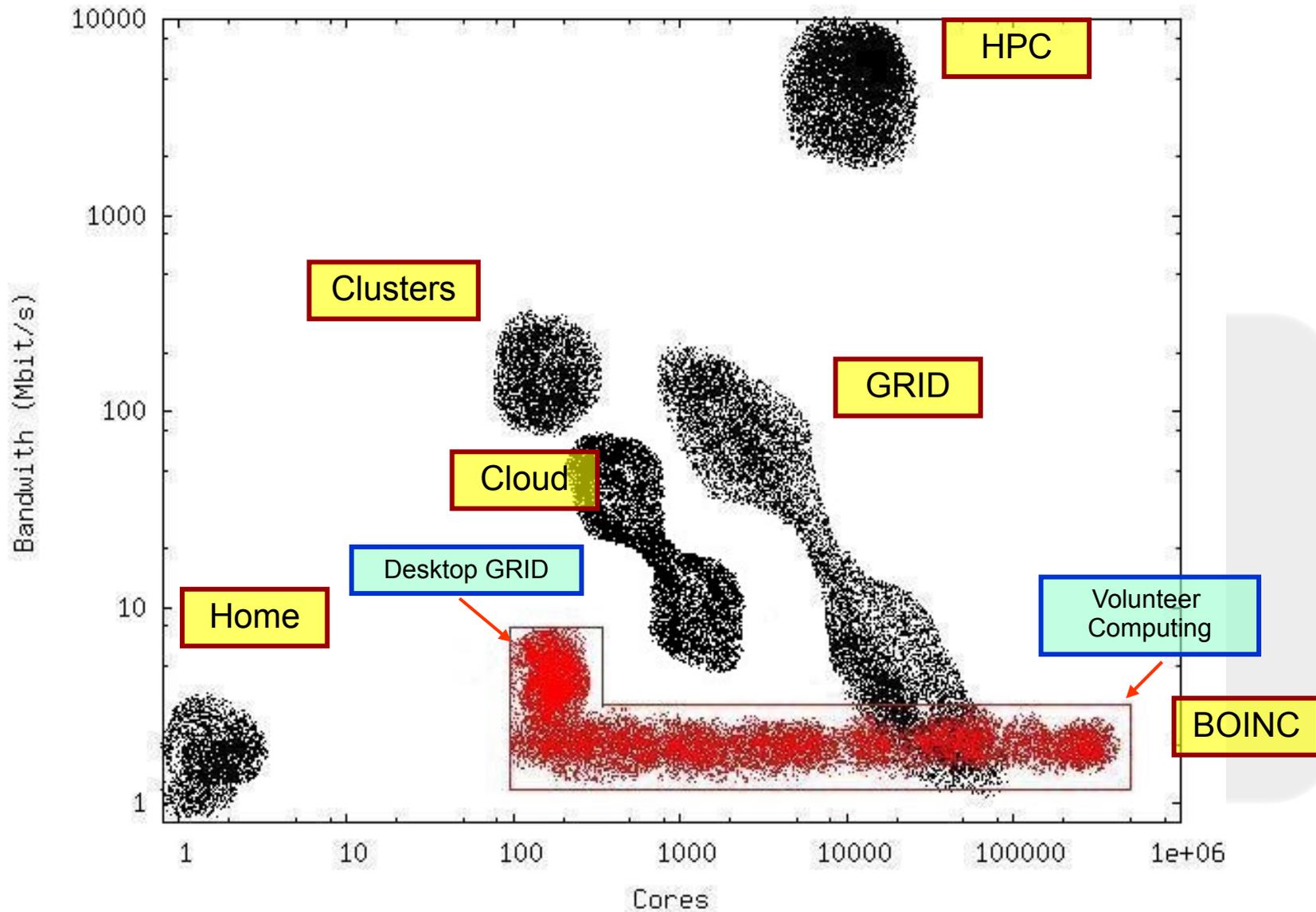
Computación Voluntaria-BOINC



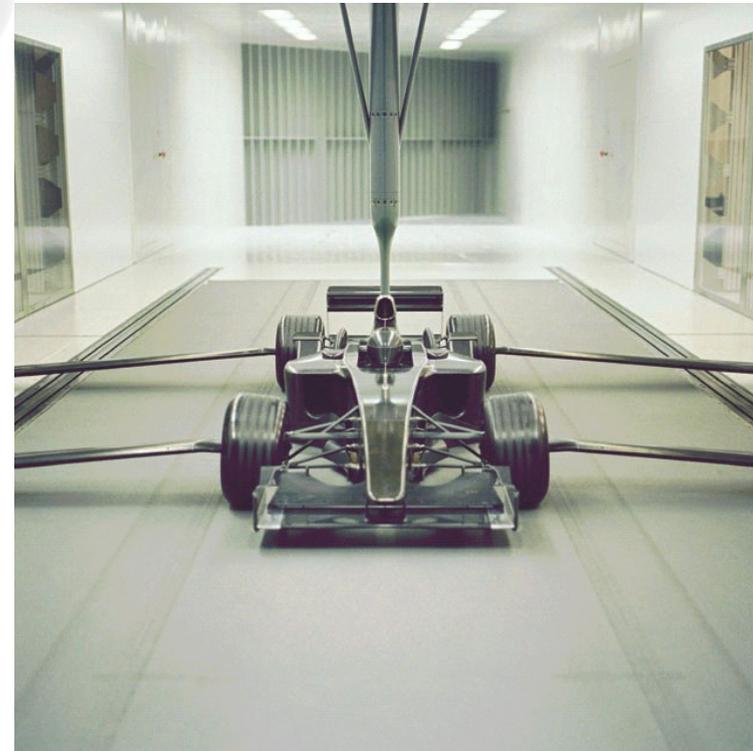
- Proyecto SETI: Search for Extra Terrestrial Intelligence
- BOINC
- Ibercivis



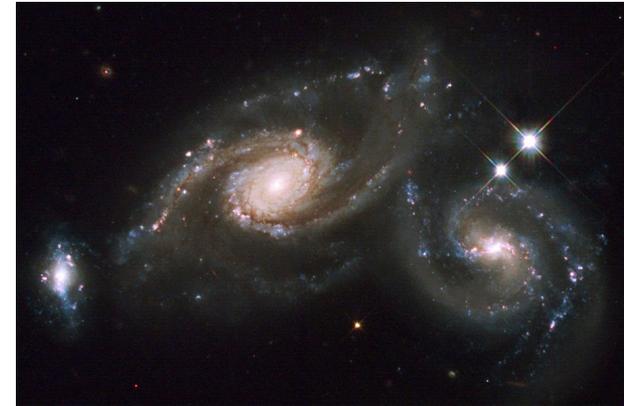
Diagrama Cores-BandWidth



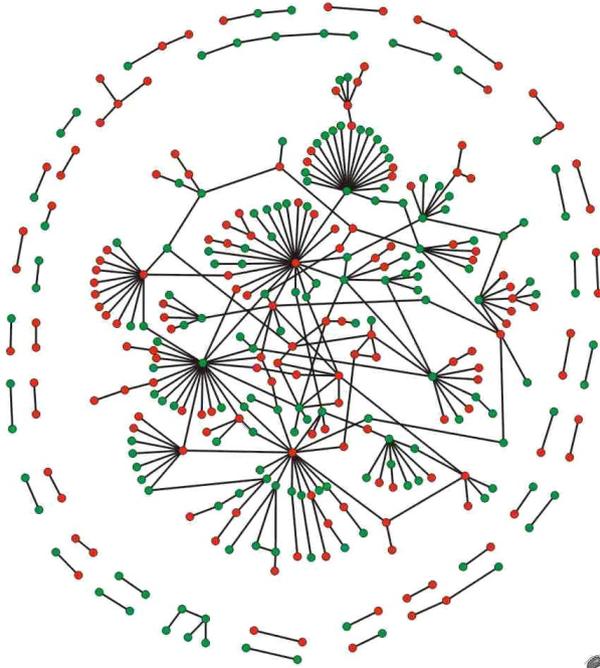
**Necesidad de Computación:
Fluidodinámica**



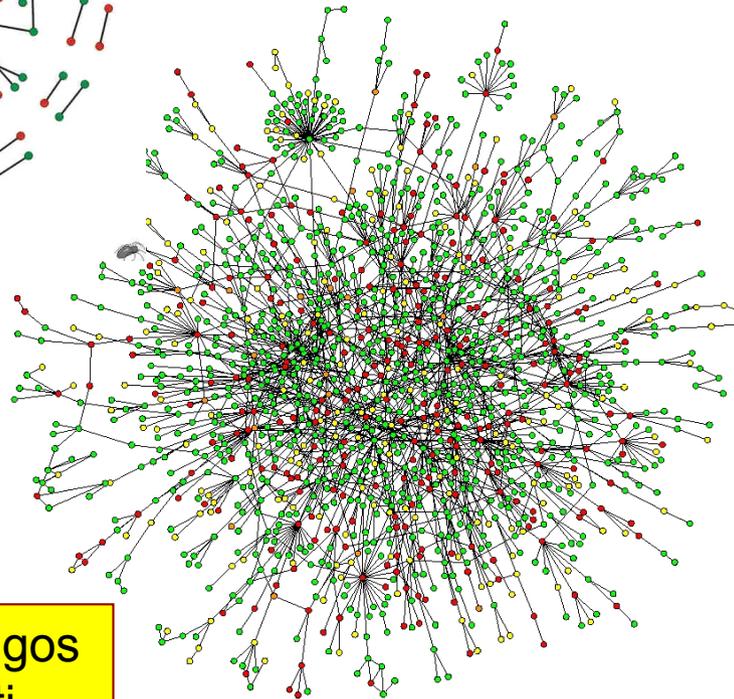
Choque de Galaxias



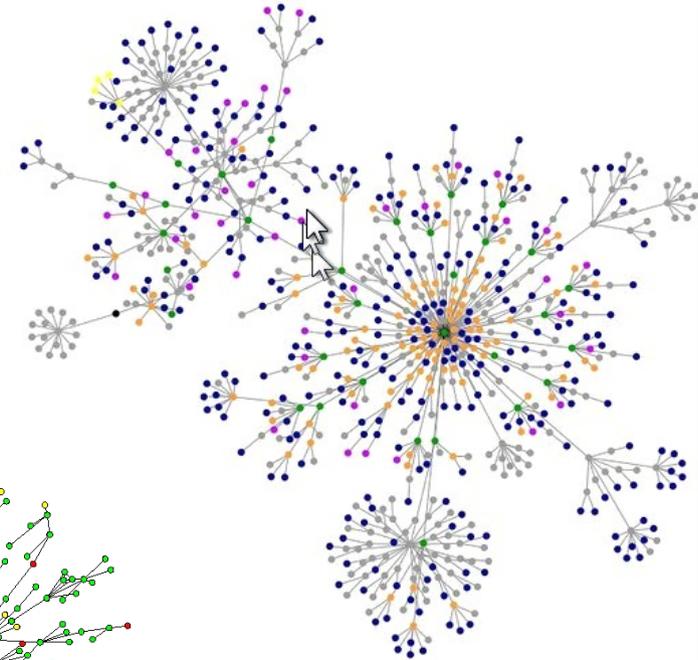
Redes Naturales y Sociales



Red de Interacción de Proteínas

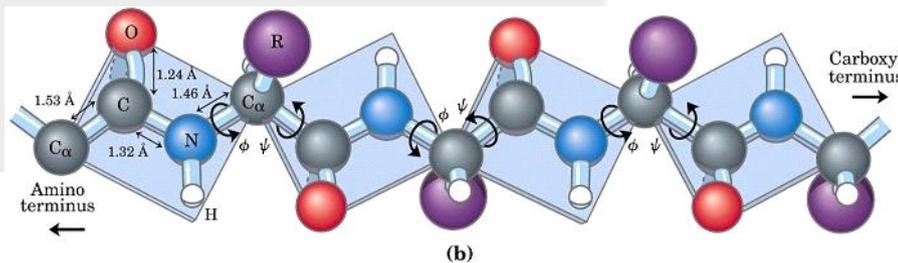


Red de Amigos en Tuenti

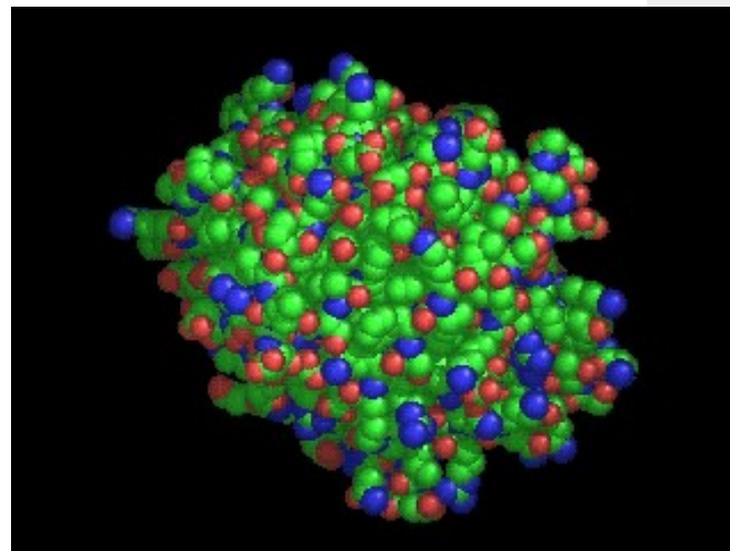
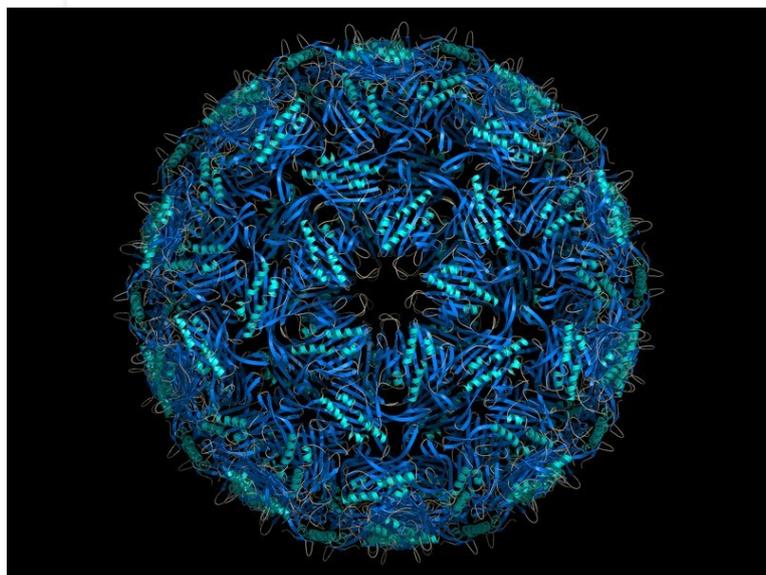
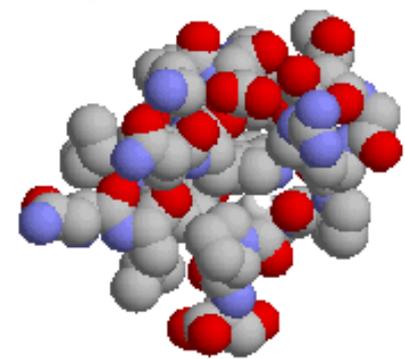
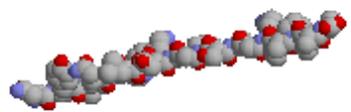


Red de Referencias de Artículos Científicos

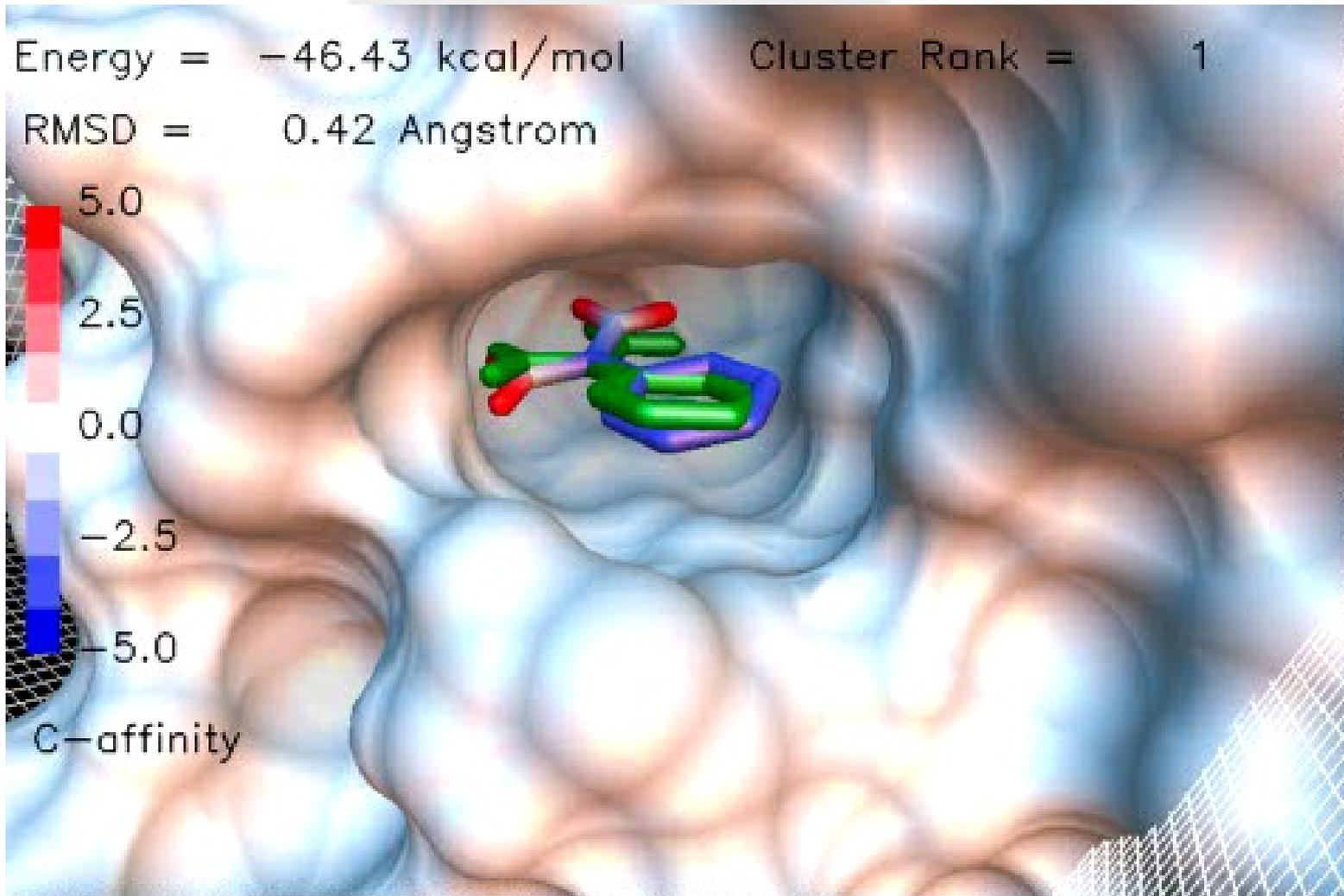
Proteínas



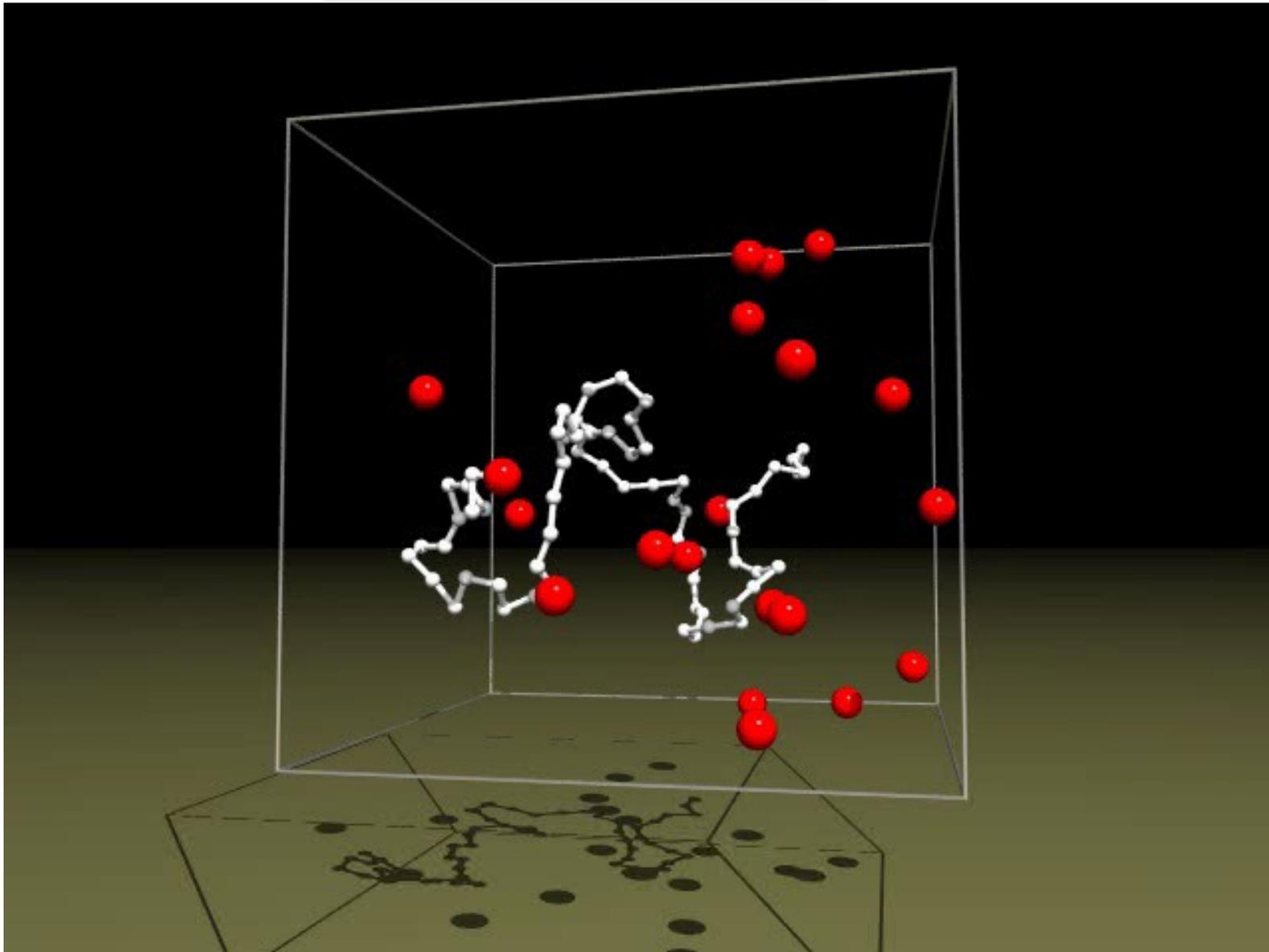
Proteína : Cadenas de Aminoácidos



Docking



Dinámica Molecular



Plegamiento de Proteínas



Y quién sabe ...



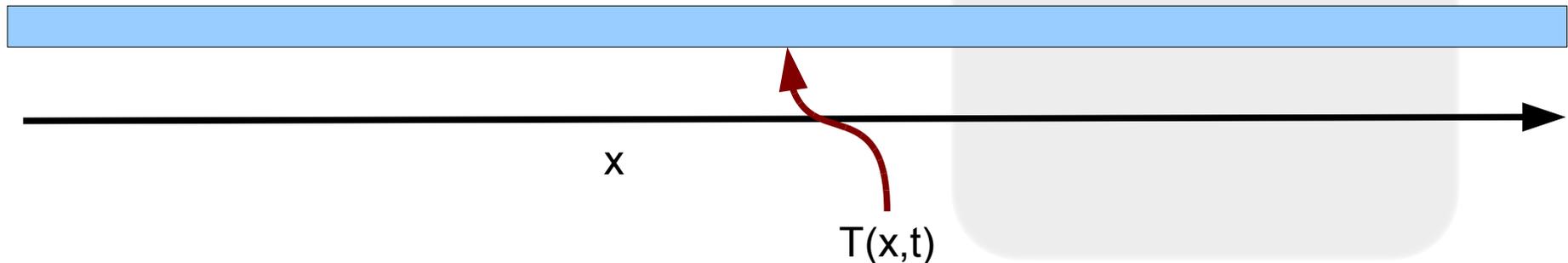
Interludio: Estudio de Aplicación Paralela

Aplicación Paralela



- Difusión de Calor sobre una barra unidimensional.
- Dada T al inicio en cada punto, calcular T para cualquier punto en cualquier instante
- Dado $T(x,0)$, calcular $T(x,t)$

Instante t



- *La temperatura de cada punto varia proporcionalmente al Laplaciano en dicho punto.*

$$\frac{\partial T(\vec{r}, t)}{\partial t} = k \left(\frac{\partial^2 T(\vec{r}, t)}{\partial x^2} + \frac{\partial^2 T(\vec{r}, t)}{\partial y^2} + \frac{\partial^2 T(\vec{r}, t)}{\partial z^2} \right)$$

$$\frac{\partial T(x, t)}{\partial t} = k \frac{\partial^2 T(x, t)}{\partial x^2}$$

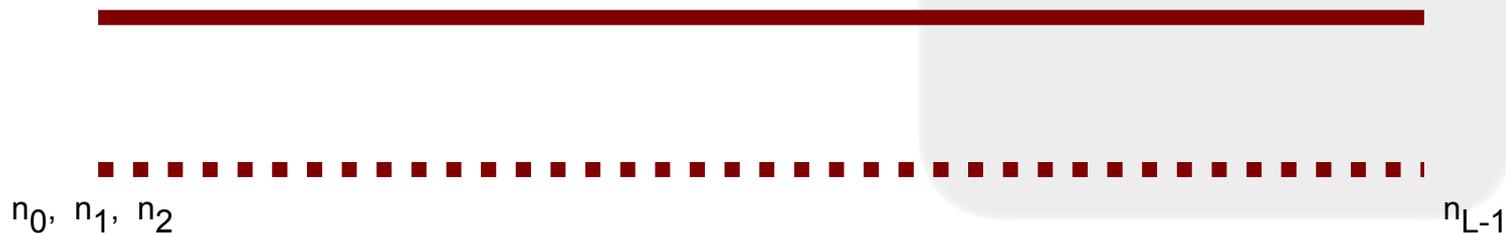
Del Continuo al Discreto



$$\frac{d f(x)}{d x} \approx \frac{f(x+h) - f(x)}{h}$$

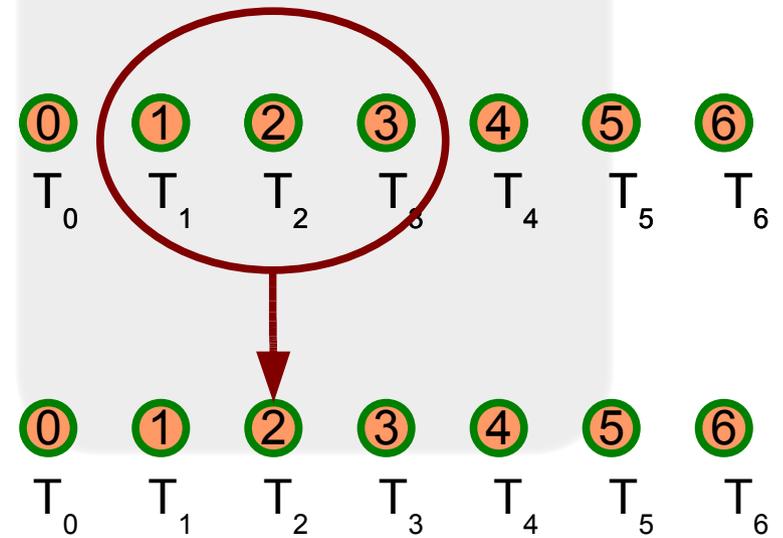
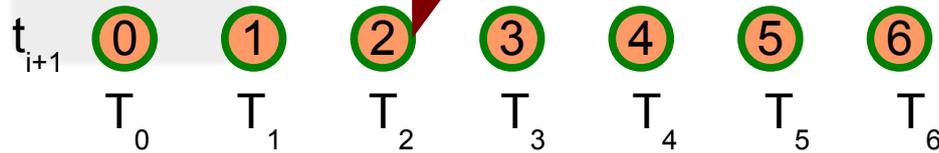
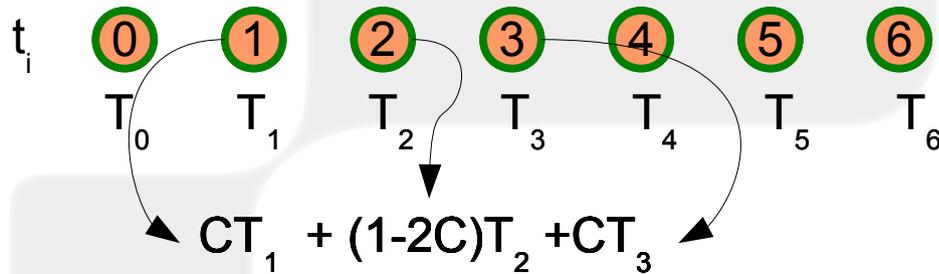
$$\frac{\partial T(x, t)}{\partial t} \approx \frac{T(x, t + \delta) - T(x, t)}{\delta}$$

$$\frac{\partial T(x, t)}{\partial x} \equiv \frac{T(n_{i+1}, t_k) - T(n_i, t_k)}{a}$$



Calculo de $T(x,t+1)$ conocido $T(x,t)$

$$T(n_i, t_{k+1}) = (1 - 2C)T(n_i, t_k) + C(T(n_{i+1}, t_k) + T(n_{i-1}, t_k))$$



Código secuencial



```
double evolucion(double *uin, double *uout)
{
    int i;

    for (i=1;i<L-1;i++)
        uout[i]=C0*uin[i]+C*(uin[i-1]+uin[i+1]);
}
```

- Memoria RAM (floats): $2*L*4$ bytes
- Operaciones por paso: $4*L$

Tiempo de Ejecución



- Tiempo de ejecución de 1 iteración en un core a 2 GHz

$$\tau = \frac{4L}{2 * 10^9} = 2L \times 10^{-9} \text{ seg.}$$

Definición de eficiencia Paralela



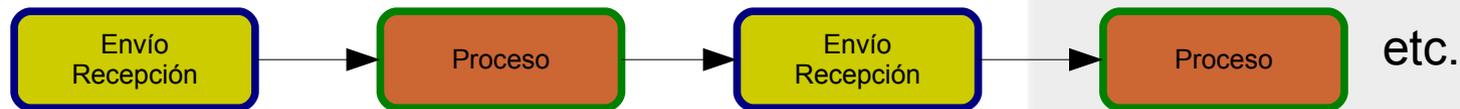
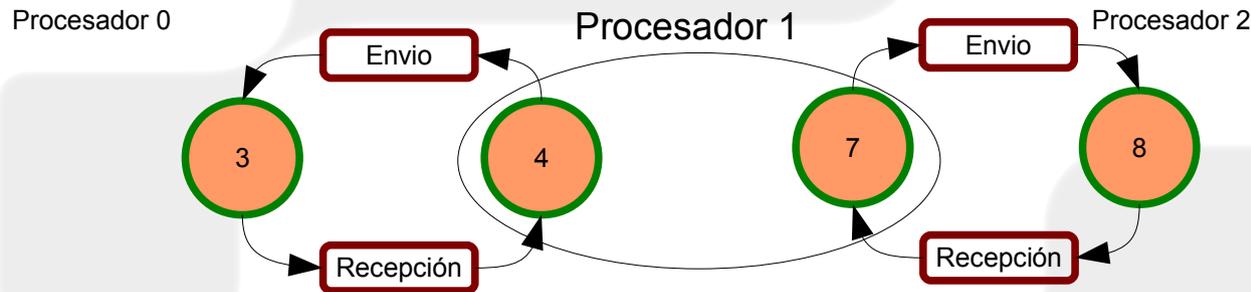
- Llamemos τ al tiempo de ejecución cuando usamos un solo procesador
- Caso Ideal: Si uso N procesadores, el tiempo de ejecución se divide por N ($\varepsilon=100\%$)
- Llamemos ρ al tiempo de ejecución cuando usamos N procesadores
- Definición: $\varepsilon=100*\tau / (N*\rho)$

Paralelismo en modo granja



- Simulamos M barras con Temperaturas iniciales diferentes, una en cada procesador.
- No existe comunicación entre cores
- Eficiencia paralela: 100%

Modo Paralelo



- $4 \cdot L/N$ operaciones aritméticas
- Enviar 4 bytes a la Izquierda
- Enviar 4 bytes a la Derecha
- Recibir 4 bytes de la Izquierda
- Recibir 4 bytes de la Derecha

Tiempo ejecución paralela



$$T = t_{\text{proceso}} + t_{\text{comunicacion}} \equiv \tau + \mu$$

$$\tau = \frac{4L}{N} \frac{1}{2 * 10^9} = \frac{L}{N} 2 \times 10^{-9} \text{ seg.}$$

$$\mu = 2 * t_{\text{latencia}} + t_{\text{transferencia}}$$

$$t_{\text{latencia}} = 5 \mu \text{ segs} , v_{\text{transferencia}} = 100 \text{ Mbits}$$

Eficiencia [0,1]



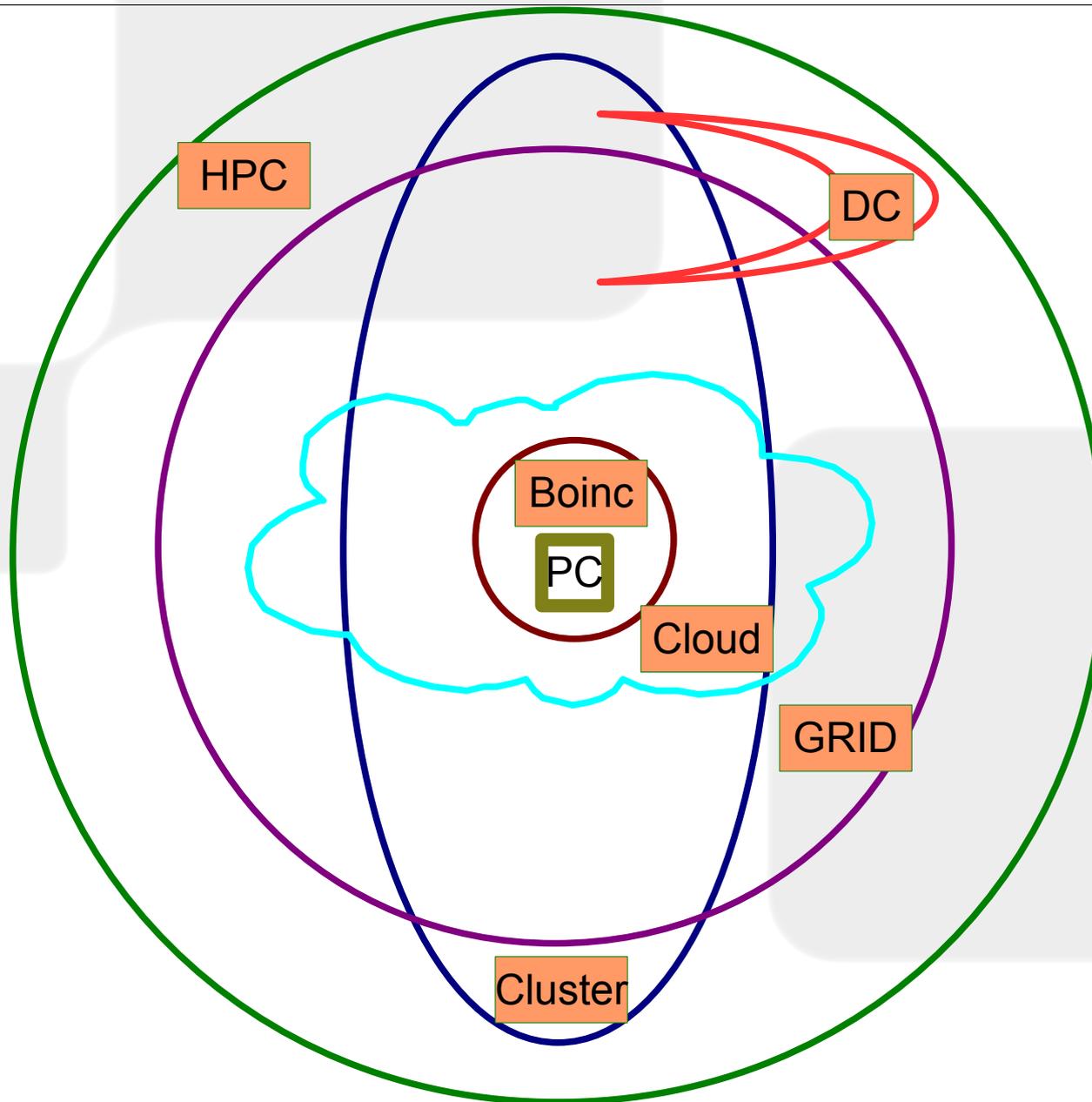
$$T = \tau + \mu = \frac{L}{N} 2 \times 10^{-9} + 10^{-5} + 6.4 * 10^{-7}$$

$$\epsilon = \frac{t_{\text{espera con un solo procesador}}}{N * t_{\text{espera con N procesadores}}} = \frac{L \times 2 \times 10^{-9}}{L \times 2 \times 10^{-9} + N \times (10^{-5} + 6.4 * 10^{-7})}$$

$$\epsilon = \frac{1}{1 + \frac{N}{L} \times 5.3 \times 10^3} \approx \frac{C}{N}$$

Desastre !
Cuesta igual usando
1000 que 2000 cores

Diagrama Aplicaciones-Arquitecturas



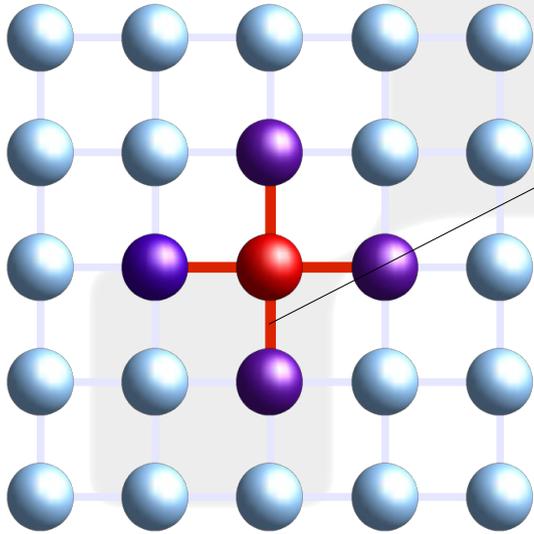
Ordenadores Dedicados: Janus

Ordenadores Dedicados



- Se diseñan ad hoc para una aplicación y un código concreto
- Son órdenes de magnitud más veloces que los procesadores convencionales
- Sólo pueden ejecutar unos pocos programas

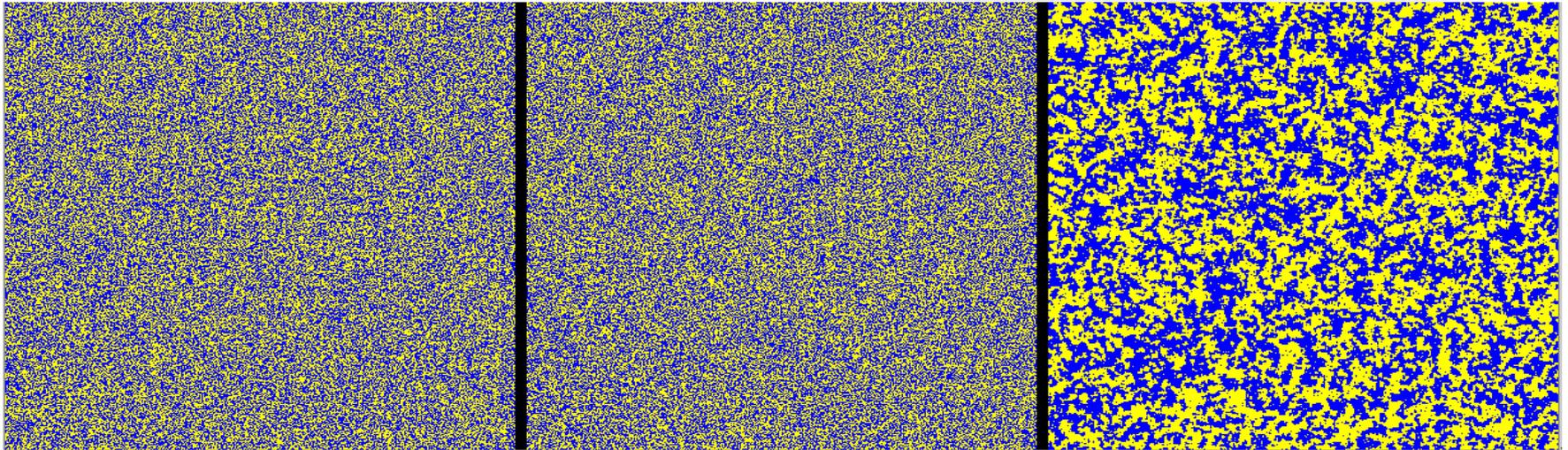
Vidrios de Spin (Spin Glass)



$$J_{n, n+\mu} \in \mathbb{Z}_2$$

$$H = \sum_{n, \mu} \sigma_n J_{n, n+\mu} \sigma_{n+\mu} + h \sum_n \sigma_n$$

$$Z_{[J]} = \sum_C e^{\beta H}$$

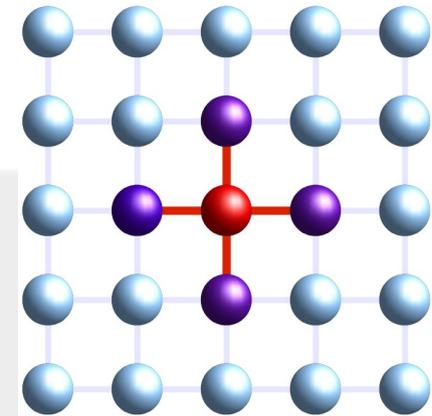


- Variables de pequeño tamaño (incluso binarias)
- Operaciones booleanas y con enteros
- Alto grado de paralelismo
- “RAM” pequeña

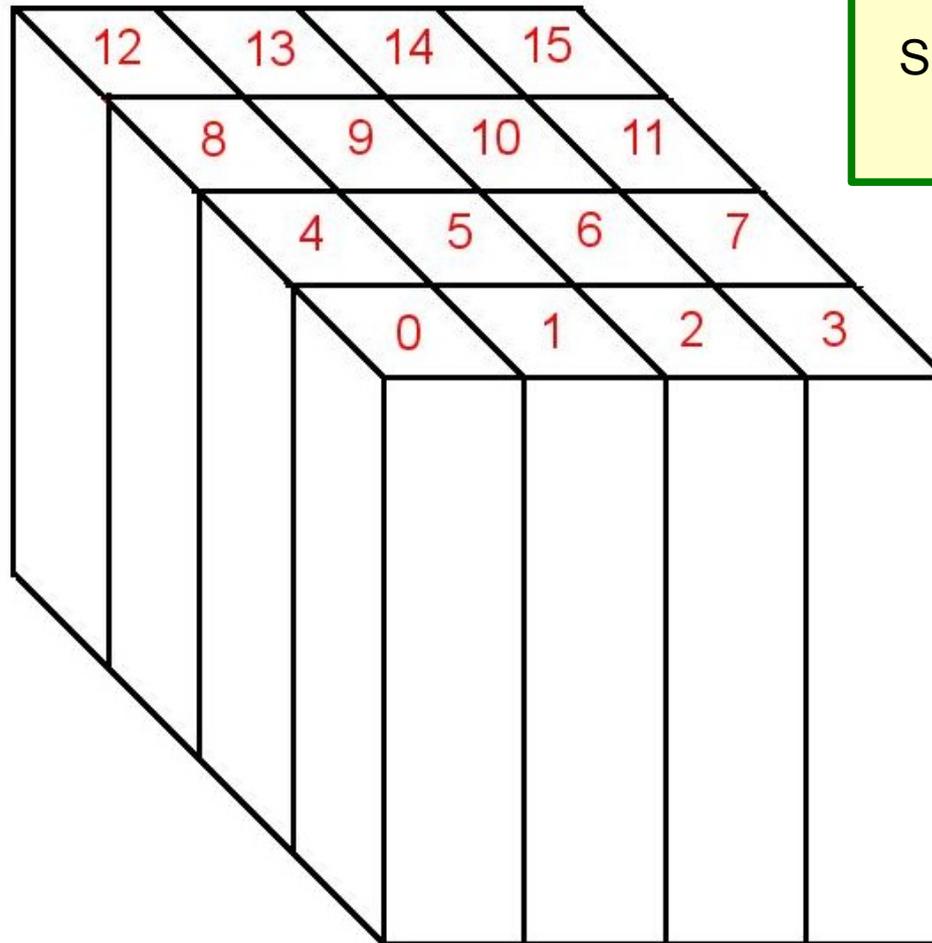
Tamaño típico de la red $< 100^3 = 1$ Mbit
(2 réplicas + 3 enlaces + 1 campo) * 1 Mbit = 6 Mbit

En cada Ciclo (punto de la Red)

- Generar un número Random de 32 bits
- Leer 12 bits
- Realizar 6 xor y 6 sumas
- Leer de una LUT
- Comparar Lut con Random
- Escribir 1 bit



Paralelismo a Nivel de Placa



Se transmiten bloques pequeños
Pero con mucha frecuencia

Latencia Cero
Eficiencia Paralela:90%

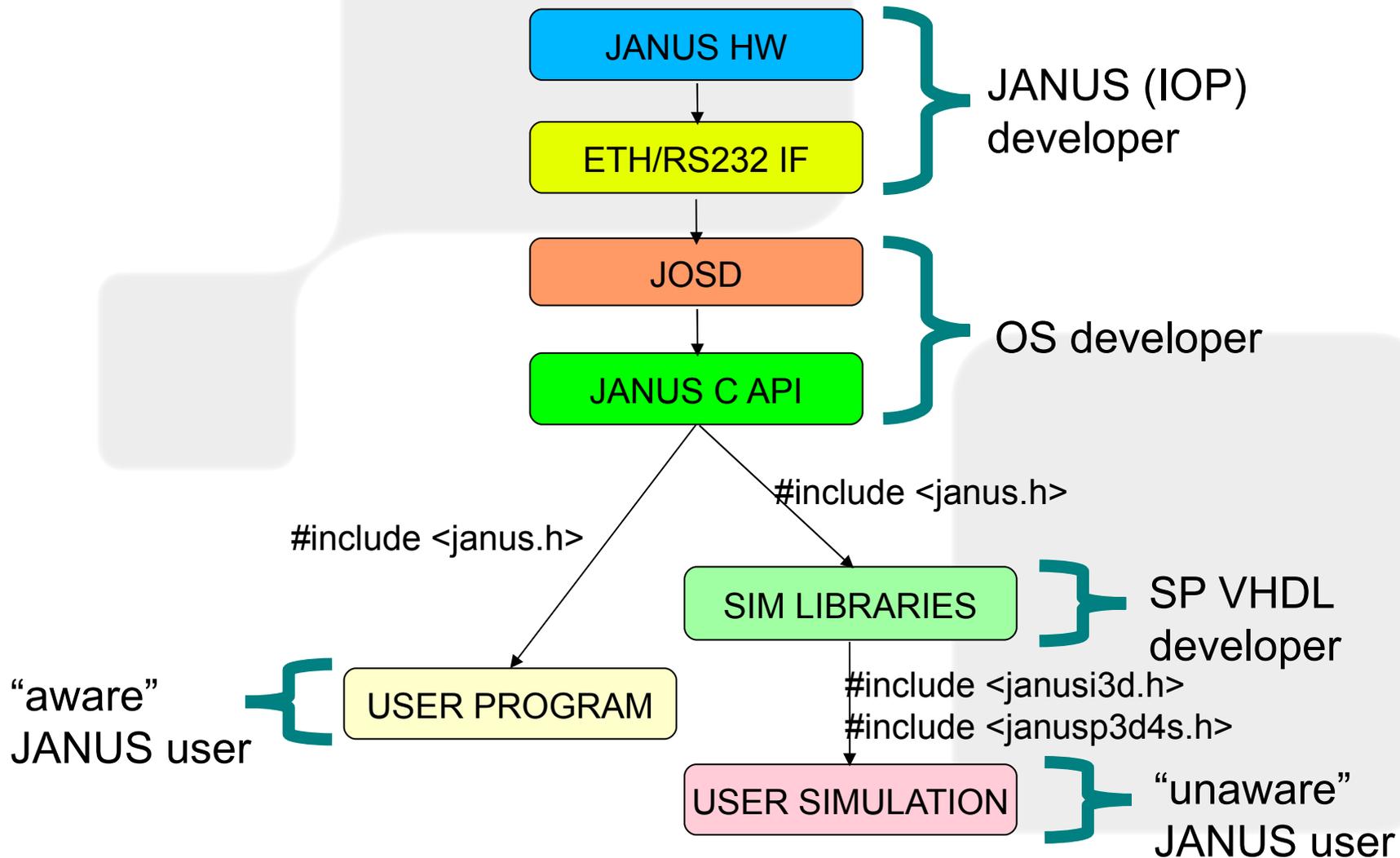
ESQUEMA FUNCIONAL

- Matriz 2-D de 4 x 4 FPGAs
- Conexiones de 32 bits @ 62.5 MHz (2 Gbps) con los SP vecinos.
- Condiciones de contorno periódicas
- 1 FPGA de control (IOP) con 2 gigabit-ethernet

A.

B.

Janus Software Layers



Prestaciones



Modelo	FPGA	PC	RATIO
EA	16 ps/spin	3 ns/spin	180
Potts 3D p=4	64ps/spin	117 ns/spin	1800
Potts 4D p=4	64 ps/spin	150 ns/spin	2300

FPGA=Virtex 4 62.5 MHz 50 Watios

PC= Intel Core 2 Duo 2.0 GHz 65 Watios

Janus = 256 FPGA x 50 Watios = 12,8 kW

Equivalente PCs=(256x180) x 65 Watts = 2,99 MW

Motherboard

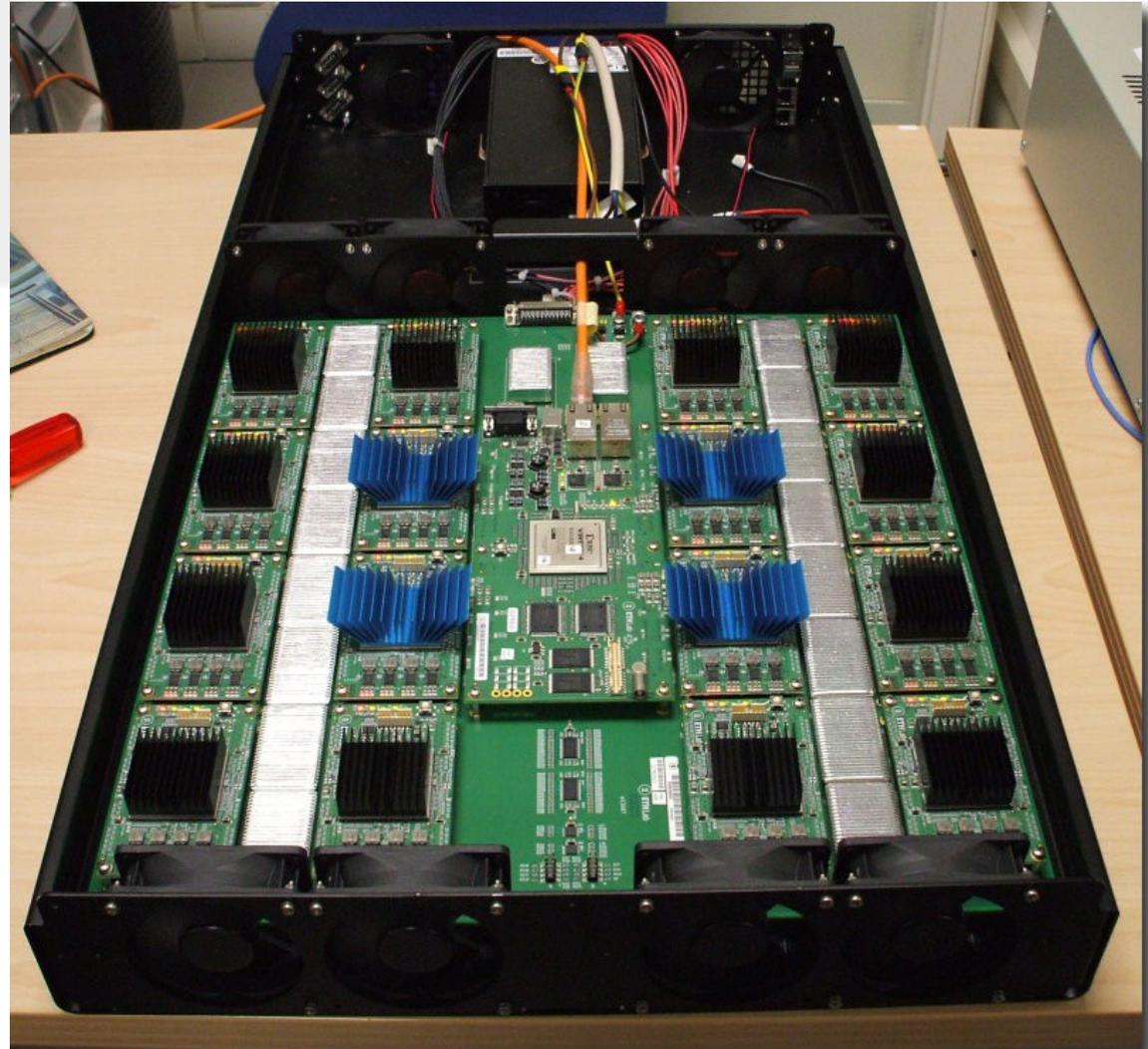
16 Nodos de calculo
(SP)

1 Procesador de I/O
(IOP)

2 GIGABIT ETHERNET

1 USB

1 RS232



Janus



JANUS RACK

8 PCs

1 ROUTER

16 Placas

17 FPGA por placa

1 PC (host) controla 2
placas

272 VIRTEX4-LX200

10-15 KWatts

180 ventiladores



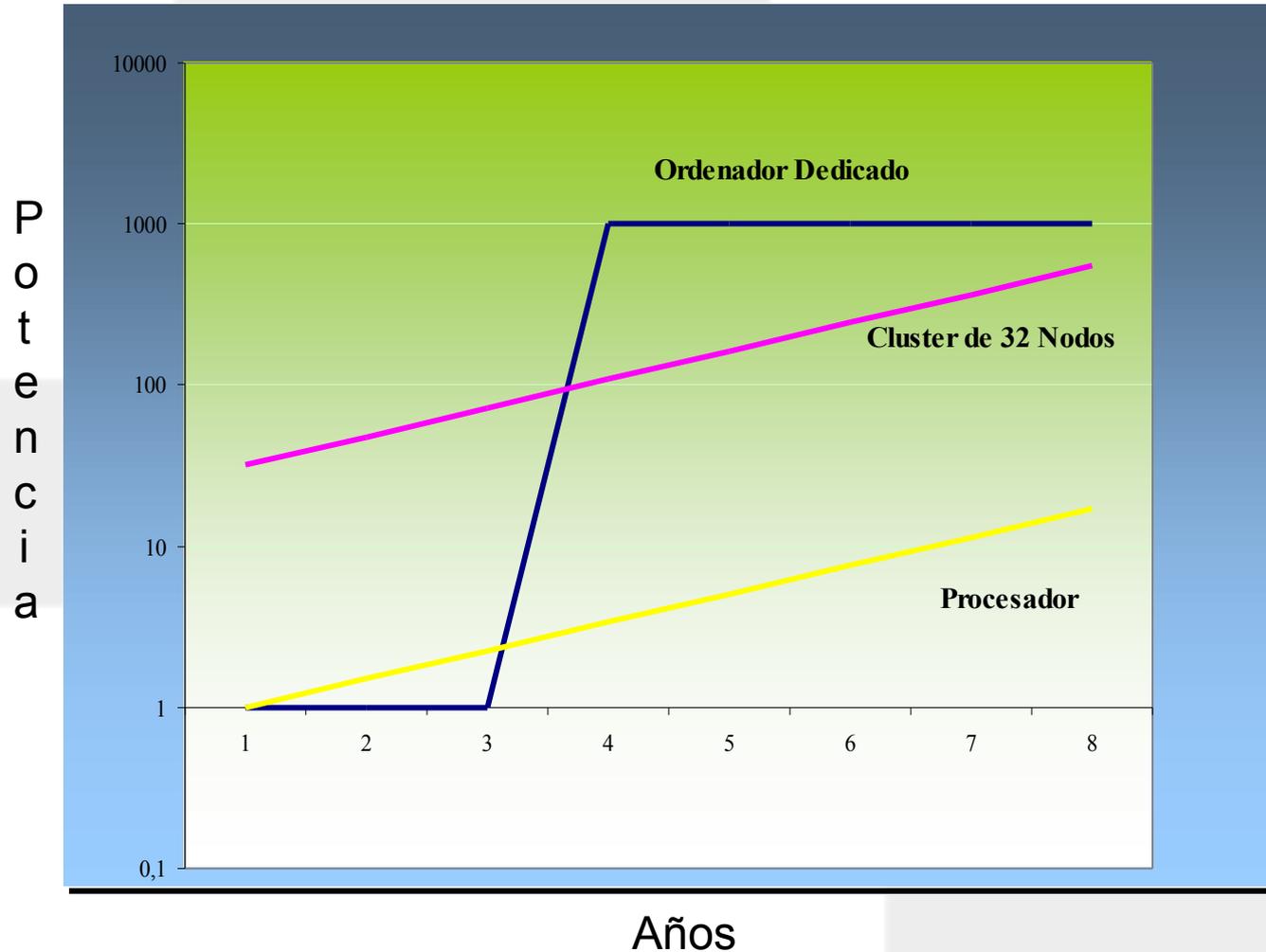
Conclusiones



- 4 AÑOS DE TRABAJO
- EL MEJOR RENDIMIENTO PARA SPIN GLASS
- TRABAJO INTERDISCIPLINAR
- DIFÍCIL ENCONTRAR NUEVOS ALGORITMOS

Janus II: Nuevas FPGAs + ASIC + Paralelismo
Futuro: Dinámica Molecular, Floating Point...

Conclusiones II

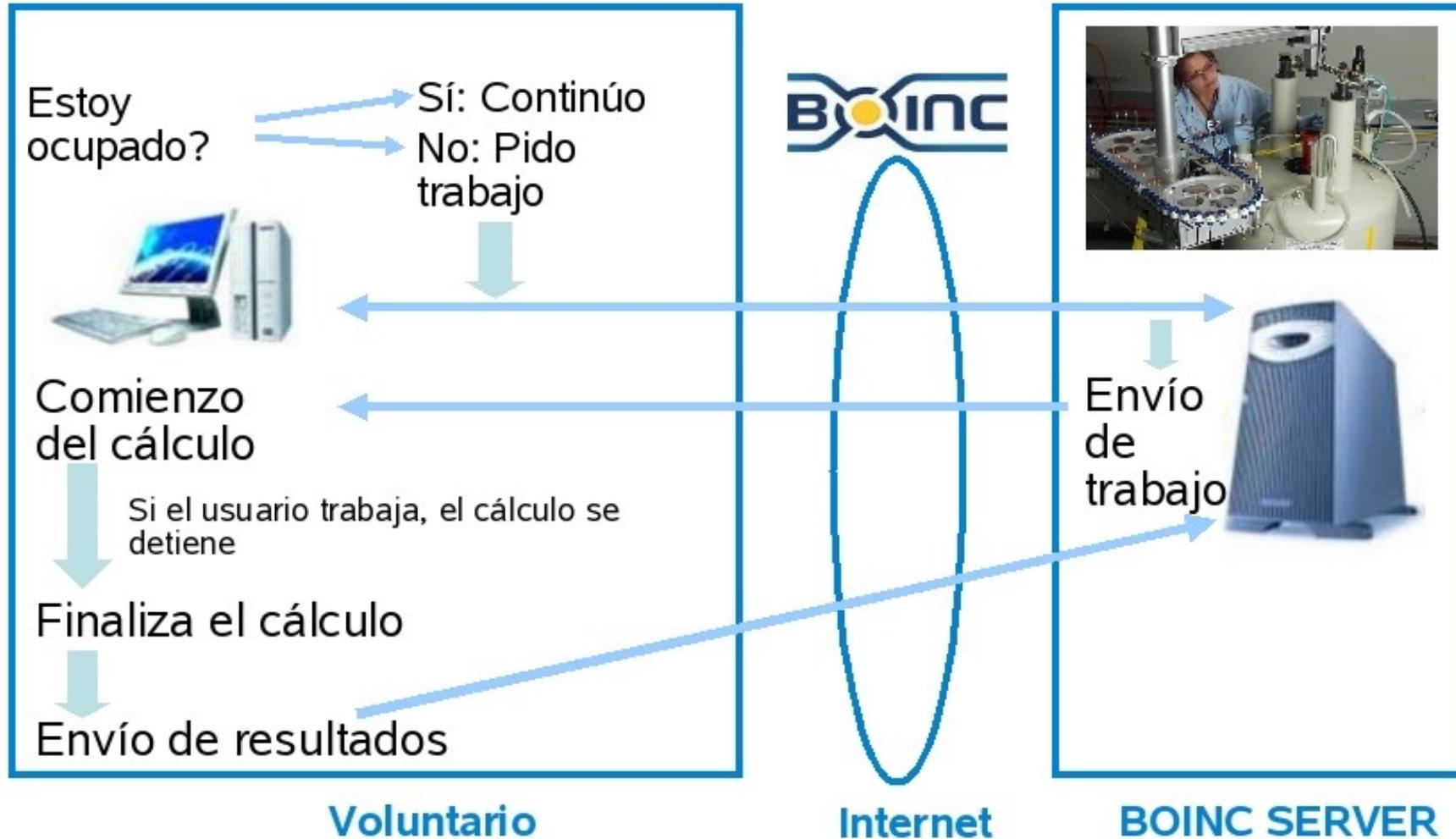


El ordenador dedicado debe ser al menos 1000 veces más potente que un procesador comercial

Debe abordar problemas tales que su Interés científico se mantenga al menos a 8 años vista

Computación voluntaria: Boinc e IberoCivis

BOINC



Origen: SETI@HOME



- Proyecto SETI: Search for Extra Terrestrial Intelligence



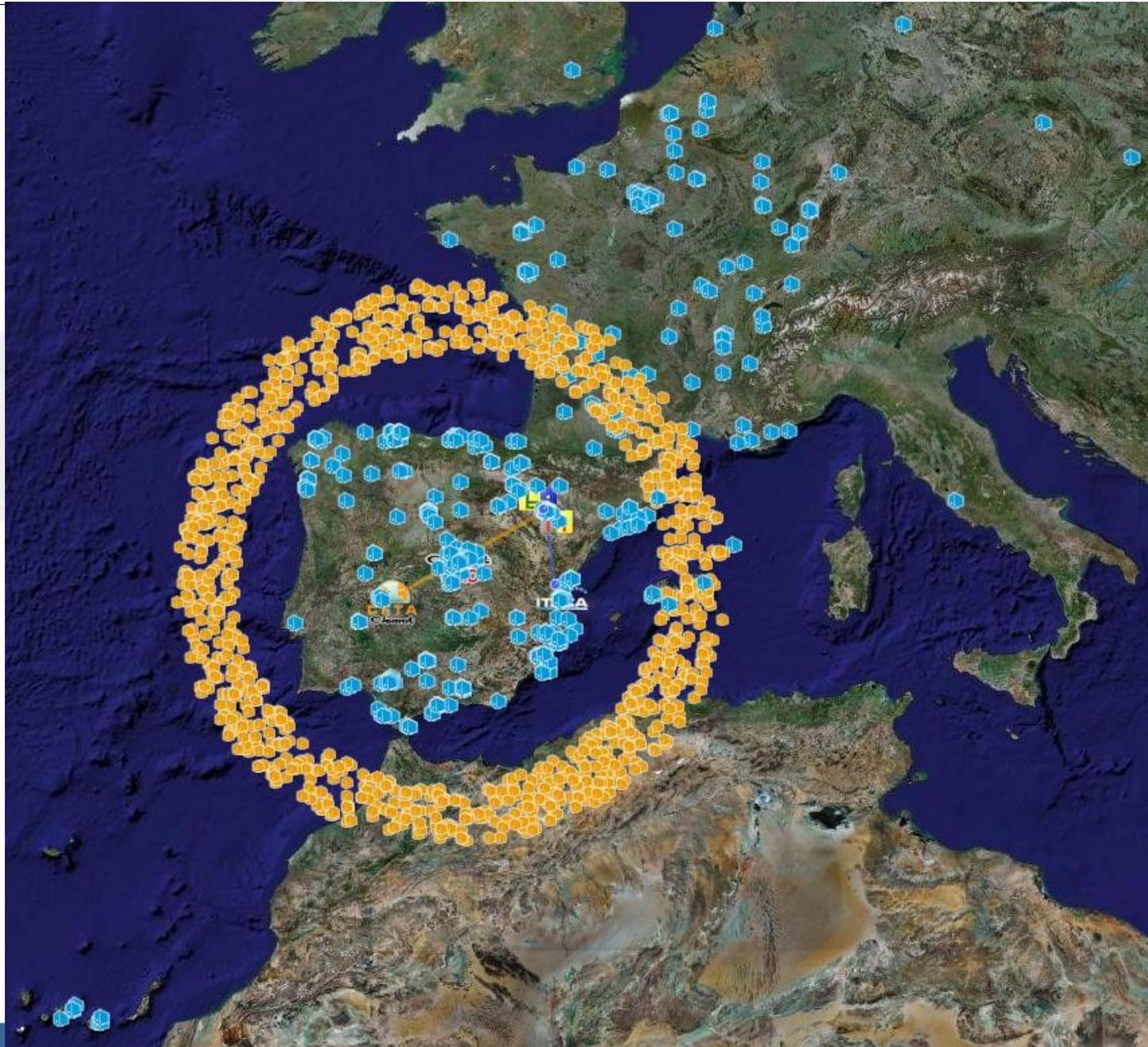
Infraestructura Hispano-Lusa



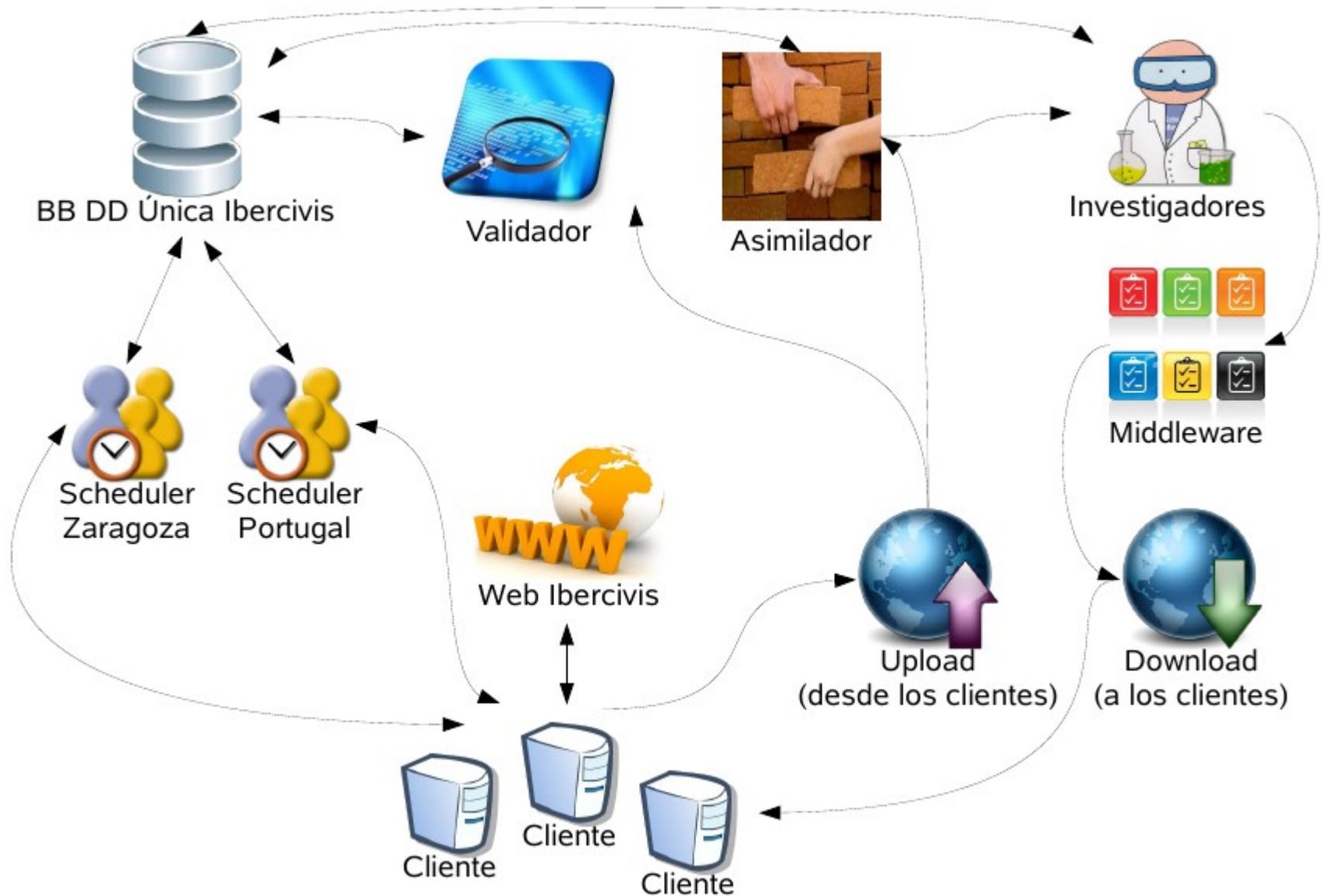
- Scheduling
- Servidores upload / download
- Servidor web
- Aplicación amiloide



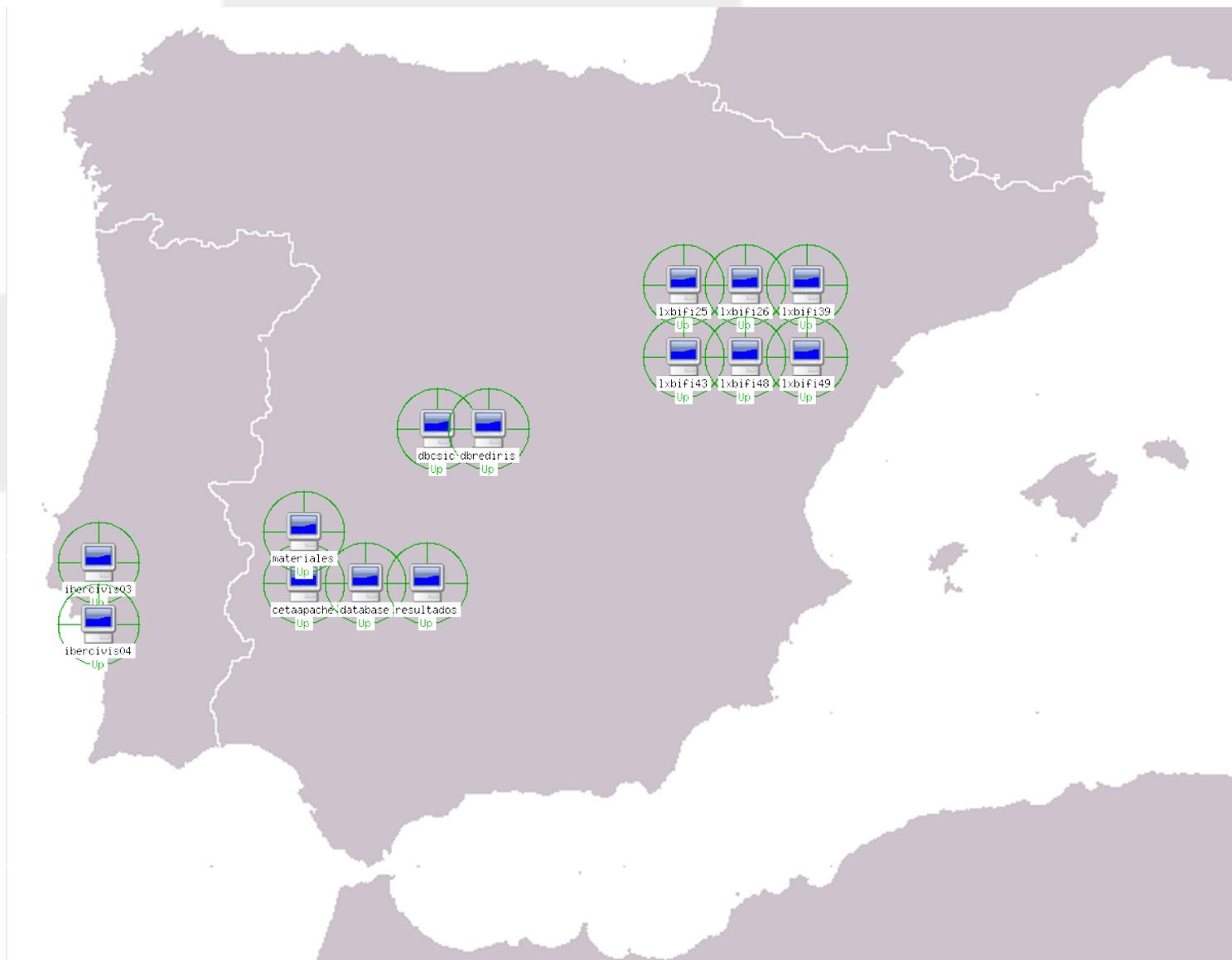
Actividad Instantánea



Infraestructura software



Infraestructura (Nagios)



Limitación en las Aplicaciones



- Segmentable en runs cortos
- Duración Job del orden de 30 minutos
- Uso de RAM máximo de 500 Mbytes
- Paralelismo nulo
- Datos enviados por Job: orden de MBytes
- Datos devueltos por Job: orden de MBytes
- Programas que NO necesiten licencias

Cuestionario para descargar en www.ibercivis.es en la sección IBERCIVIS → lanza tu aplicación.
Enviar a: info-ibercivis@bifi.es



Lanzamiento de una aplicación

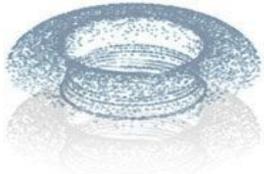
■ Qué hace Ibercivis

- Adaptación de la aplicación
- Compilación contra API Boinc
- Introducción de checkpoints
- Compilación multiplataforma
- Empaquetado librerías multiplataforma
- Verificar corrección respecto aplicación local
- Establecer mecanismo validación resultados
- Diseño y uso de la API para gráficos aplicación
- Registro en el servidor de Boinc

■ Qué hace el investigador:

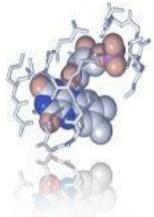
- Adaptación de la aplicación
- Verificar corrección respecto aplicación local
- Establecer mecanismo validación resultados
- Diseño y uso de la API para gráficos aplicación

Aplicaciones



Fusión

Una estrella en tu ordenador



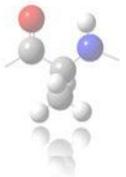
Docking de proteínas

En busca de fármacos contra el cáncer



Materiales

Simulación de sistemas magnéticos



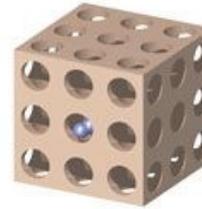
Neurosim

Una inmersión en la estructura molecular de la memoria



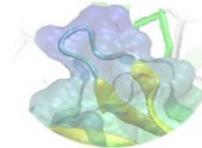
Nanoluz

Luz a escala nanométrica



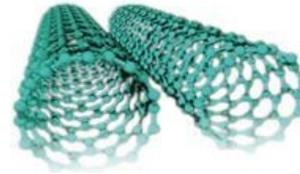
Adsorción

Simulación de fluidos moleculares confinados



Amiloide

Búsqueda de fármacos para enfermedades amiloides neurodegenerativas



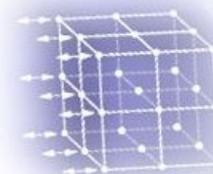
Cuanticables

Simulación de cables cuánticos



Sanidad

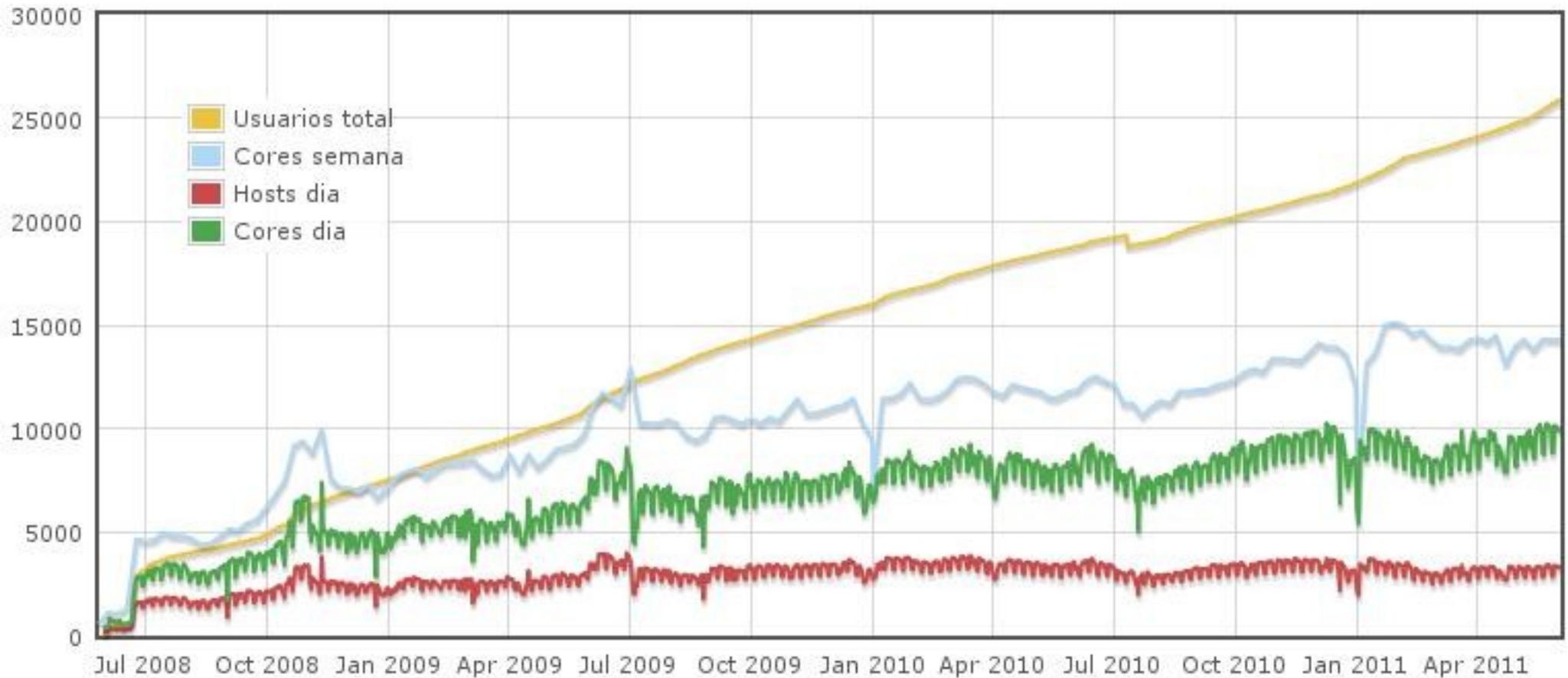
Simulación del transporte de partículas



Criticalidad

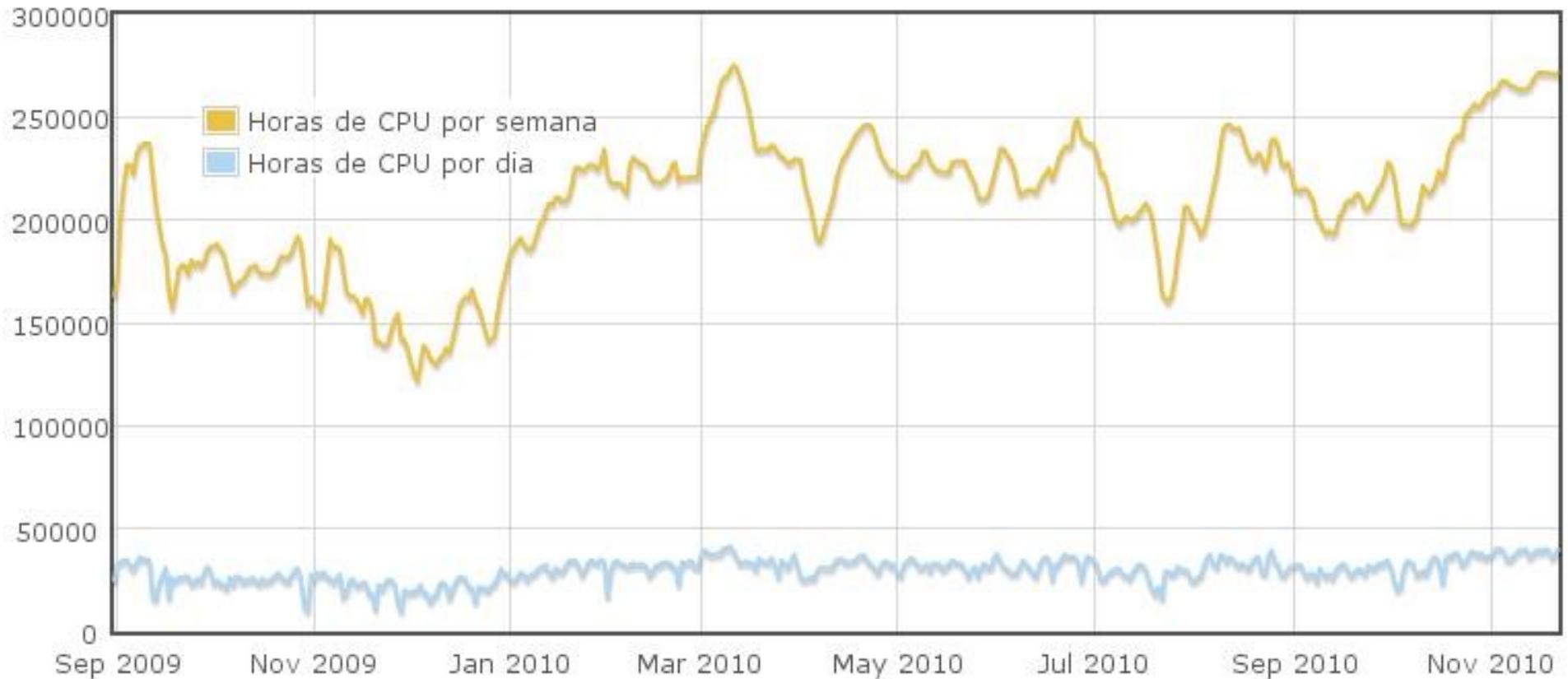
transporte electrónico en sistemas desordenados con propiedades fractales

Ibercivis en cifras



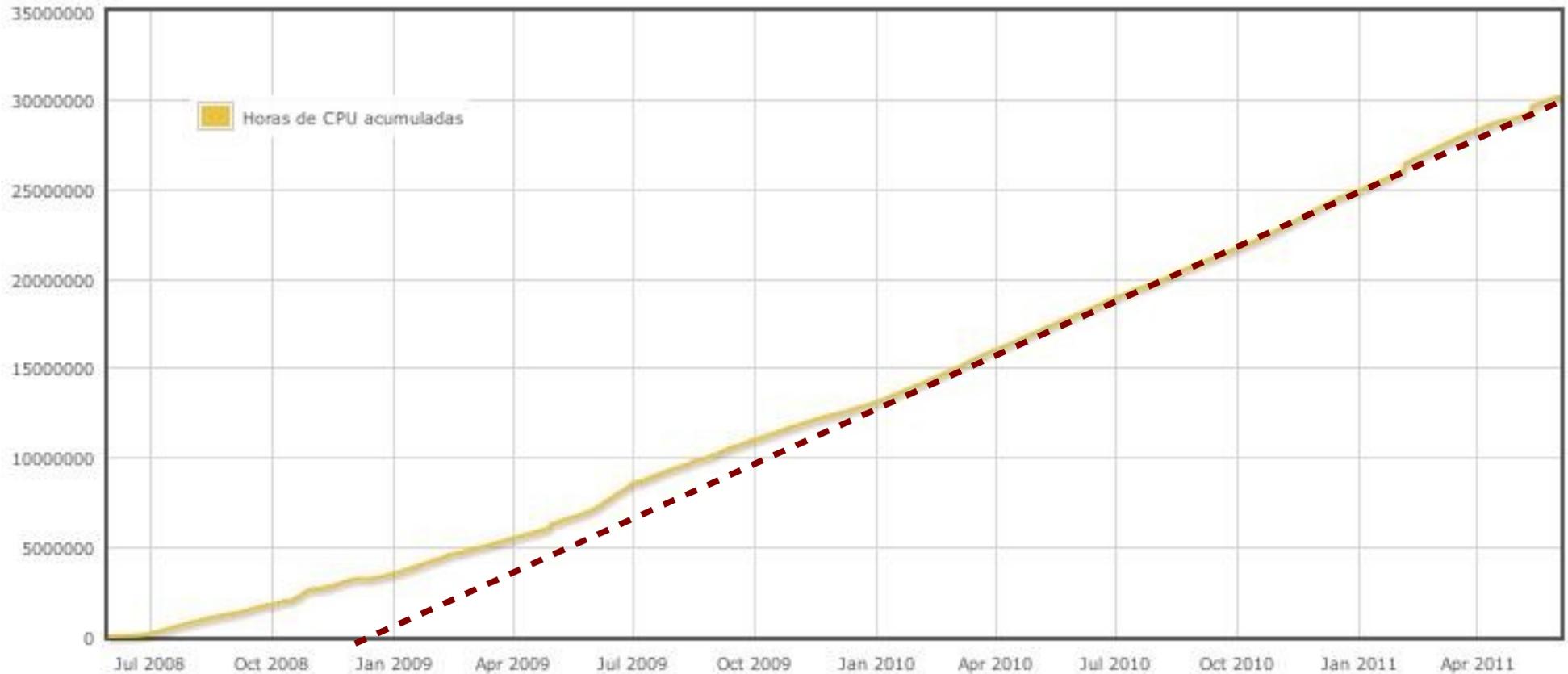
- 27.000 usuarios, 3000 usuarios/día, 10000 cores/día
- Investigadores de Portugal, Argentina, México, Cuba y España

Potencia Equivalente Ibercivis



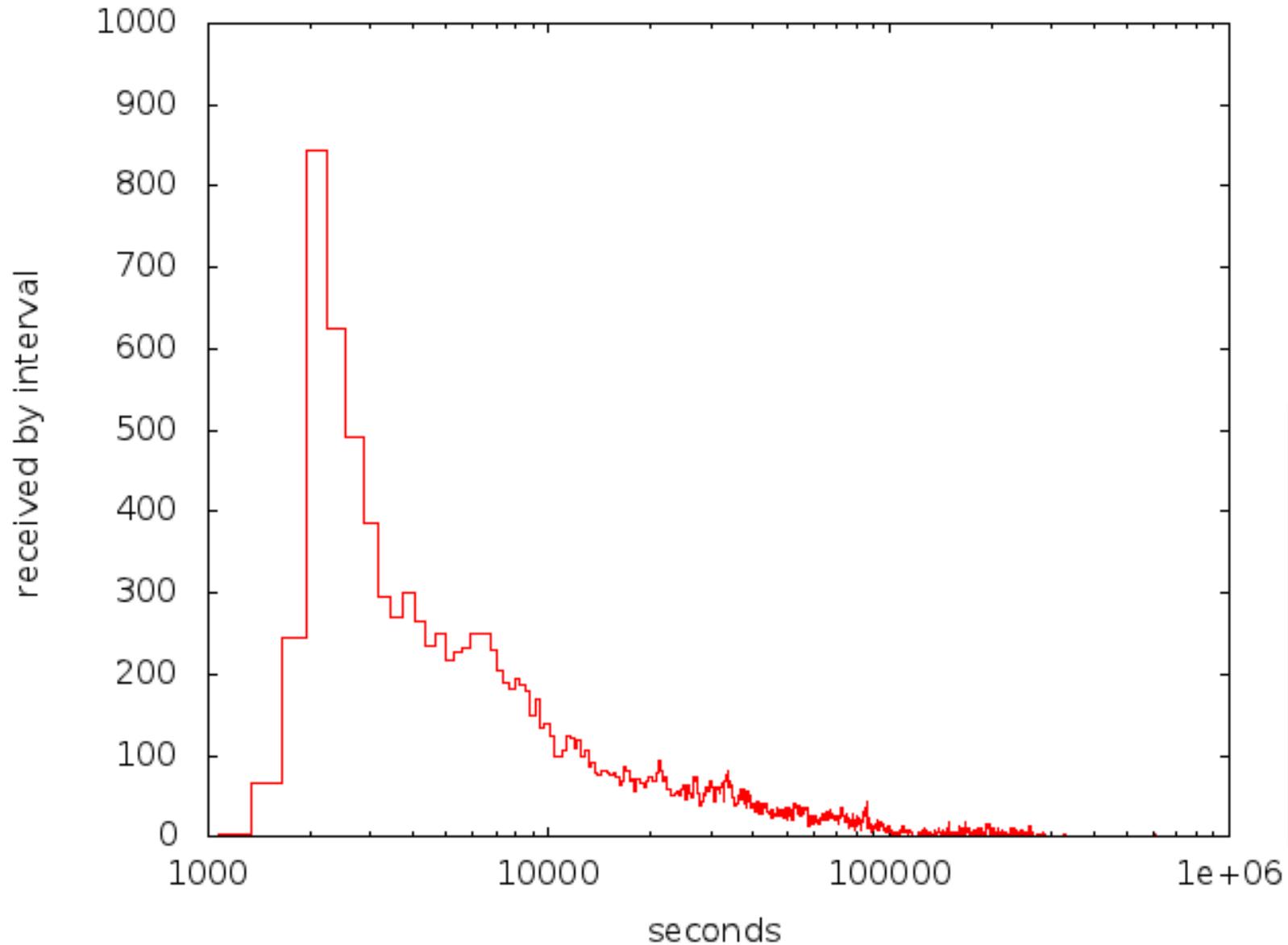
- Fluctuaciones de un factor 4 diario, un factor 2 semanal
- Difícil evaluar potencia equivalente instantánea

Potencia Media Equivalente



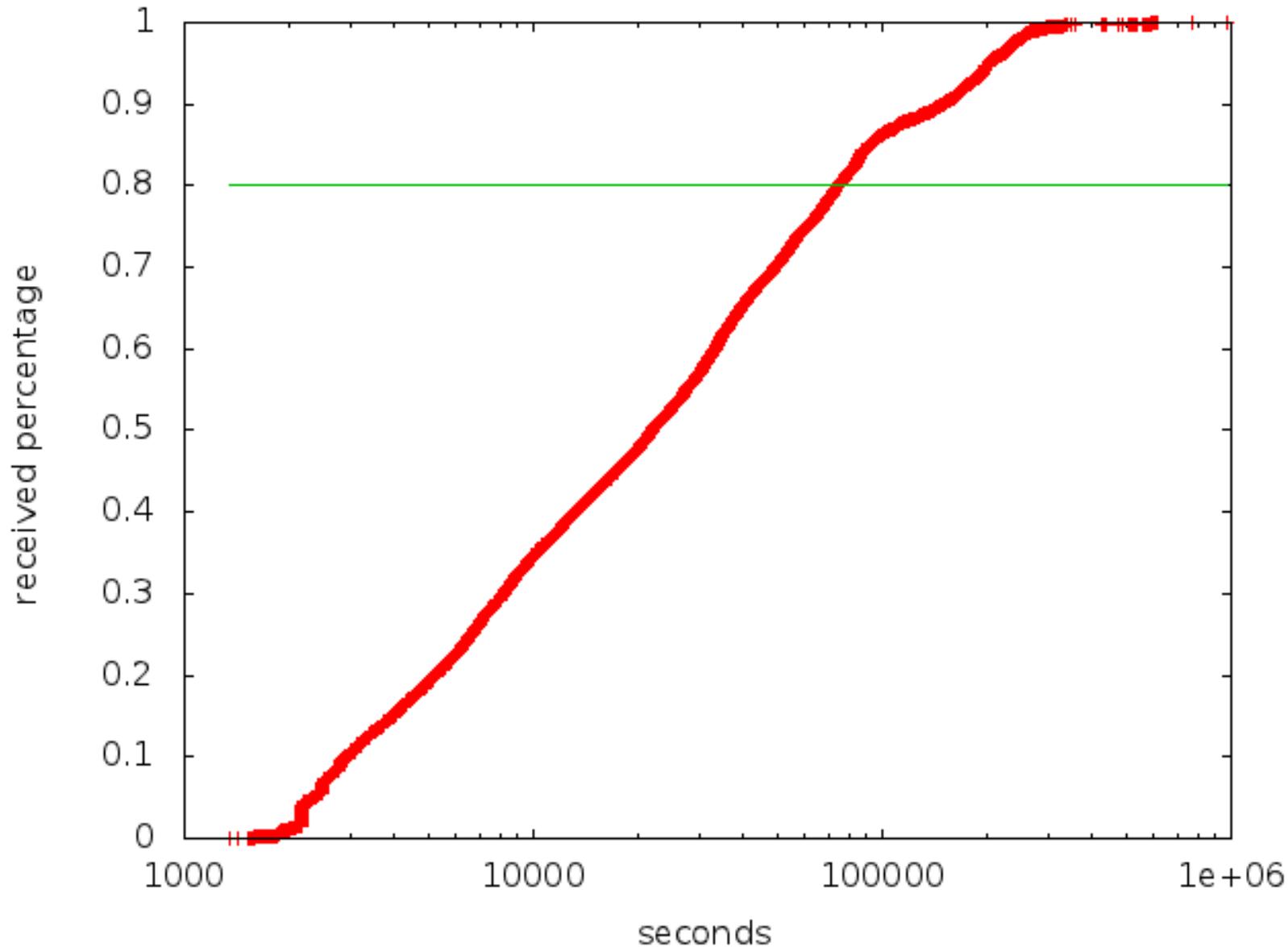
- Pendiente de la recta: 12.5 Millones horas/año
- Cluster Equivalente: 1500 Cores

Eficiencia de Ibercivis I



- Número de resultados recibidos por intervalo de tiempo

Eficiencia de Ibercivis II



- Tiempo de Latencia de Ibercivis: Espera adicional para recibir el 80% de los trabajos: **23 horas**

Desarrollo del proyecto



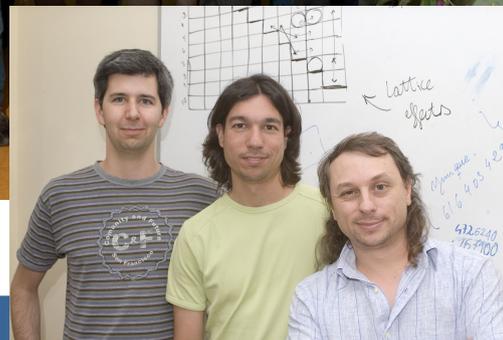
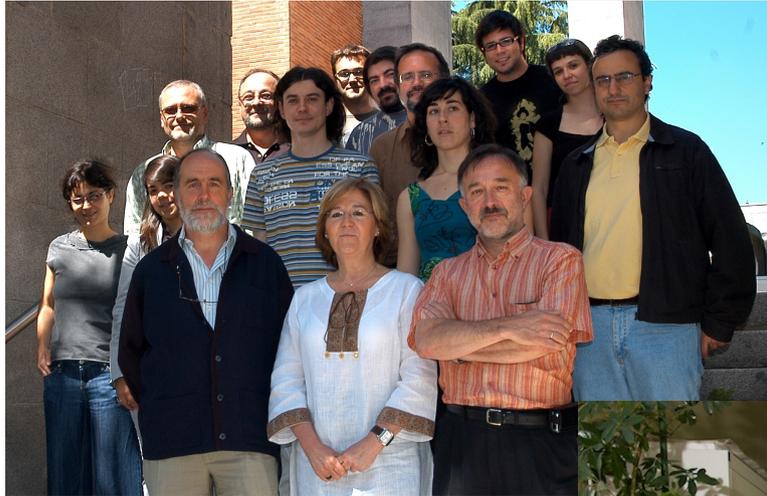
- Actividades de Divulgación
- Proyecto Hispano-Luso
- Aplicaciones de Argentina, Brasil, Méjico y Cuba
- Infraestructura en Brasil e iniciando en Cuba
- Creación Fundación Ibercivis
- Proyectos Europeos



Ciemat
Centro de Investigaciones
Energéticas, Medioambientales
y Tecnológicas



Parte del equipo de trabajo





Gracias por su atención

