

Computación paralela y clúster de la Universidad de Oriente

Dr.C. Fernando J. Artigas Fuentes
Conferencia divulgativa
UO, diciembre 2019

Ideas básicas

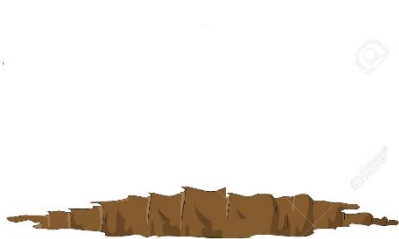
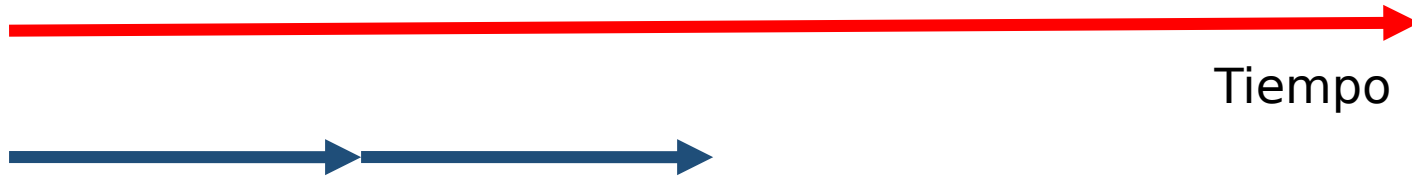
1h



1h

Ideas básicas

2h

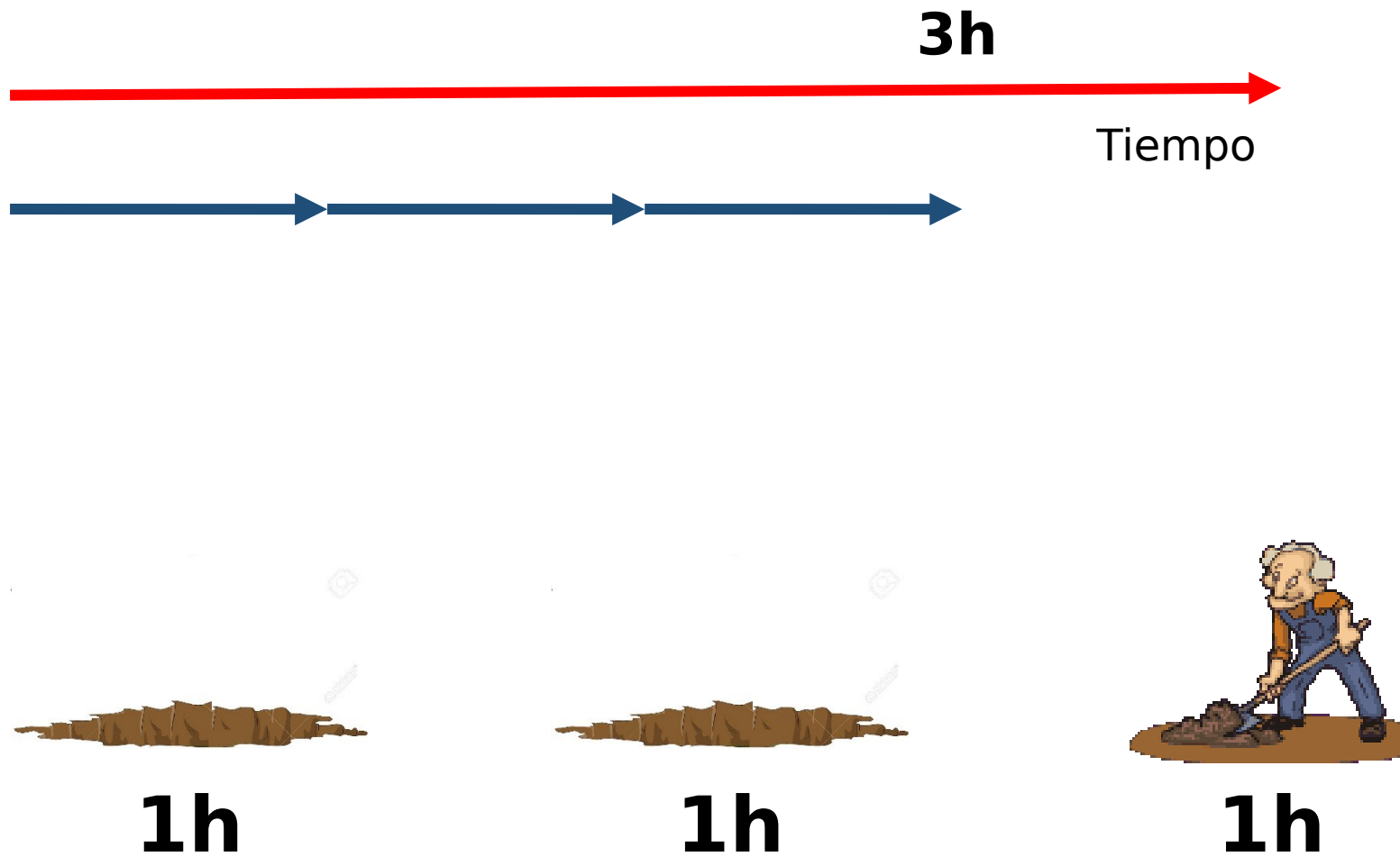


1h

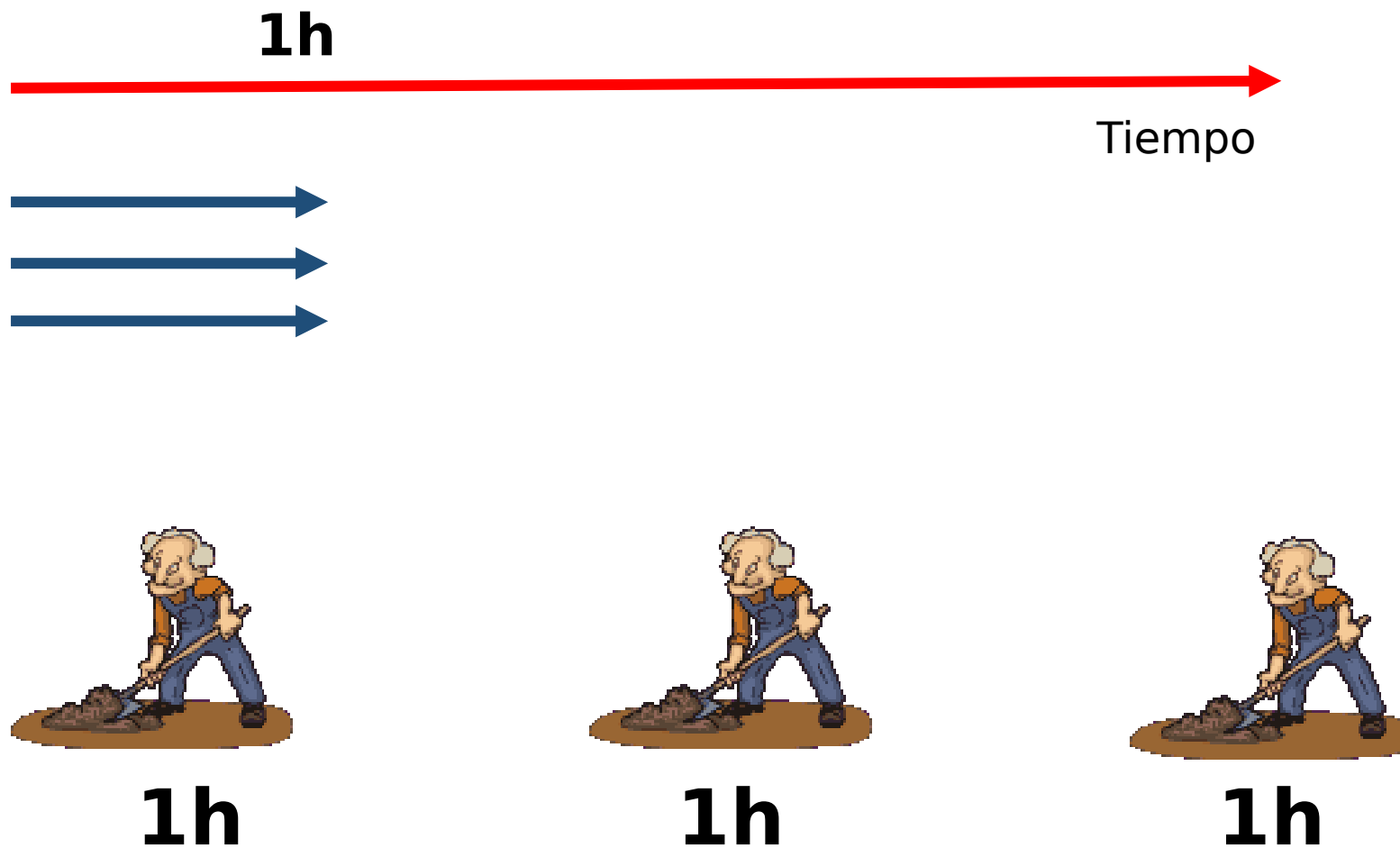


1h

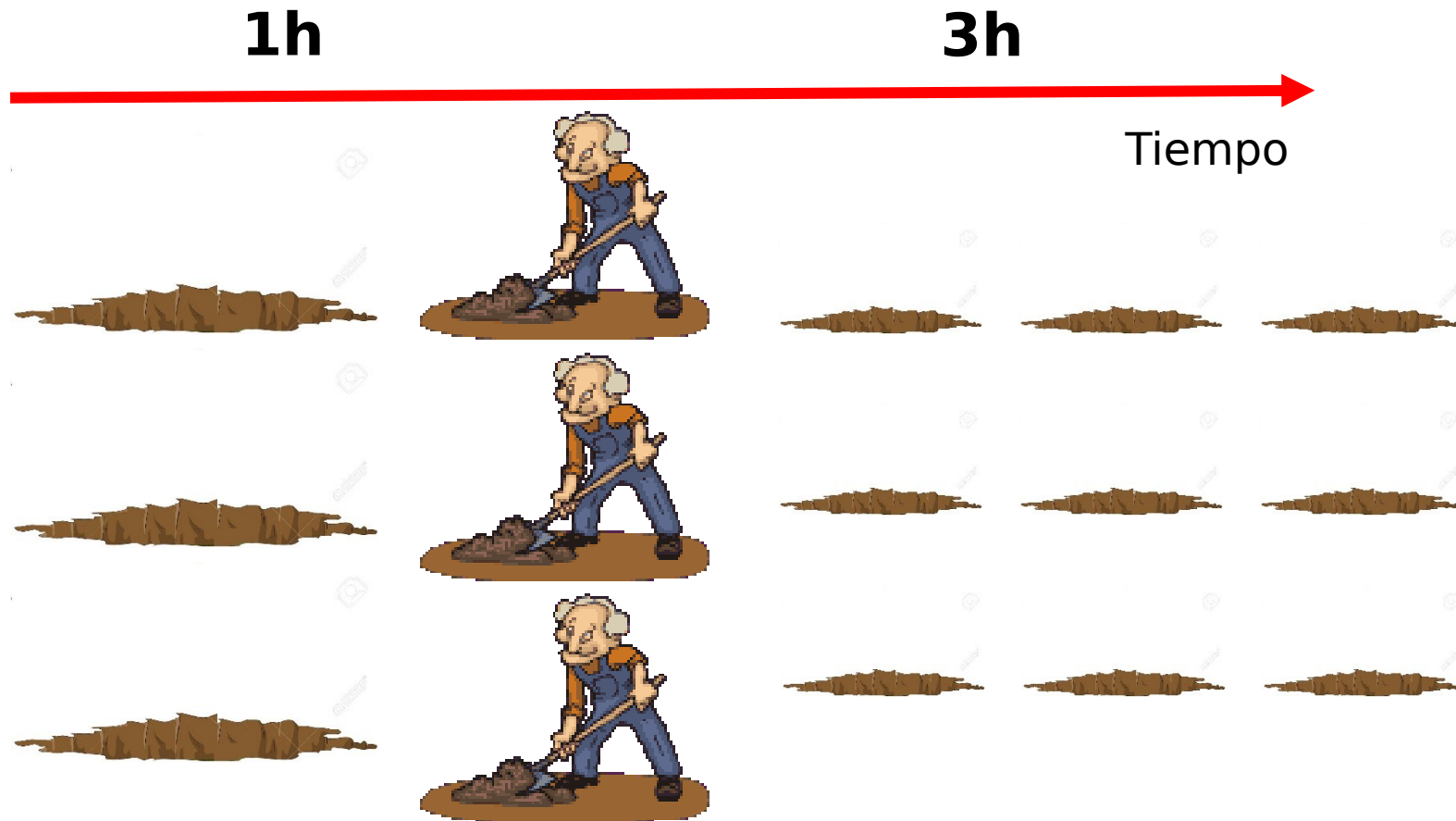
Ideas básicas



Ideas básicas



Ideas básicas



El mismo resultado en menos tiempo

Ahorro tiempo

Más resultados en el mismo tiempo

Trato un problema mayor

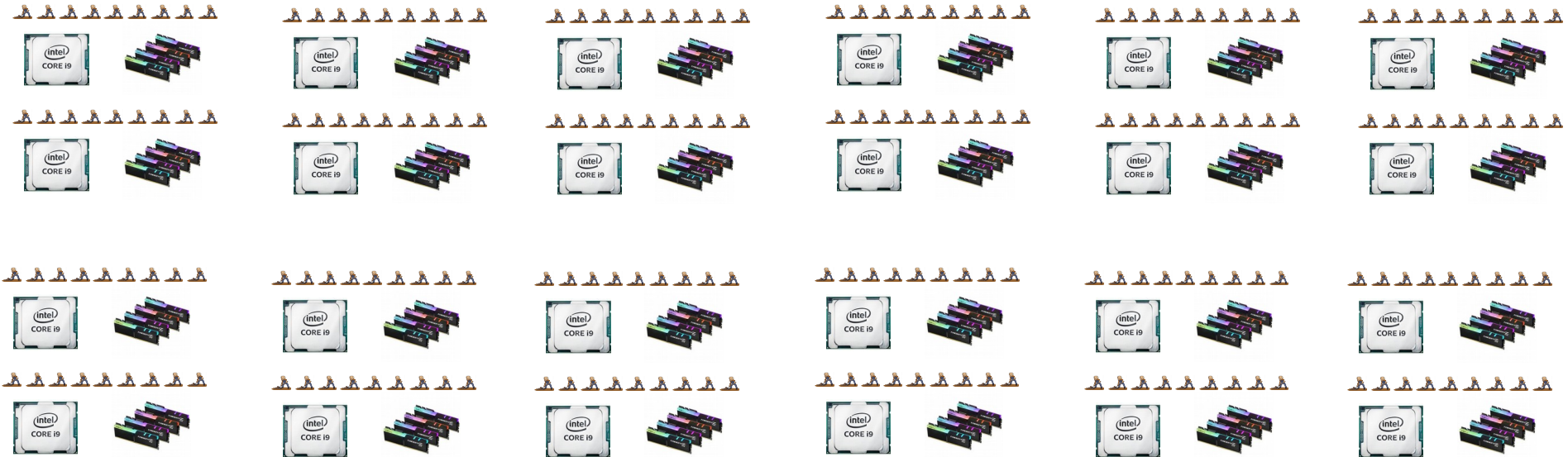
Pc multiprocesadores

9 procesadores = “trabajadores”



32 GB RAM

Clúster de computadoras

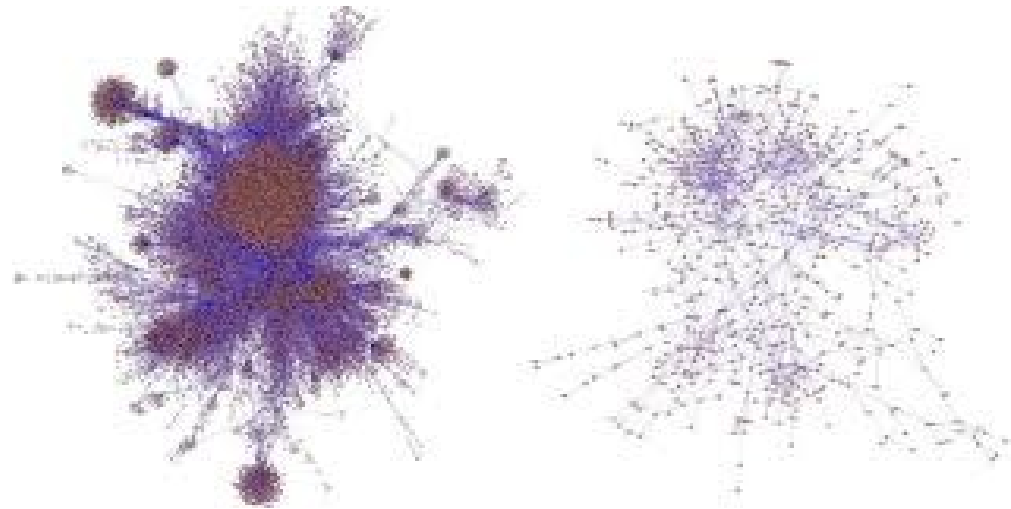


13 nodos, 416 cores, 832 GB RAM

¿Para cuáles tipos de problemas?

Aquellos cuyo tratamiento...

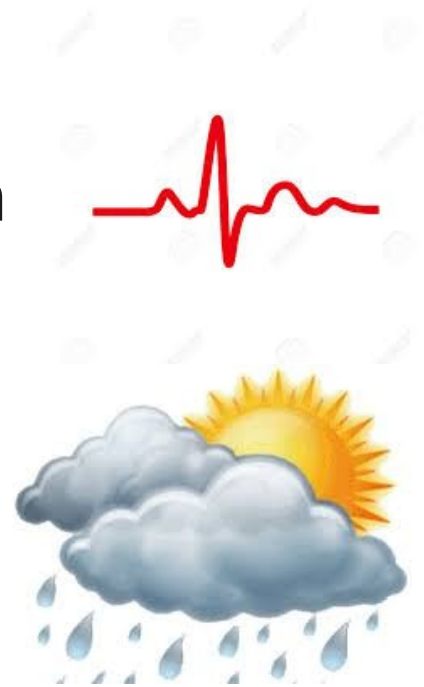
- **Demora mucho tiempo en una Pc.**
- **Maneja muchos datos de entrada.**
- **Ocupa mucho espacio en memoria.**



¿Qué significa demorar mucho tiempo?

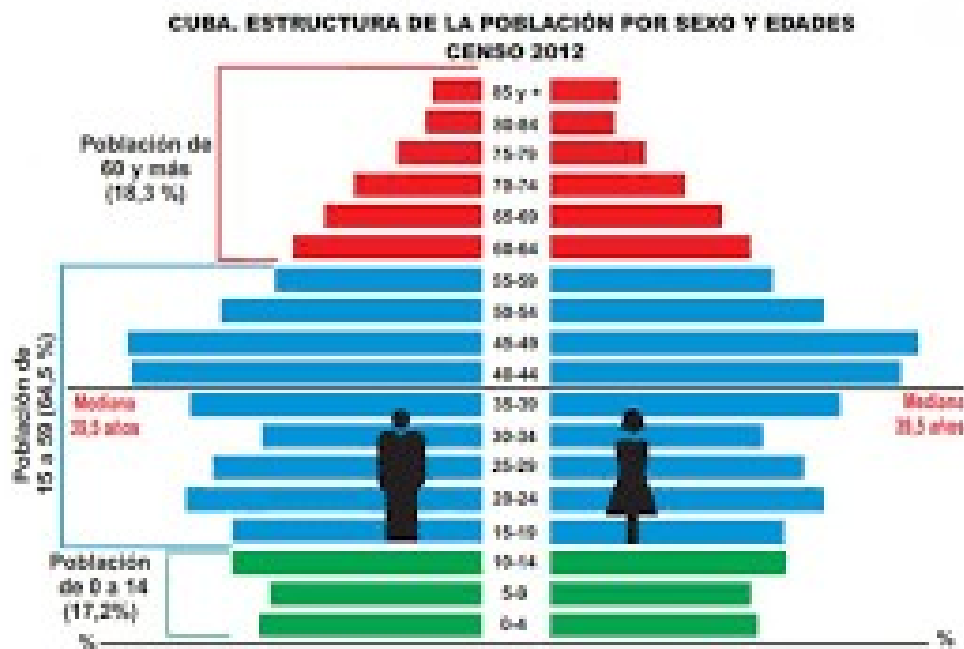
Depende de los requerimientos...

- **Capturar un simple latido del corazón en 3 minutos **es inadmisibile.****
- **Calcular el pronóstico del clima de mañana en 48h **no tiene sentido.****



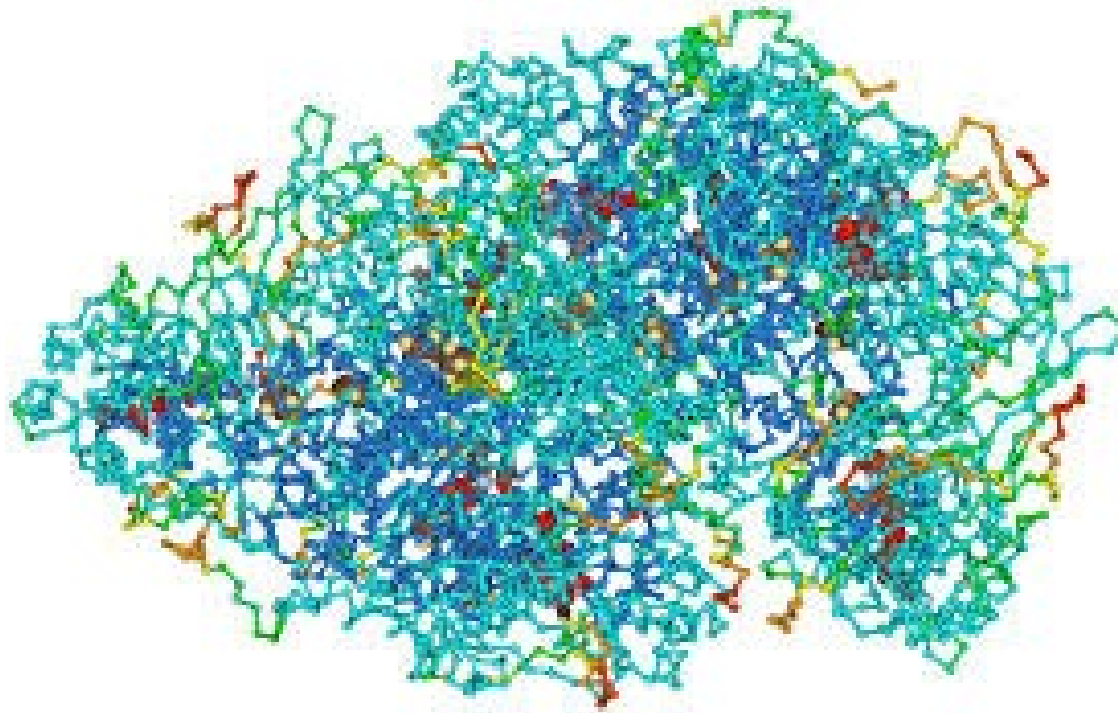
¿Qué significa procesar muchos datos?

- **Analizar la población de Cuba requiere procesar bases de datos de 11 millones de registros.**



¿Ocupar espacio en memoria?

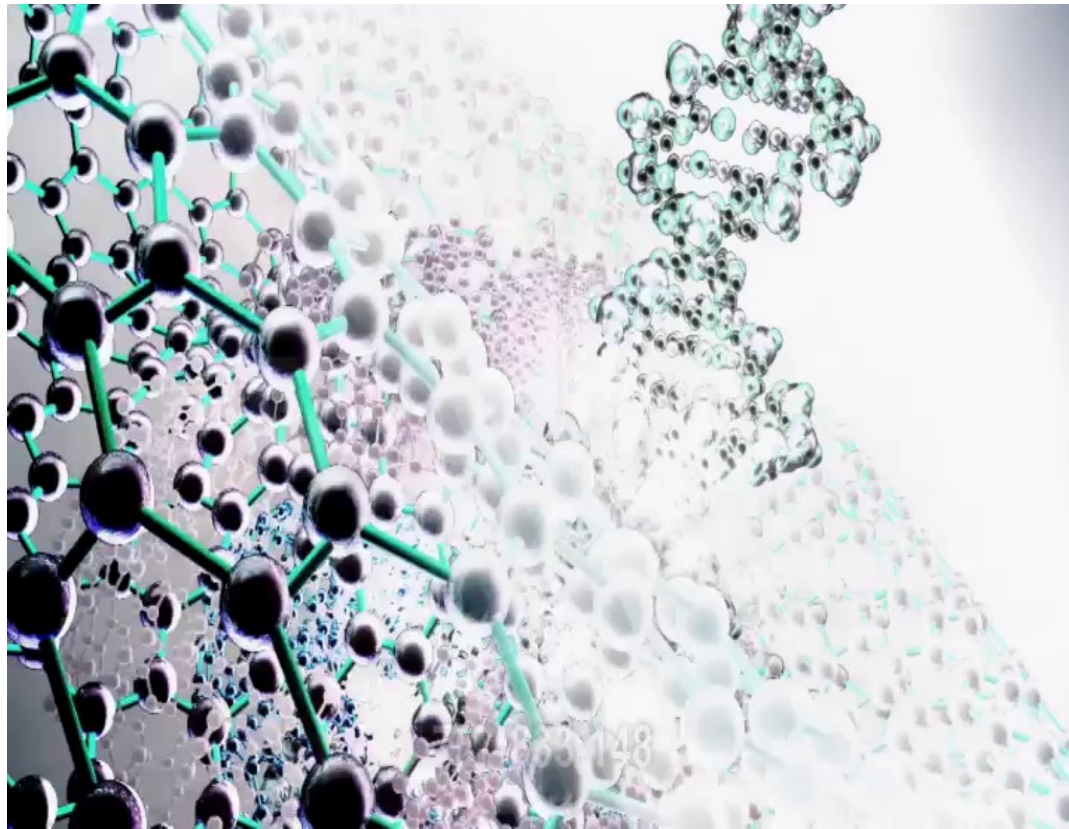
- **Representar una estructura molecular compleja requiere almacenar mucha información en RAM.**



¿Y si quisiera hacer algo como esto?



¿O algo así?



¿Necesito problemas tan enormes para usar el clúster?

- **Problemas simples, secuenciales que usen un solo procesador.**
- **Problemas paralelos que usen 2 o más procesadores.**

- **¿Qué beneficios aporta?**

Ventajas inmediatas

- **Sistema activo las 24h del día el año completo.**
- **Respaldo eléctrico, no debe fallar la ejecución.**
- **Respaldo de datos garantizado.**
- **Salvo raras excepciones, debe ir más rápido que en una PC.**
- **No se necesita mantener encendida la computadora de la oficina.**
- **El sistema avisa cuando acaba mi tarea u ocurre un error.**



MENSAJE



Ventajas inmediatas

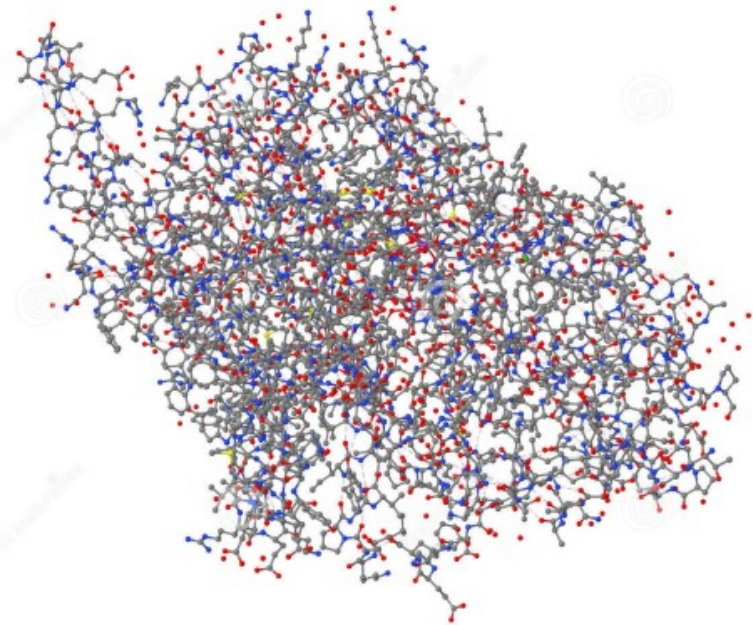
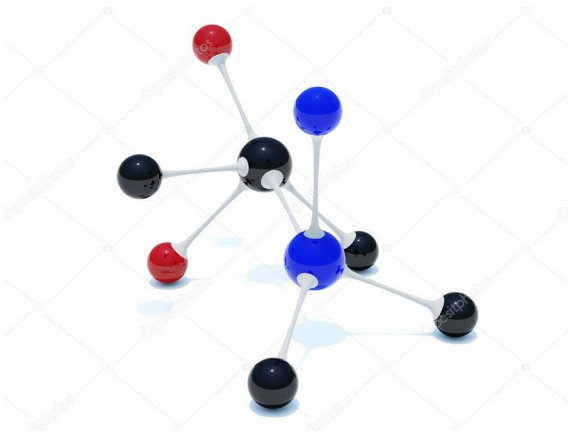
- **Posibilidad de crear nuevas tareas y acceder a los resultados a través de la red.**
- **Si las salidas son simples podría accederse desde casa con un módem (el mismo que se usa para correo electrónico).**
- **¿Y si tuviera Nauta Hogar? ... ¡VPN!.**
- **¿A través del móvil? ... también.**

Ventajas especiales

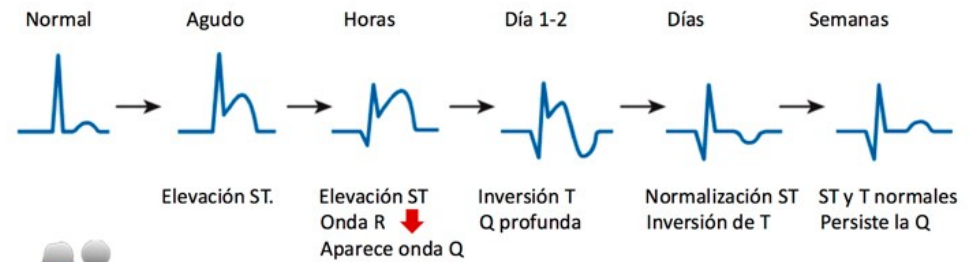
- **Abordar problemas más complejos.**
- **Acelerar la obtención de los resultados finales**
- **Dar fuerza al entorno experimental.**
- **Publicar en revistas de impacto.**



¿Quién me impone límites?... (escalar el problema)

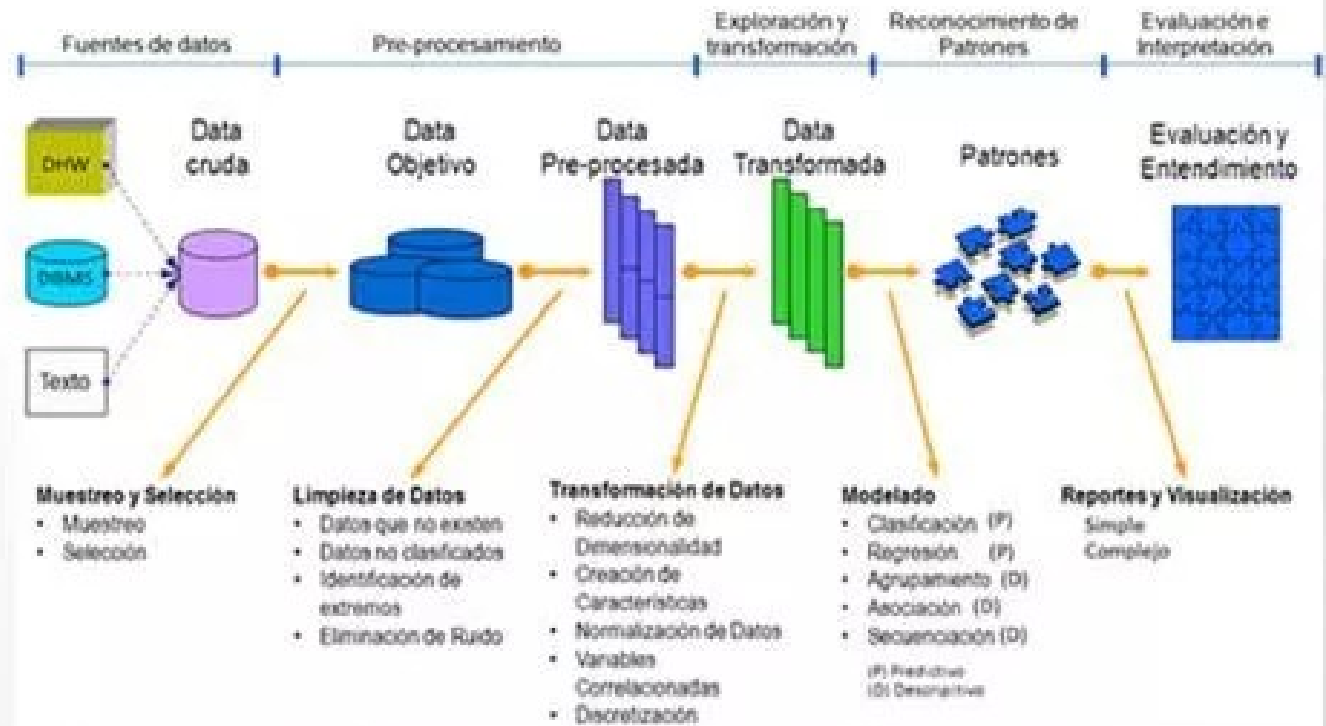


Análisis de electrocardiogramas... (acelerar procesamiento)



Minería de datos

El Proceso de la Minería de Datos



Posibilidad de colaborar...

La UE impulsa la computación paralela para aplicaciones biomédicas



Me gusta 0 Tweet

Un proyecto europeo, en el que participan la Universidad Carlos III de Madrid y la multinacional informática IBM, estudia cómo mejorar el desarrollo de sistemas de computación avanzados para conseguir *software* más rápido. Los investigadores están colaborando con el Hospital Gregorio Marañón de Madrid para acelerar el procesamiento de imágenes capturadas mediante resonancia magnética del cerebro, que en la actualidad tarda horas, para que pueda ser realizado en minutos.

Más información sobre: software diagnóstico cerebro computación paralela Rephrase sistemas heterogéneos Horizonte 2020

UC3M / SINC | Seguir a @ciencia_uc3m | 18 noviembre 2015 09:30

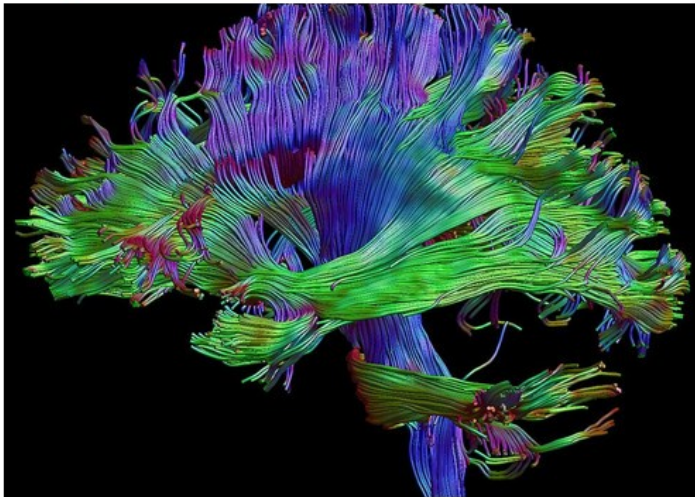


Imagen de cerebro generada a partir de resonancia magnética. / UC3M Ver más

12th RES Users' Conference
+ 7th HPC Advisory Council Spain Conference

20th and 21th of September 2018 - Valencia (Spain)

Information and registration: www.res.es/users-conference-2018

QCHPC2019

Workshop
**Quantum Computers and
High Performance Computing**

November 27, 2019
9:00 - 19:00

CITUS
Campus Vida (USC)
15782 Santiago de Compostela

**Latin American
Introductory School to
Parallel Programming and
Parallel Architecture for High
Performance Computing**

ICTP

Further information:
<http://indico.ictp.it/event/8344/>
smc3187@ictp.it

12 to 23 February 2018
CINVESTAV and ININ, Ocoyoacac, Mexico

¿Qué debo saber para usar el clúster?

¿Cuál es el tipo de problema que abordo?

- El programa demora mucho tiempo en mi computadora.
- Mi computadora tiene pocos recursos.
- Tengo que procesar muchos datos.
- Quiero escalar mi problema.



¿Qué debo saber para usar el clúster?

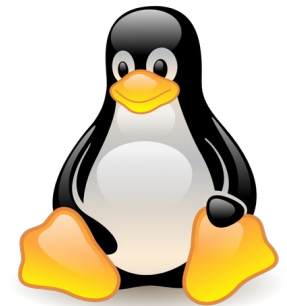
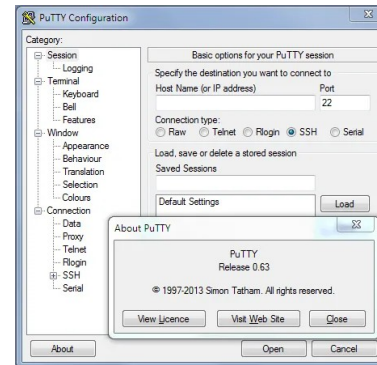
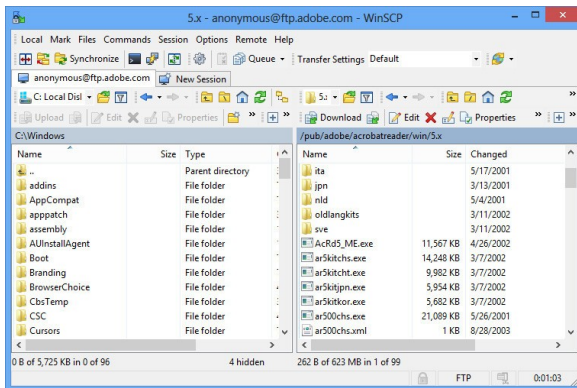
¿Qué tipo de usuario soy?

- Soy experto en computación y paralelismo.
- Soy experto en computación, pero no en paralelismo.
- Se usar la PC, no se programar pero uso programas de terceros.
- Le tengo odio a la computadora.



Para usuarios expertos...

- Usar el sistema operativo Linux.
- Usar programa de conexión remota, como el ssh, WinSCP, PuTTY.
- Conocer el uso de SLURM y comandos básicos de ejecución de tareas en el clúster.
- Solicitar una cuenta de usuario en el clúster.



Para usuarios intermedios

- **Solicitar una cuenta de usuario en el clúster.**
- **Aprender a usar la interfaz web.**
- **Disponer de los programas a ejecutar (versiones para Linux) y datos a procesar.**
- **Acceder a través de:**
 - **<http://portal.uo.hpc.cu>**

Para usuarios intermedios (II)

<http://portal.uo.hpc.cu>

The image shows a web browser window displaying the 'Solicitar "LDAP usuario"' form. The form is overlaid on a dark background of the website. The form fields are as follows:

- Nombre:** - Nombre -
- Apellidos:** - Apellidos -
- Identificador:** - Identificador -
- Correo:** - artigas
- Contraseña:** - [masked]
- Confirmación de contraseña:** - Confirmación de contraseña -
- Breve resumen:** - Breve resumen -
- Instituto:** - Instituto -
- Área de investigación:** - Área de Investigación -
- Grupo de investigación:** - [blank]

Below the 'Breve resumen' field, there is a note: *"Escribe un breve resumen explicando su solicitud."*

The background website shows a navigation menu with 'Inicio' and 'Red Académica de Supercómputo'. There are also links for 'Iniciar sesión' and 'Español'.

Para usuarios intermedios (III)

<http://portal.uo.hpc.cu>

The screenshot displays the HPC-UI portal in a Mozilla Firefox browser. The address bar shows the URL <https://portal.uo.hpc.cu/hpc/>. The page features a dark red navigation bar with links for 'Sitio web', 'HPC', 'BIGDATA', and 'Administración'. A user profile for 'Fernando Jose Artigas Fuentes' is visible in the top right corner, along with a language selector set to 'Español'.

The main content area is titled 'Generador de Scripts' and is divided into several sections:

- Encabezado:** A header section for the script generator.
- Opciones generales:** A section for general options, including:
 - 'Nombre del script *': A text input field with a placeholder '- Insertar nombre del script -' and a '.sl' suffix.
 - 'Nombre del trabajo': A text input field with a placeholder '- Insertar nombre del trabajo -'.
 - 'Exportar variables': A checkbox that is checked.
 - 'Partición': A dropdown menu currently set to 'test'.
- Manejo de recursos:** A section for resource management, including:
 - 'Lista de recursos': Two input fields for '- Cantidad de nodos -' and '- Tareas totales (en todos los nodos) -'.
 - 'Lista de recursos': Two input fields for '- Tareas por nodos -' and '- Núcleos de CPU por tarea -'.
 - 'Memoria': Two input fields for '- RAM por nodo -' (set to 'M') and '- RAM por núcleo de CPU -' (set to 'M').
 - 'Tiempo de ejecución': An input field for '- Límite de tiempo (mm o [[dd-]hh:]mm:ss o dd-hh[mm[ss]]) -'.
 - 'Test': A checkbox that is unchecked.
 - 'Requiere GPUs': A checkbox that is unchecked.
 - 'Lista de nodos': An input field for '- Nodos preferidos -'.
 - 'Excluir': An input field for '- Nodos a evitar -'.
- Opciones de flujo de salida:** A section for output flow options, currently empty.

Para usuarios intermedios (IV)

<http://portal.uo.hpc.cu>

The screenshot shows the HPC-UI portal interface. The top navigation bar includes links for 'Sitio web', 'HPC', 'BIGDATA', and 'Administración'. The user is identified as 'Fernando Jose Artigas Fuentes' with a language dropdown set to 'Español'. The main content area is titled 'Trabajos Activos' and displays a table of active jobs. The table has columns for 'Id de trabajo', 'Estado', 'Nombre', 'Usuario', 'Partición', 'Tiempo', 'Límite de tiempo', 'Nodos', and 'Lista de nodos'. Below the table, there is a section titled 'Códigos de estado de trabajo' which provides a legend for various job states.

Id de trabajo	Estado	Nombre	Usuario	Partición	Tiempo	Límite de tiempo	Nodos	Lista de nodos
4214	RUNNING	CdLSO500	uo_yzulueta	public	2:04:18	Unlimited	4	nodo[002-005]
4215	RUNNING	CdLSO600	uo_yzulueta	public	2:04:13	Unlimited	4	nodo[006-009]
4216	RUNNING	CdLSO700	uo_yzulueta	public	2:04:09	Unlimited	2	nodo[010-011]
4217	PENDING	CdLSO900	uo_yzulueta	public	0:00	Unlimited	4	(Resources)
4218	RUNNING	CdLSO1000	uo_yzulueta	public	2:03:48	365-00:00:00	2	nodo[012-013]
4219	PENDING	CdLSO1100	uo_yzulueta	public	0:00	Unlimited	4	(Priority)

Mostrando registros del 1 al 6 de un total de 6 registros

Códigos de estado de trabajo

Los trabajos generalmente pasan por varios estados en el curso de su ejecución. Una explicación de cada estado sigue.

CA	CANCELLED	El trabajo fue cancelado explícitamente por el usuario o el administrador del sistema. El trabajo puede o no haber sido iniciado.
CD	COMPLETED	El trabajo ha terminado todos los procesos en todos los nodos.
CG	COMPLETING	El trabajo está en proceso de completar. Algunos procesos en algunos nodos aún pueden estar activos.
F	FAILED	El trabajo terminó con un código de salida distinto de cero u otra condición de falla.
NF	NODE_FAIL	El trabajo finalizó debido a la falla de uno o más nodos asignados.
PD	PENDING	El trabajo está esperando la asignación de recursos.
R	RUNNING	El trabajo actualmente tiene una asignación.

Para los que odian...

- **Buscar colaboración con expertos.**



Para los que odian... (II)

Sistema a la carta... futuro asistente inteligente.



Desarrollo de aplicaciones paralelas

Bibliotecas de alto desempeño:

- **Álgebra lineal densa**: BLAS, LAPACK, MAGMA, PLASMA, ScaLAPACK, CUBLAS, CULA, ATLAS, GotoBLAS, MKL.
- **Álgebra lineal dispersa**: ARPACK, SPARSE, SPARSE-BLAS, ITPACK, SVDPACK, SuperLU, Trilinus, PETSc.
- **Optimización**: HeO, ParadisEO, MALLBA, GitHub.
- **Transformada de Fourier**: FFTPACK, P3DFFT, FFTW.
- **Científicas**: MOOSE, COOLFluid, OpenFVM, PyClaw, PetIGA, ROMS, ...

Ejemplo del beneficio de BLAS

Multiplicación de matrices 1000×1000 en sistema con 12 núcleos, tiempo en segundos

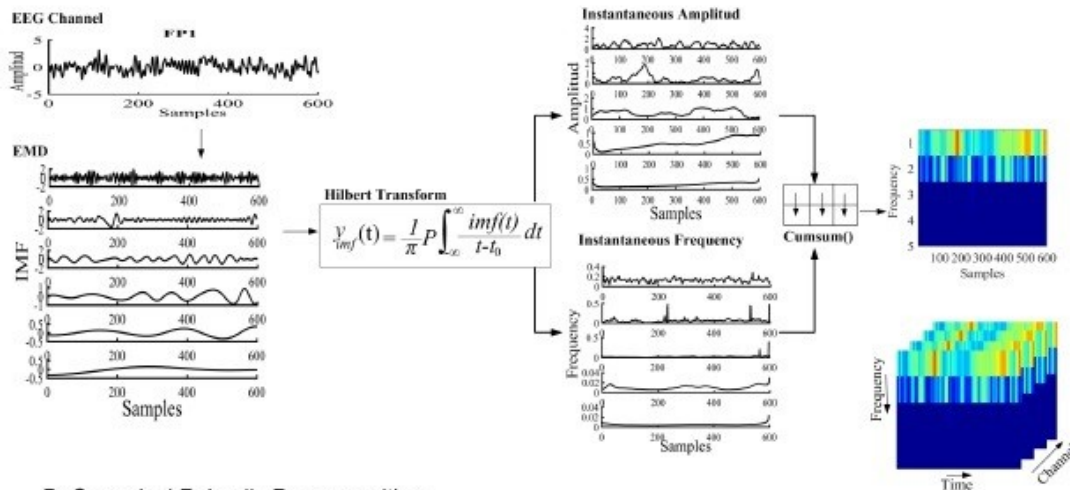
núcleos:	1	2	4	8	12
BLAS 1	11.15	10.66	10.64	10.63	10.63
BLAS 2	0.61	0.30	0.17	0.099	0.11
BLAS 3	0.14	0.072	0.044	0.030	0.055

de 0.25 Gflops, con uso de librería 20 Gflops, con paralelismo implícito a 100 Gflops

Resultados obtenidos en la UO

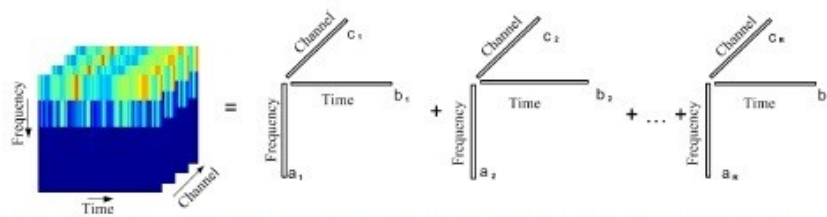
Detección de crisis epilépticas no convulsivas usando Mínimos Cuadrados Parciales

A: Tensorization Process



Dr.C. Yissel Rodriguez Aldana

B: Canonical Polyadic Decomposition



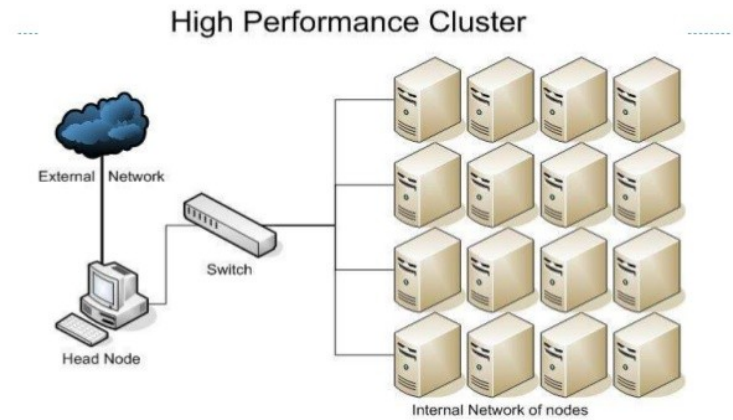
¿A quién contactar? (I)

Miembros del grupo HPC

- **Administración general:** Prof. Oscar Au Álvarez:
oau@uo.edu.cu
- **Tecnología:** Ing. Ernesto Eduardo Diaz Conde:
eediaz@uo.edu.cu
- **Ciencia:** Dr.C. Fernando J. Artigas Fuentes:
artigas@uo.edu.cu

¿A quién contactar? (II)





¿Preguntas?

Dr.C. Fernando J. Artigas Fuentes
Conferencia divulgativa
UO, diciembre 2019