

Computación Científica de Alto Desempeño

Un acercamiento pragmático

(sólo es verdadero aquello que funciona)

Edison Montoya

Benemérita Universidad Autónoma de Puebla

Febrero 4 de 2015

Computación en Paralelo

“Parallel computing is a form of computation in which many calculations are carried out simultaneously”

Gottlieb, Almasi. Highly parallel computing.

Hardware

- Procesadores multicore
- Coprocesadores
- Tarjetas de Video
- Clusters
- Computación en la nube

Software

- MPI
- OpenMP
- CUDA

De ahora en adelante seremos pragmáticos y solo vamos a hablar de lo que ha sido útil para el expositor, no haremos un recorrido exhaustivo por ninguno de los temas a tratar.

Temas a tratar

Vamos a hablar (un poco) sobre:

- Nociones **básicas** de arquitectura de procesadores (intel, amd), coprocesadores (intel) y tarjetas de video (Nvidia).
- Métodos y lenguajes de paralelización.

Nos enfocaremos solo en plataformas linux.

No vamos a hablar sobre:

- Las distintas redes de intercomunicación que se pueden utilizar en una infraestructura en paralelo.
- Tipos de software utilizados para controlar el hardware.
- Entrada y salida de datos (lectura y escritura a disco duro).
 - **Nota:** El programador debe olvidarse de obtener la entrada de datos por teclado y la salida por pantalla.
- Optimización.

“Lenguajes” de Programación

Los “lenguajes” de paralelización que vamos a discutir son:

- **MPI (Message Passing Interface)**

Es un estándar que define la sintaxis y la semántica de las funciones contenidas en una [biblioteca](#) de paso de mensajes diseñada para ser usada en programas que exploten la existencia de múltiples procesadores.

MPI es un estandar en el cual los nodos de computo en un cluster no comparten memoria, todo el intercambio de datos e interacciones se dan a traves de comunicaciones explicitas

- **OpenMP (Open Multi-Processing)**

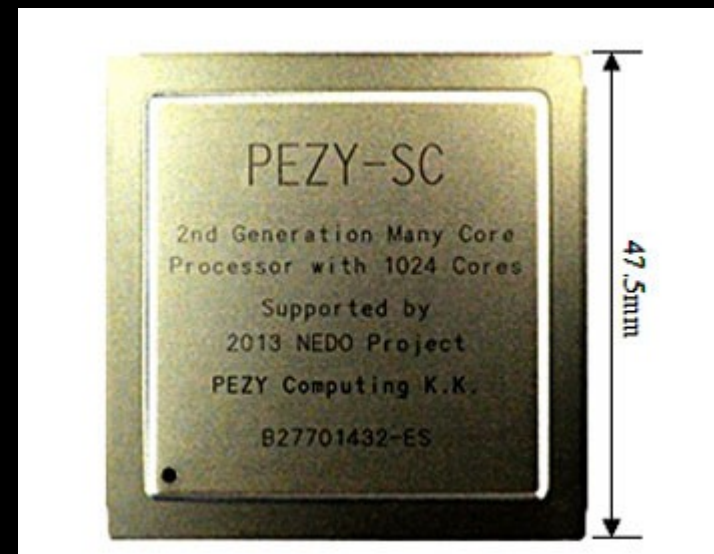
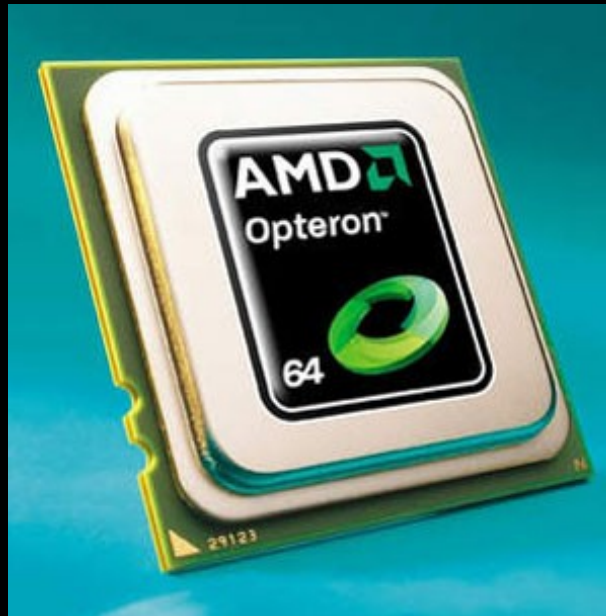
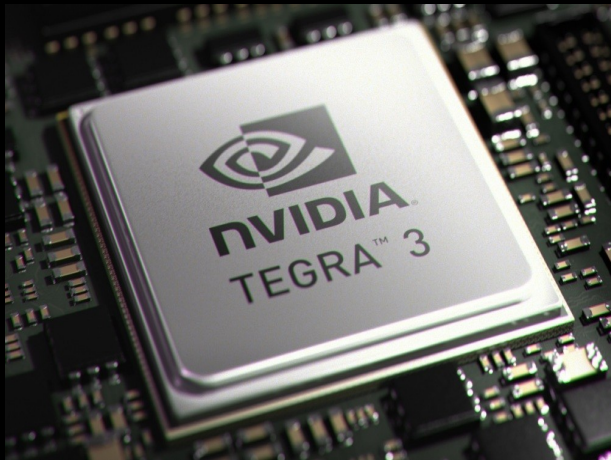
Es una interfaz de programación de aplicaciones ([API](#)) para la programación multiproceso de memoria compartida en múltiples plataformas.

OpenMP es para sistemas con multiples procesadores que comparten memoria, todas las comunicaciones son implicitas.

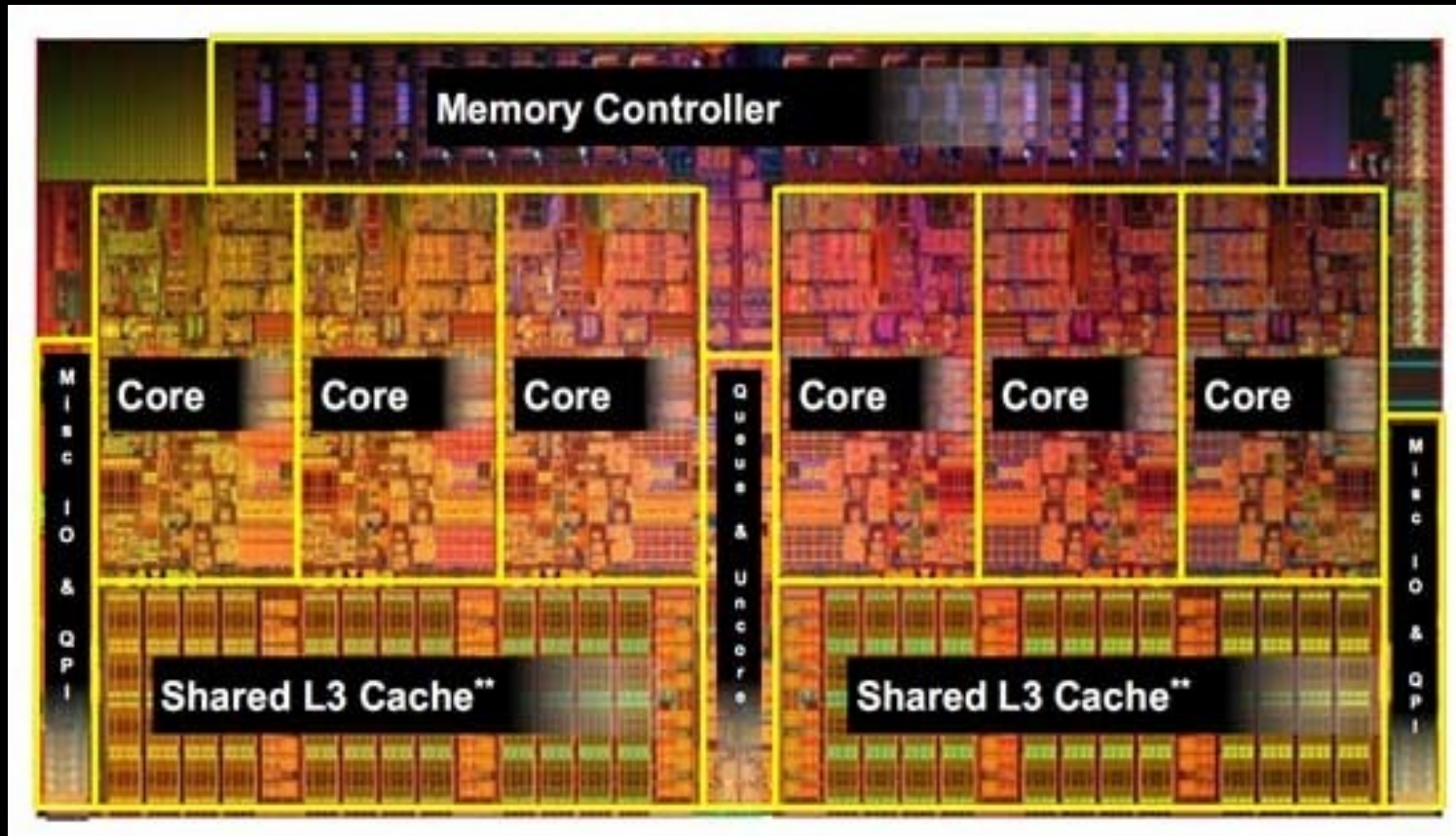
- **CUDA C (Compute Unified Device Architecture)**

Hace referencia tanto a un compilador (nvcc) como a un conjunto de herramientas de desarrollo creadas por Nvidia que permiten a los programadores usar una variación del [lenguaje de programación C](#) para codificar algoritmos en GPU (Graphics Processing Unit) de Nvidia.

Procesadores



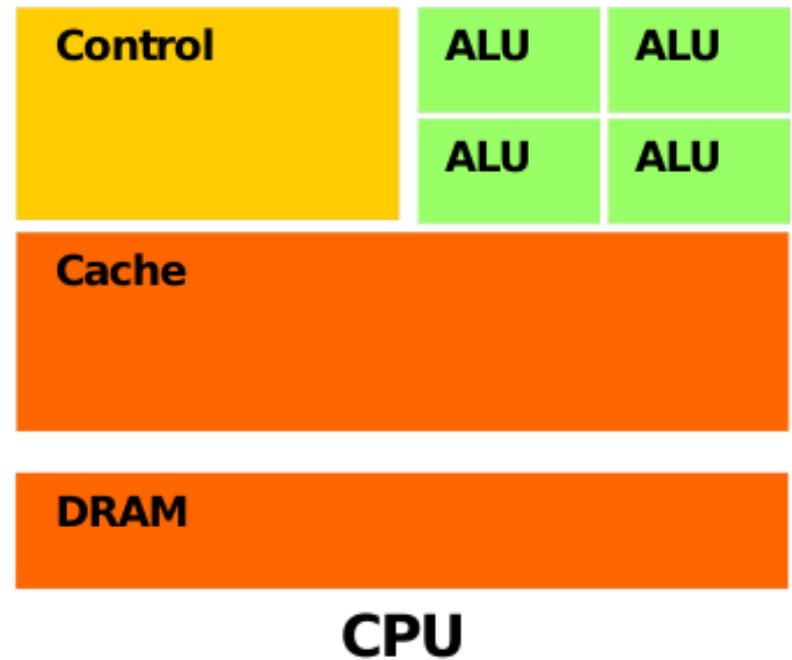
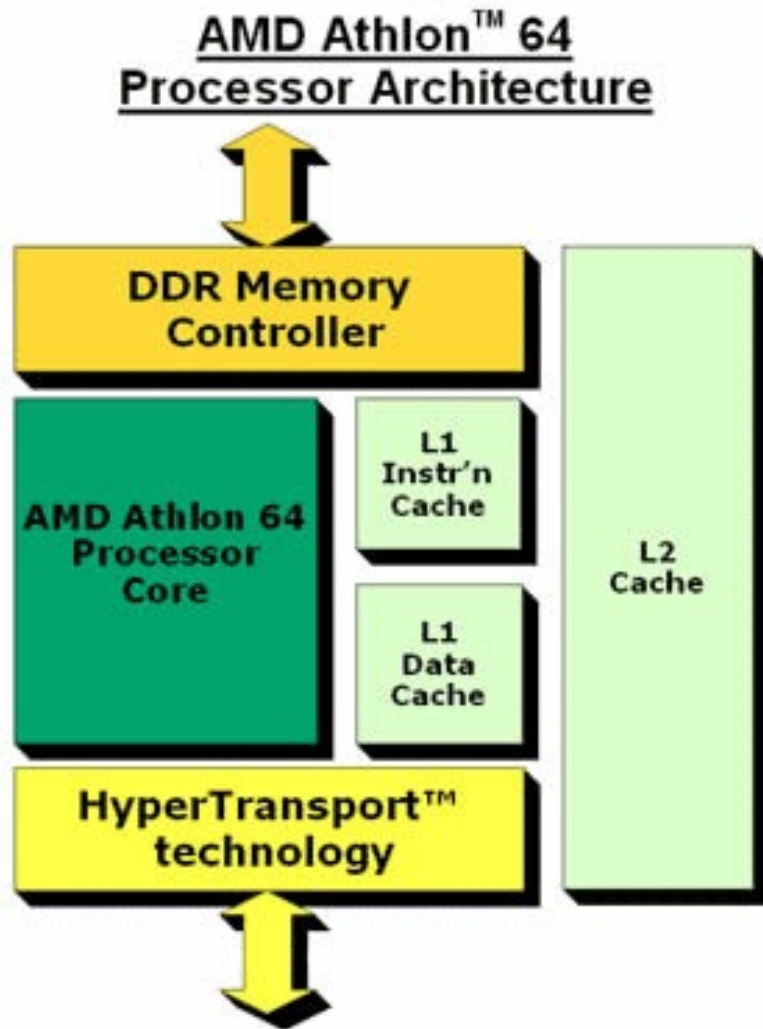
Intel® Core™ i7-990X Processor Extreme Edition



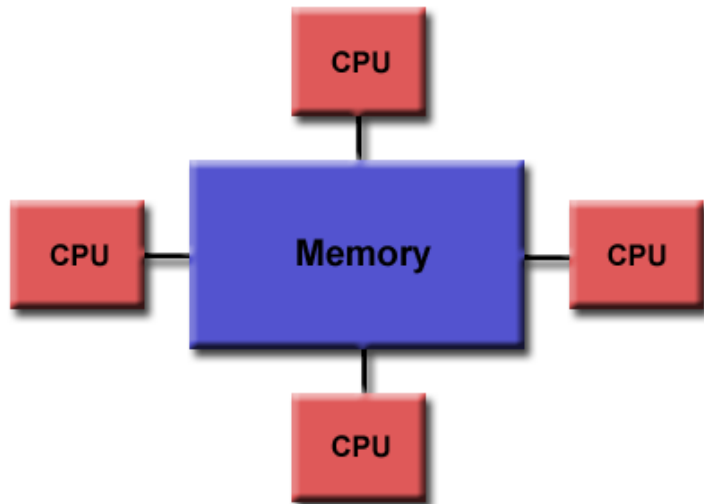
6 Cores , 12 Threads, 3.46 GHz Clock Speed, 24 GB Max
Memory Size, Memory Type DDR3-1066

Lenguaje: Nucleo=Thread=logical core=CPU

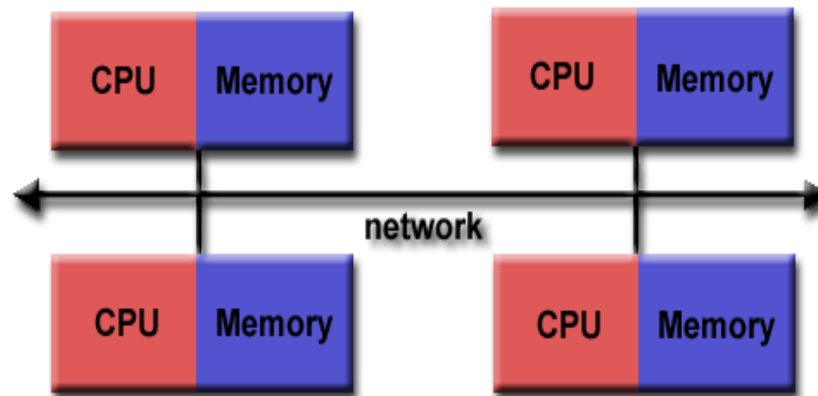
Procesadores



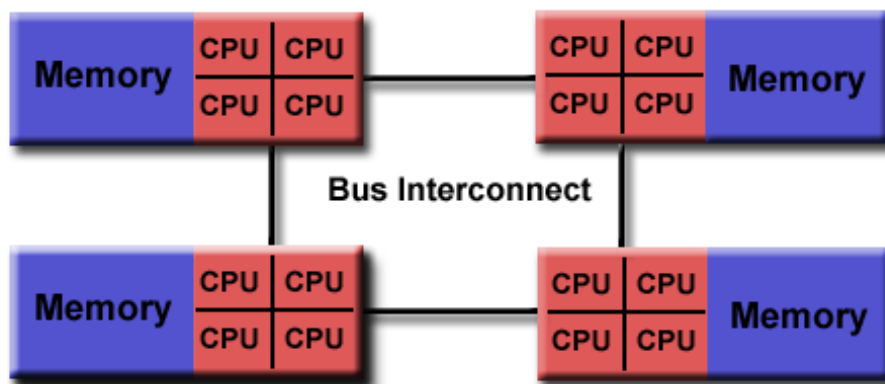
Procesadores en Paralelo



← OpenMP



← MPI

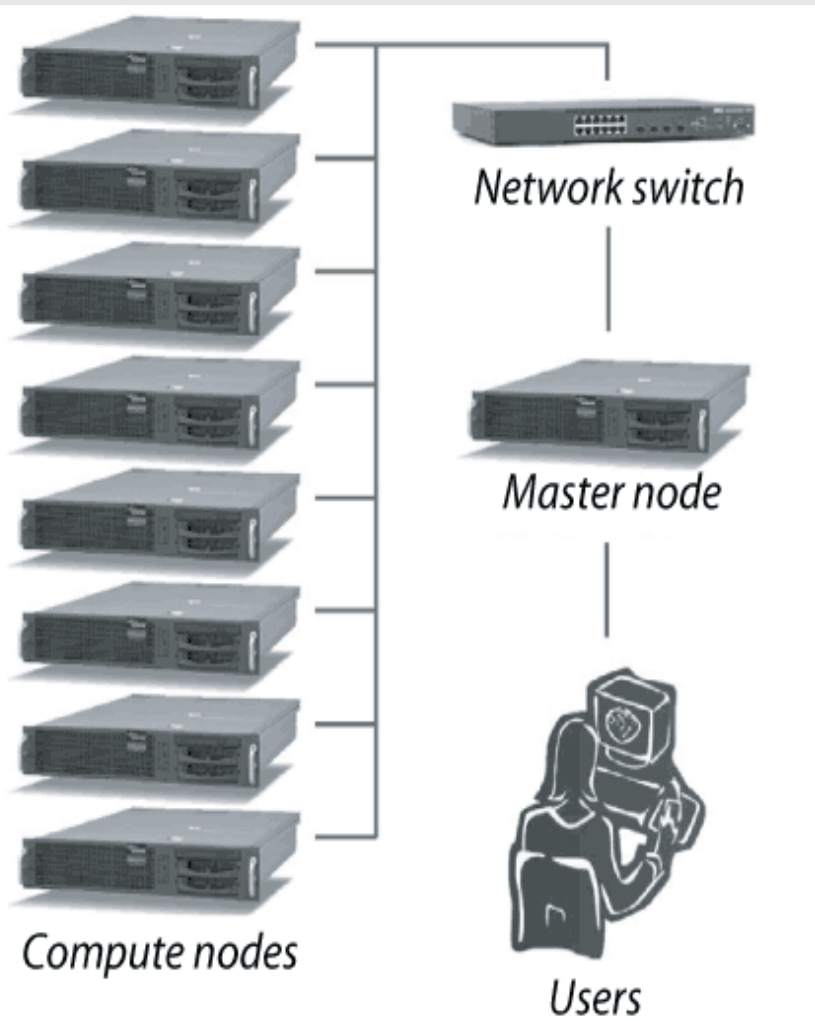


← MPI

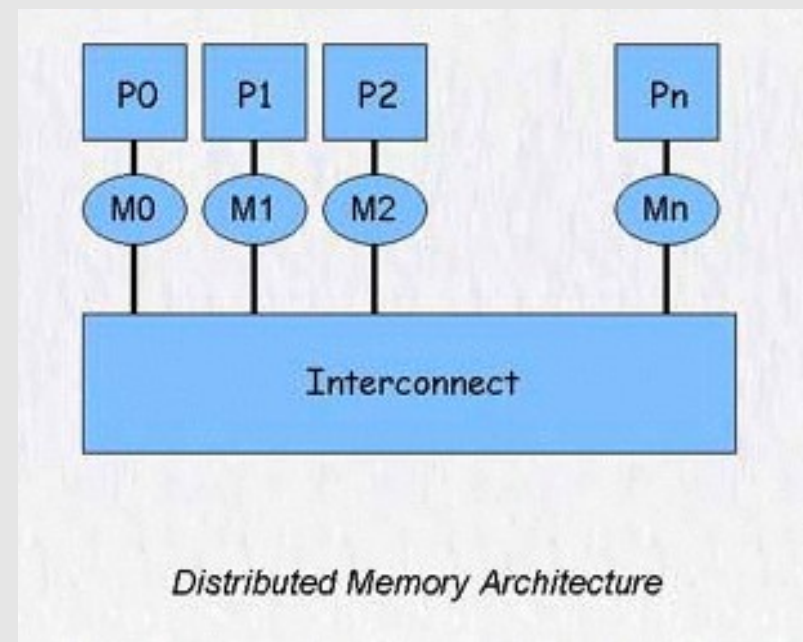
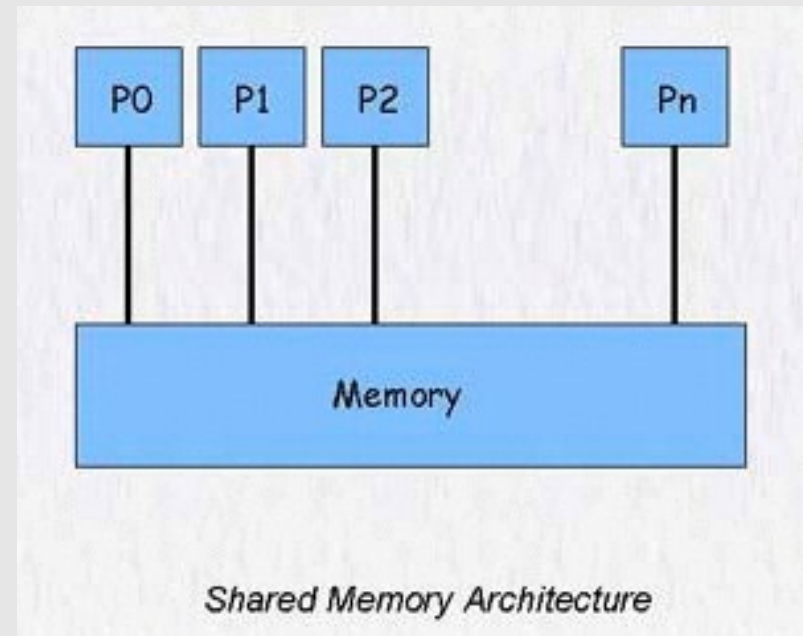
ó

← MPI + OpenMP

Cluster

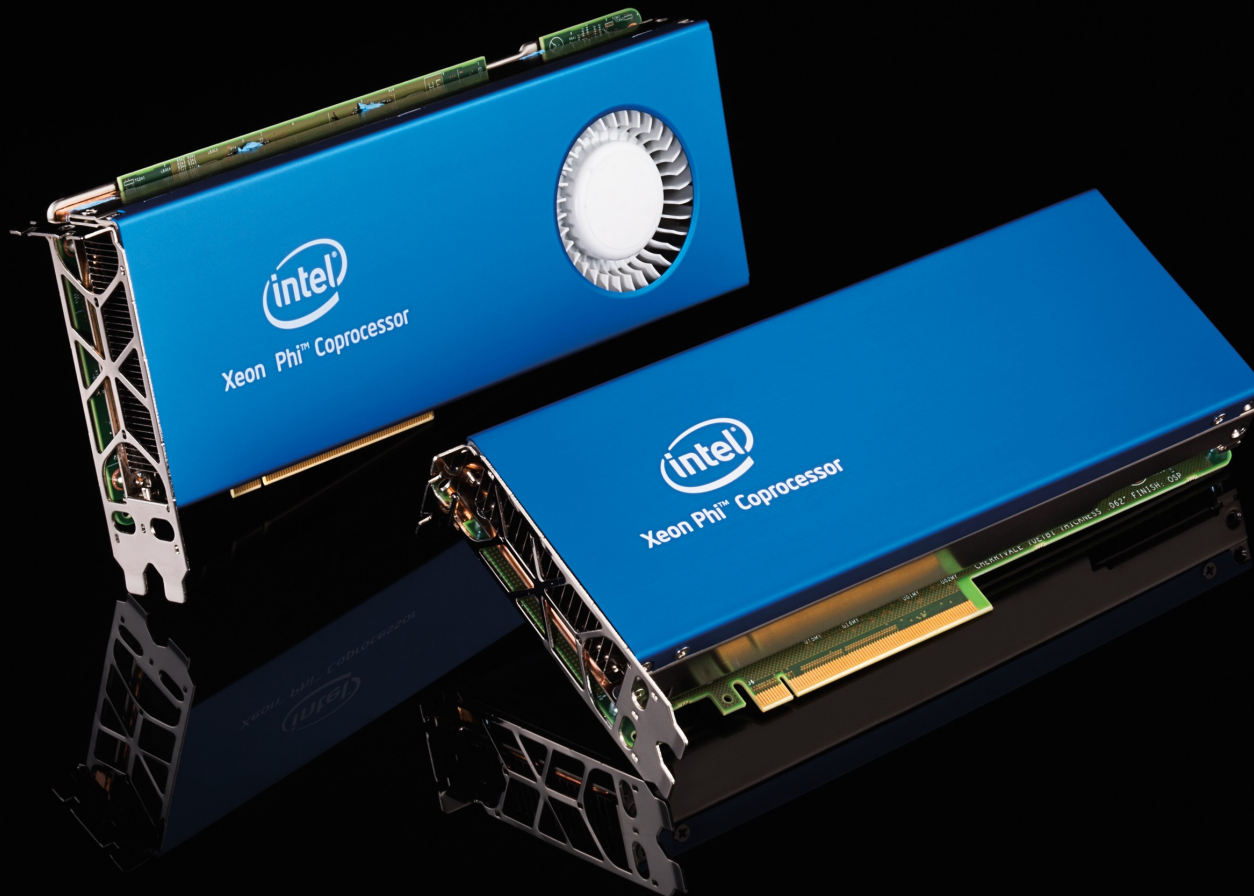


Source: ILRI High Performance Computing

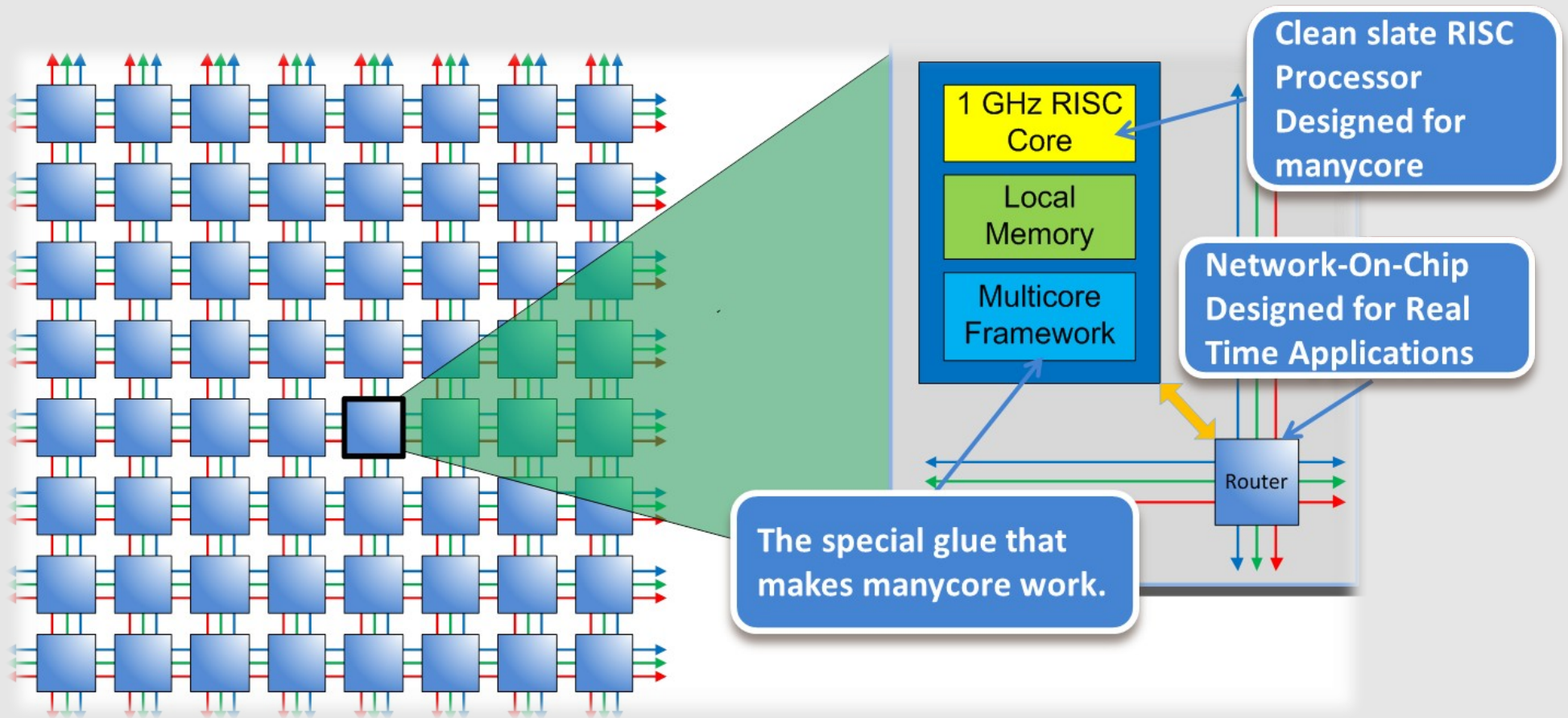


Coprocesadores

Un coprocesador es un microprocesador de un ordenador utilizado como suplemento de las funciones del procesador principal (CPU). Wikipedia



Intel® Xeon Phi Coprocessor



Coprocessor to
ARM/Intel CPU

25mW per core

Ease To Use

Coprocesadores

Fabricante: Intel

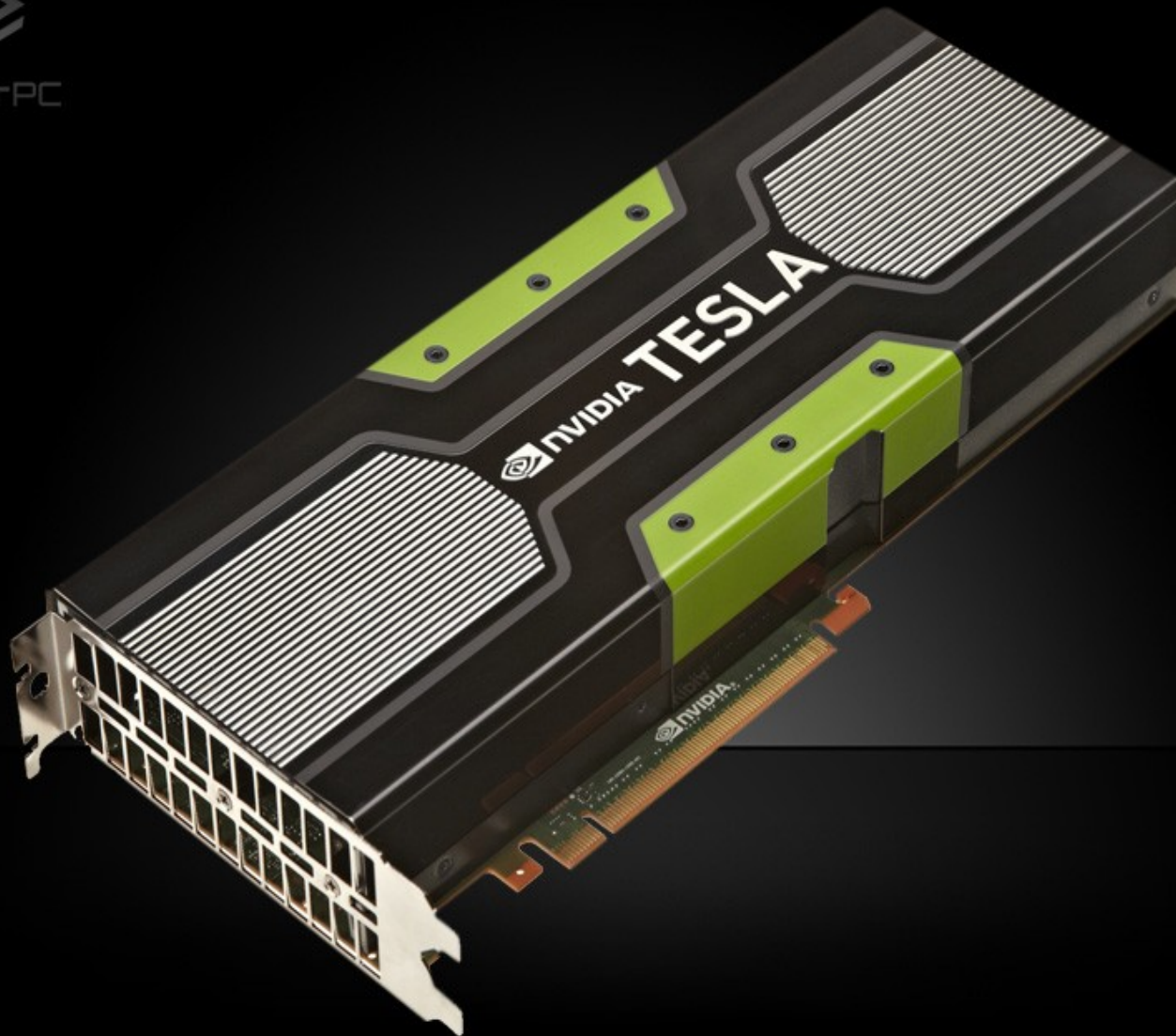
Arquitectura: MIC (Many Integrated Cores)

Marca de Fabrica: Xeon Phi

Modelos: 7120X, 7120P, 5120D, 5120P, 3120P, 3120A

	7120X, 7120P	5120D, 5120P	3120P, 3120A
Memoria	16GB	8GB	4GB
Cores	61	60	57
Threads	244	240	228
Frecuencia	1.2 GHz	1.0 GHz	1.1 GHz
Precio	\$4100	\$2700	\$1700

Tarjetas de Video (Nvidia)



Tarjetas Video (Nvidia)

Familias

GeForce

Video Juegos

Quadro

Diseño gráfico y visualización (CAD software)

Tesla

Computación Científica

Modelos Tesla

C2050/C2070/C2075/M2050/M2070/M2070Q/M2090/K10/K20/K20X/K40

C20../M20..

Arquitectura Fermi. Enfocado en el rendimiento.

K10/ K20/ K20X/ K40

Arquitectura Kepler. Enfocado en rendimiento, eficiencia y programabilidad.

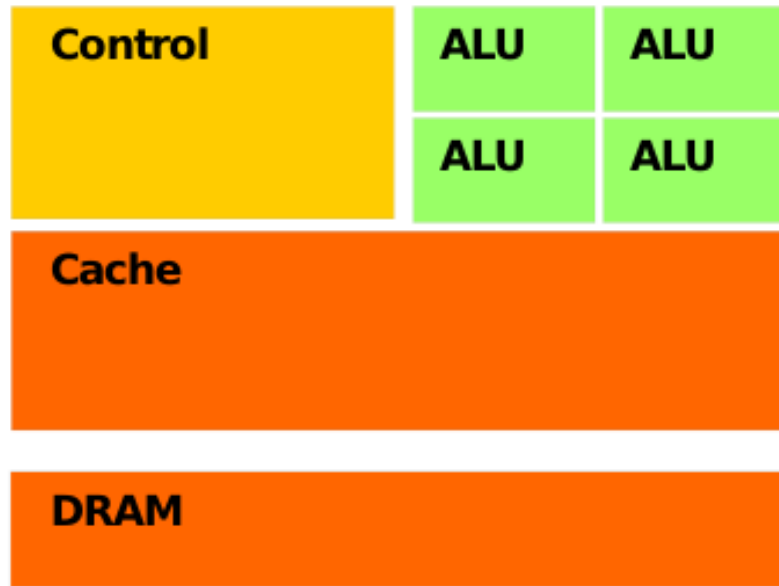
Nvidia Tesla

Features	Tesla K40	Tesla K20X	Tesla K20	Tesla K10
Number and Type of GPU	1 Kepler GK110B	1 Kepler GK110		2 Kepler GK104s
Peak double precision floating point performance	1.43 Tflops	1.31 Tflops	1.17 Tflops	0.19 Tflops
Peak single precision floating point performance	4.29 Tflops	3.95 Tflops	3.52 Tflops	4.58 Tflops
Memory bandwidth (ECC off)	288 GB/sec	250 GB/sec	208 GB/sec	320 GB/sec
Memory size (GDDR5)	12 GB	6 GB	5 GB	8 GB
<u>CUDA</u> cores	2880	2688	2496	2 x 1536

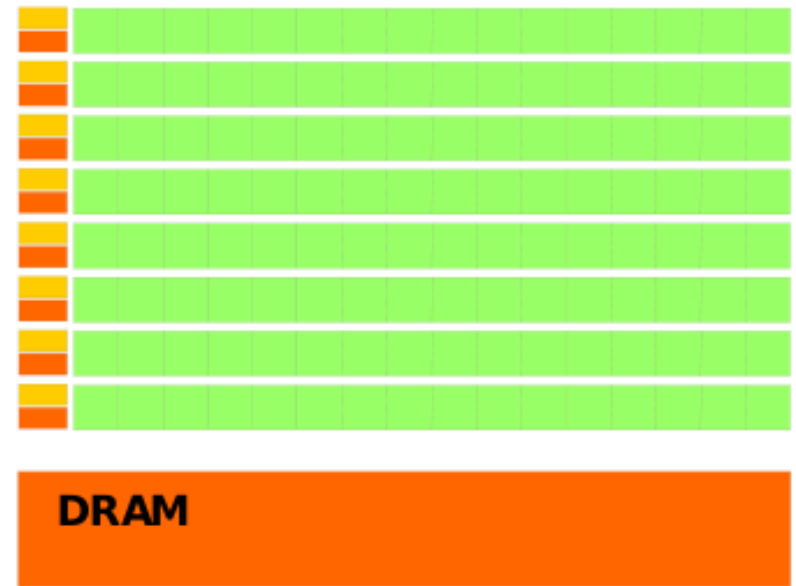
Tesla K20X

Chip	GK110
Processor clock	732 MHz
Memory clock	2.6 GHz
Memory size	6 GB
Memory I/O	384-bit GDDR5
Memory configuration	24 pieces of 64M ×16 GDDR5 SDRAM
Display connectors	None

Tarjetas de Video



CPU



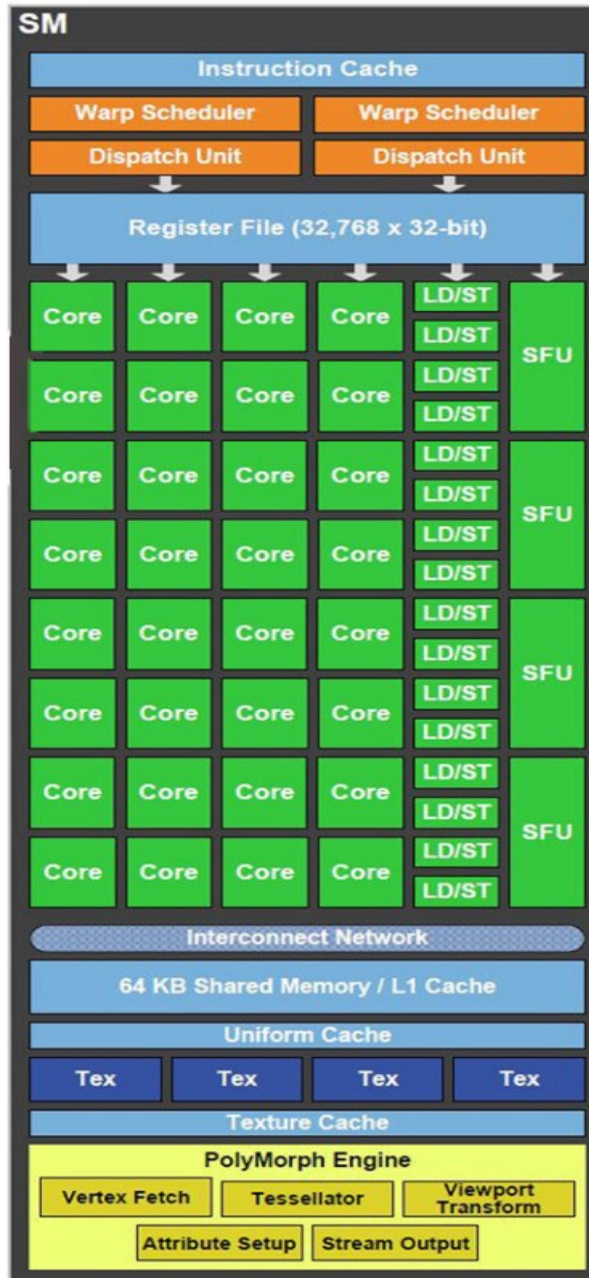
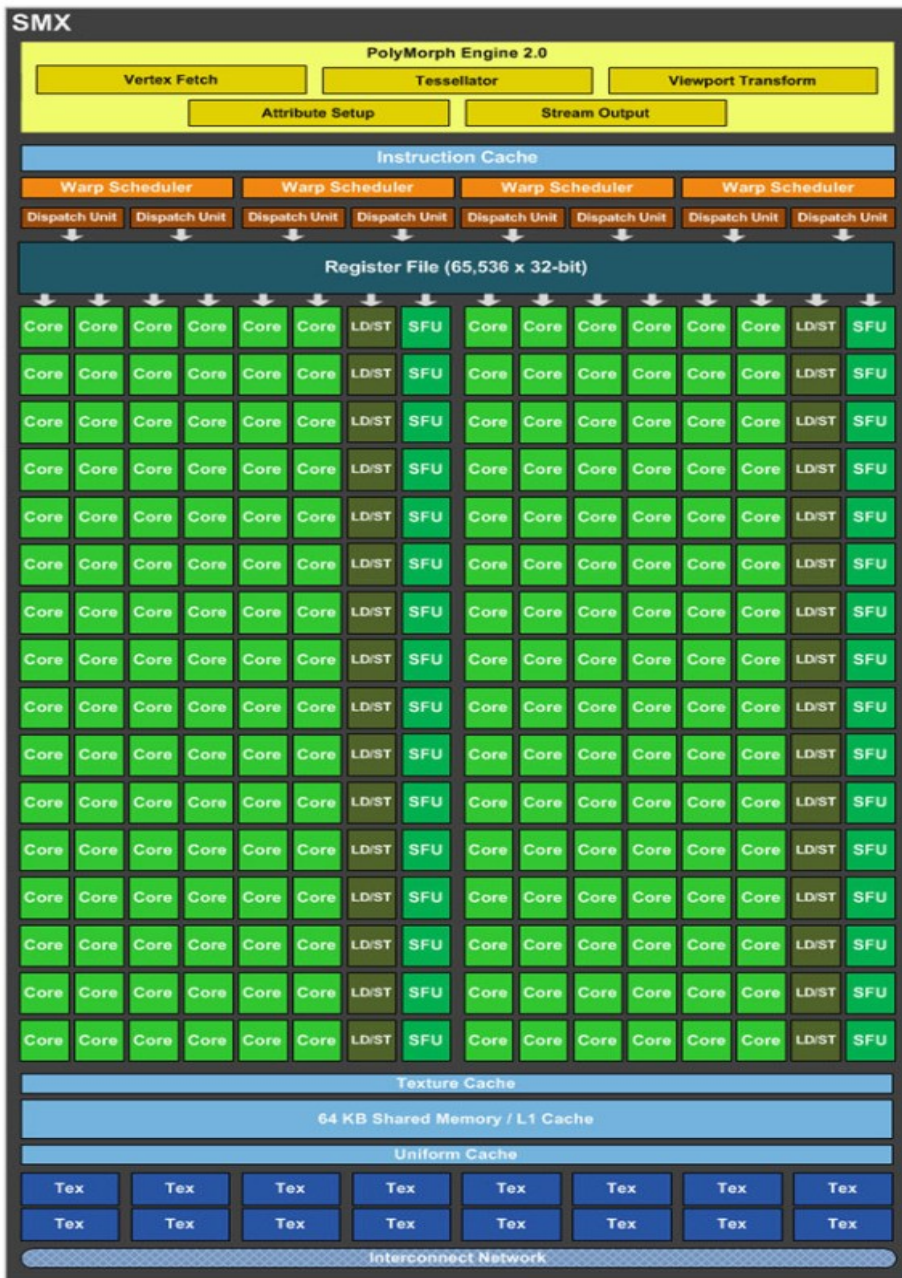
GPU

Nos enfocaremos en las tarjetas de video Nvidia, en particular la familia **Tesla**, diseñada para computación de alto desempeño.

Dentro de la familia Tesla hay dos tipos de arquitectura:

Fermi y Kepler

“Stream Multiprocessors”



Comparación
entre Fermi y
Kepler

SM: 32 Cuda cores
SMX: 192 Cuda cores

Intel Xeon vs. Intel Xeon Phi vs. NVIDIA Tesla GPU

	Xeon E5-2670	Xeon Phi 5110P	Tesla K20X
Cores	8	60	14 SMX
Logical Cores	16	240	2,688 CUDA cores
Frequency	2.60GHz	1.053GHz	735MHz
Memory	~16-128GB	8GB	6GB
Price	\$1556	\$2700	\$3350

Benchmark

	Xeon E5-2670	Xeon Phi 5110P	Tesla K20X
GFLOPs (double)	333	1,010	1,317

<http://blog.xcelerit.com/intel-xeon-phi-vs-nvidia-tesla-gpu/>

Paralelización

<i>Hardware</i>	Instrucciones Individuales	Instrucciones Múltiples
Datos Múltiples (MD)	Tarjetas de Video	Múltiples CPU + Memoria Compartida
	Coprocesadores	Múltiples CPU + Memoria Distribuida

<i>Software</i>	Instrucciones Individuales	Instrucciones Múltiples
Datos Múltiples (MD)	CUDA	OpenMP
	OpenMP + opciones Intel	MPI

Rank	Site	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	National Super Computer Center in Guangzhou China	Tianhe-2 (MilkyWay-2) - 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	17,808
2	DOE/SC/Oak Ridge National Laboratory United States	Titan - Cray XK7 , Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x Cray Inc.	560,640	17,590.0	27,112.5	8,209
3	DOE/NNSA/LLNL United States	Sequoia - BlueGene/Q Power BQC 16C 1.60 GHz, Custom IBM	1,572,864	17,173.2	20,132.7	7,890
4	RIKEN Advanced Institute for Computational Science (AICS) Japan	K computer SPARC64 VIIIfx 2.0GHz, Tofu interconnect Fujitsu	705,024	10,510.0	11,280.4	12,660
5	DOE/SC/Argonne National Laboratory United States	Mira - BlueGene/Q Power BQC 16C 1.60GHz, Custom IBM	786,432	8,586.6	10,066.3	3,945
6	Swiss National Supercomputing Centre (CSCS) Switzerland	Piz Daint - Cray XC30, Xeon E5-2670 8C 2.600GHz, Aries interconnect, NVIDIA K20x Cray Inc.	115,984	6,271.0	7,788.9	2,325
7	Texas Advanced Computing Center/Univ. of Texas United States	Stampede - 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
8	Forschungszentrum Juelich (FZJ) Germany	JUQUEEN - BlueGene/Q Power BQC 16C 1.600GHz, Custom Interconnect IBM	458,752	5,008.9	5,872.0	2,301
9	DOE/NNSA/LLNL United States	Vulcan - BlueGene/Q Power BQC 16C 1.600GHz, Custom Interconnect IBM	393,216	4,293.3	5,033.2	1,972
10	Leibniz Rechenzentrum Germany	SuperMUC - iDataPlex DX360M4 Xeon E5-2680 8C 2.70GHz, Infiniband FDR IBM	147,456	2,897.0	3,185.1	3,423

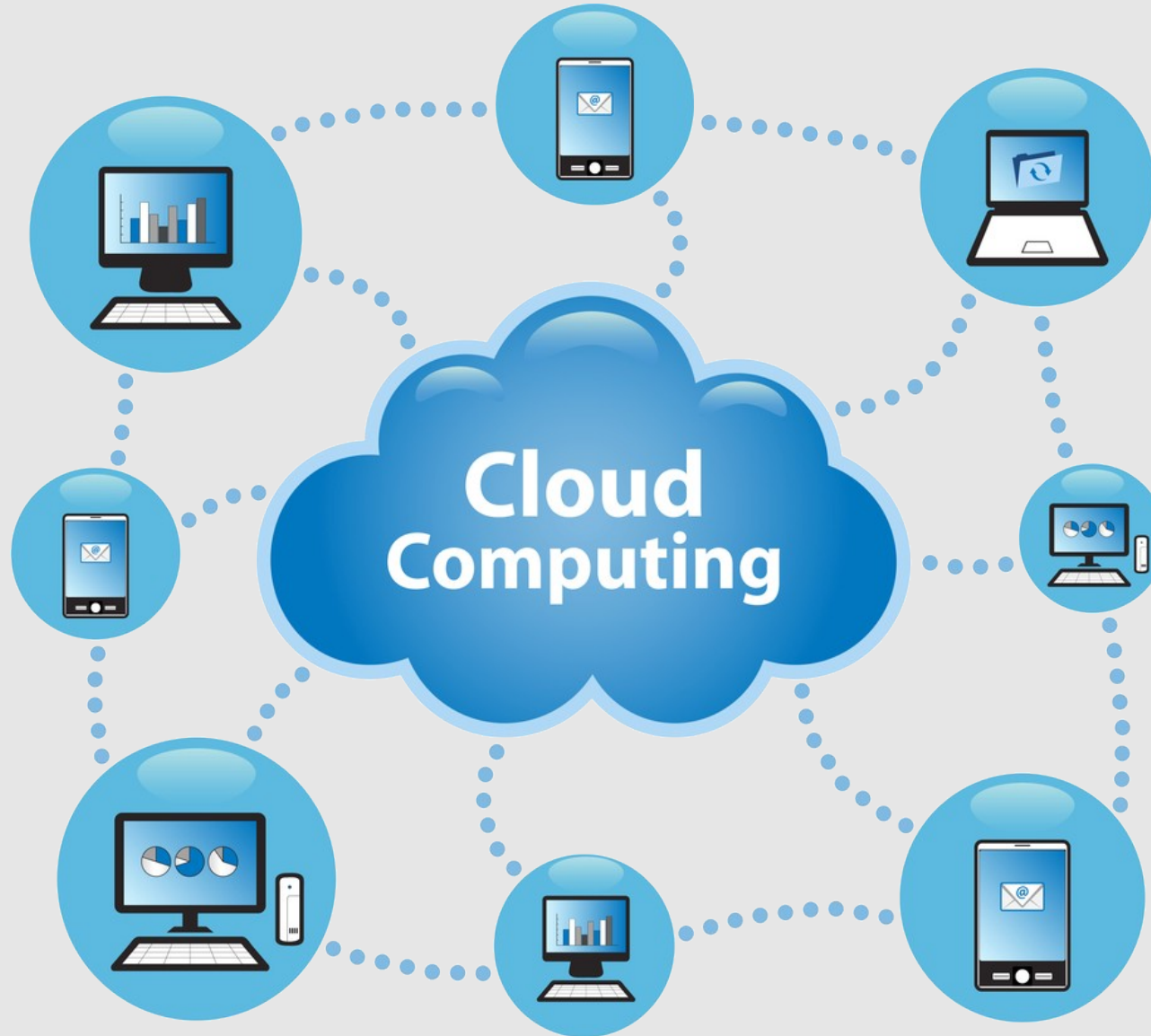
The Green500 List

Listed below are the November 2014 The Green500's energy-efficient supercomputers ranked from 1 to 10.

Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)
1	5,271.81	GSI Helmholtz Center	L-CSC - ASUS ESC4000 FDR/G2S Intel Xeon E5-2690v2 10C 3GHz, Infiniband FDR, AMD FirePro S9150 Level 1 measurement data available	57.15
2	4,945.63	High Energy Accelerator Research Organization /KEK	Suiren - ExaScaler 32U256SC Cluster Intel Xeon E5-2660v2 10C 2.2GHz, Infiniband FDR, PEZY-SC	37.83
3	4,447.58	GSIC Center, Tokyo Institute of Technology	TSUBAME-KFC - LX 1U-4GPU/104Re-1G Cluster Intel Xeon E5-2620v2 6C 2.100GHz, Infiniband FDR, NVIDIA K20x	35.39
4	3,962.73	Cray Inc.	Storm1 - Cray CS-Storm Intel Xeon E5-2660v2 10C 2.2GHz, Infiniband FDR, Nvidia K40m Level 3 measurement data available	44.54
5	3,631.70	Cambridge University	Wilkes - Dell T620 Cluster Intel Xeon E5-2630v2 6C 2.600GHz, Infiniband FDR, NVIDIA K20	52.62
6	3,543.32	Financial Institution	iDataPlex DX360M4 Intel Xeon E5-2680v2 10C 2.800GHz, Infiniband, NVIDIA K20x	54.60
7	3,517.84	Center for Computational Sciences, University of Tsukuba	HA-PACS TCA - Cray CS300 Cluster Intel Xeon E5-2680v2 10C 2.800GHz, Infiniband QDR, NVIDIA K20x	78.77
8	3,459.46	SURFsara	Cartesius Accelerator Island - Bullx B515 cluster, Intel Xeon E5-2450v2 8C 2.5GHz, InfiniBand 4x FDR, Nvidia K40m	44.40
9	3,185.91	Swiss National Supercomputing Centre (CSCS)	Piz Daint - Cray XC300 Xeon E5-2670 8C 2.600GHz, Aries interconnect, NVIDIA K20x Level 3 measurement data available	1,753.66
10	3,131.06	ROMEIO HPC Center - Champagne-Ardenne	romeo - Bull R421-E3 Cluster Intel Xeon E5-2650v2 8C 2.600GHz, Infiniband FDR, NVIDIA K20x	81.41

* Performance data obtained from publicly available sources including TOP500

Computación en la Nube



Cluster en la Nube



Computación en la Nube

- Seguridad.
- Costos: "Se paga lo que se utiliza".
- Accesibilidad.
- Fácil prestación de servicios.

Google Compute Engine

Permite correr trabajos sobre máquinas virtuales montadas sobre la infraestructura de Google.

- Se puede escoger la máquina virtual que se ajuste a las necesidades propias (hasta 16 cores virtuales por máquina).

Sistemas operativos:

- Debian 6.0, 7.0
- CentOS 6.2

<https://cloud.google.com/products/compute-engine/>

Computación en la Nube

Amazon Elastic Compute Cloud (Amazon EC2),

“Is a web service that provides resizable compute capacity in the cloud. It is designed to make web-scale computing easier for developers. EC2 allows users to rent virtual computers on which to run their own computer applications.”

Sistemas operativos:

Amazon tiene su propia distribución de Linux (basada en RHEL). Además ofrece:

- RHEL(Red Hat Enterprise Linux)
- SLES (SUSE Linux Enterprise Server)
- Microsoft Windows Server® (2003 R2, 2008, 2008 R2 and 2012)

Hasta 32 cores virtuales por máquina virtual

Máquinas virtuales con tarjetas Nvidia:

- M2050 con 448 CUDA cores y 3GB de memoria.
- GK104 con 1,536 CUDA cores y 4GB de memoria.

<http://aws.amazon.com/ec2/>

¿Que es mejor?

Procesadores, tarjetas de video, coprocesadores, máquinas virtuales, ...

"Eso depende", depende del problema que se quiera solucionar.

Necesidades:

- Mucha memoria y procesamiento: *Cluster de procesadores + MPI.*
- Memoria y procesamiento "medio": *Un computador "poderoso" + OpenMP.*
- Poca memoria y mucho procesamiento:
 - *Una tarjeta de video Nvidia + CUDA C.*
 - *Un coprocesador intel + usar OpenMP + compilador de intel.*
- Un cluster por poco tiempo: *Comprar tiempo en internet.*

Problemas (sociales) con las Nuevas Tecnologías

Barreras

- **Falta de acceso.** Las instituciones de financiamiento necesitan un elemento "físico" para justificar los gastos.
- **Resistencia al cambio.** Viejas prácticas que no se quieren cambiar: desarrollo de software, lenguajes de programación, uso individual de recursos, ...
- **Falta de tiempo.** Las instituciones no consideran tiempo para que los profesores y estudiantes adquieran nuevos conocimientos.
- **Falta de capacitación.** Compra de equipos que nadie sabe usar.
- **Falta de soporte técnico.** No proveer soporte técnico constante implica pérdida de recursos.

Algunas Aplicaciones a Cómputo Científico

Dinámica molecular

“Accelerating molecular dynamic simulation on graphics processing units”, M. S. Friedrichs et al., J. Computational Chemistry, 30(6):864-72, 2009.

Simulaciones sísmicas

“GPU and CPU cooperation parallel visualisation for large seismic data”, K. Xie, IET Electronics Letters vol. 46 , no. 17, 2010 , pp 1196 – 1197.

Modelos climáticos

“GPU acceleration of numerical weather prediction”, J. Michalakes and M. Vachharajani, Paralle Processing Letters, Aug 2008, pp. 1-18.

Electromagnetismo

“How to Render FDTD Computations More Effective Using a Graphics Accelerator”, P. Sypek et al., IEEE Transactions on Magnetics, vol. 45 , no. 3, 2009 , pp. 1324 - 1327

Sistemas eléctricos de potencia

“SIMD-based Large-Scale Transient Stability Simulation on the Graphics Processing Units”, V. Dinavahi et al., IEEE Trans. on Power Systems, vol. 25, no. 3, 2010.

Cosmología y Astronomía

“Cosmological calculations on the GPU”, D. Bard, M. Bellis, M. T. Allen, H. Yepremyan, J. M. Kratochvil, Astronomy and Computing, vol 1, 2013.

“Adaptive mesh fluid simulations on GPU”, P. Wang, T. Abel, R. Kaehler, New Astronomy, vol 15, Issue 7, 2010.

Algunos ejemplos de problemas en física computacional

Problema de N-cuerpos

Considere N cuerpos bajo la interacción gravitacional. La aceleración sobre cada uno de los cuerpos esta dada por

$$\vec{a}(\vec{r}_j) = -G \sum_{i=1}^N m_i \frac{\vec{r}_j - \vec{r}_i}{(|\vec{r}_j - \vec{r}_i|^2 + \epsilon^2)^{3/2}},$$

Donde epsilon es una longitud de suavizado que se introduce para reducir los efectos de relajación binaria ficticia a la vez que trata de reducir el error sistemático en cálculo de fuerzas debido a la modificación de la fuerza.

Hidrodinámica

Las ecuaciones que rigen la dinámica de los fluidos son:

- La ecuación de continuidad

$$\frac{d\rho}{dt} + \rho \nabla \cdot \vec{v} = 0.$$

- La ecuación de conservación del momento para fluidos no viscosos

$$\frac{d\vec{v}}{dt} = -\frac{\nabla P}{\rho} + \vec{F},$$

donde \vec{F} codifica las fuerzas no hidrodinámicas sobre el sistema, tales como la fuerza gravitacional.

- La ecuación de conservación de la energía (para procesos adiabáticos)

$$\frac{du}{dt} = \frac{P}{\rho^2} \frac{d\rho}{dt} = -\frac{P}{\rho} \nabla \cdot \vec{v}.$$

- La ecuación de estado de gas ideal

$$P = (\gamma - 1)\rho u.$$

Hidrodinámica

Después de discretizar las ecuaciones hidrodinámicas se obtiene:
Una ecuación algebraica para la densidad que reemplaza a la ecuación de continuidad

$$\rho_a = \sum_b m_b W_{ab}.$$

Las ecuaciones que se deben integrar

$$\begin{aligned}\frac{d\vec{r}_a}{dt} &= \vec{v}_a, \\ \frac{d\vec{v}_a}{dt} &= - \sum_b m_b \left(\frac{P_a}{\rho_a^2} + \frac{P_b}{\rho_b^2} + \Pi_{ab} \right) \nabla_a W_{ab} + \vec{F}_a, \\ \frac{du_a}{dt} &= \sum_b m_b \left(\frac{P_a}{\rho_a^2} + \frac{P_b}{\rho_b^2} + \Pi_{ab} \right) \vec{v}_{ab} \cdot \nabla_a W_{ab},\end{aligned}$$

y finalmente la ecuación de estado

$$P_a = (\gamma - 1)\rho_a u_a.$$

Estas ecuaciones describen completamente la hidrodinámica del sistema.

Relatividad General

Consideremos las ecuaciones de Einstein-Yang-Mills en un espacio-tiempo esféricamente simétrico tipo Anti-de Sitter, con grupo de norma SU(2). La métrica es de la forma

$$ds^2 = L^2 \sec^2 \Psi \left[-d\tau^2 + d\Psi^2 + \sin^2 \Psi d\Omega^2 \right]$$

La ecuación de movimiento para el campo de Yang-Mills (ω) es

$$\partial_\tau^2 \omega - \partial_\psi^2 \omega + \frac{(\omega^2 - 1)\omega}{\sin^2 \psi} = 0$$

Esta ecuación se reescribe como ($\tau \rightarrow t$, $\psi \rightarrow x$)

$$\partial_t \omega = \pi$$

$$\partial_t \Sigma = \partial_x \pi$$

$$\partial_t \pi = \partial_x \Sigma - \frac{(\omega^2 - 1)\omega}{\sin^2 x}$$

Con $\Sigma = \partial_x \omega$

Relatividad General

Las derivadas espaciales se aproximan por medio del metodo de diferencias finitas

$$\partial_x f(t, x) = \frac{f(t, x + \Delta x) - f(t, x - \Delta x)}{2\Delta x}$$

Por lo tanto el sistema queda descrito sobre una malla discreta.

$$\partial_t \omega(t, x) = \pi(t, x)$$

$$\partial_t \Sigma(t, x) = \frac{\pi(t, x + \Delta x) - \pi(t, x - \Delta x)}{2\Delta x}$$

$$\partial_t \pi(t, x) = \frac{\Sigma(t, x + \Delta x) - \Sigma(t, x - \Delta x)}{2\Delta x} - \frac{(\omega^2(t, x) - 1)\omega(t, x)}{\sin^2 x}$$

Cosmología Cuántica de Lazos

La **Gravedad Cuántica de Lazos** es una teoría que trata de cuantizar la gravedad, entre sus resultados mas significativos esta la descripción estadística de la entropía de los agujeros negros y el espectro discreto de los operadores de volumen y area, que nos hablan de un espacio-tiempo discreto.

La **Cosmología Cuántica de Lazos** es un modelo reducido que utiliza las tecnicas de gravedad cuántica de lazos, con el fin de cuantizar modelos cosmológicos. Entre los resultados mas significativos esta la resolución de la singularidad del big bang, la cual es reemplazada por un rebote cuántico cuando la densidad del universo es del orden de la densidad de Planck.

Por ejemplo, para el modelo que vamos a exponer, la acción del operador de volumen sobre los estados esta dada por

$$\hat{V}|v\rangle = \left(\frac{8\pi\gamma}{6}\right)^{3/2} \frac{|v|}{K} l_{\text{Pl}}^3 |v\rangle, \quad K = \frac{2}{3\sqrt{3\sqrt{3}}},$$

Con $\gamma \approx 0.2375$ el parámetro de Barbero-Immirzi.

Cosmología Cuántica de Lazos

Consideremos un universo homogéneo, isotrópico, plano, con contenido de materia un campo escalar sin masa. La función de onda de este universo esta descrita por la ecuación de evolución

$$\frac{\partial^2}{\partial \phi^2} \Psi(v, \phi) = -\hat{\Theta} \Psi(v, \phi)$$

$$\hat{\Theta} := -\frac{1}{B(v)} \left[C^+(v) \Psi(v+4, \phi) + C^0(v) \Psi(v, \phi) + C^-(v) \Psi(v-4, \phi) \right]$$

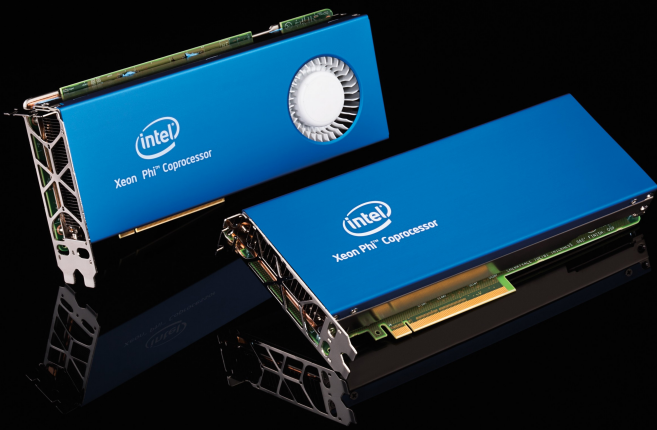
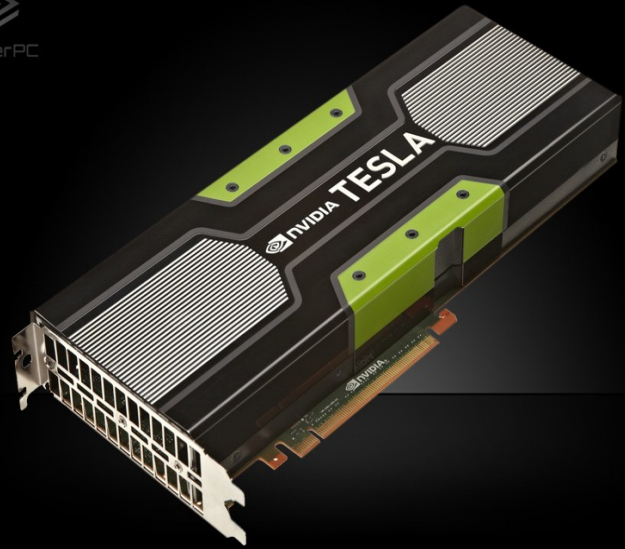
Donde los coeficientes son

$$C^+(v) = \frac{3\pi K G}{8} |v+2| \left(|v+1| - |v+3| \right),$$

$$C^-(v) = C^+(v-4) = \frac{3\pi K G}{8} |v-2| \left(|v-3| - |v-1| \right),$$

$$C^0(v) = -C^+(v) - C^-(v).$$

$$B(v) = \frac{27}{8} K |v| \left(|v+1|^{1/3} - |v-1|^{1/3} \right)^3$$



Gracias !!!

