



# ANÁLISIS Y DISEÑO DE ALGORITMOS PARALELOS

SANDRA LUZ MORALES GUITRON

ESCOM



Un clúster Beowulf



*Supercomputadora Blue Gene de IBM*



# ¿QUE ES UNA SUPERCOMPUTADORA?

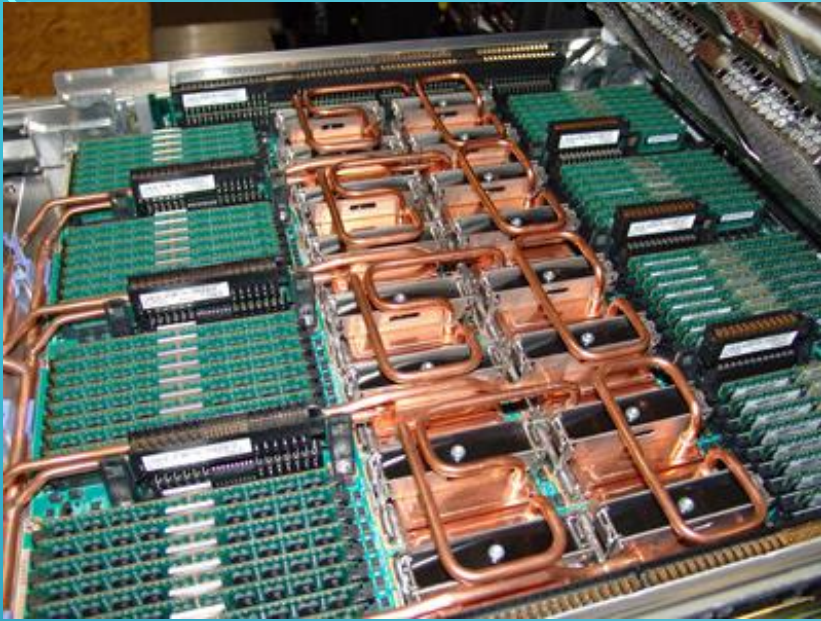
- Una **supercomputadora** o un **superordenador** es aquella con capacidades de cálculo muy superiores a las computadoras comunes y de escritorio y que son usadas con fines específicos. Hoy día los términos de supercomputadora y superordenador están siendo reemplazados por **computadora de alto rendimiento y ambiente de cómputo de alto rendimiento**, ya que las supercomputadoras son un conjunto de poderosos ordenadores unidos entre sí para aumentar su potencia de trabajo y rendimiento.

# ¿QUE ES UNA SUPERCOMPUTADORA?

- ¿Como es posible obtener un poder computacional cientos o miles de veces mayor que un PC de escritorio con la tecnología de microprocesadores existente? Esto se puede lograr mediante el procesamiento masivo en paralelo. Una supercomputadora esta compuesta por cientos o miles de computadoras mas pequeñas interconectadas entre si. Cada una de estas computadoras mas pequeñas recibe el nombre de nodo. Cada nodo a su vez esta compuesto por múltiples procesadores multi-núcleo con una gran cantidad de memoria. Los nodos de una supercomputadora se diseñan por lo general en un formato delgado y se encuentran organizados en clusters que se acomodan físicamente en racks verticales.



*Nodo de supercomputadora*



*Nodos organizados en Racks, Supercomputadora*



- Una supercomputadora requiere un sistema de refrigeración especial que básicamente consiste de un sistema complejo e aire acondicionado que debe mantener los nodos a cierta temperatura para que puedan funcionar de manera optima. Es evidente que la potencia consumida por el sistema de refrigeración contribuye a aumentar de forma significativa el consumo de potencia total del sistema.
- Las supercomputadoras se utilizan mas que todo en el area de investigación, por ejemplo, en la simulación de sistemas físicos y biológicos complejos que permiten ser procesados de forma masivamente paralela y que aprovechan este poder de computo tan grande.

# HTTP://TOP500.ORG/LISTS/2015/11/

Hogar > Liza > de noviembre de el año 2015

## DE NOVIEMBRE DE EL AÑO 2015

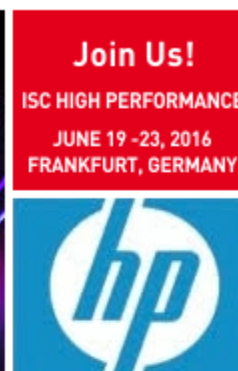
Por sexta vez consecutiva, Tianhe-2, un superordenador desarrollado por la Universidad Nacional de Tecnología de Defensa Nacional de China, ha mantenido su posición como sistema de Nº 1 del mundo, de acuerdo a la 46ª edición de la lista TOP500 dos veces al año de la más potente del mundo superordenadores. En general, el cambio en la parte superior de la lista es de nuevo menor, con sólo dos nuevos sistemas en el Top 10 - El superordenador Trinidad construida por Cray y en forma conjunta desplegada por el Departamento de Energía (DOE) laboratorios nacionales de Los Alamos y Sandia, y el sistema de Hazel-gallina construida por Cray e instalado en el HLRS - Höchstleistungsrechenzentrum Stuttgart en Alemania.

En el cuadro más grande, China casi se triplicó el número de sistemas en la última lista, mientras que el número de sistemas en los Estados Unidos ha caído al punto más bajo desde la lista TOP500 fue creado en 1993. China también se está forjando una mayor participación como un fabricante de ordenadores de alto rendimiento con múltiples fabricantes chinos cada vez más activa en este campo.

LANZAMIENTO

La lista

Presione soltar



### Tweets

Seguir



ISC High Performance  
@ISChpc

7h

#ISC16 proclaims women's excellence in  
#HPC & introduces Jackie Chen of  
@SandiaLabs as Tue keynoter  
[bit.ly/1LtOx1k](http://bit.ly/1LtOx1k).

Retwitteado por TOP500

Tianhe-2, lo que significa Vía Láctea-2, condujo la lista con un rendimiento de 33,86 petaflop / s (trillones de cálculos por segundo o Pflop / s) en el benchmark Linpack. Mantener su posición en el puesto número 2 es Titan, un sistema Cray Xk7 instalado en el Laboratorio Nacional de Oak Ridge del Departamento de Energía. Titán, el sistema superior en los Estados Unidos y uno de los sistemas más eficientes energéticamente en la lista, logró 17,59 petaflop / s en el benchmark Linpack.

Las únicas nuevas entradas en el top 10 en la lista de los superordenadores más reciente son la Trinidad en el No. 6 y Hazel-Gallina en el No. 8. Trinidad es un sistema Cray XC, que tiene 301,056 núcleos y alcanzó 8.1 PFLOPS / s. Trinidad es administrado y operado por el Laboratorio Nacional de Los Alamos y Sandia National Laboratories bajo la Alianza para la computación en la Escala de asociación extrema (ACES). Hazel-gallina es también un sistema Cray XC instalado en Alemania en el HLRS - Höchstleistungsrechenzentrum Stuttgart - y cuenta con 185,088 núcleos y alcanzó 5,6 Pflop / s.

Seis de los 10 mejores sistemas fueron instalados en 2011 o 2012, Tianhe-2 en 2013 y sólo Trinidad, Hazel-Hen, y Shaheen II en Arabia Saudita fueron instalados en 2015. Este bajo nivel de rotación entre los superordenadores refleja una tendencia a la desaceleración que comenzó en 2008.

#### actualizaciones

- 17/11/2015: Varias actualizaciones de rendimiento

Presione sobre los  
aspectos más destacados de la lista

#### DESCARGAS

Lista TOP500 (XML)  
Lista TOP500 (Excel)  
TOP500 cartel  
Cartel en PDF

Abrir




 **TOP500** 15 feb  
@top500supercomp  
New @ThisWeekinHPC podcast out on  
#TOP500.org: Gravity Makes Waves; @HP  
Takes SGI Technology to Market:  
[bit.ly/1TIJQy8](http://bit.ly/1TIJQy8) #HPC



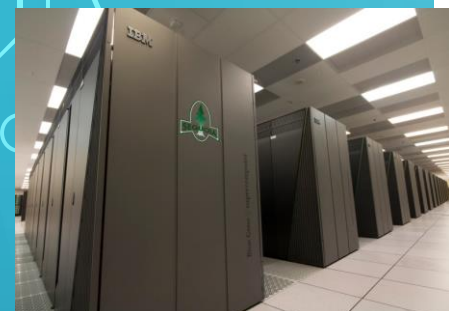
 **TOP500** 15 feb  
@top500supercomp  
For those catching up on last week's  
#HPC news...  
[twitter.com/ThisWeekinHPC/...](http://twitter.com/ThisWeekinHPC/)



 **insideHPC.com** 10 feb  
@insideHPC  
SC16 Workshop Proposals Due Feb. 14  
[insidehpc.com/2016/02/sc16-n...](http://insidehpc.com/2016/02/sc16-n...) #HPC  
[pic.twitter.com/LFiJwNTOT4](http://pic.twitter.com/LFiJwNTOT4)  
Retwitteado por TOP500

Twittear a @top500supercomp





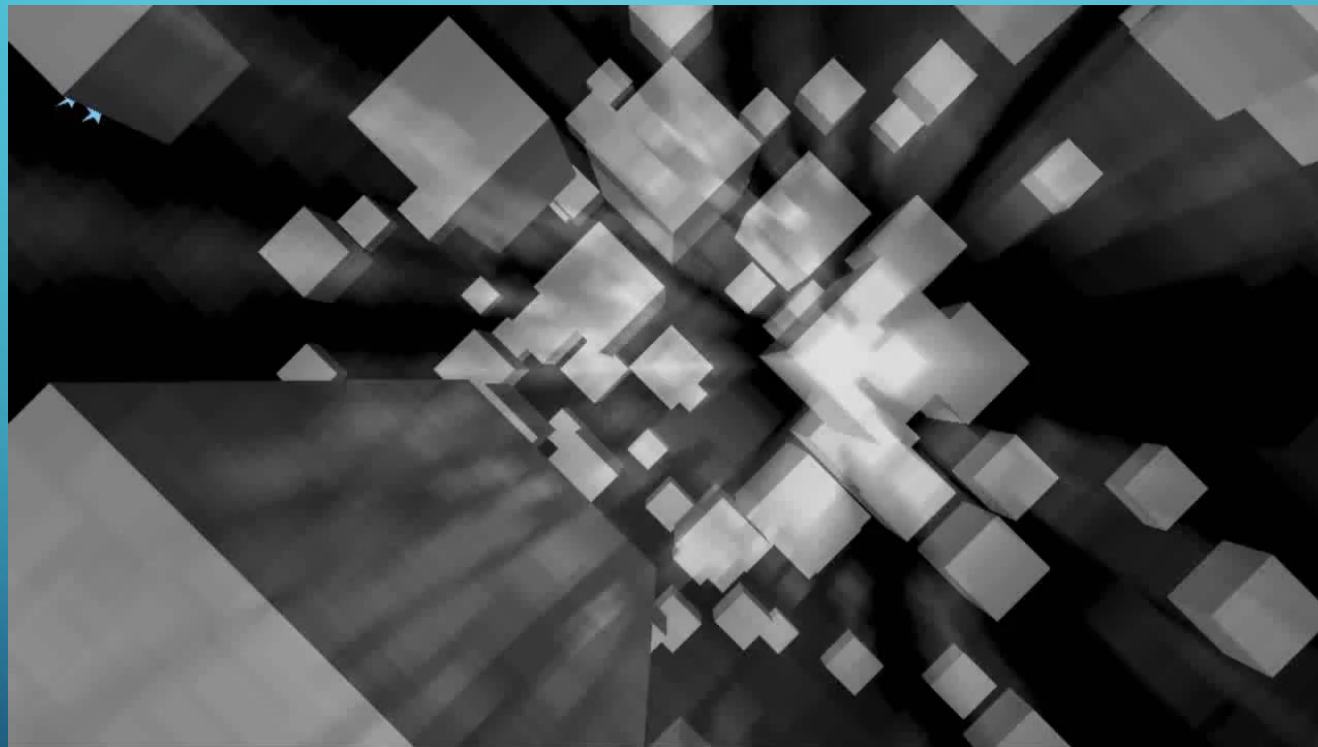
RANGO	SITIO	SISTEMA	NÚCLEOS	RMAX (TFLOP / S)	RPEAK (TFLOP / S)	POTENCIA (KW)
1	Centro Nacional de Super Computadora en Guangzhou China	<b>Tianhe-2 (MilkyWay-2)</b> - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P NUDT	3.120.000	33,862.7	54,902.4	17808
2	/ SC / Oak Ridge National Laboratory DOE Estados Unidos	<b>Titan</b> - Cray Xk7, Opteron 6274 2.200GHz 16C, Cray Gemini interconexión, NVIDIA K20x Cray Inc.	560.640	17,590.0	27,112.5	8209
3	DOE / NNSA / LLNL Estados Unidos	<b>Sequoia</b> - Blue Gene / Q, Power BQC 16C 1,60 GHz, personalizada IBM	1,572,864	17,173.2	20,132.7	7,890
4	Instituto Avanzado RIKEN de Ciencias de la Computación (AICS) Japón	K ordenador, SPARC64 VIIIfx 2,0 GHz, interconexión Tofu Fujitsu	705.024	10,510.0	11,280.4	12.660
5	DOE / SC / Argonne National Laboratory Estados Unidos	<b>Mira</b> - Blue Gene / Q, Power BQC 16C 1,60 GHz, personalizada IBM	786.432	8,586.6	10,066.3	3.945
6	DOE / NNSA / LANL / SNL Estados Unidos	<b>Trinidad</b> - Cray XC40, Xeon E5-2698v3 16C 2,3 GHz, interconexión Aries Cray Inc.	301.056	8,100.9	11,078.9	
7	Suizo Centro Nacional de Supercomputación (CSCS) Suiza	<b>Piz Daint</b> - Cray XC30, Xeon E5-2670 8C 2.600GHz, interconexión Aries, NVIDIA K20x Cray Inc.	115.984	6,271.0	7,788.9	2.325
8	HLRS - Höchstleistungsrechenzentrum Stuttgart Alemania	<b>Hazel gallina</b> - Cray XC40, Xeon a 2,5 GHz E5-2680v3 12C, interconexión Aries Cray Inc.	185.088	5,640.2	7,403.5	
9	Universidad Rey Abdullah de Ciencia y Tecnología Arabia Saudita	<b>Shaheen II</b> - Cray XC40, Xeon E5-2698v3 16C 2,3 GHz, interconexión Aries Cray Inc.	196.608	5,537.0	7,235.2	2.834
10	Computación Avanzada de Texas Centro / Univ. de Tejas Estados Unidos	<b>Estampida</b> - PowerEdge C8220, Xeon E5-2680 8C 2.700GHz, Infiniband FDR, Intel Xeon Phi SE10P Dell	462.462	5,168.1	8,520.1	4,510





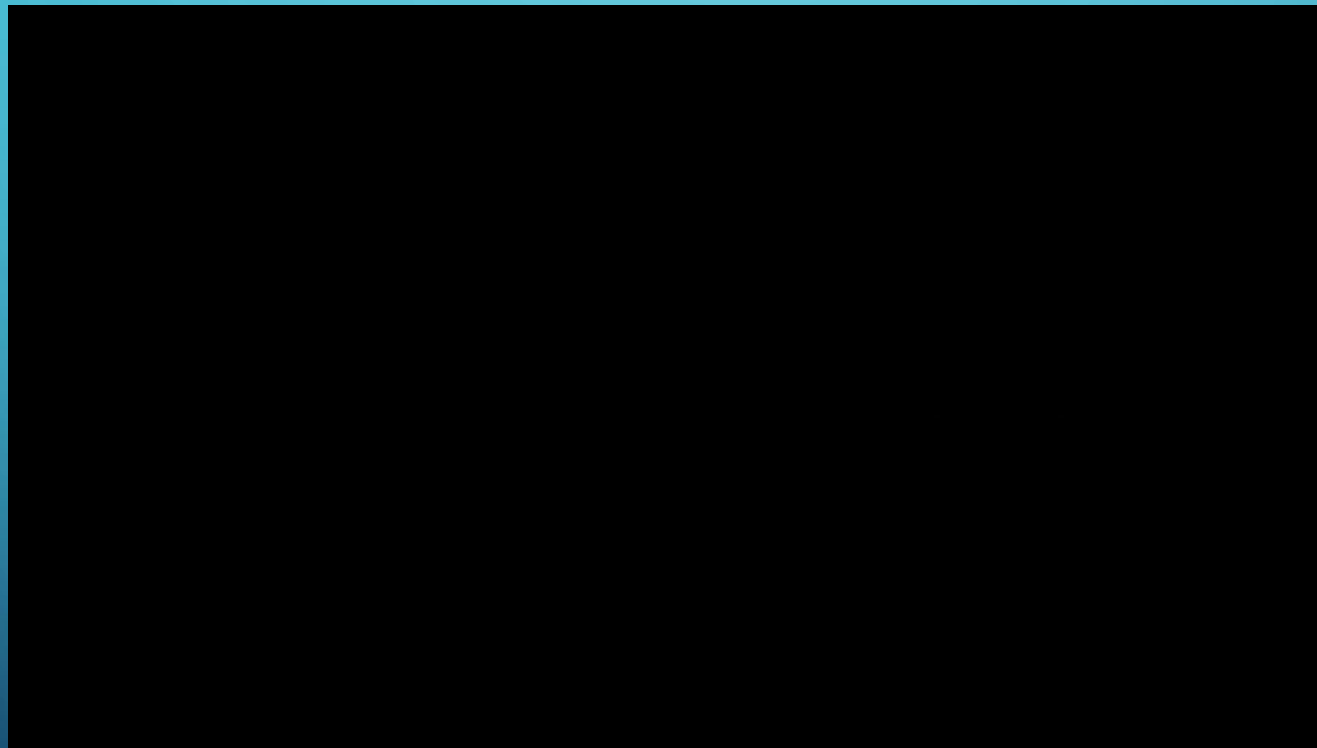
# ARMADO DE UN SUPERCOMPUTADORA

<https://www.youtube.com/watch?v=ZSb4ePL4hjo>



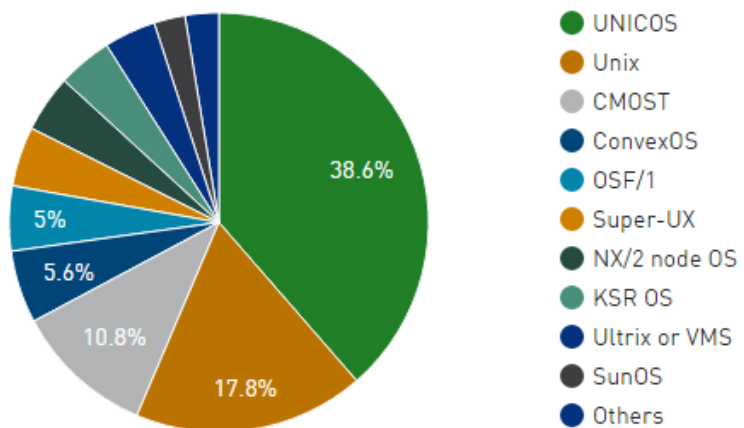
# SIMULACIONES

<https://www.youtube.com/watch?v=FMpuujaldZk>



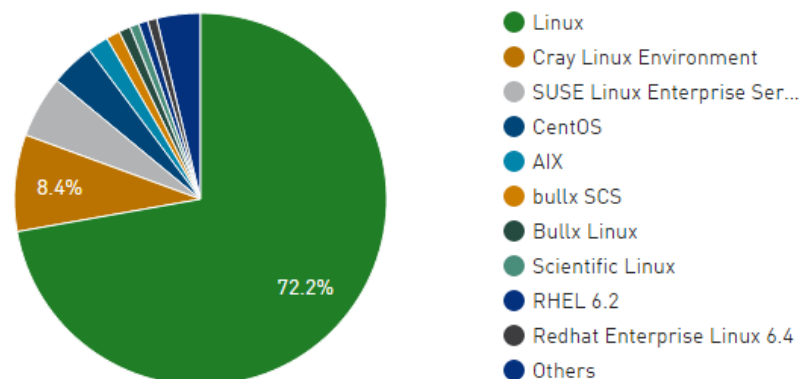
# SISTEMAS OPERATIVOS

Operating System System Share



Junio 1993

Operating System System Share

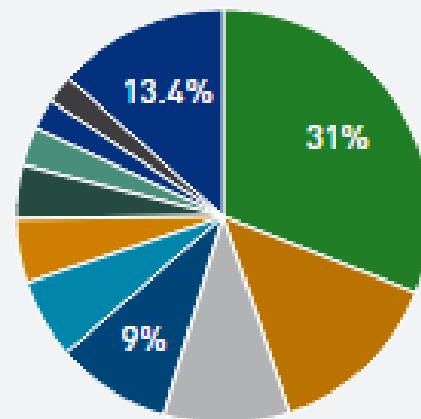


Junio 2015



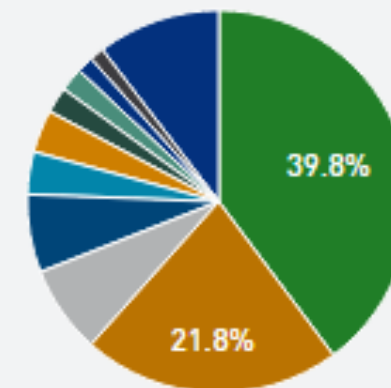
## Vendors System Share

- HP
- Cray Inc.
- Sugon
- IBM
- SGI
- Lenovo
- Bull
- Inspur
- Fujitsu
- Dell
- Others



## Country System Share

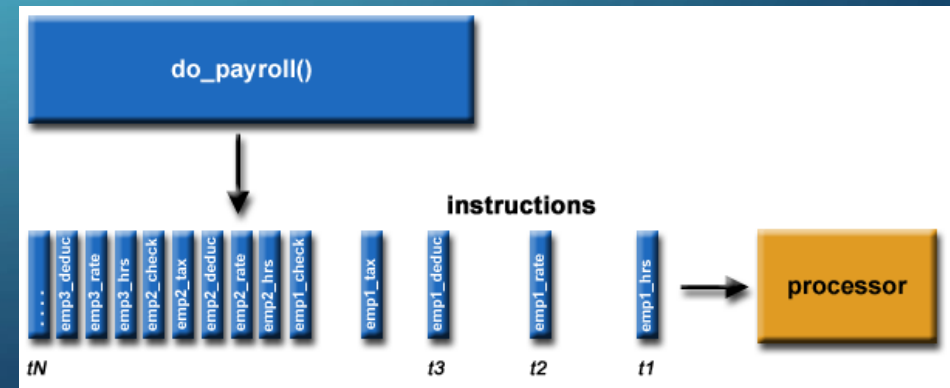
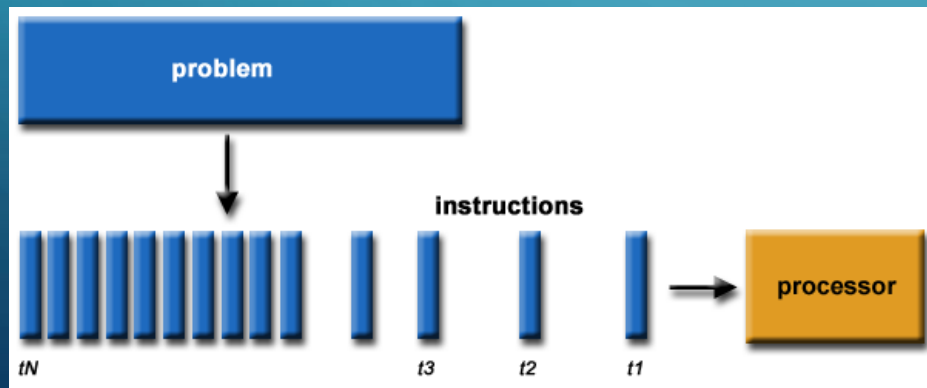
- United States
- China
- Japan
- Germany
- United K...
- France
- India
- Korea, S...
- Russia
- Saudi Ar...
- Others



# ¿QUÉ ES LA COMPUTACIÓN PARALELA?

- **Computación de serie:**

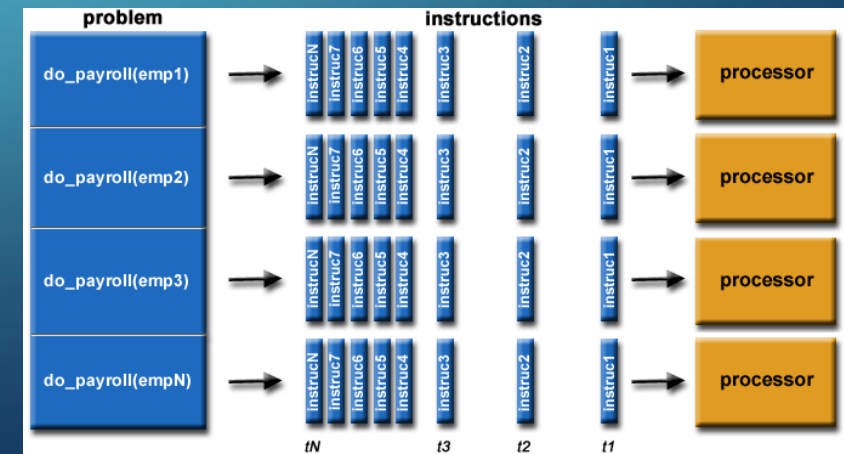
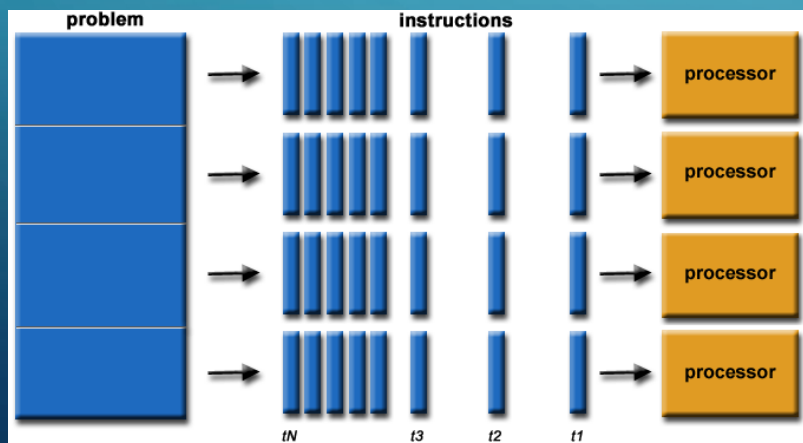
- Tradicionalmente, el software ha sido escrito para el cálculo **de serie**:
  - Un problema se divide en una serie discreta de instrucciones
  - Las instrucciones se ejecutan secuencialmente uno tras otro
  - Ejecutado en un único procesador
  - Sólo una instrucción puede ejecutar en cualquier momento en el tiempo



# ¿QUÉ ES LA COMPUTACIÓN PARALELA?

- **Computación paralela:**

- En el sentido más simple, *la computación paralela* es el uso simultáneo de múltiples recursos informáticos para resolver un problema de cálculo:
  - Un problema se divide en partes discretas que se pueden resolver simultáneamente
  - Cada parte se subdivide en una serie de instrucciones
  - Las instrucciones de cada parte se ejecutan de forma simultánea en diferentes procesadores
  - Se emplea un mecanismo general de control / coordinación



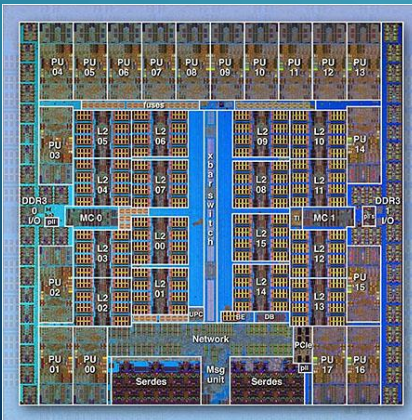


# ¿QUÉ ES LA COMPUTACIÓN PARALELA?

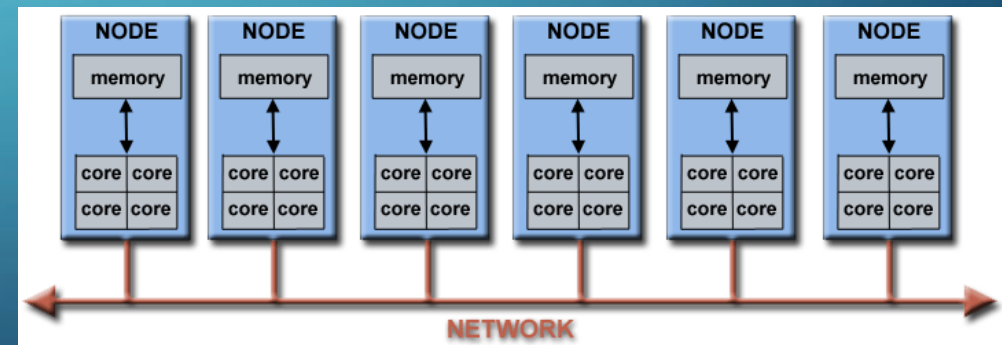
- Los recursos informáticos suelen ser: Un solo equipo con múltiples procesadores / núcleos
- Un número arbitrario de este tipo de ordenadores conectados por una red

# ¿QUÉ ES LA COMPUTACIÓN PARALELA?

- **Las computadoras paralelas:**
- Prácticamente todos los ordenadores autónomos hoy son paralelos desde la perspectiva del hardware:
  - unidades funcionales múltiples (caché L1, L2 cache, rama, captación previa, decodificar, de punto flotante, de procesamiento de gráficos (GPU), número entero, etc.)
  - múltiples unidades de ejecución / núcleos
  - hilos de hardware múltiples



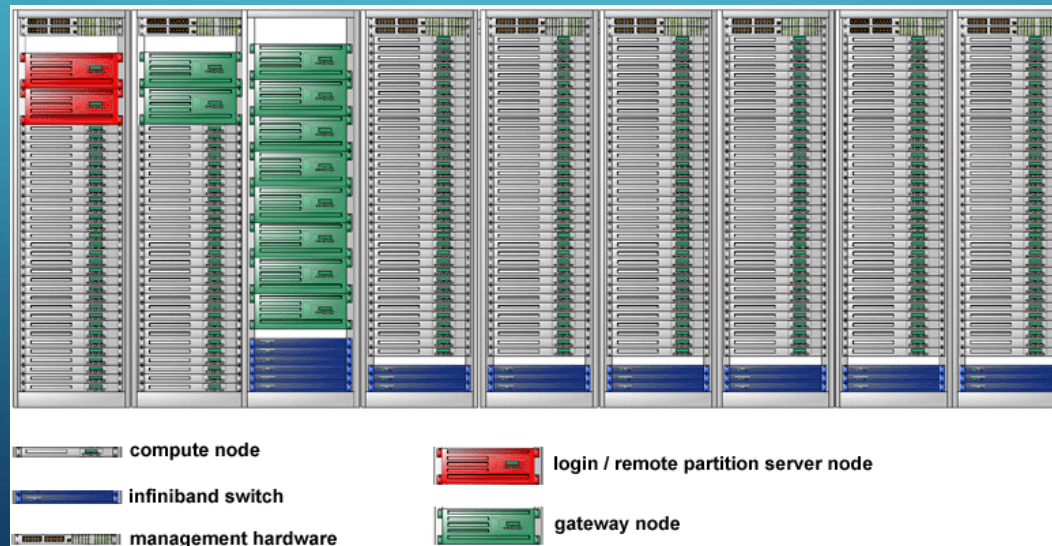
IBM BG / Q Calcular la viruta con 18 núcleos (PU) y 16 unidades de caché L2 (L2)



Las redes se conectan varios equipos independientes (nodos) para hacer clusters de ordenadores paralelos de mayor tamaño.

# ¿QUÉ ES LA COMPUTACIÓN PARALELA?

- Por ejemplo, el esquema de abajo muestra un cluster de ordenadores en paralelo **Lawrence Livermore National Laboratory** (LLNL) típica:
  - Cada nodo de cálculo es un ordenador paralelo multi-procesador en sí mismo
  - múltiples nodos de cómputo están conectados en red, junto con una red Infiniband
  - nodos de propósito especial, también con varios procesadores, se utilizan para otros fines





# ¿POR QUÉ UTILIZAR LA COMPUTACIÓN EN PARALELO?

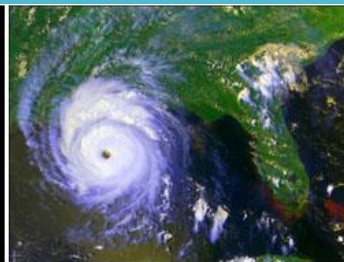
- El mundo real es paralelo masivo:
  - En el mundo natural, muchos, eventos interrelacionados complejos están sucediendo al mismo tiempo, sin embargo, dentro de una secuencia temporal.
  - En comparación con la computación en serie, la computación paralela es mucho más adecuada para la modelización, la simulación y la comprensión compleja, fenómenos del mundo real.
  - Por ejemplo, imagina el modelado de éstos en serie:



Galaxy Formation



Planetary Movments



Climate Change



Rush Hour Traffic



Plate Tectonics



Weather



Auto Assembly



Jet Construction



Drive-thru Lunch

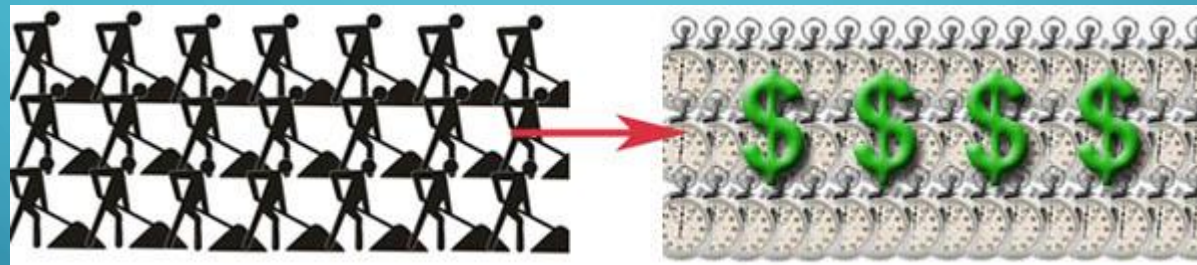
# ¿POR QUÉ UTILIZAR LA COMPUTACIÓN EN PARALELO?

- **Las razones principales:**

- Ahorrar tiempo y / o dinero.
- **RESOLVER LOS PROBLEMAS MAYORES / más complejo.**
- **SUMINISTRO DE CONCURRENCIA.**
- Aprovechar los recursos no locales.
- Hacer un mejor uso del hardware **PARALELO** subyacente.

# ¿POR QUÉ UTILIZAR LA COMPUTACIÓN EN PARALELO?

- **Ahorrar tiempo y / o dinero:**
  - En teoría, lanzando más recursos a una tarea acortará su tiempo de finalización, con un ahorro potencial de costes.
  - ordenadores paralelos se pueden construir a partir de las materias primas, componentes baratos





# ¿POR QUÉ UTILIZAR LA COMPUTACIÓN EN PARALELO?

- **RESOLVER LOS PROBLEMAS MAYORES / más complejo:**

- Muchos problemas son tan grandes y / o complejos que es imposible o poco práctico para resolverlos en un solo equipo, especialmente teniendo en cuenta la memoria del ordenador limitado.
- Ejemplo: "Gran Desafío de problemas" ( [en.wikipedia.org/wiki/Grand\\_Challenge](https://en.wikipedia.org/wiki/Grand_Challenge) ) que requiere PetaFLOPS y petabytes de los recursos informáticos.
- Ejemplo: motores de búsqueda web / bases de datos procesar millones de transacciones por segundo



# ¿POR QUÉ UTILIZAR LA COMPUTACIÓN EN PARALELO?

- **SUMINISTRO DE CONCURRENCIA:**

- Un recurso de cómputo única sólo puede hacer una cosa a la vez. Múltiples recursos informáticos pueden hacer muchas cosas a la vez.
- Ejemplo: Las redes de colaboración proporcionan un lugar mundial donde la gente de todo el mundo pueden reunirse y llevar a cabo el trabajo "virtualmente".



# ¿POR QUÉ UTILIZAR LA COMPUTACIÓN EN PARALELO?

- **Aprovechar los recursos no locales :**

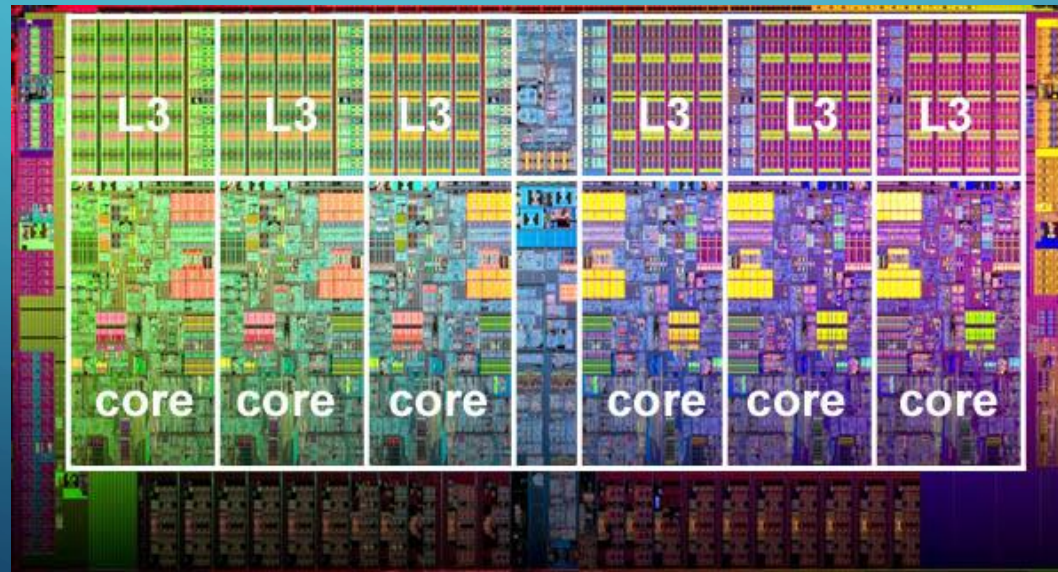
- El uso de los recursos informáticos en una red de área amplia, o incluso por Internet cuando los recursos informáticos locales son escasos o insuficientes.
- Ejemplo: SETI @ home ( [setiathome.berkeley.edu](http://setiathome.berkeley.edu) ) más de 1,5 millones de usuarios en casi todos los países del mundo. Fuente: [www.boincsynergy.com/stats/](http://www.boincsynergy.com/stats/) (junio de 2015).
- Ejemplo: Folding @ home ( [folding.stanford.edu](http://folding.stanford.edu) ) utiliza más de 160.000 ordenadores en todo el mundo (junio de 2015)



# ¿POR QUÉ UTILIZAR LA COMPUTACIÓN EN PARALELO?

- **Hacer un mejor uso del hardware PARALELO subyacente:**

- Las computadoras modernas, incluso ordenadores portátiles, son paralelos en la arquitectura con múltiples procesadores / núcleos.
- software paralelo está destinado específicamente para el hardware paralelo con múltiples núcleos, hilos, etc.
- En la mayoría de los casos, los programas de serie se ejecutan en equipos modernos "desperdician" potencial de potencia de cálculo.



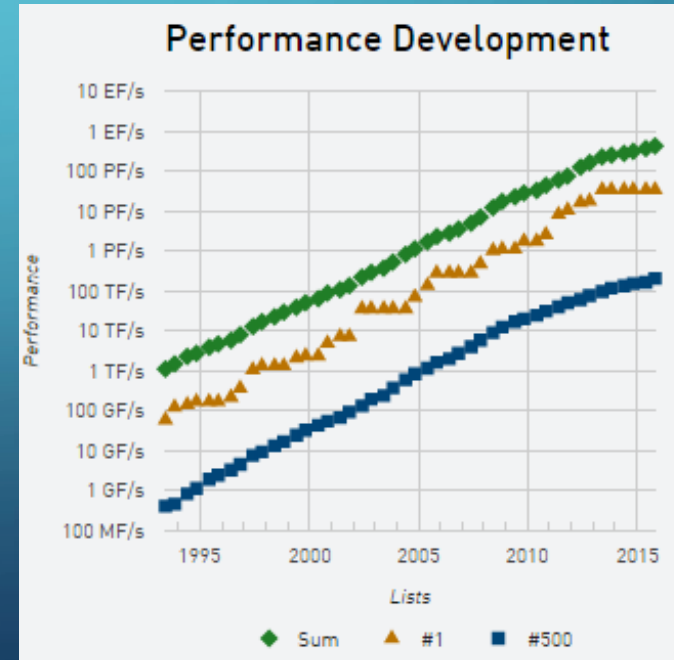
procesador Intel Xeon de 6 núcleos y 6 unidades de caché L3



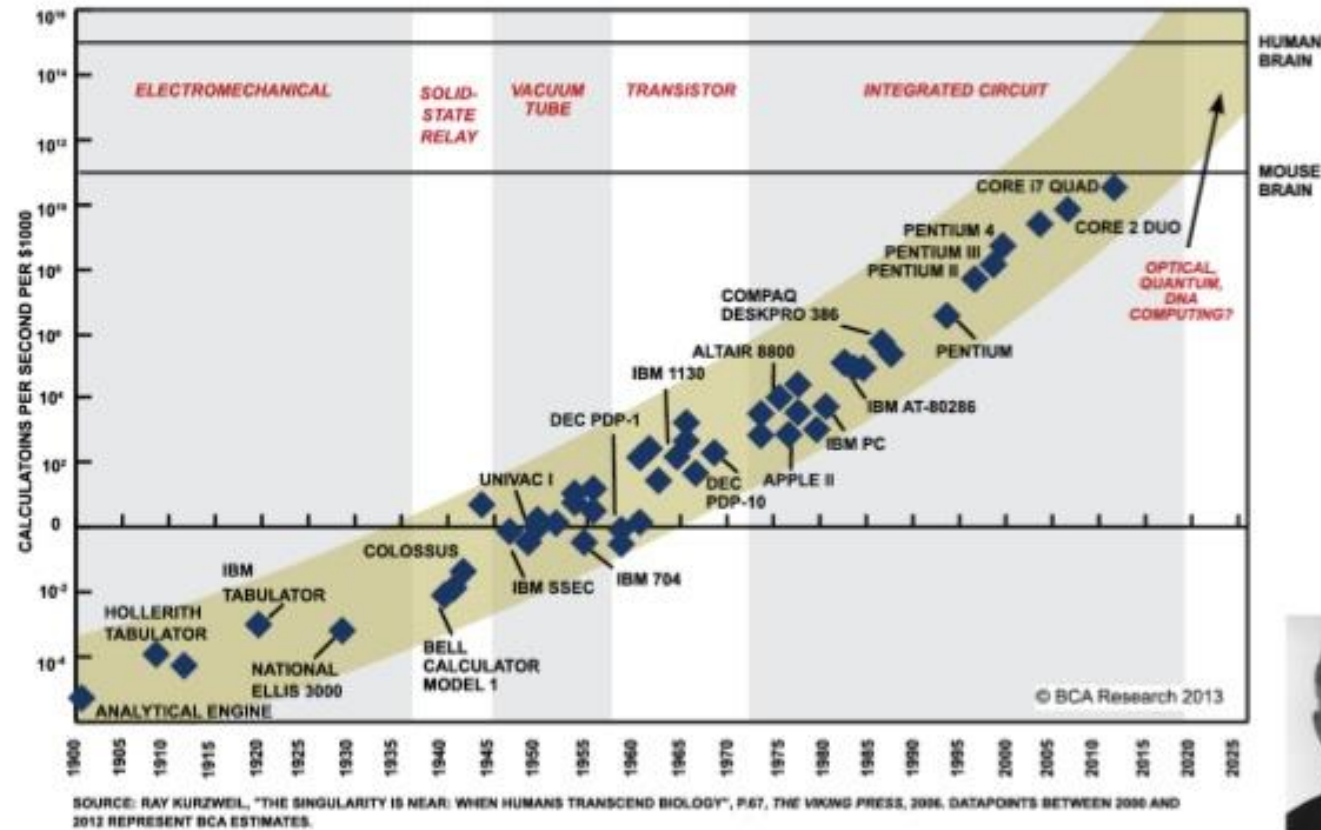
# ¿POR QUÉ UTILIZAR LA COMPUTACIÓN EN PARALELO?

- **El futuro:**

- Durante los últimos 20 años, las tendencias indicadas por las redes cada vez más rápido, sistemas distribuidos y arquitecturas de computadora con múltiples procesadores (incluso a nivel de escritorio) muestran claramente *que el paralelismo es el futuro de la computación.*
- En este mismo período de tiempo, se ha producido un incremento mayor que **500,000x** en el rendimiento superordenador, actualmente sin un final a la vista.
- *La carrera ya está en Exaescala de Informática!*
- Exaflop =  $10^{18}$  cálculos por segundo



# La Ley de Moore mueve el mundo ...



- <http://elrincondelacienciaytecnologia.blogspot.mx/2011/10/el-fin-de-la-ley-de-moore.html>

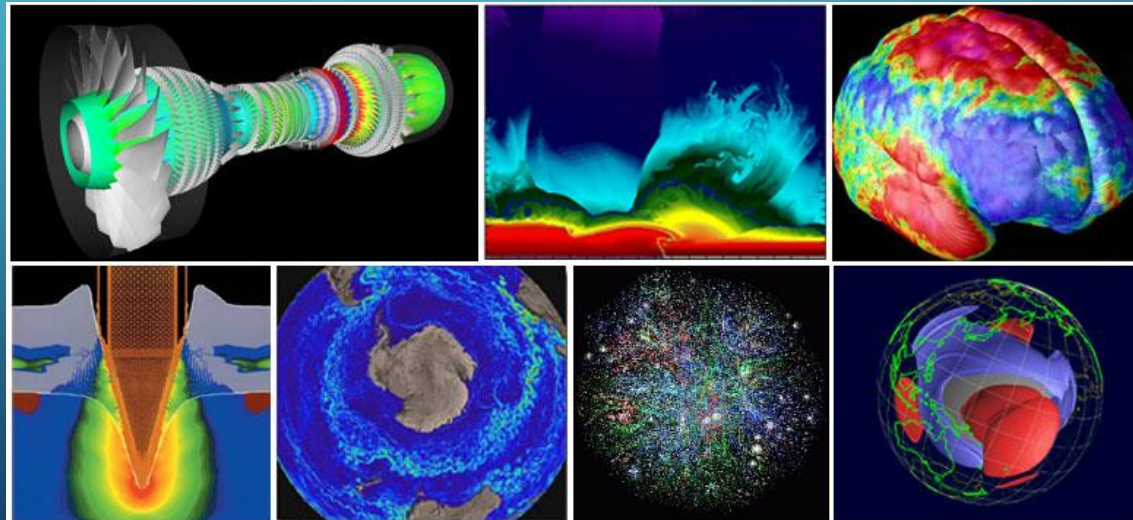
# ¿QUIÉN ESTÁ UTILIZANDO LA COMPUTACIÓN EN PARALELO?

## Ciencia e Ingeniería:

- Históricamente, la computación paralela se ha considerado como "el extremo superior de la computación",
- y se ha utilizado para modelar problemas difíciles en muchas áreas de la ciencia y la ingeniería:

- Atmósfera, la Tierra, Medio Ambiente
- Física - aplicado, nuclear, partículas, la materia condensada, de alta presión, la fusión, la fotónica
- Bioscience, Biotecnología, Genética
- Química, Ciencias Moleculares
- Geología, sismología

- Ingeniería Mecánica - de prótesis para las naves espaciales
- Ingeniería Eléctrica, Diseño de Circuitos, Microelectrónica
- Informática, matemáticas
- Las armas de defensa,



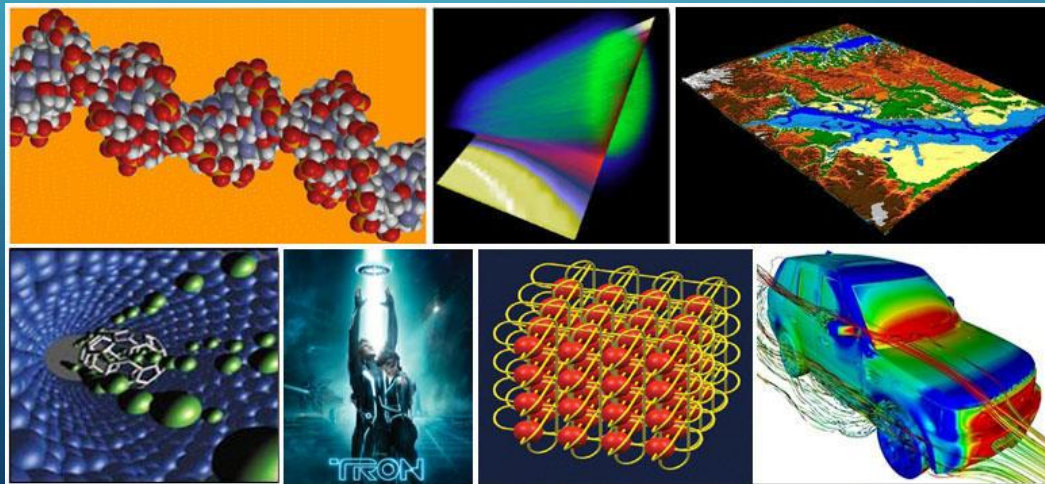


# ¿QUIÉN ESTÁ UTILIZANDO LA COMPUTACIÓN EN PARALELO?

## Industrial y Comercial:

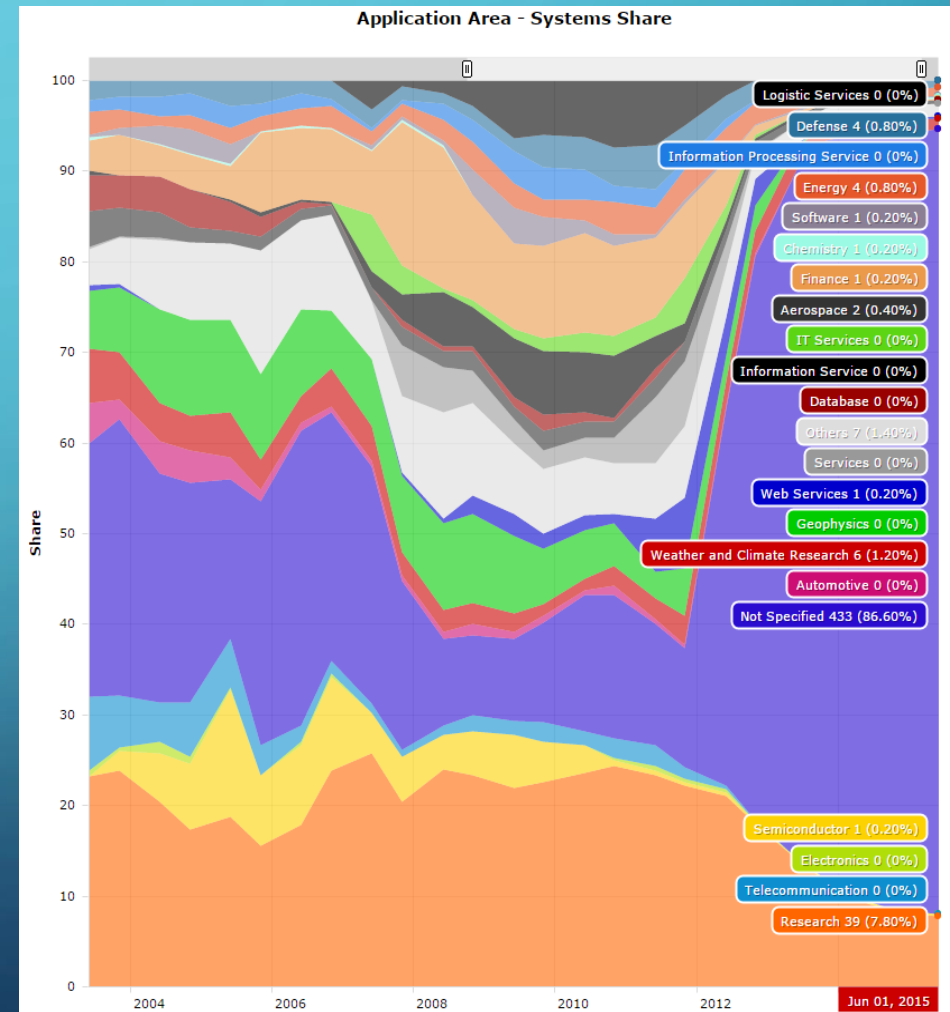
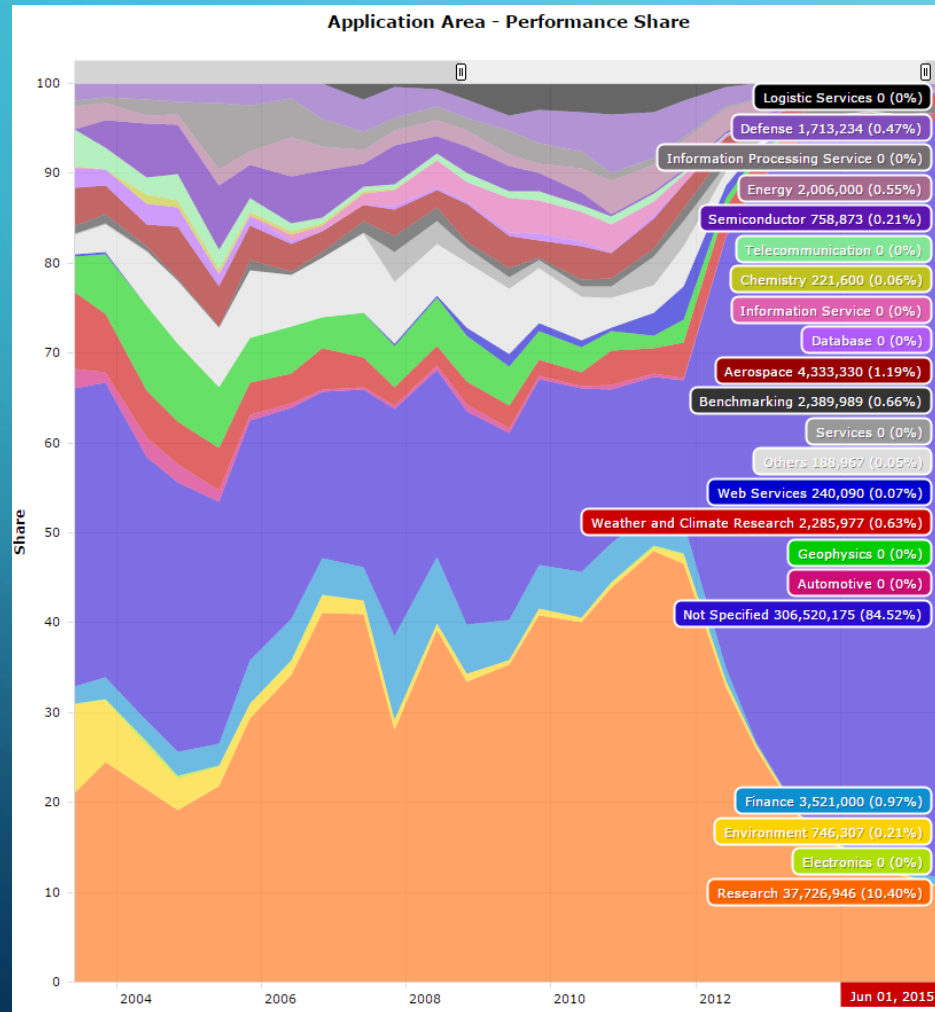
•Hoy en día, las aplicaciones comerciales proporcionan una fuerza motriz igual o mayor en el desarrollo de ordenadores más rápidos. Estas aplicaciones requieren el procesamiento de grandes cantidades de datos en formas sofisticadas. Por ejemplo:

- la minería "Big Data", bases de datos, los datos
- Exploración de aceite
- Los motores de búsqueda en Internet, servicios de negocios basados en la web
- imágenes médicas y diagnóstico
- Pharmaceutical Design
- modelado financiero y económico
- La gestión de las empresas nacionales y multinacionales
- Gráficos avanzados y la realidad virtual, en particular en la industria del entretenimiento
- tecnologías de vídeo en red y multimedia
- entornos de trabajo colaborativo





# ¿QUIÉN ESTÁ UTILIZANDO LA COMPUTACIÓN EN PARALELO?



# LIGAS IMPORTANTES.

- Introducción a la Programación Paralelas
  - [https://computing.llnl.gov/tutorials/parallel\\_comp/](https://computing.llnl.gov/tutorials/parallel_comp/)
- Ejemplo con MPI
  - [http://lsi.ugr.es/jmantas/pdp/tutoriales/tutorial\\_mpi.php](http://lsi.ugr.es/jmantas/pdp/tutoriales/tutorial_mpi.php)
- <http://top500.org/timeline/>