



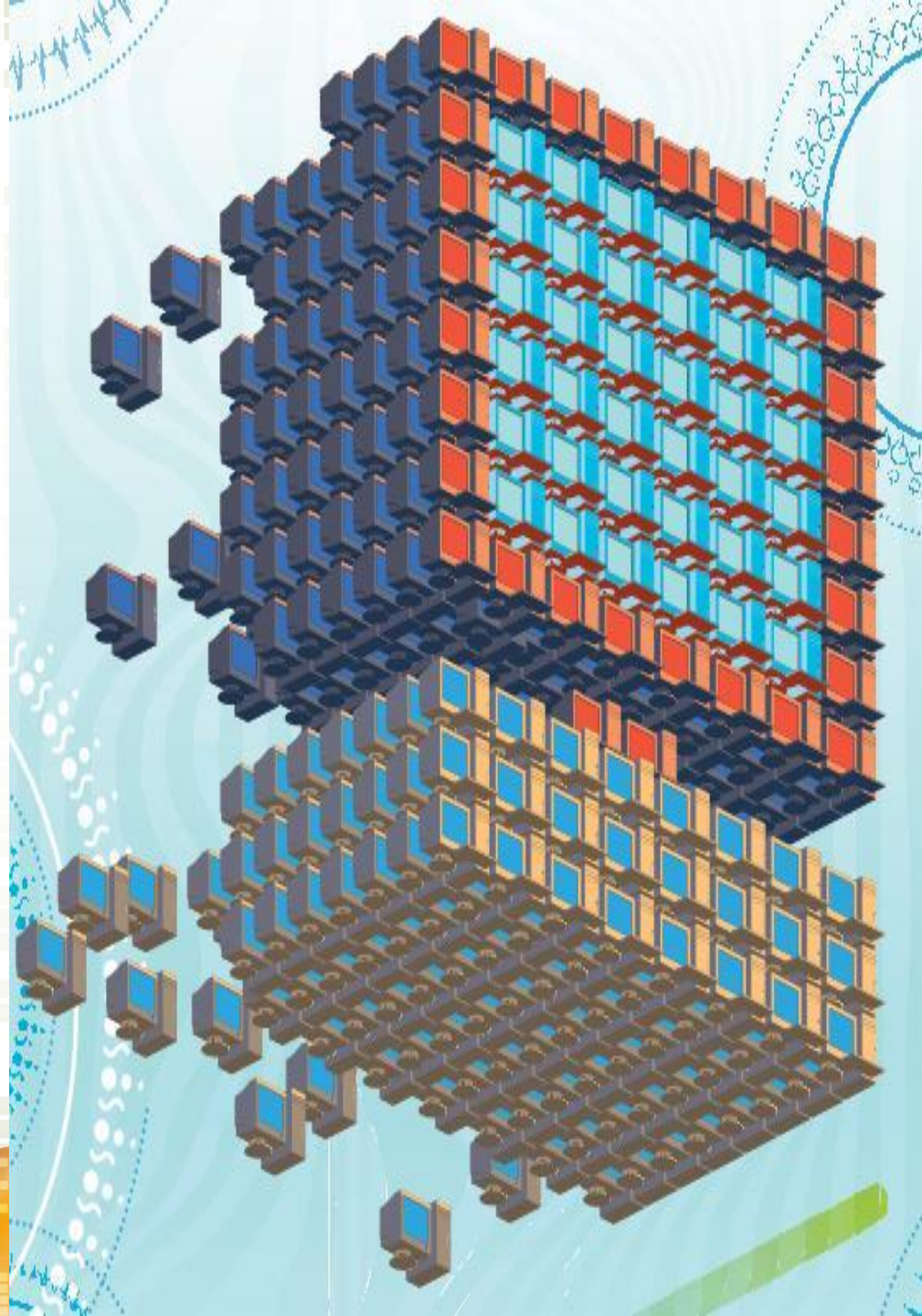
PUCP

Legión: Proyecto de desktop grid computing aprovechando las aulas informáticas



AGENDA

- Supercómputo en el mundo
- ¿Qué es BOINC?
- ¿Qué es Legión?
- Uso en la PUCP
- Difusión
- Conclusiones



Supercómputo en el mundo



<http://www.conceivablytech.com/wp-content/uploads/2010/05/supercomputer2.jpg>

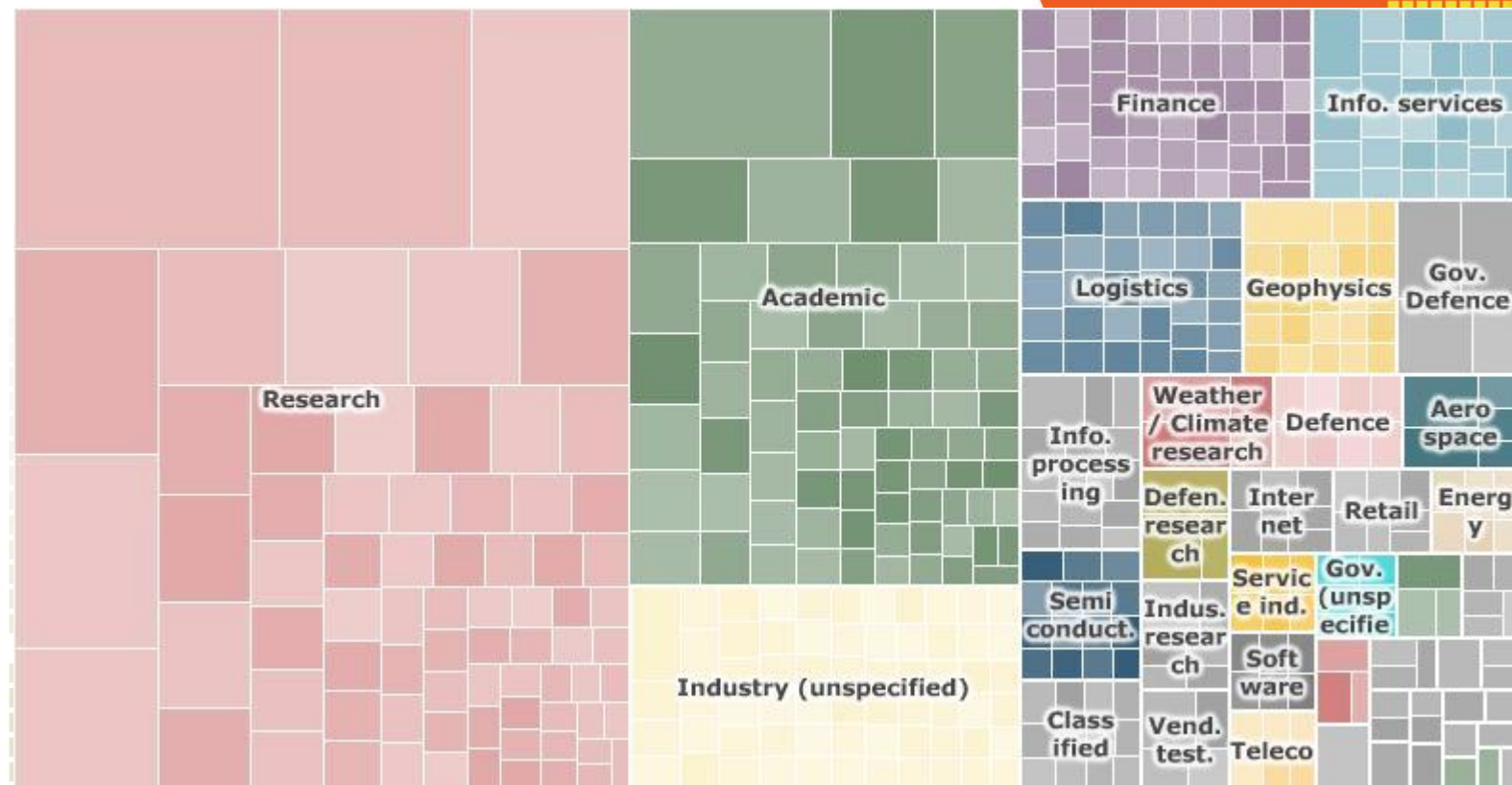


Supercómputo en el mundo

- **Diversos campos de investigación no pueden ser abordados de manera satisfactoria sin la ayuda de una elevada capacidad de cómputo.**
- **La web <http://www.top500.org/> lista a las 500 supercomputadoras más potentes del mundo.**
- **El portal de la BBC clasificó el listado de los 500 sistemas según las áreas de estudio, sistemas operativos y país de origen.**

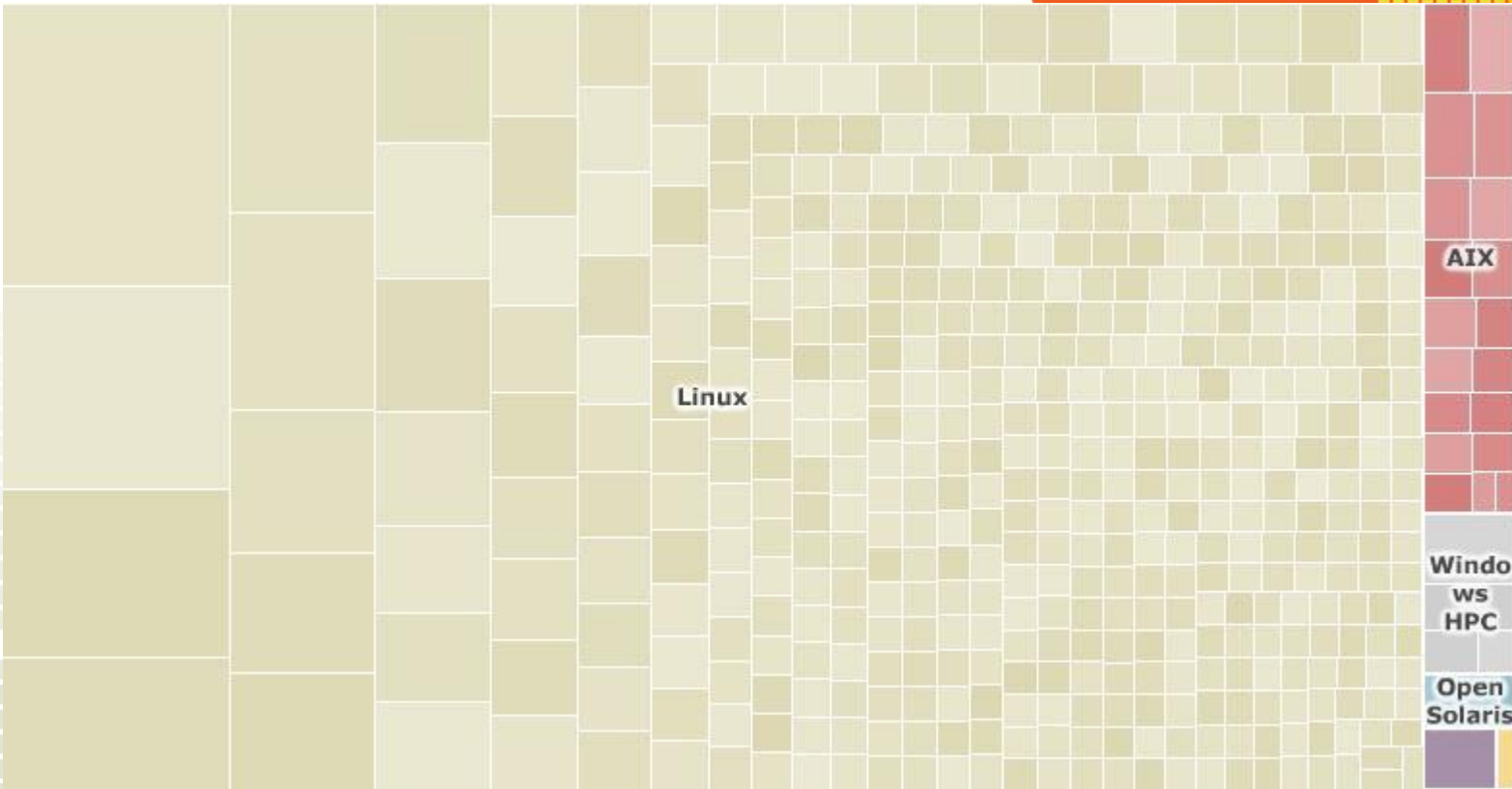


Agrupación por área de estudio



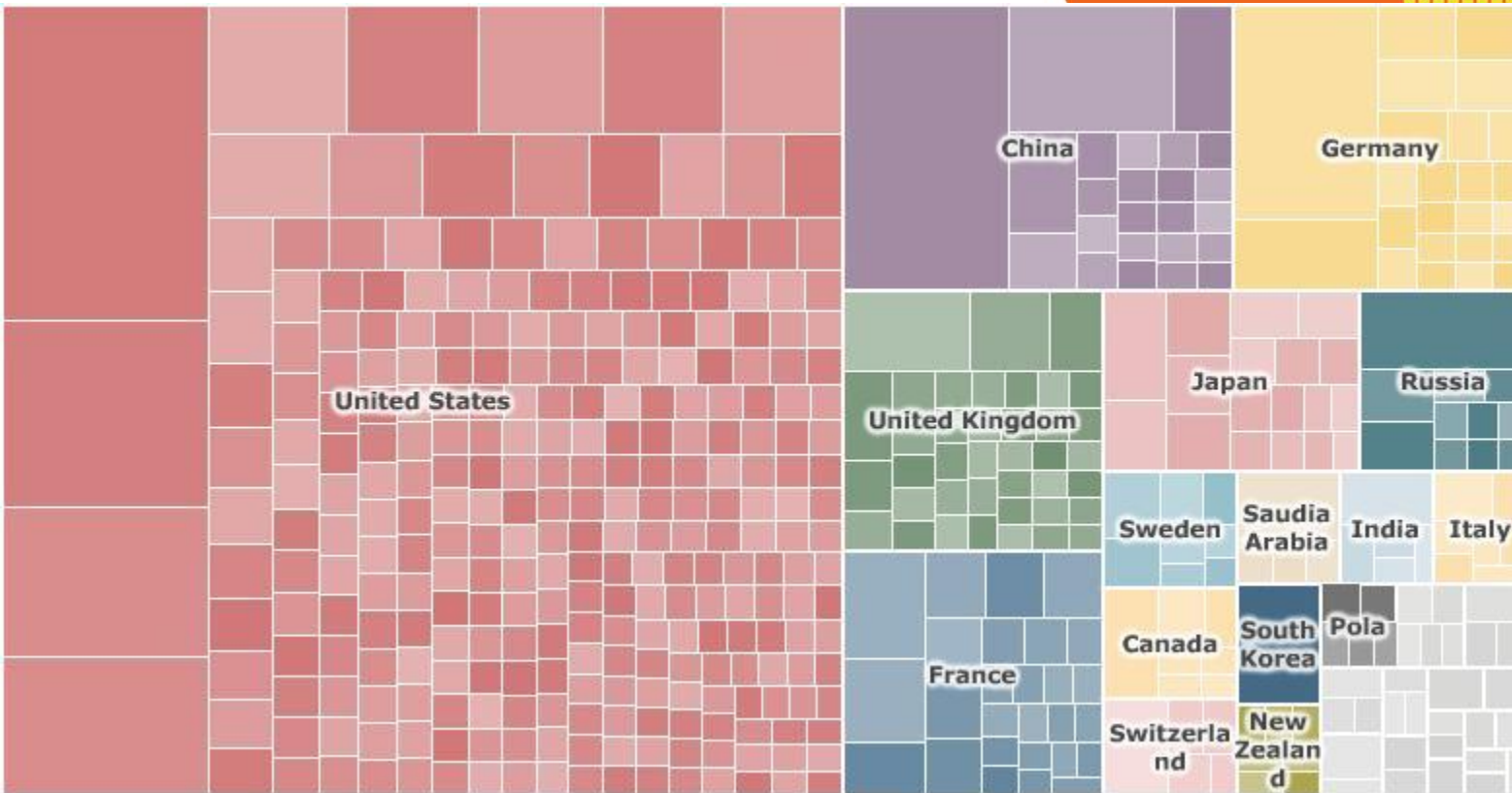
Fuente: <http://www.bbc.co.uk/news/10187248>

Agrupación por Sistema Operativo



Fuente: <http://www.bbc.co.uk/news/10187248>

Agrupación por país



Fuente: <http://www.bbc.co.uk/news/10187248>

Texas Advanced Computing Center (TACC)



System Name: **Ranger**
Operating System: Linux
of Processing Cores: 62,976
Total Memory: 123TB
Peak Performance: 579.4TFlops
Total Disk: 1.73PB
(shared)31.4TB (local)

US
\$59M

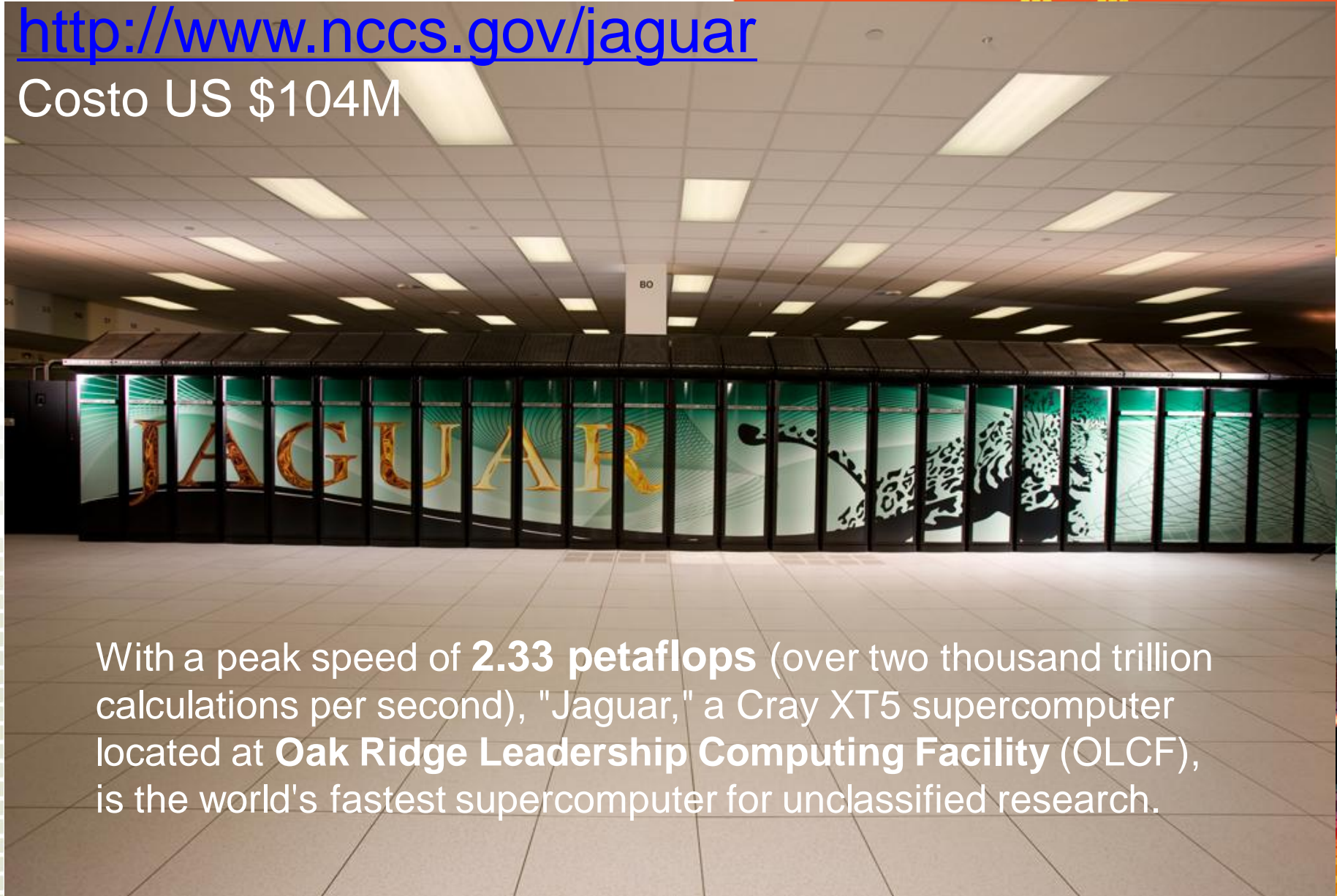


System Name: **Lonestar 4**
Operating System: Linux
of Processors: 22,656 (compute)
Total Memory: 44 TB
Peak Performance: 302 TFLOPS
Total Disk: 276TB(local)
1000TB(global)

US
\$12M

<http://www.nccs.gov/jaguar>

Costo US \$104M



With a peak speed of **2.33 petaflops** (over two thousand trillion calculations per second), "Jaguar," a Cray XT5 supercomputer located at **Oak Ridge Leadership Computing Facility (OLCF)**, is the world's fastest supercomputer for unclassified research.

K-Computer (Japón)

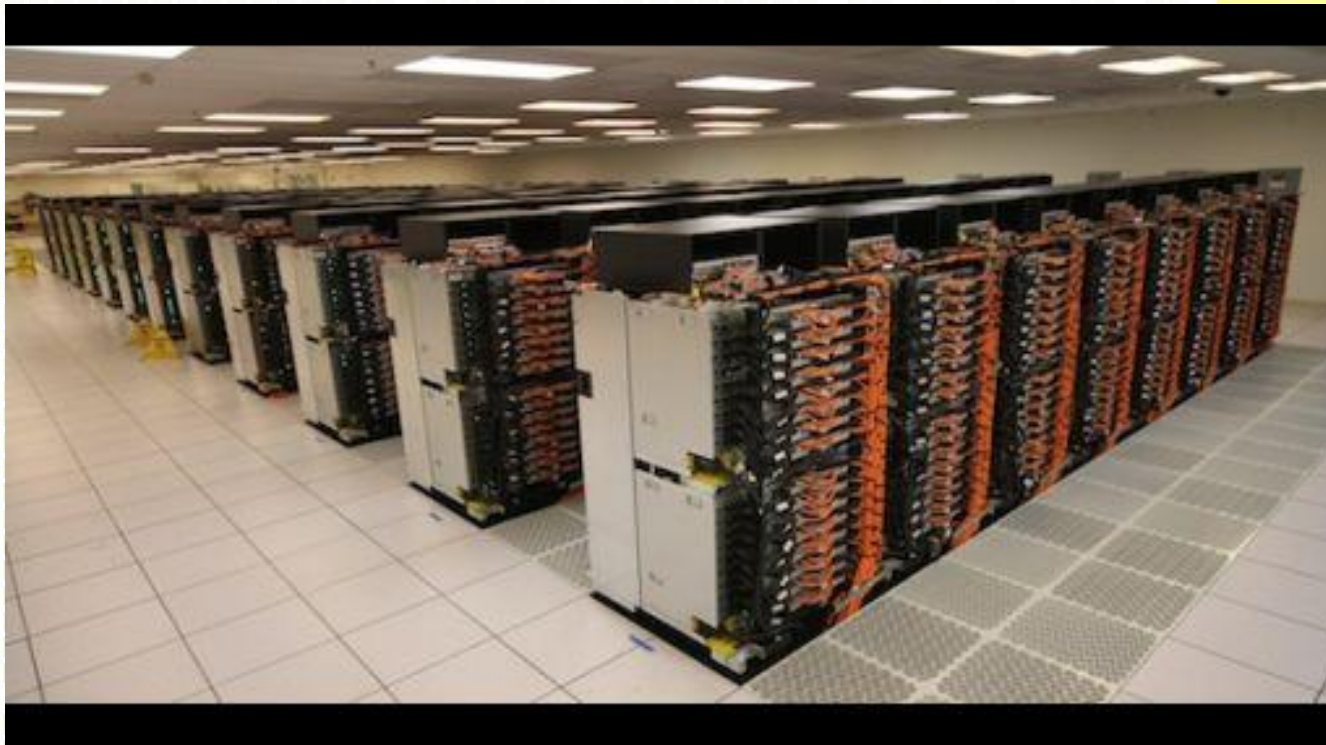
10.51 PetaFLOPs,
68.544 procesadores de 8 núcleos cada uno (548.352
núcleos)


Costo: 975M euros



#1 Top 500 List: IBM Sequoia (US)

16.32 PetaFLOPs,
98,304 nodos (1,572,864 núcleos)
Costo: \$250M





• **Tradicionalmente se utiliza un cluster de computadoras dedicado, el cual puede representar un costo muy elevado en términos de adquisición, espacio, instalación y mantenimiento.**

• **Por otro lado muchas universidades en Latinoamérica cuentan con laboratorios equipados con computadoras de última generación, cuya potencia generalmente excede las necesidades del estudiante.**

• **La capacidad de cómputo ociosa puede ser aprovechada por el sistema BOINC.**



¿Qué es BOINC?



.Berkeley Open Infrastructure for Network Computing.

.Sistema de grid computing voluntario.

.Proyectos de cómputo intensivo reciben apoyo de voluntarios que ofrecen sus computadoras.

.BOINC toma recursos de cómputo no usado por los voluntarios.

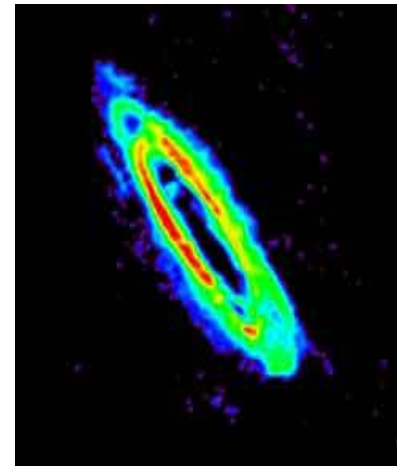


Algunos proyectos que usan BOINC



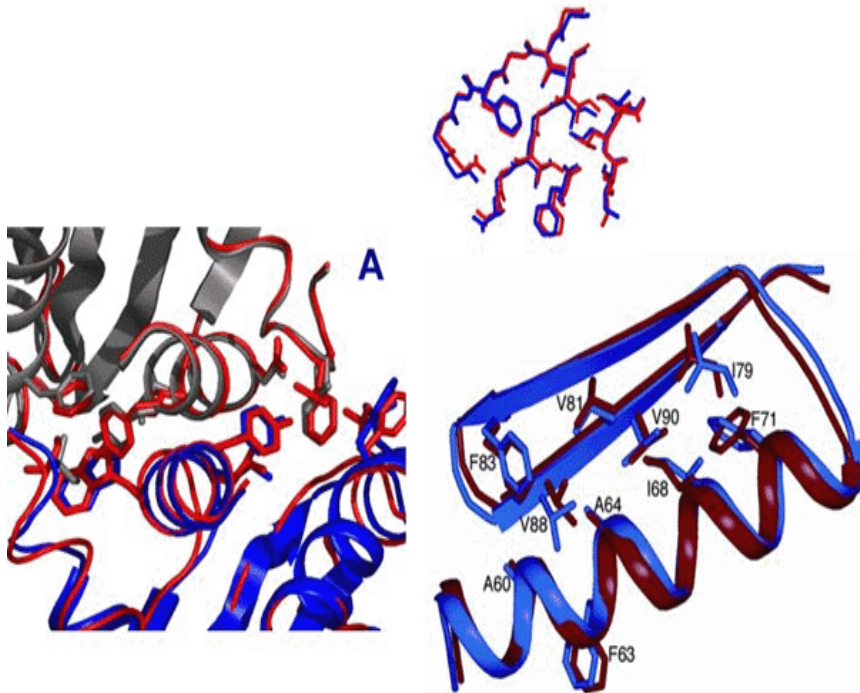
Proyecto SETI

Búsqueda de Inteligencia extraterrestre a través de una gran red de radiotelescopios.



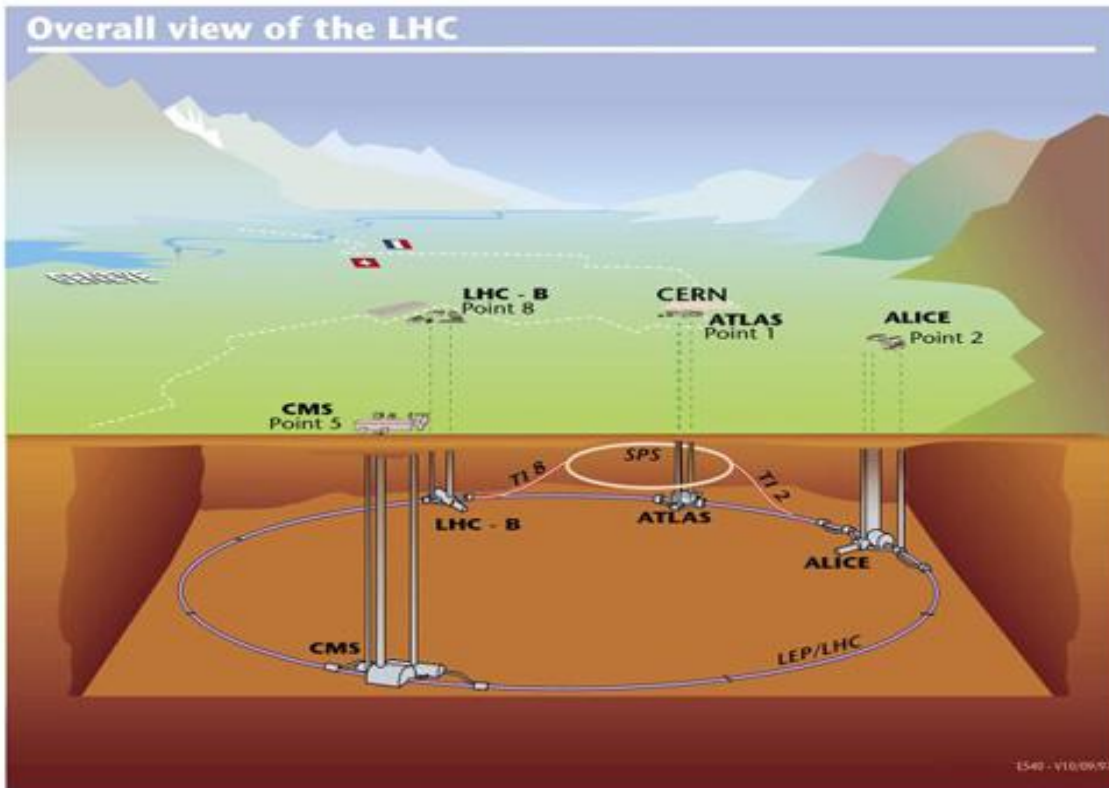
Proyecto Rosetta

Determinación de las formas tridimensionales de las proteínas.



Proyecto LHC

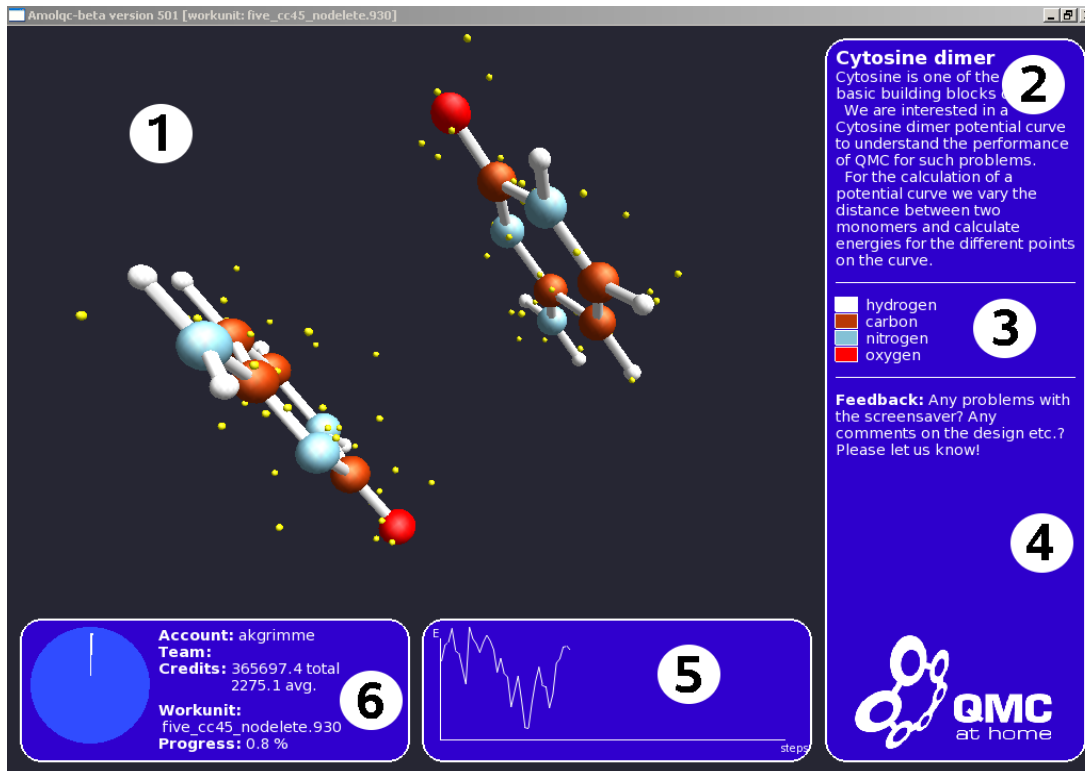
Simulación de partículas circulando por el LHC (Large Hadron Collider) del CERN.



Fuente <http://lhcat.home.cern.ch/lhcat/home/>

Proyecto QMC

Aplicación de la física cuántica a la química mediante simulaciones de Monte Carlo.





world
community
grid™



600,000
volunteers

in 80
countries



using 2
million PCs

By pooling resources, IBM
World Community Grid PCs
have performed computations
in the equivalent of



600,000
years

to help
process

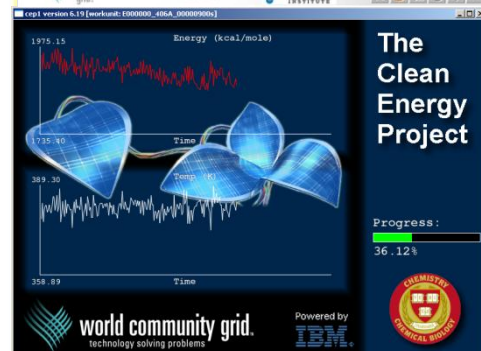
21
research projects



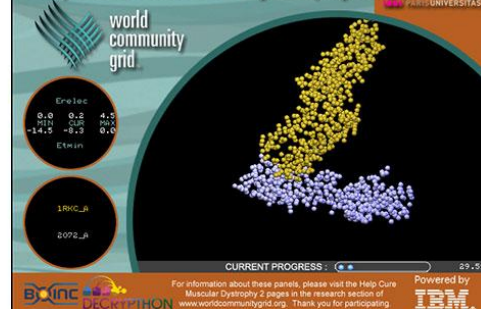
<http://www.worldcommunitygrid.org/>



Powered by IBM. The Scripps Research Institute



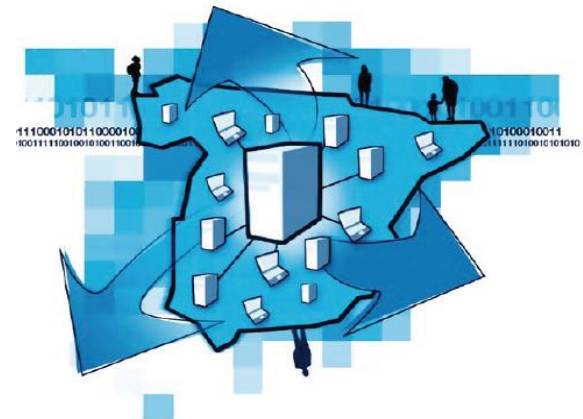
Help Cure Muscular Dystrophy 2 UPMC



IBERCIVIS



- **Creado en España por el Instituto de Biocomputación y Física de Sistemas Complejos de la Universidad de Zaragoza, CIEMAT, CSIC y RedIRIS**
- **Pretende involucrar al máximo número posible de ciudadanos en la computación voluntaria con BOINC.**
- **IBERCIVIS acerca a la ciudadanía investigaciones punteras y la hace partícipe de la generación de conocimiento científico**



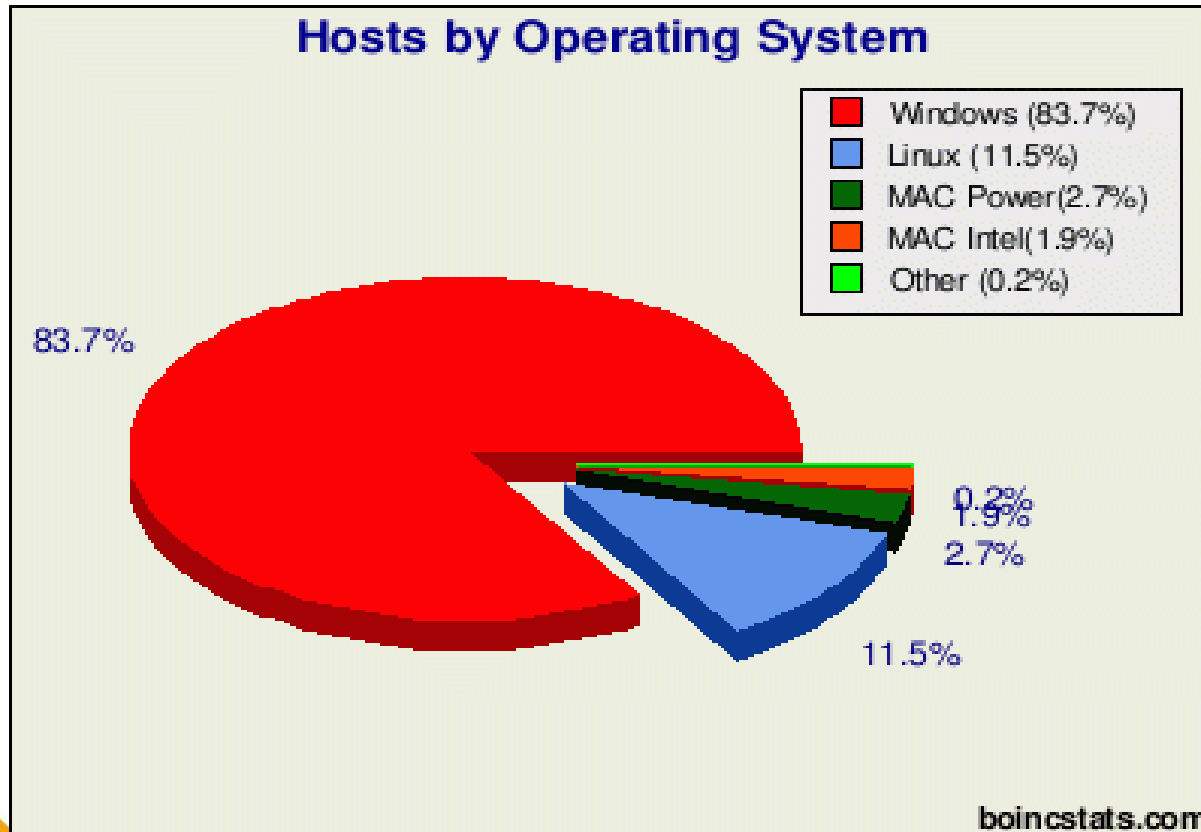
Fuente: <http://www.ibercivis.es>

¿Qué ofrece BOINC?

- **Es software libre.**
- **Asignación flexible de recursos.**
- **Múltiple plataforma:**
 - **Windows,**
 - **Linux,**
 - **Unix,**
 - **Mac Os X.**
- **Alto desempeño y escalabilidad.**



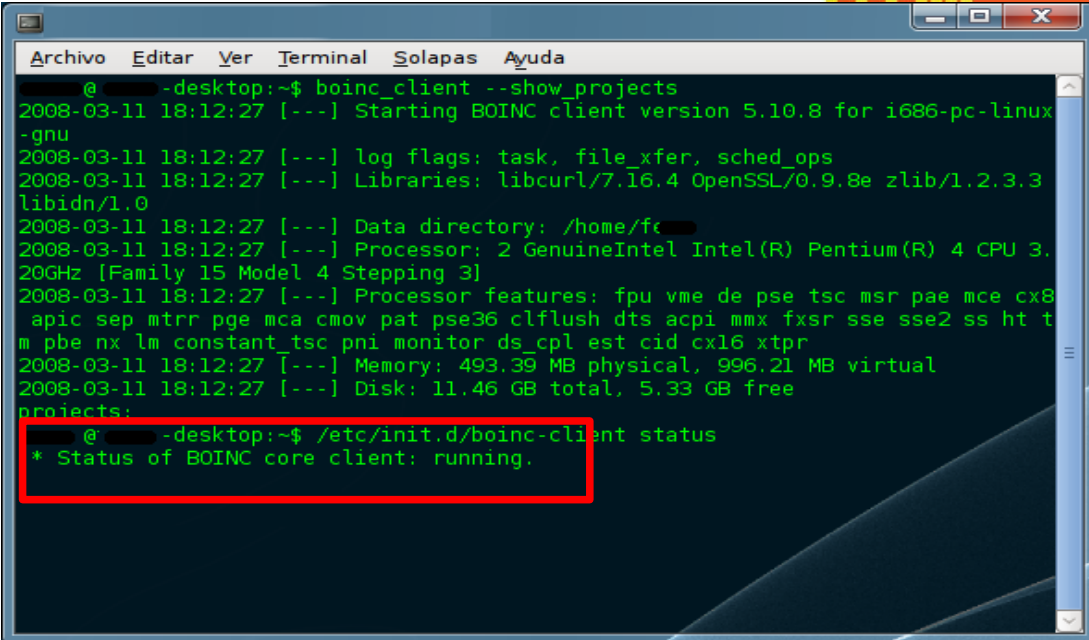
Estadísticas de uso por Sistema Operativo



Fuente: <http://boincstats.com>

Cliente BOINC

Linux:

A terminal window titled "Archivo Editor Ver Terminal Solapas Ayuda" showing the output of the command "boinc_client --show_projects". The output displays system information including date, time, log flags, libraries, data directory, processor details, and memory/disk status. The last line shows the command "/etc/init.d/boinc-client status" and its output: "* Status of BOINC core client: running.", which is highlighted with a red rectangle.

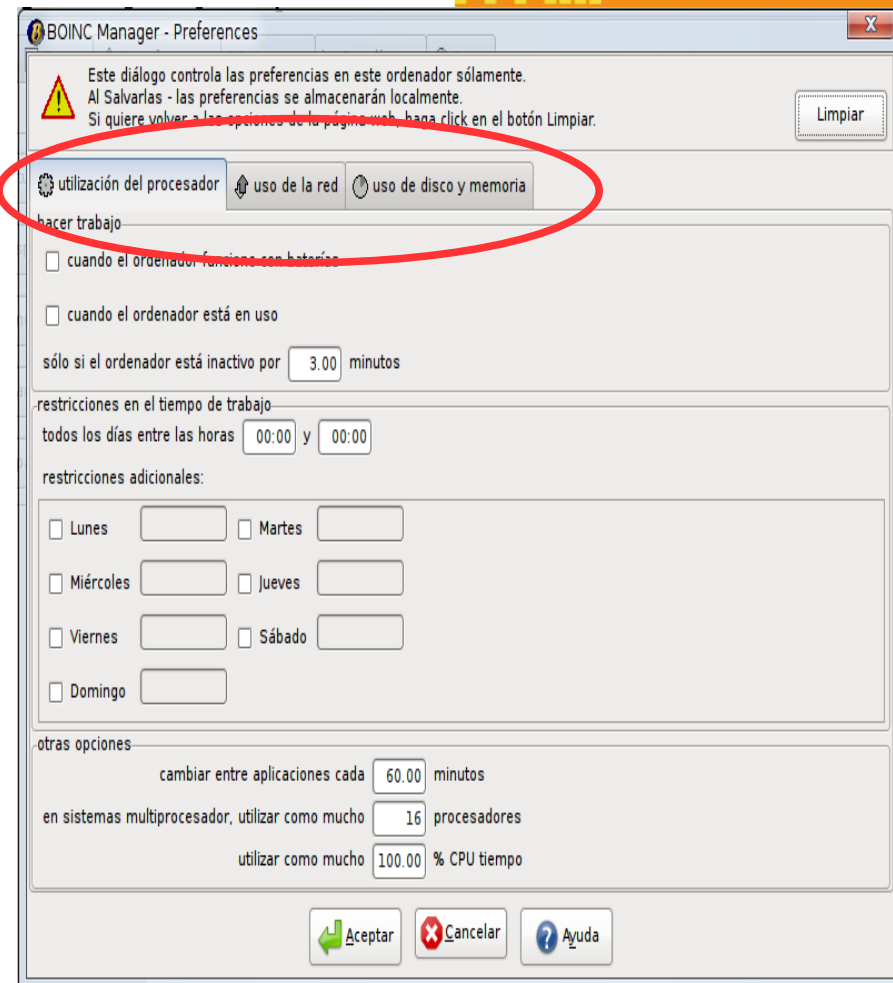
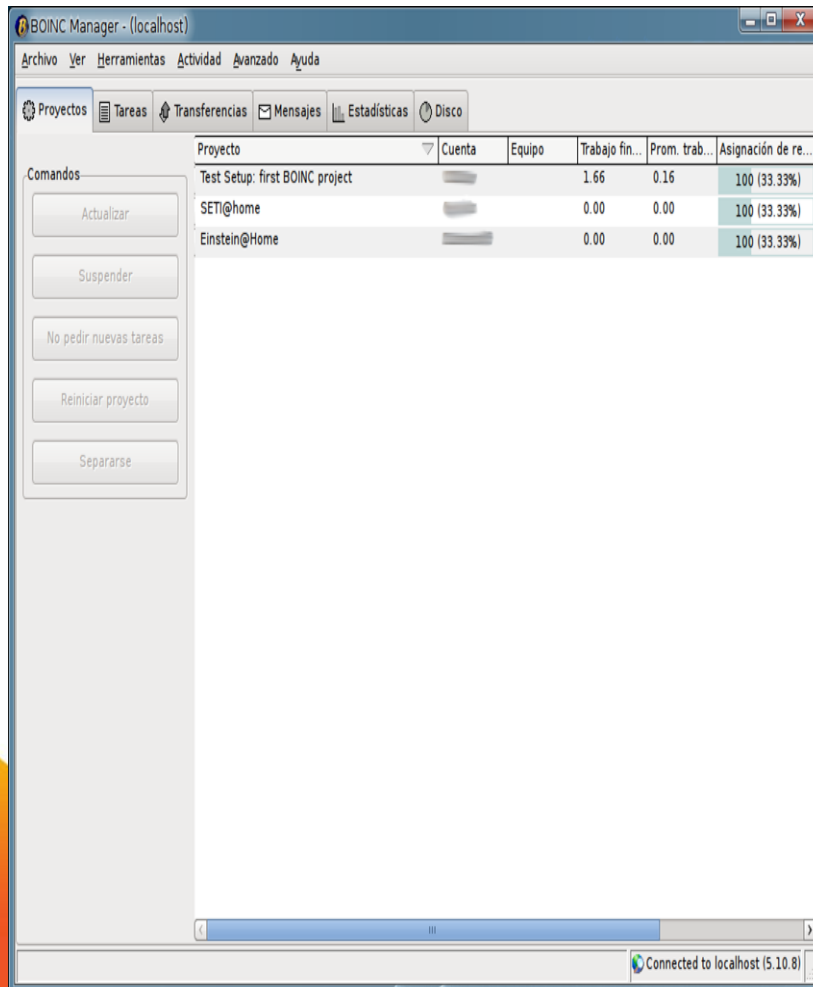
```
#####@#####-desktop:~$ boinc_client --show_projects
2008-03-11 18:12:27 [---] Starting BOINC client version 5.10.8 for i686-pc-linux-gnu
2008-03-11 18:12:27 [---] log flags: task, file_xfer, sched_ops
2008-03-11 18:12:27 [---] Libraries: libcurl/7.16.4 OpenSSL/0.9.8e zlib/1.2.3.3 libidn/1.0
2008-03-11 18:12:27 [---] Data directory: /home/fe#####
2008-03-11 18:12:27 [---] Processor: 2 GenuineIntel Intel(R) Pentium(R) 4 CPU 3.20GHz [Family 15 Model 4 Stepping 3]
2008-03-11 18:12:27 [---] Processor features: fpu vme de pse tsc msr pae mce cx8 apic sep mtrr pge mca cmov pat pse36 clflush dts acpi mmx fxsr sse sse2 ss ht tm pbe nx lm constant_tsc pni monitor ds_cpl est cid cx16 xtpr
2008-03-11 18:12:27 [---] Memory: 493.39 MB physical, 996.21 MB virtual
2008-03-11 18:12:27 [---] Disk: 11.46 GB total, 5.33 GB free
projects:
#####@#####-desktop:~$ /etc/init.d/boinc-client status
* Status of BOINC core client: running.
```

Windows:



Ciente BOINC

BOINC Manager: GUI de configuración del cliente

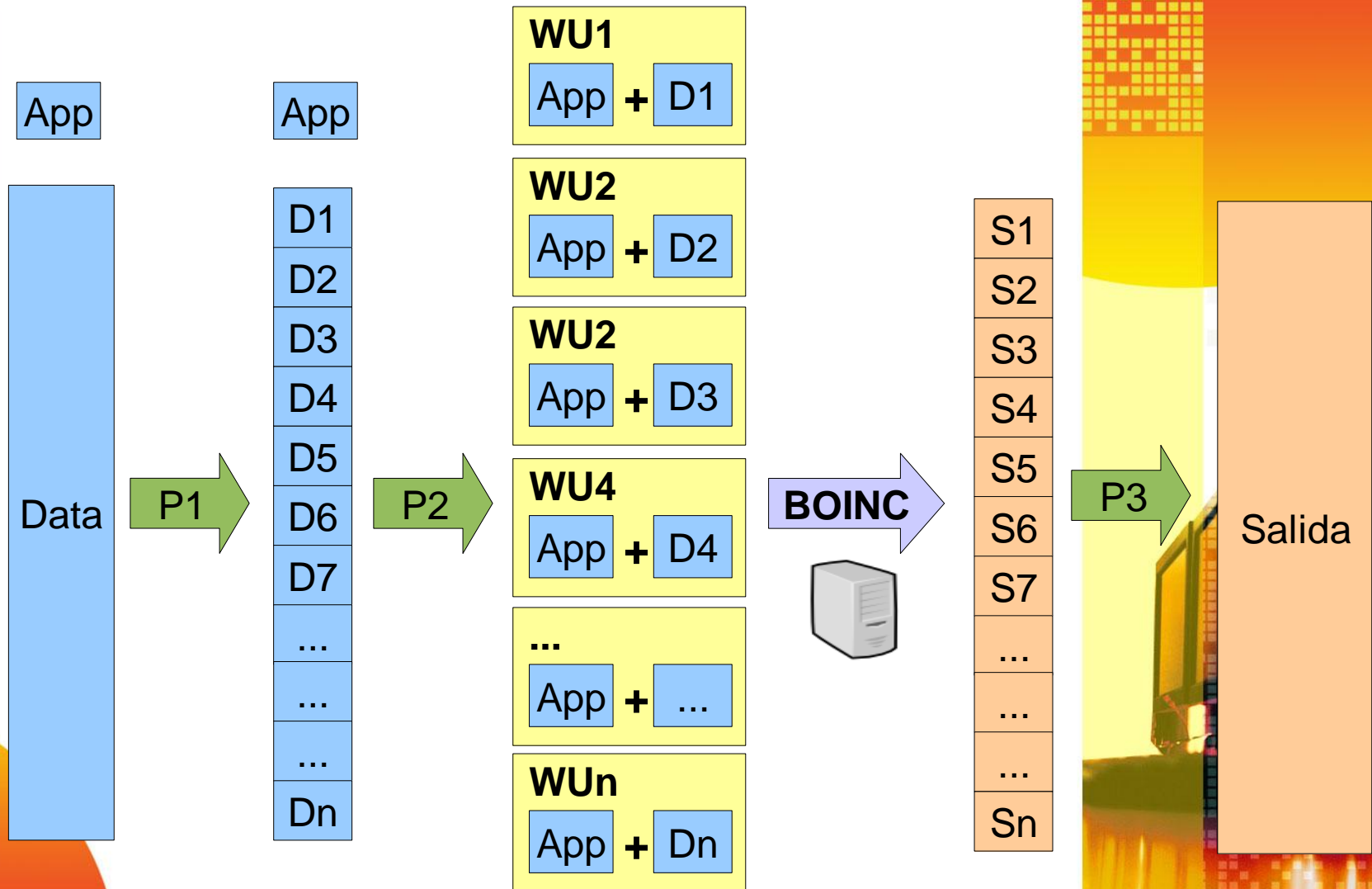


¿Cómo funciona?

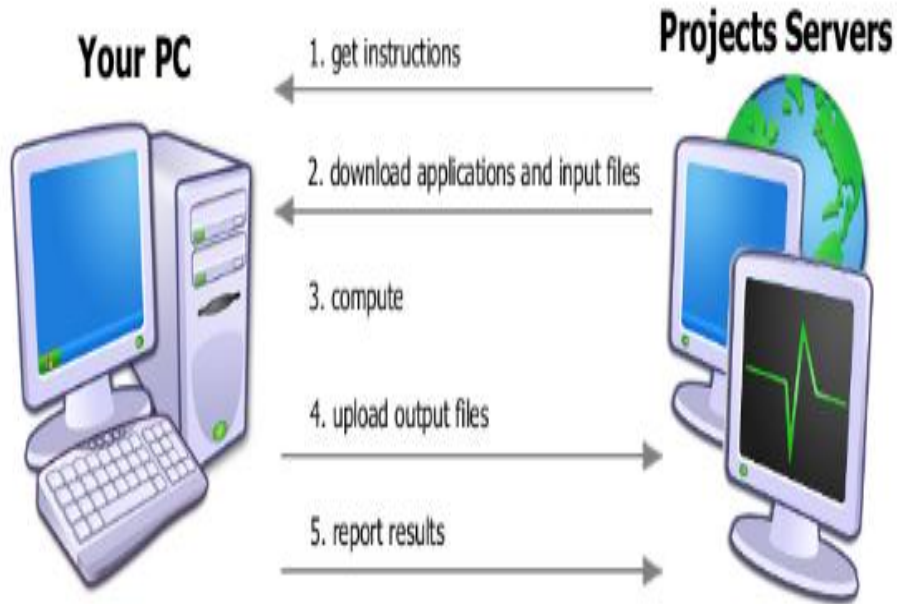
- Un proyecto BOINC ejecuta múltiples unidades de trabajo (workunits) (WU)
- Cada WU ejecuta una misma aplicación.
- Un WU recibe datos desde **archivos** y almacena los resultados en otros archivos.



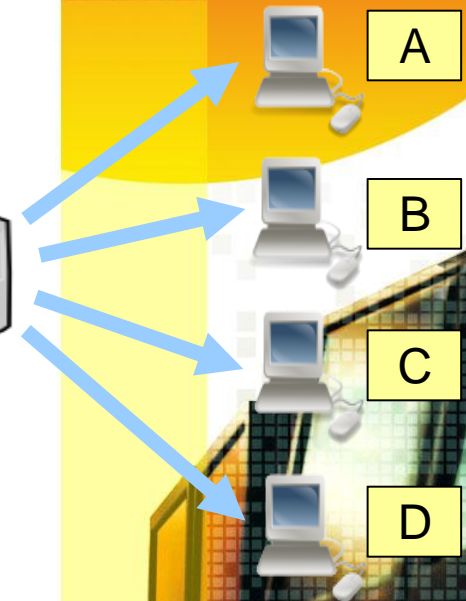
Desarrollos P1, P2, P3



¿Cómo funciona?



A
B
C
D



Arquitectura BOINC

Servidor BOINC

- **Controla ejecución de proyectos.**
- **Define cuentas de usuarios.**

Cliente BOINC

- **Ejecuta aplicaciones específicas a un proyecto.**
- **Envía reportes al servidor.**



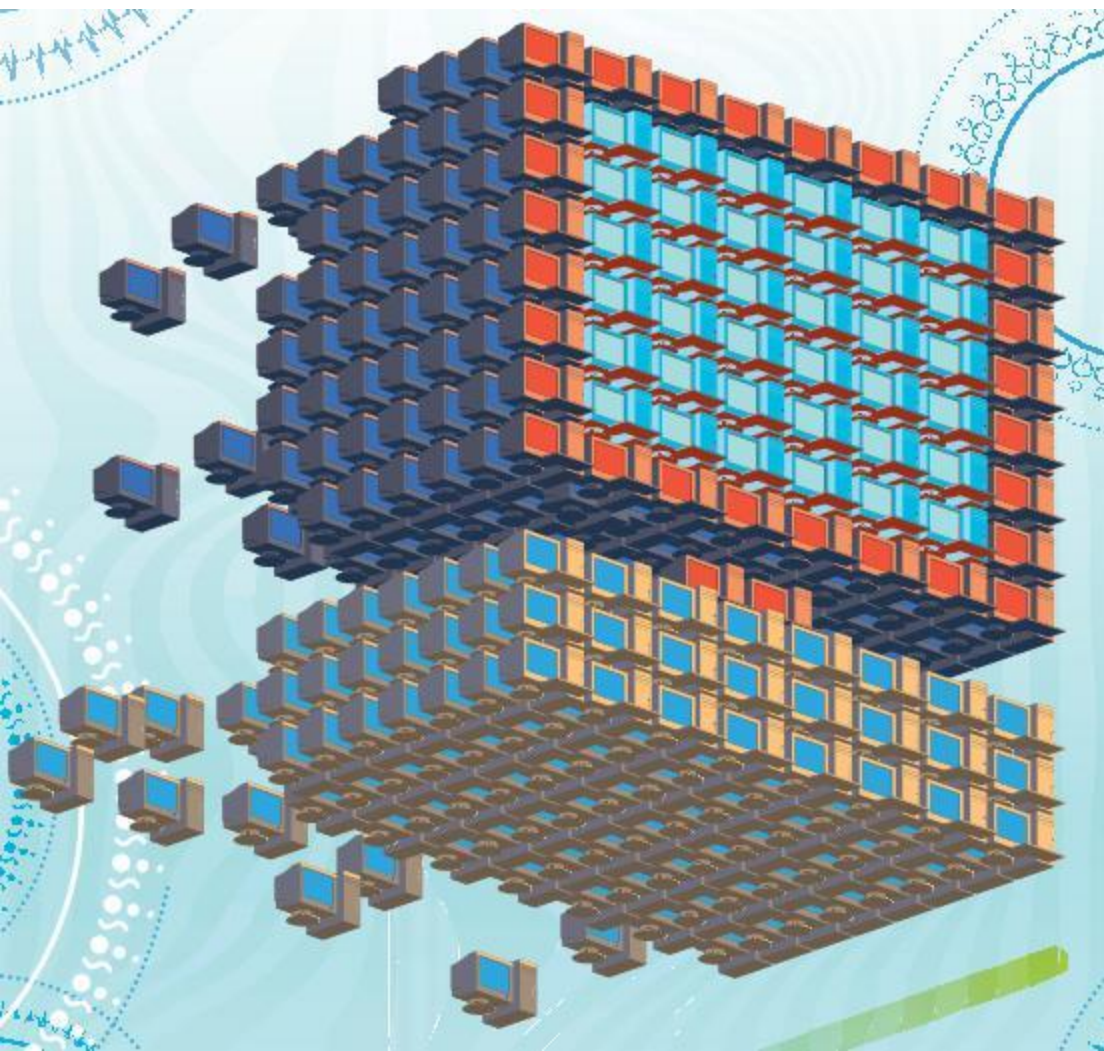
Interacción con BOINC

.La grid BOINC es administrada por expertos en sistemas operativos e informática.

.Los investigadores dependen de los administradores para correr cada nueva simulación.



¿Qué es **LEGIÓN**?



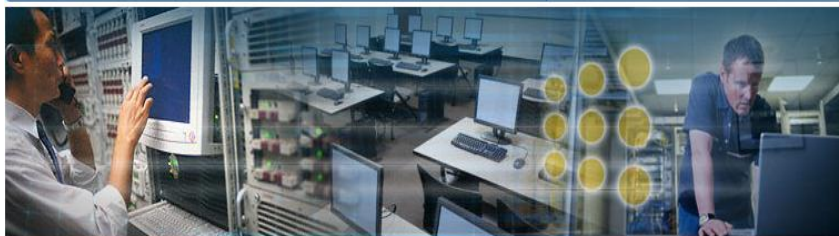
¿Qué es **LEGIÓN**?

- Es el servicio de **Grid Computing** de la **PUCP** basado en **BOINC**.
- Brinda al investigador una **amigable interfaz web**, para el envío de tareas.
- Configura los **nodos remotamente** para asociarlos a un proyecto.



<http://legion.pucp.edu.pe>

http://legion.pucp.edu.pe



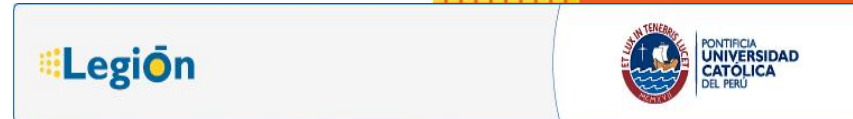
[Inicio](#) > [Proyecto Estadística](#) > Simulación - Detalle

Estadística

<< < 1 > >> Nueva Tarea | Eliminar - Cancelar | Actualizar

Fecha	Nombre de la Tarea	Progreso	Data Set's	Resultado	
10:42:31 22-09-08	Simulacion Inicial	100%	1	↓	□
12:55:29 07-10-08	Prueba 100 7-10	59%	100	↓	□
15:41:40 02-10-08	Prueba Simulacion ah	100%	1	↓	□
10:58:31 02-10-08	Prueba Simulacion ag	100%	1	↓	□
10:40:05 02-10-08	Prueba Simulacion ad	100%	1	↓	□
10:24:17 02-10-08	Prueba Simulacion ab	100%	1	↓	□
15:49:03 01-10-08	Prueba Simulacion xv	100%	1	↓	□
15:27:32 01-10-08	Prueba Simulacion xr	100%	1	↓	□
11:25:56 01-10-08	Prueba Simulacion xo	100%	1	↓	□
11:31:26 30-09-08	Prueba Simulacion xn	100%	1	↓	□

Servicio ofrecido por la Dirección de Informática Académica (DIA)
Diseñado y Modificado por DIA | Contactarse con el administrador del servicio
Av. Universitaria N° 1801, San Miguel, Lima - Perú | Teléfono: (511) 626-2000 | © 2008 Pontificia Universidad Católica del Perú



[Inicio](#) > [Proyecto Estadística](#) > Simulación - Detalle

Estadística

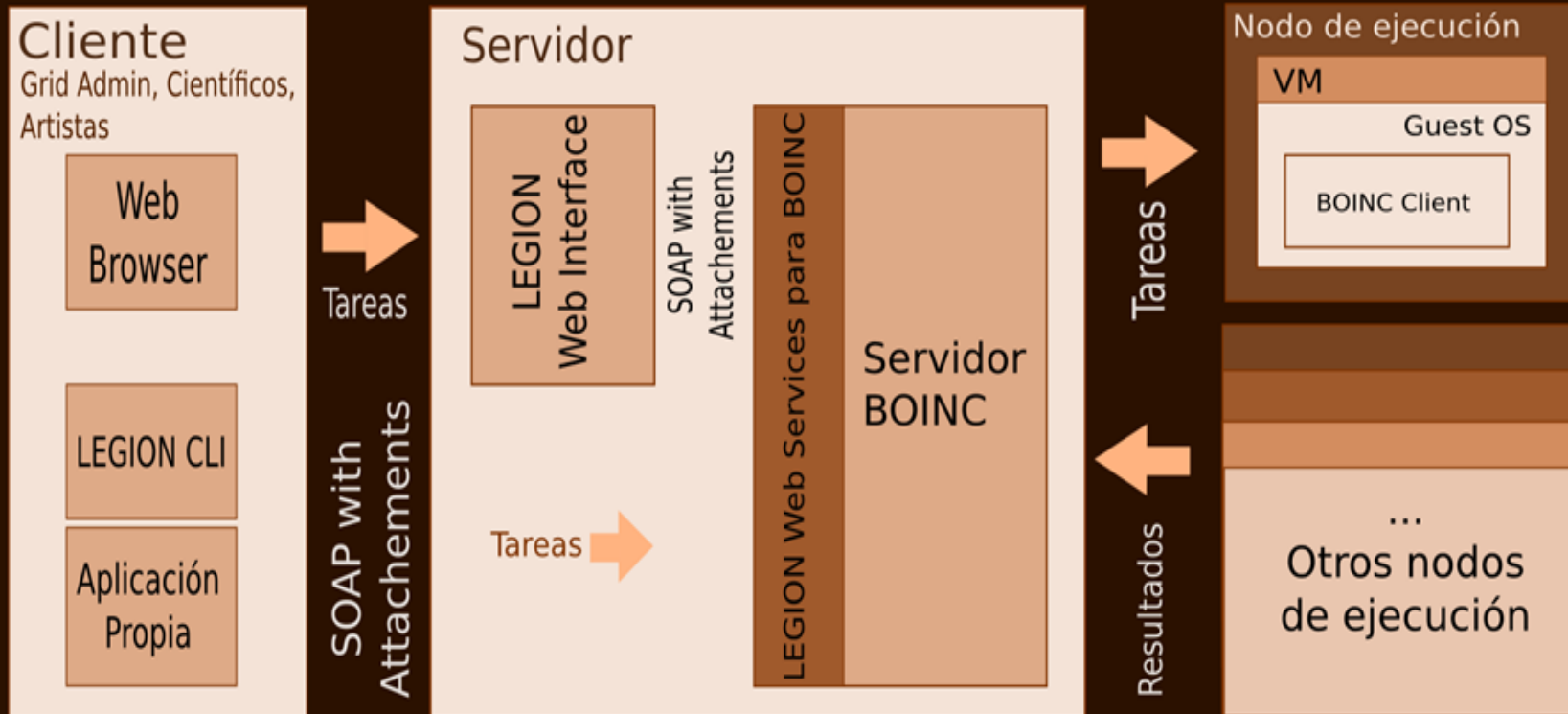
<< < 1 > >> Nueva Tarea | Eliminar - Cancelar | Actualizar

Fecha	Nombre de la Tarea	Progreso	Data Set's	Resultado	
10:42:31 22-09-08	Simulacion Inicial	100%	1	↓	□
12:55:29 07-10-08	Prueba 100 7-10	100%	100	↓	□
15:41:40 02-10-08	Prueba Simulacion ah	100%	1	↓	□
10:58:31 02-10-08	Prueba Simulacion ag	100%	1	↓	□
10:40:05 02-10-08	Prueba Simulacion ad	100%	1	↓	□
10:24:17 02-10-08	Prueba Simulacion ab	100%	1	↓	□
15:49:03 01-10-08	Prueba Simulacion xv	100%	1	↓	□
15:27:32 01-10-08	Prueba Simulacion xr	100%	1	↓	□
11:25:56 01-10-08	Prueba Simulacion xo	100%	1	↓	□
11:31:26 30-09-08	Prueba Simulacion xn	100%	1	↓	□

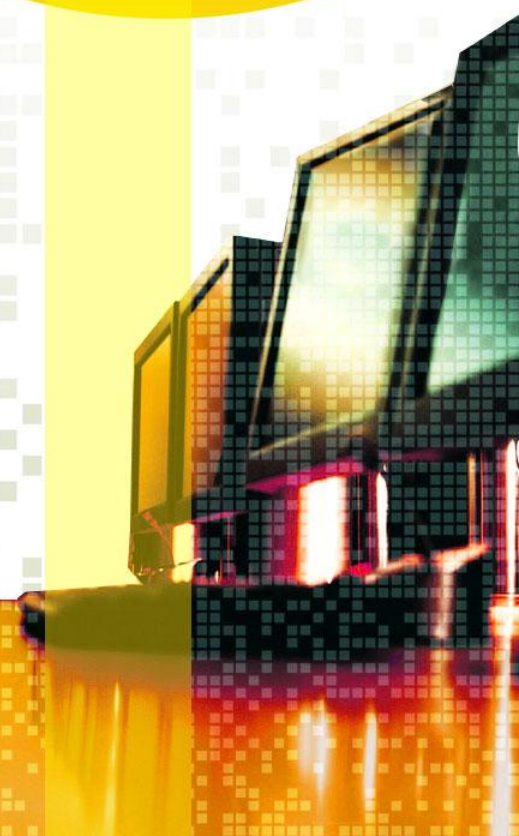
Servicio ofrecido por la Dirección de Informática Académica (DIA)
Diseñado y Modificado por DIA | Contactarse con el administrador del servicio
Av. Universitaria N° 1801, San Miguel, Lima - Perú | Teléfono: (511) 626-2000 | © 2008 Pontificia Universidad Católica del Perú

<http://legion.pucp.edu.pe>

ARQUITECTURA DE LEGION FRAMEWORK



Uso en la PUCP



BOINC en la PUCP

- **465** computadoras distribuídas en 15 laboratorios a lo largo del campus universitario.
- Computadoras **Quad Core, 4GB RAM.**
- Usuarios utilizan básicamente Windows 7 para ofimática y la web.

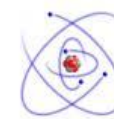


BOINC en la PUCP

- Windows ofrece un soporte nulo en herramientas de software para uso científico
- Se optó por el uso de **VMware con Scientific Linux (CERN)** en las estaciones.
- **VMWare** tomar el **50%** de la RAM y CPU.
- Potencia de cálculo estimada: **1.6 TeraFLOP.**



Scientific Linux



Uso de Software Libre

- Es importante usar software libre para un despliegue efectivo a gran escala de una aplicación.
- Usamos **Scientific Linux** en todos los nodos

Campo	Software Propietario	Software Libre
Simulación	<i>Matlab</i>	<i>Scilab, Octave</i>
Matemáticas	<i>Mathematica</i>	<i>Maxima</i>
Estadística	<i>SPSS, Minitab</i>	<i>PSPP, R, Openstat</i>
Animación 3D	<i>3D Studio</i>	<i>Blender</i>
Diseño 3D	<i>Autocad</i>	<i>Salome</i>
Programación	<i>MS Visual C</i>	<i>Eclipse</i>
GIS	<i>IDRISI</i>	<i>Grass</i>
Circuitos	<i>ORCAD</i>	<i>Kicad</i>

Consideraciones en consumo de energía

Computer state (24 hrs/day)	Typical power usage	Energy per month	Cost per month (USA)	Cost per month (Europe)
Idle	100 watts	73 kWh	\$5.84	€14,60
Active	150 watts	110 kWh	\$8.80	€22,00

Fuente: http://boinc.berkeley.edu/wiki/Heat_and_energy_considerations

En general puede haber un incremento del consumo eléctrico del orden del 50% ejecutando aplicaciones con BOINC.



Trabajos desarrollados

- **Validación del modelo de regresión binaria BBB Skew Probit. Dr. Jorge Bazán.**
- **Simulación del número de alumnos en pregrado en la PUCP empleando Cadenas de Markov Absorbentes. Ing. Eduardo Carbajal López.**



Trabajos GAE PUCP

- **Bombara et al. (E. Calvo, Y. Delgado, A. Gago) AD detector array: Diffractive and Photon-induced physics with ALICE - ALICE-INT-2010-014 v.1. (Trabajo en ALICE)**
- **C. Arguelles, M. Bustamante, A. M. Gago, IceCube expectations for two high-energy neutrino production models at active galactic nuclei, Journal of Cosmology and Astroparticle Physics 1012, 005, 2010 (Trabajo en Astrofísica)**



Trabajos GAE PUCP

- **Presentaciones de resultados en ALICE meetings:**
 - **Monte Carlos studies: for a Diffractive Gap Tagging Detector, E. Calvo, A. Gago, R. Helaconde**
 - **A study of the efficiency of the Bench test-ADA1, E. Calvo, A. Gago, R. Helaconde**
 - **Selecting diffractive events using the ADA1 and ADD1 detectors, E. Calvo, A. Gago y R.Helaconde.**

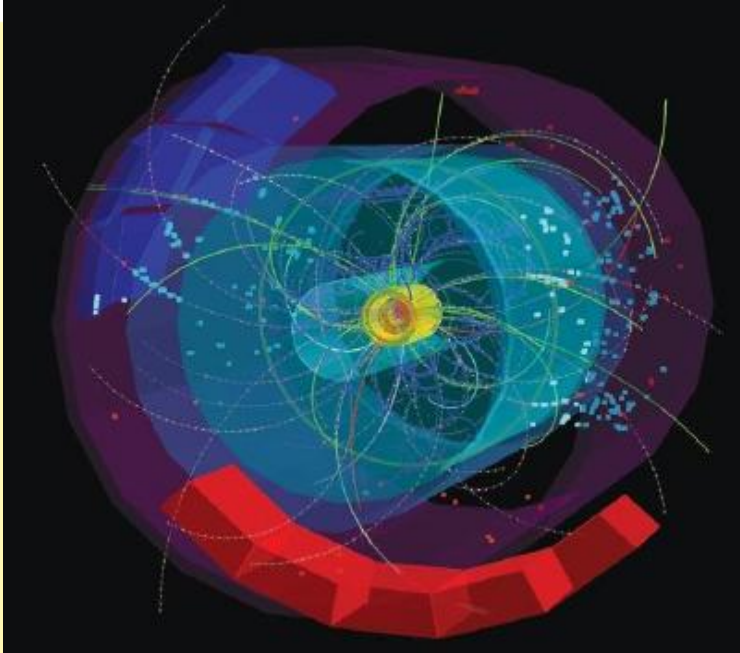


Ejemplo Caso Grupo AE

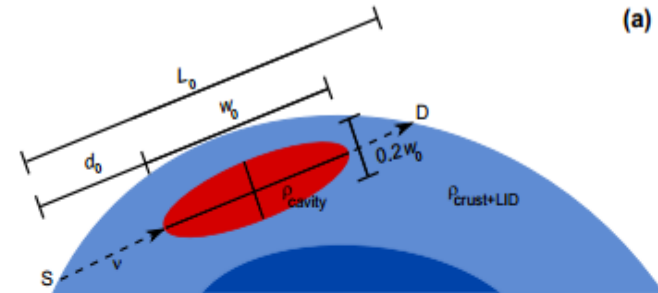
Legión		 PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ				
Inicio > Proyecto Física GAE1						
Física GAE1						
Inicio Perfil Cerrar		Nueva Tarea Eliminar				
Fecha	Nombre de la Tarea	Progreso	PC's	Horas/CPU	↓	□
06:32:56 22-05-09	BB loopclass=1, Nalpha=Ninner=400, Emin=1.06	19.81 %	412	-	↓	□
16:32:16 19-05-09	BB loopclass=1, Nalpha=Ninner=400, Emin=1.05	100 %	-	15236.83	↓	□
14:59:11 19-05-09	KT con evolucion Nalpha = 400, Emin = 1.06	100 %	-	38.26	↓	□
14:43:52 19-05-09	KT con evolucion Nalpha = 400, Emin = 1.05	100 %	-	38.26	↓	□
14:26:04 19-05-09	KT sin evolucion Nalpha = 400, Emin = 1.06	100 %	-	38.25	↓	□
13:26:18 19-05-09	KT sin evolucion Nalpha = 400, Emin = 1.05	100 %	-	39.03	↓	□

15236 Horas/CPU (aprox. 635 días)
de procesamiento en aprox. **60 horas**.

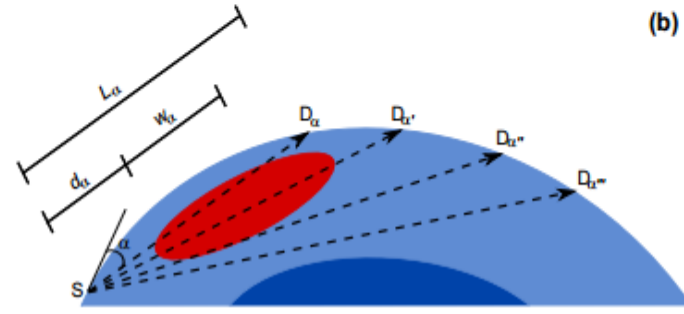
Ejemplo Caso Grupo AE PUCP



Optimización de detectores del proyecto ALICE. Con ALICE se producen colisiones de los haces de protones que son acelerados en el Large Hadron Collider

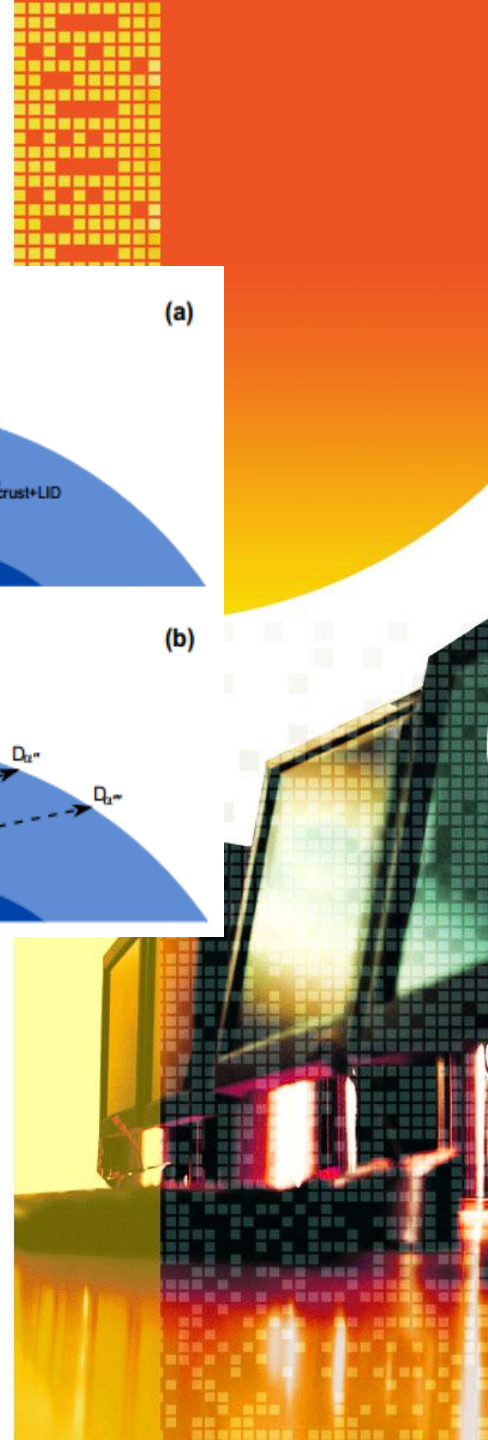


(a)



(b)

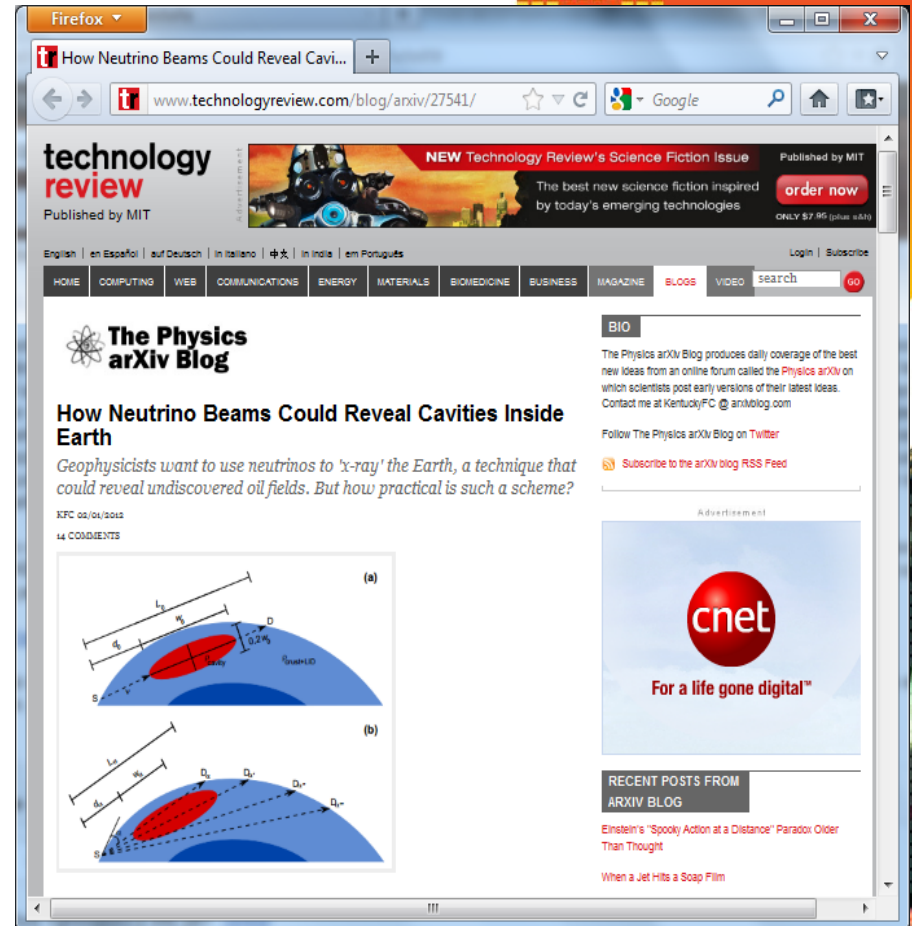
Buscar cavidades de distintas densidades en la corteza terrestre mediante el análisis del paso de los neutrinos



Ejemplo Caso Grupo AE PUCP

Publicación en el portal Technology Review del MIT la investigación sobre análisis de densidad de la corteza terrestre al paso de neutrinos.

Dicho trabajo utilizó durante el 2011 el sistema Legión para las simulaciones, las cuales hubieran tomado **26.67 años** de haberse empleado una sola computadora.

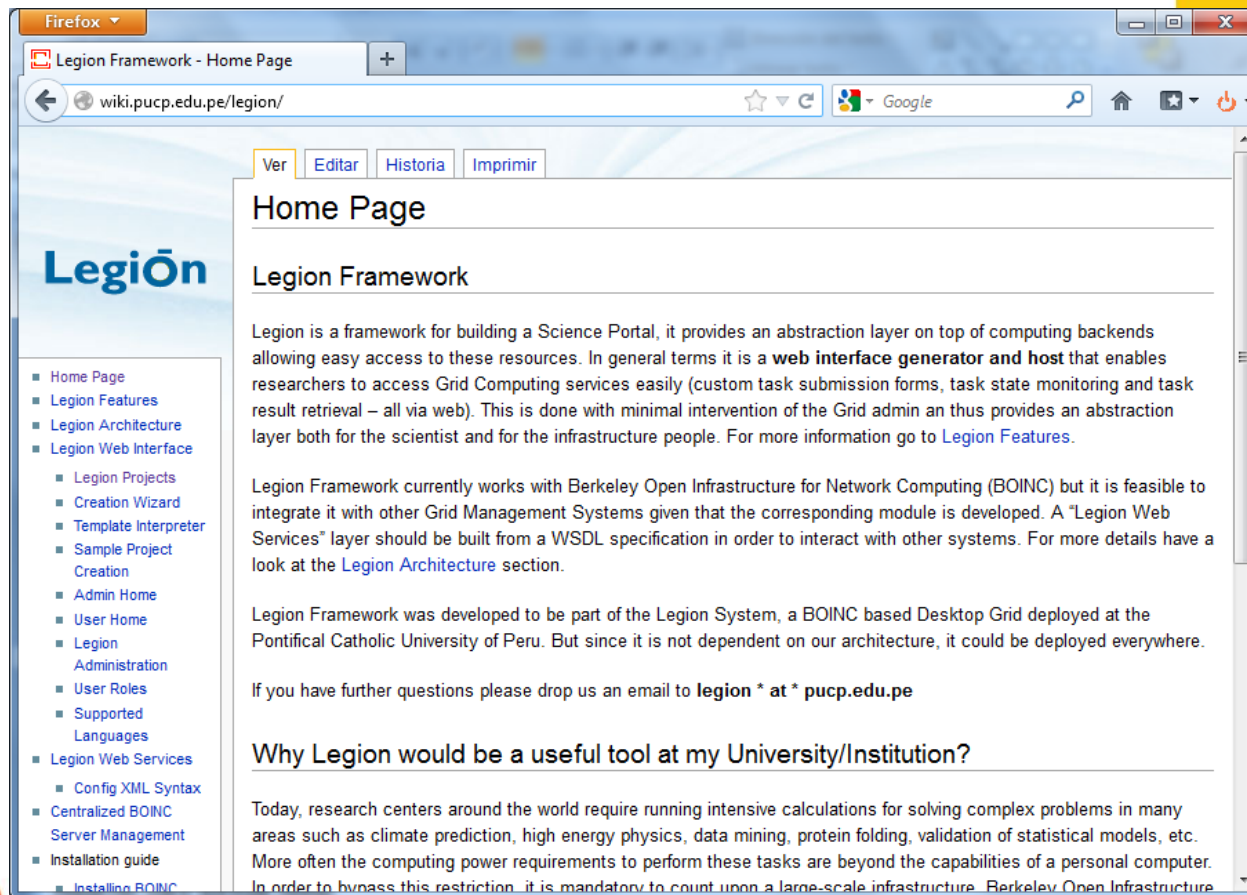


<http://www.technologyreview.com/blog/arxiv/27541>

Difusión



- **El sistema Legión ha sido liberado como software libre bajo licencia GPLv2**
- **<http://wiki.pucp.edu.pe/legion/>**



Presentaciones internacionales de Legión

- **CLCAR 2009 (Mérida, Venezuela)**
- **Workshop BOINC 2009 (Barcelona, España)**
- **CICIC 2010 (Florida, EEUU)**
- **CLCAR 2011 (Colima, México).**
- **Workshop BOINC 2011 (Hannover, Alemania)**
- **CLCAR 2012 (Panamá)**
- **STIC-Amsud 2012 (Uruguay)**



Taller Legión/BOINC

- Se realizó el 22 de Junio en la PUCP.
- Se enseñó la instalación y configuración de un servidor y cliente BOINC.
- Se enseñó la instalación y configuración del servicio Legión para BOINC.
- Se replicó en Agosto en el CLCAR 2012 (Panamá)



Mención a Legión en el CLCAR 2012 como caso de éxito a nivel de RedCLARA



Promoting Successful Stories



Observatorio Latinoamericano de
Eventos Extraordinarios

<http://www.cmc.org.ve/ole2/>



<http://www.soibio.org/RedCentroamericanaDeBioinformatica>

Red Centroamericana
de Bioinformática
y Biocomputación molecular



RNP
REDE NACIONAL DE ENSINO E PESQUISA

<http://rute.rnp.br/>



<http://legion.raap.org.pe/>



PONTIFICIA
UNIVERSIDAD
CATOLICA
DEL PERU

Panamá Aug 2012

CLCAR

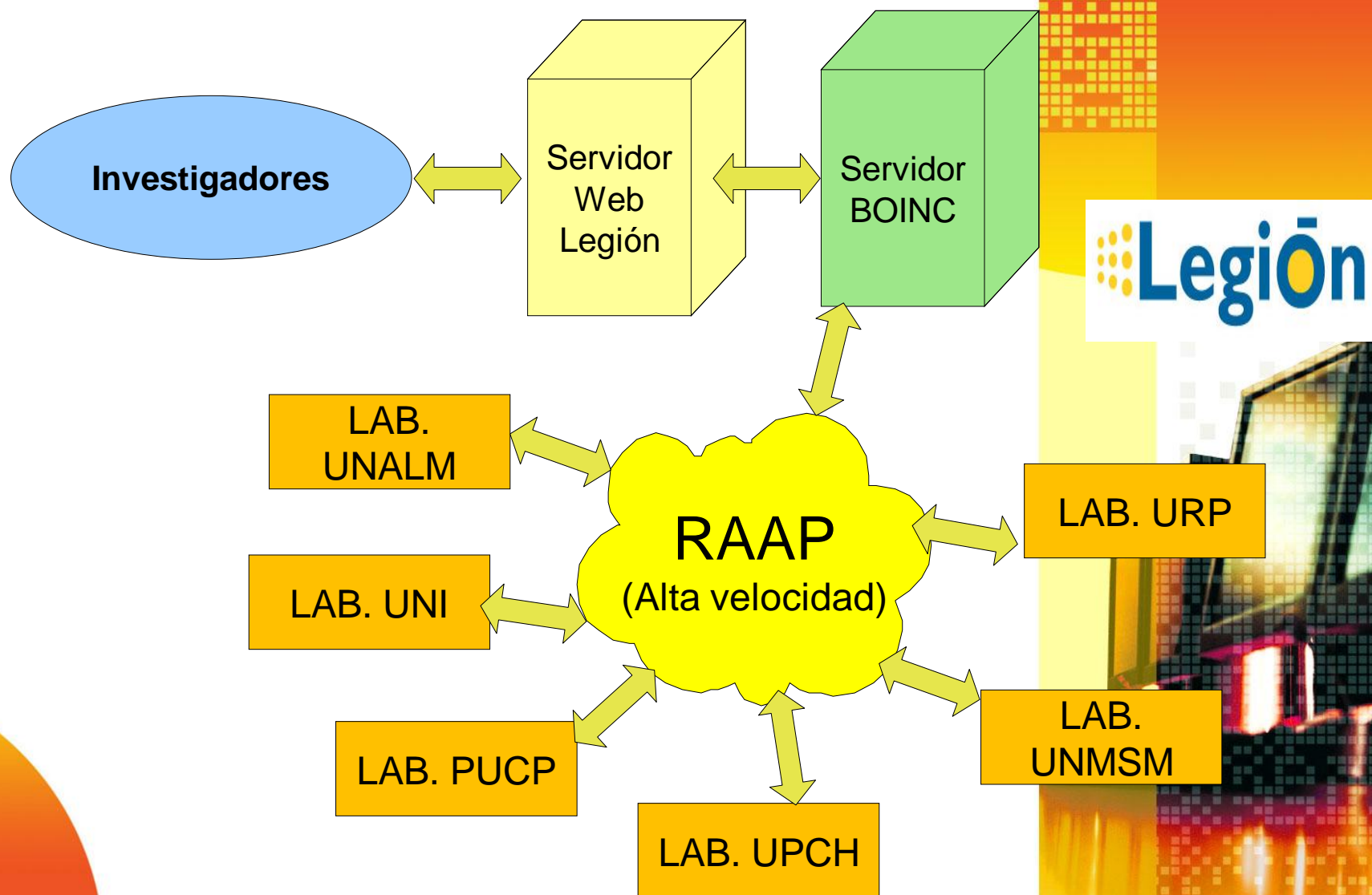
19

Coordinaciones de trabajo con otras instituciones

- **Universidad Nacional de Ingeniería – Perú**
- **Universidad Nacional Mayor de San Marcos – Perú**
- **Universidad Particular Cayetano Heredia – Perú**
- **Universidad Politécnica Salesiana – Ecuador**
- **Universidad Tecnológica de Panamá – Panamá**



Proyecto Legión RAAP



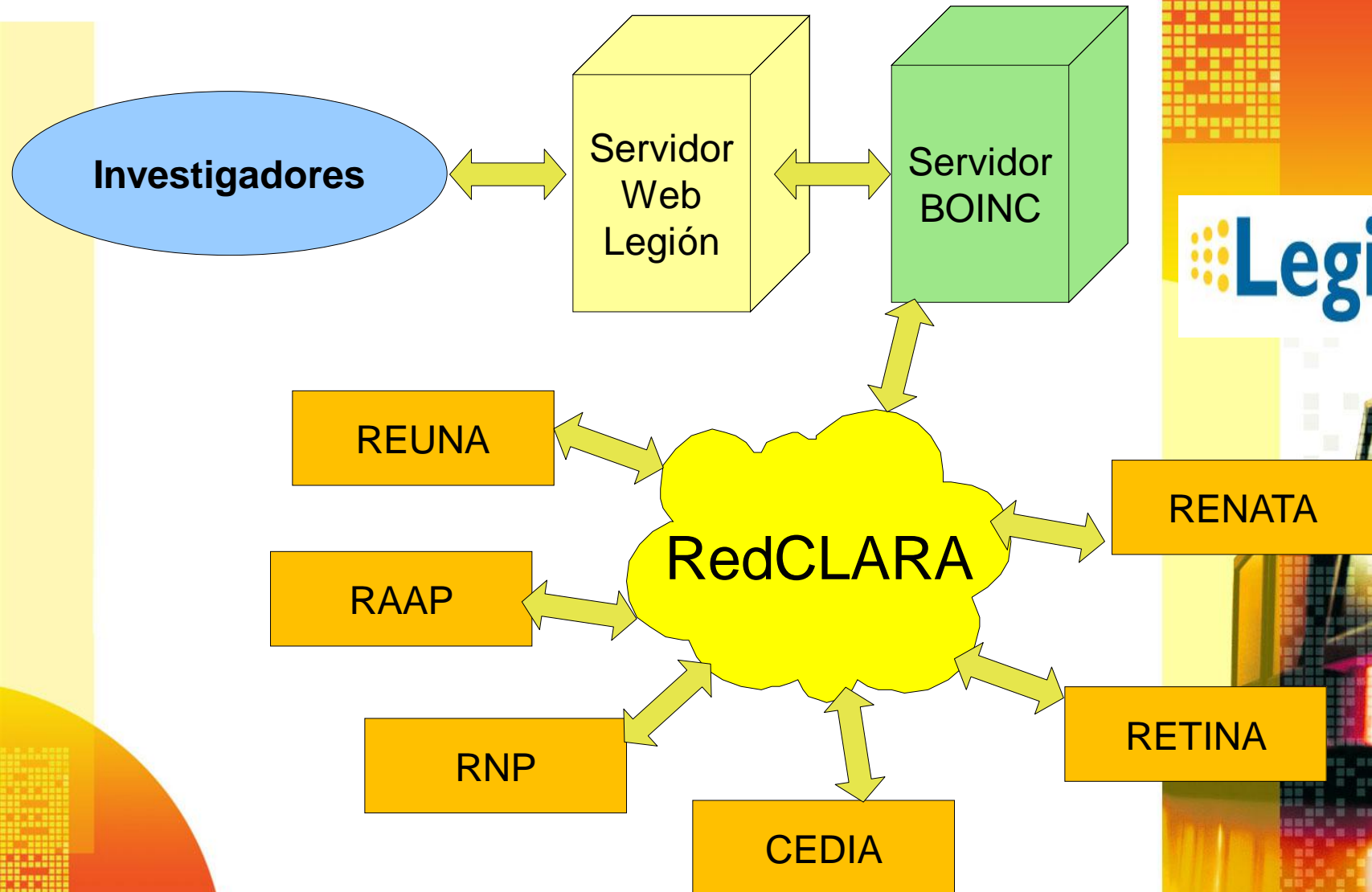
Servidor Legión RAAP

<http://legion.raap.org.pe>

- Busca replicar la experiencia de la PUCP con Legión/BOINC en la RAAP.
- Cada universidad puede brindar un laboratorio de PC's para su uso dentro de una grid computacional.
- Los laboratorios se pueden enlazar a través de la RAAP.
- El servidor BOINC tendría una alta velocidad de comunicación con los nodos de trabajo.
- La interfaz Legión brindaría facilidad de uso a los investigadores.
- Se requiere capacitación del personal técnico



Propuesta Legión RedCLARA



 **Legión**

Propuesta Legión RedCLARA

- Cada NREN y sus instituciones afiliadas pueden configurar sus nodos para apuntar al servidor BOINC de RedCLARA.
- La interfaz Legión brindaría facilidad de uso a los investigadores.
- Oportunidad para demostrar nuestra capacidad de trabajo conjunto para armar la desktop grid.
- Posibilidad de coordinar trabajos con otras iniciativas similares como **IBERCIVIS**
 - **World Community Grid.**





• **CONCLUSIONES**

Beneficios de BOINC

- **Se aprovecha al máximo la infraestructura de aulas informáticas.**
- **Los investigadores tienen acceso, de forma sencilla y económica, a un sistema de cómputo de alto rendimiento.**
- **Ingreso a nuevos campos de investigación que requieran cómputo intensivo.**
- **Incremento en la producción científica y posicionamiento académico.**



Beneficios LEGION

- **Facilita la creación de interfaces web para el monitoreo de los proyectos que se ejecutan con BOINC.**
- **Arquitectura modular en base a web services.**
- **Un portal web Legion puede apuntar a varios servidores BOINC.**
- **Posee un Wizard que acelera el despliegue de la aplicación en la web.**



Equipo de trabajo

- Coordinador:
 - Genghis Ríos Kruger
- Administrador Servidor BOINC:
 - Oscar Díaz Barriga
- Programación Sistema Legión:
 - Pablo Fonseca



¿PREGUNTAS?



Muchas Gracias

legion@pucp.edu.pe

