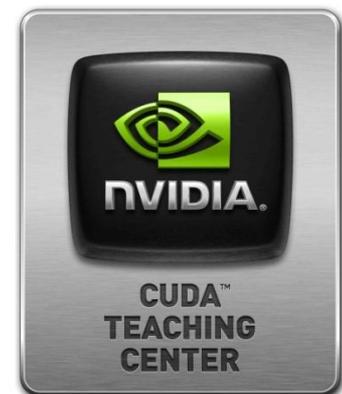


Programación masivamente paralela en procesadores gráficos (GPUs)

E. Dufrechou, M. Freire, P. Ezzatti y M. Pedemonte



Clase 1

Introducción

Contenido

- **Motivación**
- **Curso de GPGPU**
- **Un poco de historia...**
- **El pipeline gráfico**
- **Tarjetas programables**
- **¿Por qué una arquitectura unificada?**
- **CUDA**
- **Comparación de la performance entre CPU vs GPU**

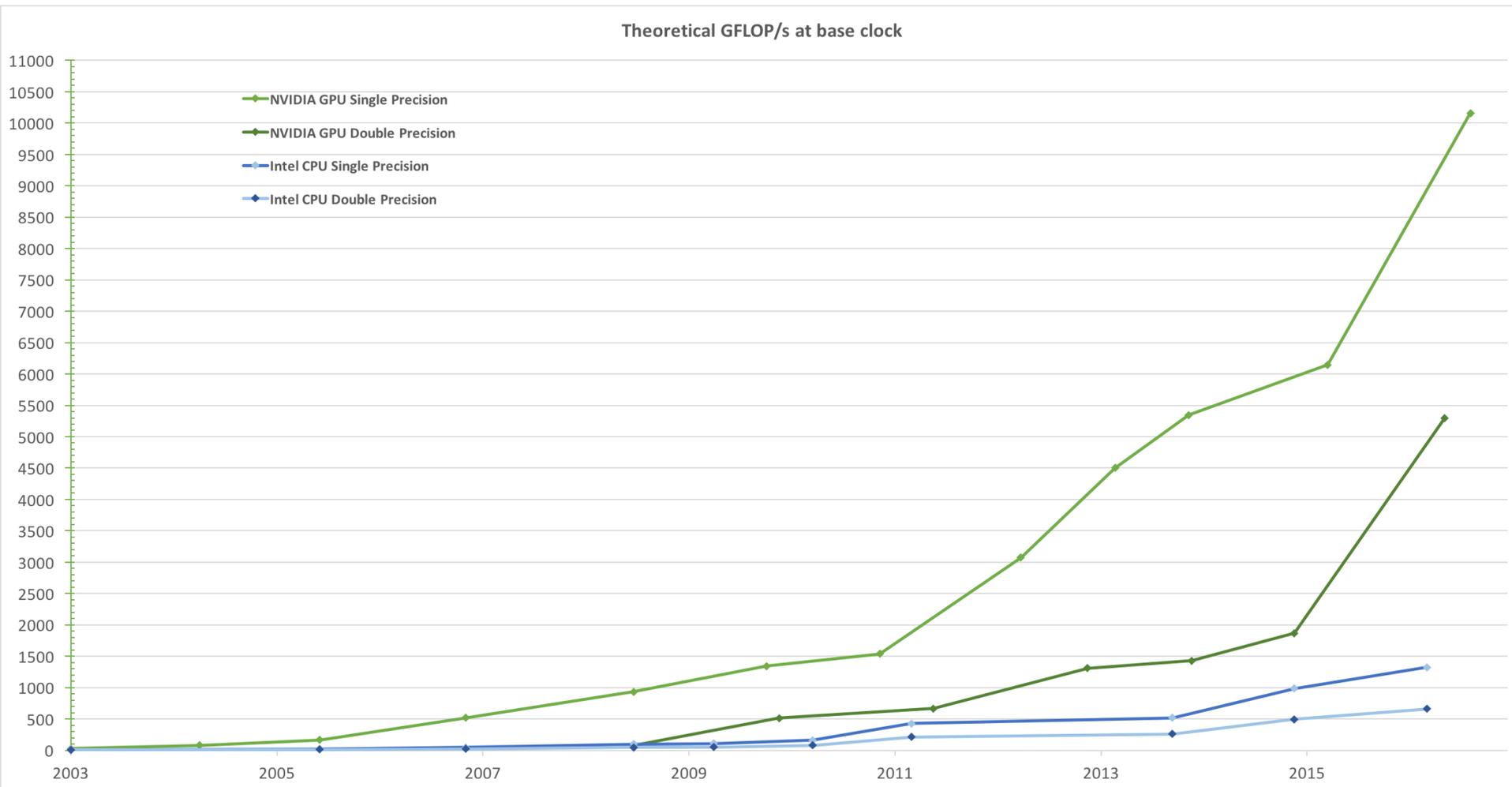
Motivación

Motivación

- **Utilización de hardware secundario para acelerar cómputos.**
- **Usar la tarjeta gráfica (GPU) para resolver problemas de propósito general (GPGPU).**
- **Se sustenta en el fuerte desarrollo tecnológico de las GPUs debido a la gran presión de la industria de los videojuegos.**
- **Surgimiento de lenguajes de programación de propósito general para programar las GPUs como CUDA y OpenCL.**

Motivación

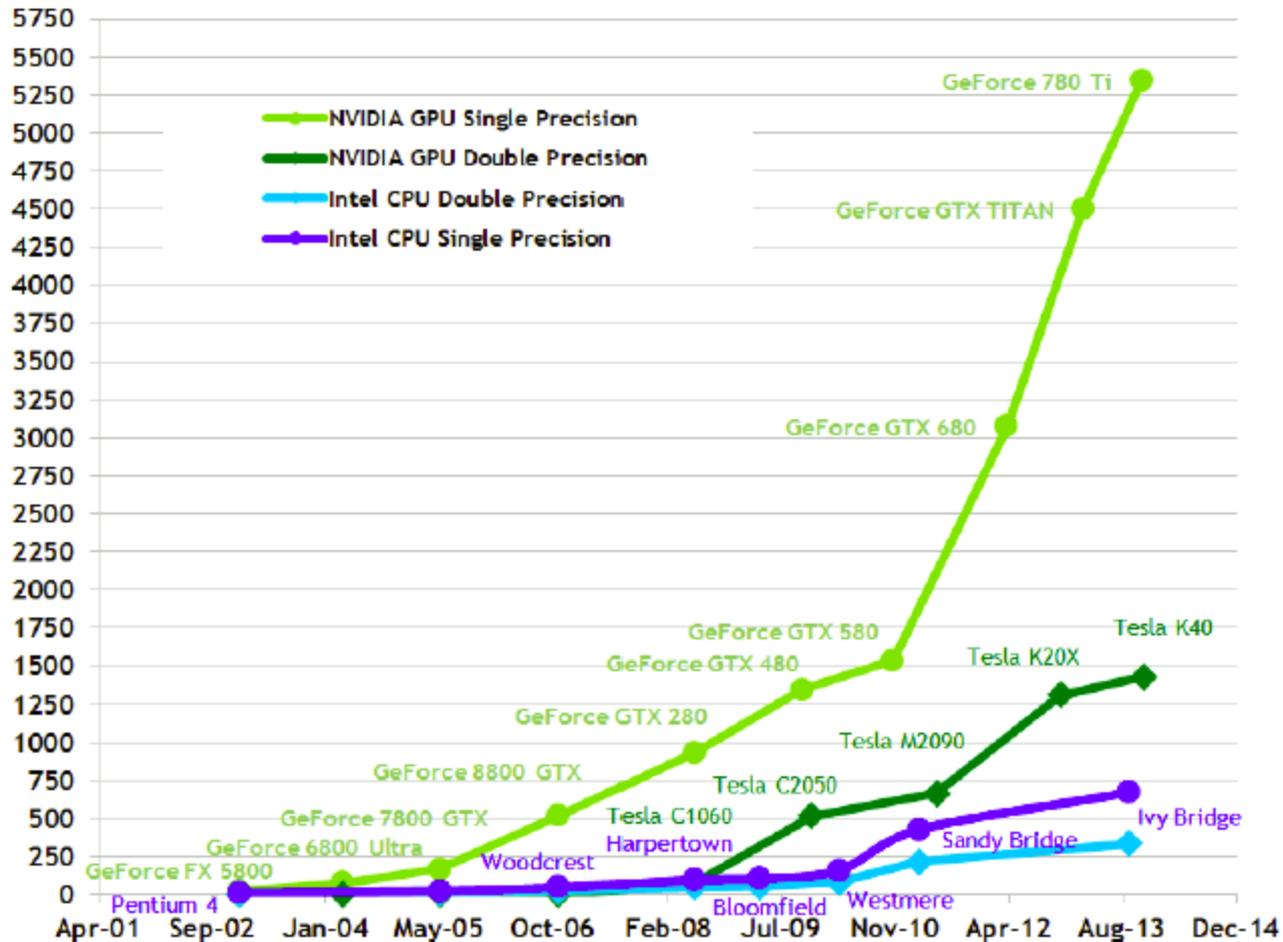
Incremento significativo del poder de cómputo de las GPUs.



Motivación

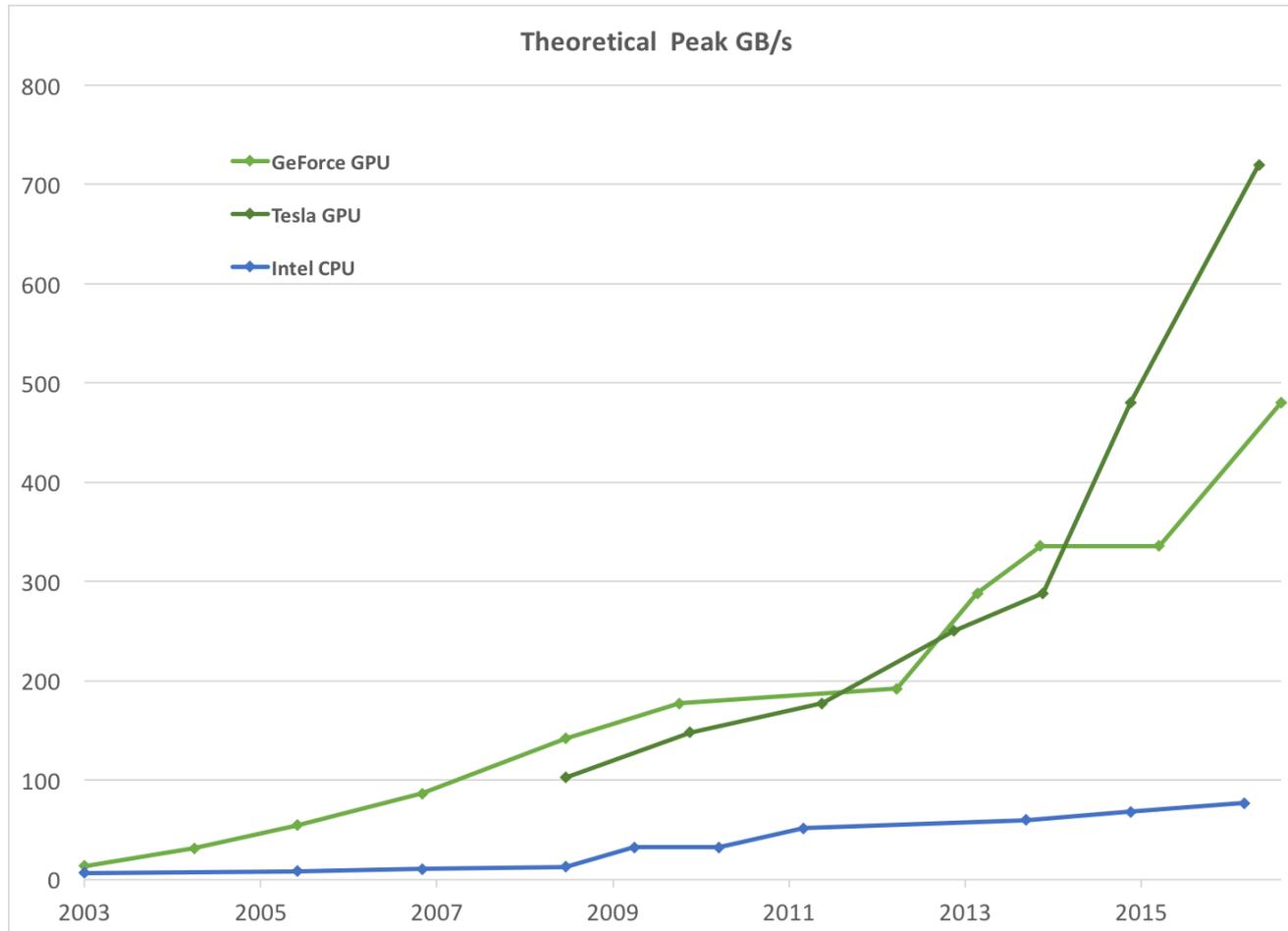
Incremento significativo del poder de cómputo de las GPUs.

Theoretical GFLOP/s



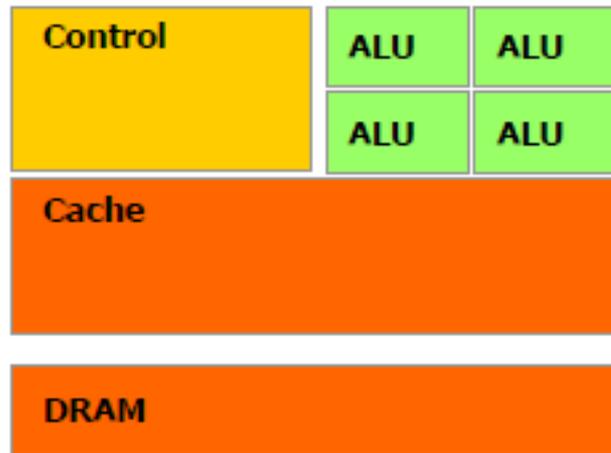
Motivación

Incremento significativo en el ancho de banda de acceso a memoria

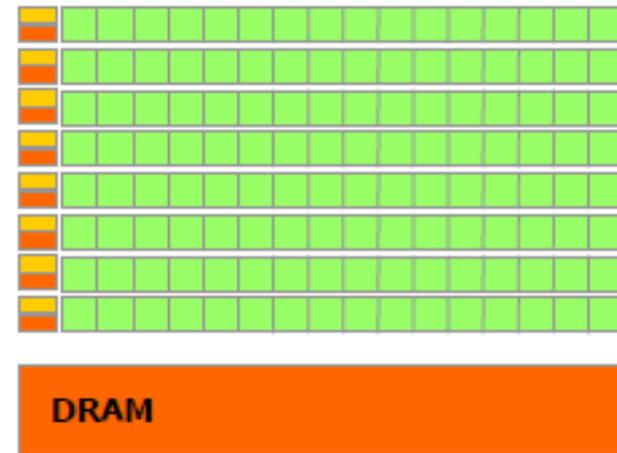


Motivación

Arquitectura intrínsecamente paralela (multiprocesadores)



CPU

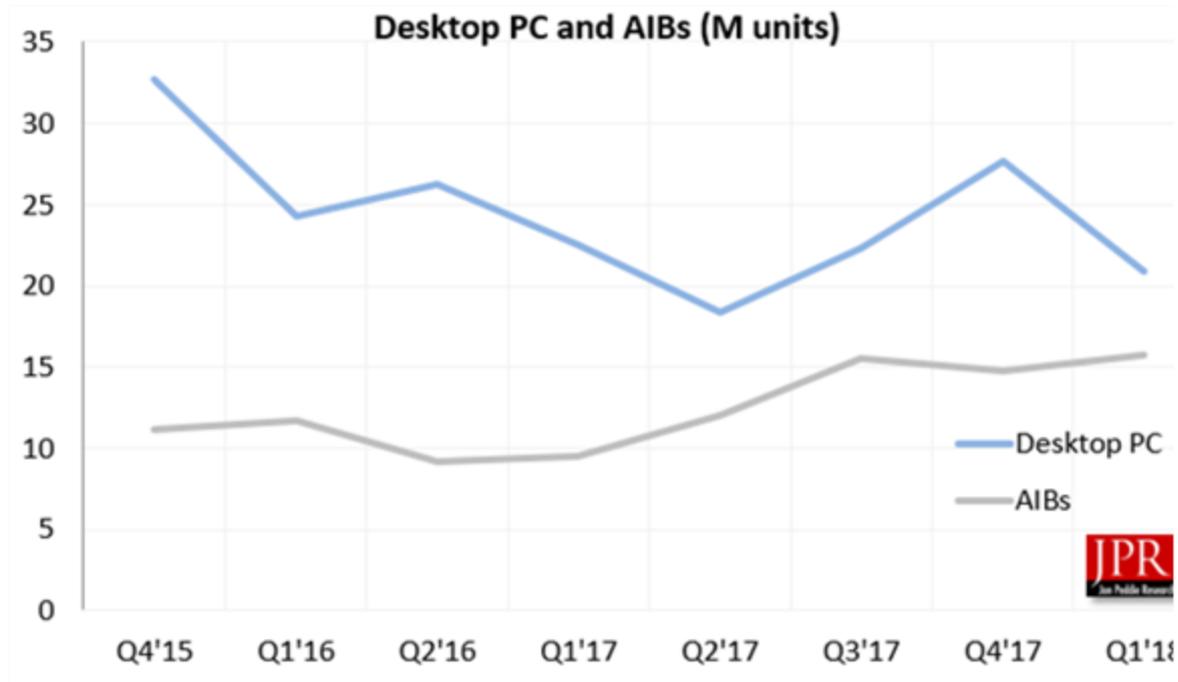


GPU

Motivación

- **Es una plataforma ampliamente disponible.**
- **Hoy en día, una gran cantidad de laptops y desktops tienen disponible una GPU de Nvidia o AMD.**
- **Los celulares y tablets tiene disponibles GPUs (si bien no son exactamente iguales).**
- **Representan una tecnología que puede incorporarse tanto a un laptop como a una supercomputadora.**

Motivación



| GPU supplier | Market share this quarter | Market share last quarter | Market share last year |
|--------------|---------------------------|---------------------------|------------------------|
| AMD | 34.9% | 33.7% | 27.5% |
| Nvidia | 65.1% | 66.3% | 72.5% |
| Total | 100% | 100% | 100% |

Motivación

- **Es una tecnología barata.**
- **Tesla P100 (Q2 2016):**
 - **Costo aproximado U\$S 5900 y U\$S 6600 en Amazon.com**
 - **Tiene 3584 CUDA cores**
 - **Pico de performance en half precision 18.7 TFlops**
 - **Pico de performance en simple precisión 9.3 TFlops**
 - **Pico de performance en doble precisión 4.7 Tflops**
 - **37.36 Gflops/W en simple**
 - **Ancho de banda de memoria de la GPU 720 GB/s**

Motivación

- **Es una tecnología barata.**
- **Tesla V100 (Q2 2017):**
 - **Costo aproximado U\$S 6600 y U\$S 7900 en Amazon.com**
 - **Tiene 5120 CUDA cores**
 - **Pico de performance en half precision 28 Tflops**
 - **Pico de performance Tensor 112 Tflops**
 - **Pico de performance en simple precisión 14 TFlops**
 - **Pico de performance en doble precisión 7 Tflops**
 - **46.67 Gflops/W en simple**
 - **Ancho de banda de memoria de la GPU 900 GB/s**

Motivación

- **Es una tecnología barata.**
- **GeForce GTX 1080 Ti (Q1 2017):**
 - **Costo aproximado U\$S 800 en Amazon.com**
 - **Tiene 3584 CUDA cores**
 - **Pico de performance en half precision 166 Gflops (1/64 de simple)**
 - **Pico de performance en simple precisión 10.6 TFlops**
 - **Pico de performance en doble precisión 332 Gflops (1/32 de simple)**
 - **42.4 Gflops/W en simple**
 - **Ancho de banda de memoria de la GPU 484 GB/s**

Motivación

"Las GPUs han evolucionado hasta un punto en que muchas de las aplicaciones industriales actuales se ejecutan en ellas con niveles de rendimiento muy superiores a los que ofrecerían si se ejecutasen en sistemas multinúcleo. Las arquitecturas informáticas del futuro serán sistemas híbridos con GPUs compuestas por núcleos de procesamiento paralelo que trabajarán en colaboración con las CPUs multinúcleo".

Prof. Jack Dongarra

Director del Innovative Computing Laboratory

Universidad de Tennessee

Motivación

- **Existe una apuesta fuerte de Nvidia por la adopción de GPUs para la computación de alta performance.**
- **Cientos de instituciones educativas enseñando GPGPU.**
- **Desarrollo de una comunidad científica y conferencias.**
- **Top 500:**
 - **Las 500 computadoras más poderosas del mundo.**
 - **<http://www.top500.org/>**
 - **Revisemos los 10 primeros puestos**

Motivación

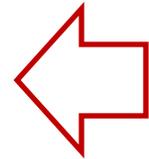
| Rank | Site | System | Cores | Rmax (TFlop/s) | Rpeak (TFlop/s) | Power (kW) |
|------|---|---|------------|----------------|-----------------|------------|
| 1 | National Supercomputing Center in Wuxi China | Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway NRPC | 10,649,600 | 93,014.6 | 125,435.9 | 15,371 |
| 2 | 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 | 3,120,000 | 33,862.7 | 54,902.4 | 17,808 |
| 3 | 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 | 27,112.3 | 8,908 |
| 4 | 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 |
| 5 | 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 |
| 6 | 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 |
| 7 | DOE/NNSA/LANL/SNL United States | Trinity - Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Aries interconnect Cray Inc. | 301,056 | 8,100.9 | 11,078.9 | |
| 8 | 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 |
| 9 | HLRS - Höchstleistungsrechenzentrum Stuttgart Germany | Hazel Hen - Cray XC40, Xeon E5-2680v3 12C 2.5GHz, Aries interconnect Cray Inc. | 185,088 | 5,648.2 | 7,403.5 | |
| 10 | King Abdullah University of Science and Technology Saudi Arabia | Shaheen II - Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Aries interconnect Cray Inc. | 196,608 | 5,537.0 | 7,235.2 | 2,834 |

Top 500
Junio 2016



Motivación

| | | | | | |
|----|--|------------|-----------|-----------|--------|
| 1 | Supercomputer Fugaku - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan | 7,299,072 | 415,530.0 | 513,854.7 | 28,335 |
| 2 | Summit - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM DOE/SC/Oak Ridge National Laboratory United States | 2,414,592 | 148,600.0 | 200,794.9 | 10,096 |
| 3 | Sierra - IBM Power System AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States | 1,572,480 | 94,640.0 | 125,712.0 | 7,438 |
| 4 | Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway, NRCPC National Supercomputing Center in Wuxi China | 10,649,600 | 93,014.6 | 125,435.9 | 15,371 |
| 5 | Tianhe-2A - TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000, NUDT National Super Computer Center in Guangzhou China | 4,981,760 | 61,444.5 | 100,678.7 | 18,482 |
| 6 | HPC5 - PowerEdge C4140, Xeon Gold 6252 24C 2.1GHz, NVIDIA Tesla V100, Mellanox HDR Infiniband, Dell EMC Eni S.p.A. Italy | 669,760 | 35,450.0 | 51,720.8 | 2,252 |
| 7 | Setene - DGX A100 SuperPOD, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Mellanox HDR Infiniband, Nvidia NVIDIA Corporation United States | 277,760 | 27,580.0 | 34,566.6 | 1,344 |
| 8 | Frontera - Dell C6420, Xeon Platinum 8280 28C 2.7GHz, Mellanox InfiniBand HDR, Dell EMC Texas Advanced Computing Center/Univ. of Texas United States | 448,448 | 23,516.4 | 38,745.9 | |
| 9 | Marconi-100 - IBM Power System AC922, IBM POWER9 16C 3GHz, Nvidia Volta V100, Dual-rail Mellanox EDR Infiniband, IBM CINECA Italy | 347,776 | 21,640.0 | 29,354.0 | 1,476 |
| 10 | Piz Daint - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect, NVIDIA Tesla P100, Cray/HPE Swiss National Supercomputing Centre (CSCS) Switzerland | 387,872 | 21,230.0 | 27,154.3 | 2,384 |



Top 500
Junio 2020

<https://www.top500.org/lists/top500/2020/06/>

Motivación

Top 500
Junio 2023

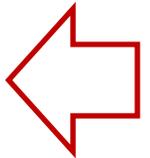
| Rank | System | Cores | Rmax (PFlop/s) | Rpeak (PFlop/s) | Power (kW) |
|------|---|-----------|----------------|-----------------|------------|
| 1 | Frontier - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE DOE/SC/Oak Ridge National Laboratory United States | 8,699,904 | 1,194.00 | 1,679.82 | 22,703 |
| 2 | Supercomputer Fugaku - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan | 7,630,848 | 442.01 | 537.21 | 29,899 |
| 3 | LUMI - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE EuroHPC/CSC Finland | 2,220,288 | 309.10 | 428.70 | 6,016 |
| 4 | Leonardo - BullSequana XH2000, Xeon Platinum 8358 32C 2.6GHz, NVIDIA A100 SXM4 64 GB, Quad-rail NVIDIA HDR100 Infiniband, EVIDEN EuroHPC/CINECA Italy | 1,824,768 | 238.70 | 304.47 | 7,404 |
| 5 | Summit - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM DOE/SC/Oak Ridge National Laboratory United States | 2,414,592 | 148.60 | 200.79 | 10,096 |

<https://www.top500.org/lists/top500/2023/06/>

Motivación

Top 500
Junio 2023

| | | | | | |
|----|---|------------|-------|--------|--------|
| 6 | Sierra - IBM Power System AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States | 1,572,480 | 94.64 | 125.71 | 7,438 |
| 7 | Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway, NRCPC National Supercomputing Center in Wuxi China | 10,649,600 | 93.01 | 125.44 | 15,371 |
| 8 | Perlmutter - HPE Cray EX235n, AMD EPYC 7763 64C 2.45GHz, NVIDIA A100 SXM4 40 GB, Slingshot-10, HPE DOE/SC/LBNL/NERSC United States | 761,856 | 70.87 | 93.75 | 2,599 |
| 9 | Selene - NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Mellanox HDR Infiniband, Nvidia NVIDIA Corporation United States | 555,520 | 63.46 | 79.22 | 2,646 |
| 10 | Tianhe-2A - TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000, NUDT National Super Computer Center in Guangzhou China | 4,981,760 | 61.44 | 100.68 | 18,482 |



<https://www.top500.org/lists/top500/2023/06/>

Motivación

Top 500
Junio 2024

| Rank | System | Cores | Rmax (PFlop/s) | Rpeak (PFlop/s) | Power (kW) |
|------|---|-----------|----------------|-----------------|------------|
| 1 | Frontier - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE DOE/SC/Oak Ridge National Laboratory United States | 8,699,904 | 1,206.00 | 1,714.81 | 22,786 |
| 2 | Aurora - HPE Cray EX - Intel Exascale Compute Blade, Xeon CPU Max 9470 52C 2.4GHz, Intel Data Center GPU Max, Slingshot-11, Intel DOE/SC/Argonne National Laboratory United States | 9,264,128 | 1,012.00 | 1,980.01 | 38,698 |
| 3 | Eagle - Microsoft NDv5, Xeon Platinum 8480C 48C 2GHz, NVIDIA H100, NVIDIA Infiniband NDR, Microsoft Azure Microsoft Azure United States | 2,073,600 | 561.20 | 846.84 | |
| 4 | Supercomputer Fugaku - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan | 7,630,848 | 442.01 | 537.21 | 29,899 |
| 5 | LUMI - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE EuroHPC/CSC Finland | 2,752,704 | 379.70 | 531.51 | 7,107 |

<https://top500.org/lists/top500/2024/06/>

Motivación

Top 500 Junio 2024

| | | | | | |
|----|---|-----------|--------|--------|--------|
| 6 | Alps - HPE Cray EX254n, NVIDIA Grace 72C 3.1GHz, NVIDIA GH200 Superchip, Slingshot-11, HPE Swiss National Supercomputing Centre (CSCS) Switzerland | 1,305,600 | 270.00 | 353.75 | 5,194 |
| 7 | Leonardo - BullSequana XH2000, Xeon Platinum 8358 32C 2.6GHz, NVIDIA A100 SXM4 64 GB, Quad-rail NVIDIA HDR100 Infiniband, EVIDEN EuroHPC/CINECA Italy | 1,824,768 | 241.20 | 306.31 | 7,494 |
| 8 | MareNostrum 5 ACC - BullSequana XH3000, Xeon Platinum 8460Y+ 32C 2.3GHz, NVIDIA H100 64GB, Infiniband NDR, EVIDEN EuroHPC/BSC Spain | 663,040 | 175.30 | 249.44 | 4,159 |
| 9 | Summit - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM DOE/SC/Oak Ridge National Laboratory United States | 2,414,592 | 148.60 | 200.79 | 10,096 |
| 10 | Eos NVIDIA DGX SuperPOD - NVIDIA DGX H100, Xeon Platinum 8480C 56C 3.8GHz, NVIDIA H100, Infiniband NDR400, Nvidia NVIDIA Corporation United States | 485,888 | 121.40 | 188.65 | |

<https://www.top500.org/lists/top500/2024/06/>

Motivación

- Desde la arquitectura Kepler uno de los objetivos del diseño ha sido la reducción del consumo energético.
- Se busca lograr aumentar la relación desempeño-consumo, es decir los Gflops efectivamente computados por Watt.
- **Green 500:**
 - Las 500 computadoras más eficientes energéticamente del mundo.
 - Se elabora usando las 500 computadoras que aparecen en el top 500.
 - <http://www.green500.org/>
 - Revisemos los 10 primeros puestos

Motivación

Green
Top 500
Junio 2016

| Green500 Rank | MFLOPS/W | Site* | Computer* | Total Power (kW) |
|---------------|----------|--|---|------------------|
| 1 | 6,673.84 | Advanced Center for Computing and Communication, RIKEN | Shoubu - ZettaScaler-1.6, Xeon E5-2618Lv3 8C 2.3GHz, Infiniband FDR, PEZY-SCnp | 149.99 |
| 2 | 6,195.22 | Computational Astrophysics Laboratory, RIKEN | Satsuki - ZettaScaler-1.6, Xeon E5-2618Lv3 8C 2.3GHz, Infiniband FDR, PEZY-SCnp | 46.89 |
| 3 | 6,051.30 | National Supercomputing Center in Wuxi | Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway | 15,371.00 |
| 4 | 5,272.09 | GSI Helmholtz Center | ASUS ESC4000 FDR/G2S, Intel Xeon E5-2690v2 10C 3GHz, Infiniband FDR, AMD FirePro S9150 | 57.15 |
| 5 | 4,778.46 | Institute of Modern Physics (IMP), Chinese Academy of Sciences | Sugon Cluster W780I, Xeon E5-2640v3 8C 2.6GHz, Infiniband QDR, NVIDIA Tesla K80 | 65.00 |
| 6 | 4,112.11 | Stanford Research Computing Center | XStream - Cray CS-Storm, Intel Xeon E5-2680v2 10C 2.8GHz, Infiniband FDR, Nvidia K80 | 190.00 |
| 7 | 3,775.45 | Internet Service (B) | Inspur TS10000 HPC Server, Intel Xeon E5-2620v2 6C 2.1GHz, 10G Ethernet, NVIDIA Tesla K40 | 110.00 |
| 8 | 3,775.45 | Internet Service (B) | Inspur TS10000 HPC Server, Intel Xeon E5-2620v2 6C 2.1GHz, 10G Ethernet, NVIDIA Tesla K40 | 110.00 |
| 9 | 3,775.45 | Internet Service (B) | Inspur TS10000 HPC Server, Intel Xeon E5-2620v2 6C 2.1GHz, 10G Ethernet, NVIDIA Tesla K40 | 110.00 |
| 10 | 3,775.45 | Internet Service (B) | Inspur TS10000 HPC Server, Intel Xeon E5-2620v2 6C 2.1GHz, 10G Ethernet, NVIDIA Tesla K40 | 110.00 |

Motivación

| Rank | TOP500 Rank | System | Cores | Rmax (PFlop/s) | Power (kW) | Energy Efficiency (GFlops/watts) |
|------|-------------|--|-----------|----------------|------------|----------------------------------|
| 1 | 27 | Frontier TDS - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE DOE/SC/Oak Ridge National Laboratory United States | 120,832 | 19.20 | 307 | 62.684 |
| 2 | 1 | Frontier - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE DOE/SC/Oak Ridge National Laboratory United States | 8,730,112 | 1,102.00 | 21,100 | 52.227 |
| 3 | 3 | LUMI - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE EuroHPC/CSC Finland | 1,110,144 | 151.90 | 2,942 | 51.629 |
| 4 | 10 | Adastra - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE Grand Equipement National de Calcul Intensif - Centre Informatique National de l'Enseignement Supérieur (GENCI-CINES) France | 319,072 | 46.10 | 721 | 50.028 |
| 5 | 146 | ATOS THX.A.B - BullSequana XH2000, Xeon Platinum 8358 32C 2.6GHz, NVIDIA A100 64GB, Infiniband HDR, Atos Atos France | 25,056 | 3.50 | 86 | 41.211 |

Green
Top 500
Junio 2022

<https://www.top500.org/lists/green500/2022/06/>

Motivación

| | | | | | | |
|----|-----|--|--------|------|-----|--------|
| 6 | 326 | MN-3 - MN-Core Server, Xeon Platinum 8260M 24C 2.4GHz, Preferred Networks MN-Core, MN-Core DirectConnect, Preferred Networks Preferred Networks Japan | 1,664 | 2.18 | 53 | 40.901 |
| 7 | 315 | SSC-21 Scalable Module - Apollo 6500 Gen10 plus, AMD EPYC 7543 32C 2.8GHz, NVIDIA A100 80GB, Infiniband HDR200, HPE Samsung Electronics South Korea | 16,704 | 2.27 | 100 | 33.983 |
| 8 | 319 | Tethys - NVIDIA DGX A100 Liquid Cooled Prototype, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100 80GB, Infiniband HDR, Nvidia NVIDIA Corporation United States | 19,840 | 2.25 | 72 | 31.538 |
| 9 | 304 | Wilkes-3 - PowerEdge XE6945, AMD EPYC 7763 64C 2.45GHz, NVIDIA A100 80GB, Infiniband HDR200 dual rail, DELL EMC University of Cambridge United Kingdom | 26,880 | 2.29 | 74 | 30.797 |
| 10 | 105 | Athena - FormatServer THOR ERG21, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100 SXM4 40 GB, Infiniband HDR, Format sp. z o.o. Cyfronet Poland | 47,616 | 5.05 | 147 | 29.926 |

Green
Top 500
Junio 2022

<https://www.top500.org/lists/green500/2022/06/>

Motivación

| TOP500 | | | | | | |
|--------|------|---|---------|----------------|------------|----------------------------------|
| Rank | Rank | System | Cores | Rmax (PFlop/s) | Power (kW) | Energy Efficiency (GFlops/watts) |
| 1 | 255 | Henri - ThinkSystem SR670 V2, Intel Xeon Platinum 8362 32C 2.8GHz, NVIDIA H100 80GB PCIe, Infiniband HDR, Lenovo Flatiron Institute United States | 8,288 | 2.88 | 44 | 65.396 |
| 2 | 34 | Frontier TDS - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE DOE/SC/Oak Ridge National Laboratory United States | 120,832 | 19.20 | 309 | 62.166 |
| 3 | 12 | Alusara - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE Grand Equipement National de Calcul Intensif - Centre Informatique National de l'Enseignement Suprieur (GENCI-CINES) France | 318,022 | 16.18 | 921 | 58.021 |
| 4 | 17 | Setonix - GPU - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE Pawsey Supercomputing Centre, Kensington, Western Australia Australia | 181,248 | 27.16 | 477 | 56.983 |
| 5 | 77 | Dardel GPU - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE KTH - Royal Institute of Technology Sweden | 52,864 | 8.26 | 146 | 56.491 |

Green
Top 500
Junio 2023

<https://www.top500.org/lists/green500/2023/06/>

Motivación



| | | | | | | |
|----|-----|--|-----------|----------|--------|--------|
| 6 | 1 | Frontier - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE DOE/SC/Oak Ridge National Laboratory United States | 8,699,904 | 1,194.00 | 22,706 | 52.592 |
| 7 | 3 | LUMI - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE EuroHPC/CSC Finland | 2,220,268 | 309.10 | 6,018 | 51.382 |
| 8 | 483 | amplitUDE (GPU Partition) - MEGWARE D50DNP, Xeon Platinum 8480+ 56C 2GHz, NVIDIA H100 80GB PCIe, Infiniband NDR, MEGWARE University of Duisburg-Essen Germany | 6,768 | 1.95 | 38 | 51.343 |
| 9 | 70 | Goethe-NHR - Supermicro AS-4124GS-TNR, AMD EPYC 7452 32C 2.35GHz, AMD Instinct MI210 64 GB, Mellanox InfiniBand EDR, MEGWARE / Supermicro Universitaet Frankfurt Germany | 96,768 | 9.09 | 196 | 46.543 |
| 10 | 187 | ATOS THX.A.B - BullSequana XH2000, Xeon Platinum 8358 32C 2.6GHz, NVIDIA A100 SXM4 64 GB, Quad-raid NVIDIA HDR100 Infiniband, EVIDEN Atos France | 25,056 | 3.50 | 86 | 41.411 |

Green
Top 500
Junio 2023

<https://www.top500.org/lists/green500/2023/06/>

Motivación

| Rank | TOP500 Rank | System | Cores | Rmax (PFlop/s) | Power (kW) | Energy Efficiency (GFlops/watts) |
|------|-------------|--|--------|----------------|------------|----------------------------------|
| 1 | 189 | JEDI - BullSequana XH3000, Grace Hopper Superchip 72C 3GHz, NVIDIA GH200 Superchip, Quad-Rail NVIDIA InfiniBand NDR200, ParTec/EVIDEN EuroHPC/FZJ Germany | 19,584 | 4.50 | 67 | 72.783 |
| 2 | 128 | Isambard-AI phase 1 - HPE Cray EX254n, NVIDIA Grace 72C 3.1GHz, NVIDIA GH200 Superchip, Slingshot-11, HPE University of Bristol United Kingdom | 34,272 | 7.42 | 117 | 68.835 |
| 3 | 55 | Helios GPU - HPE Cray EX254n, NVIDIA Grace 72C 3.1GHz, NVIDIA GH200 Superchip, Slingshot-11, HPE Cyfronet Poland | 89,760 | 19.14 | 317 | 66.948 |
| 4 | 328 | Henri - ThinkSystem SR670 V2, Intel Xeon Platinum 8362 32C 2.8GHz, NVIDIA H100 80GB PCIe, Infiniband HDR, Lenovo Flatiron Institute United States | 8,288 | 2.88 | 44 | 65.396 |
| 5 | 71 | preAlps - HPE Cray EX254n, NVIDIA Grace 72C 3.1GHz, NVIDIA GH200 Superchip, Slingshot-11, HPE Swiss National Supercomputing Centre (CSCS) Switzerland | 81,600 | 15.47 | 240 | 64.381 |

Green
Top 500
Junio 2024

<https://www.top500.org/lists/green500/2024/06/>

Motivación

| | | | | | | |
|----|-----|--|---------|-------|-------|--------|
| 6 | 299 | Horke Test - ThinkSystem SD665-N V3, AMD EPYC 9354 32C 3.25GHz, Nvidia H100 94Gb SXM5, Infiniband NDR200, Lenovo Karlsruher Institut für Technologie (KIT) Germany | 13,616 | 3.12 | 50 | 62.964 |
| 7 | 54 | Frontier TDS - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE DOE/SC/Oak Ridge National Laboratory United States | 120,832 | 19.20 | 309 | 62.684 |
| 8 | 11 | Venado - HPE Cray EX254n, NVIDIA Grace 72C 3.1GHz, NVIDIA GH200 Superchip, Slingshot-11, HPE DOE/NNSA/LANL United States | 481,440 | 98.51 | 1,662 | 59.287 |
| 9 | 20 | Adastra - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE Grand Equipement National de Calcul Intensif - Centre Informatique National de l'Enseignement Suprieur (GENCI-CINES) France | 319,072 | 46.10 | 921 | 58.021 |
| 10 | 28 | Setonix - GPU - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE Pawsey Supercomputing Centre, Kensington, Western Australia Australia | 181,248 | 27.16 | 477 | 56.983 |

Green
Top 500
Junio 2024

<https://www.top500.org/lists/green500/2024/06/>

Motivación

- **Aplicaciones novedosas e interesantes:**
 - **Deep Learning para tratamiento del cancer**
<https://news.developer.nvidia.com/improving-cancer-treatment-with-gpu-accelerated-supercomputers/>
 - **Procesamiento de imágenes en tomógrafo**
<https://www.lanacion.com.ar/2157971-que-tiene-de-singular-el-primer-tomografo-disenado-en-la-argentina>

Curso de GPGPU

Curso de GPGPU

- **Nombre de la asignatura: Programación masivamente paralela en procesadores gráficos (GPUs)**
- **Créditos: 10**
- **Materia: Arquitectura, Sistemas Operativos y Redes de Computadoras.**
- **Previaturas:**
 - **Programación 2 (Examen).**
 - **Una de las tres siguientes opciones:**
 - **Examen de Arquitectura de Computadoras**
 - **Examen de Sistemas Operativos**
 - **Curso de Arquitectura de Computadoras y Curso de Sistemas Operativos**

Curso de GPGPU

- **Objetivo de la asignatura:**
 - Introducir al estudiante en el uso de los procesadores gráficos para la resolución de problemas de propósito general.
- **Equipo docente:**
 - Pablo Ezzatti , Martín Pedemonte, Ernesto Dufrechou y Manuel Freire.

Curso de GPGPU

- **Forma de evaluación:**
 - Realización de ejercicios durante el curso.
 - Trabajo de laboratorio final.
 - Defensa o prueba escrita.
- **Para aprobar la asignatura se debe aprobar cada una de las instancias de evaluación.**
- **Calificación final:**
 - Ejercicios 40%
 - Laboratorio final 40%
 - Defensa o prueba escrita 20%

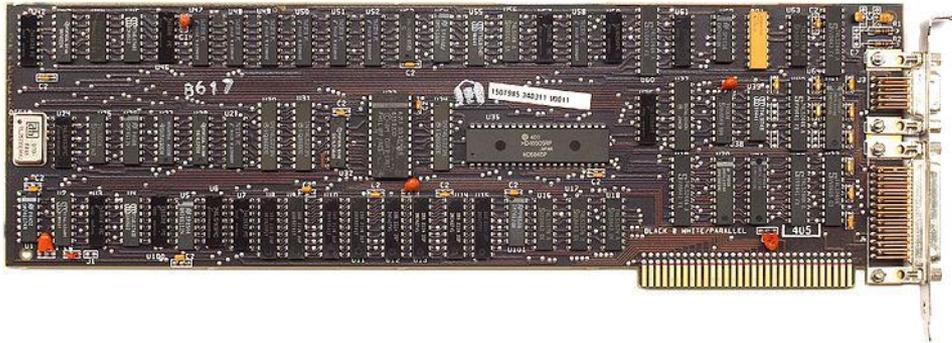
Curso de GPGPU

- El curso no tiene cupo.
- Es fundamental que se matriculen al EVA:
<https://eva.fing.edu.uy/course/view.php?id=1076>
- Toda la comunicación de información vinculada al curso se realiza a través del EVA del curso
- Dos clases por semana.

Un poco de historia...

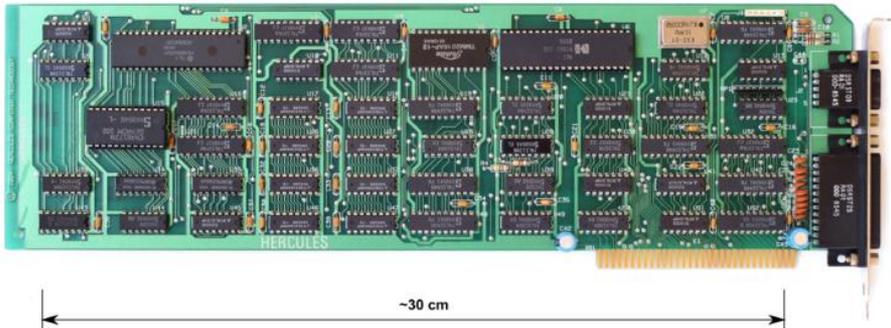
Un poco de historia...

- **1981 – Monochrome Display Adapter (MDA):**
 - No permitía el despliegue de gráficos.
 - Solamente en modo texto: 80 columnas x 25 líneas, monocromo.
 - 4 KB de memoria.



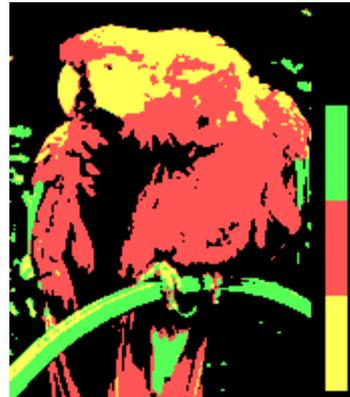
Un poco de historia...

- **1982 – Hercules Graphics Card:**
 - En modo texto: 80 columnas x 25 líneas, monocromo.
 - En modo gráfico todos los pixeles eran direccionables: 720 x 348 (mono).
 - 64 KB de memoria.



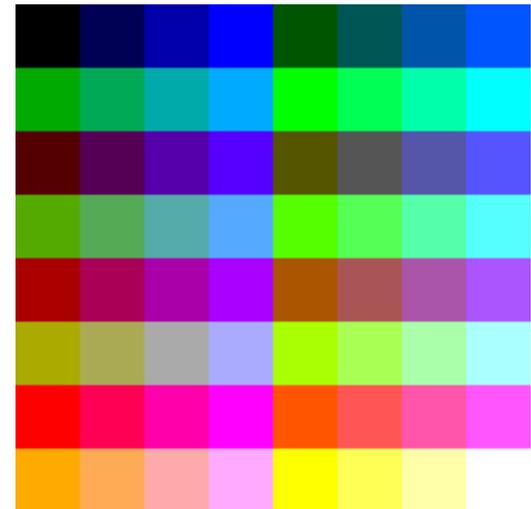
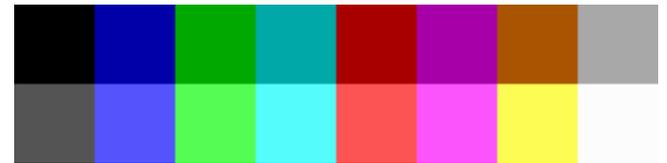
Un poco de historia...

- **1981 – Color Graphics Adapter (CGA):**
 - Estándar de facto de IBM
 - Primera que soporta color.
 - En modo texto: 80 columnas x 25 líneas (16 colores) pero con más baja resolución y 40 columnas x 25 líneas (16 colores).
 - En modo gráfico: 640 x 200 (mono), 320 x 200 (4 colores) y 160 x 100 (16 colores).
 - 16 KB de memoria.



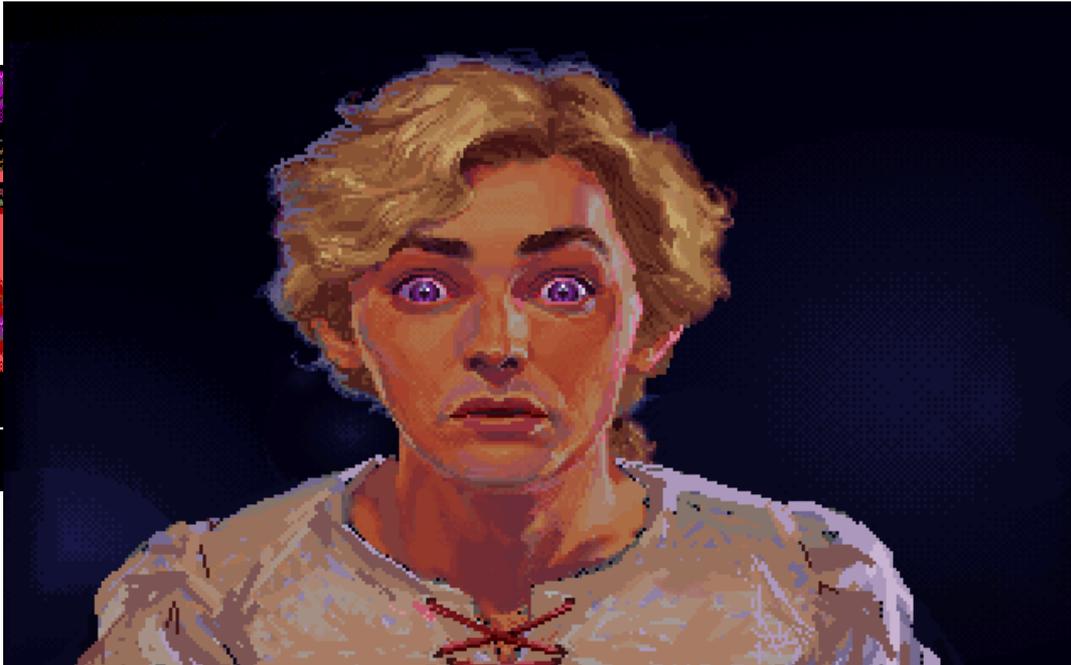
Un poco de historia...

- **1984 – Enhanced Graphics Adapter (EGA):**
 - Resolución: 640 x 350 (16 colores de una paleta de 64 colores).
 - 64 KB y 256 KB de memoria.



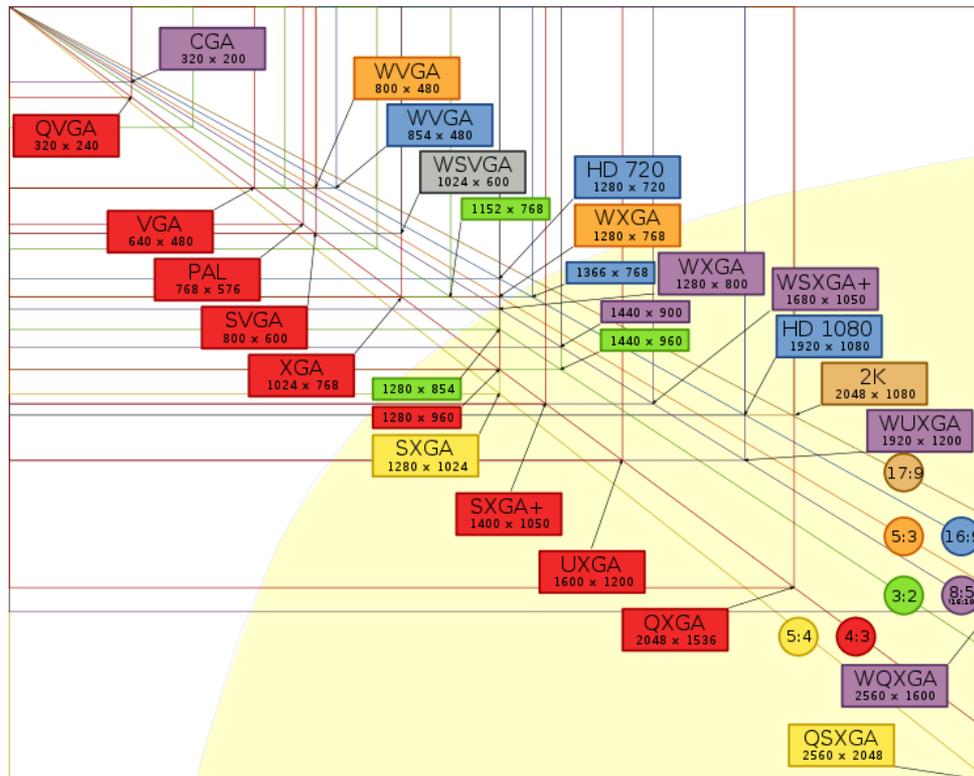
Un poco de historia...

- **1987 – Video Graphics Array (VGA):**
 - Último estándar exitoso de IBM.
 - Diseñado como un single chip.
 - Resoluciones: 640 x 480 (16 colores) y 320 x 200 (256 colores).
 - 256 KB de memoria.



Un poco de historia...

- Durante la década del '90 se mejoran las resoluciones:
 - Super VGA: 800 x 600.
 - XGA: 1024 x 768 (256 colores, 8 bits por pixel) y 800 x 600 (65536 colores, 16 bits por pixel).



Un poco de historia...

- Durante la década del '90 se presentan las primeras tarjetas con capacidad de reproducir efectos en 2D/3D:
 - 3dfx Interactive (1996-2000):
 - Línea Voodoo
 - ATI (1995-2004):
 - Línea Rage
 - Nvidia (1998-2005):
 - Líneas TNT y GeForce



¿Dónde estamos hoy?

Videos en tiempo actuales:

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

<https://www.youtube.com/watch?v=AV279wThmVU>

<https://www.youtube.com/watch?v=tjf-1BxpR9c>

El pipeline gráfico

El pipeline gráfico

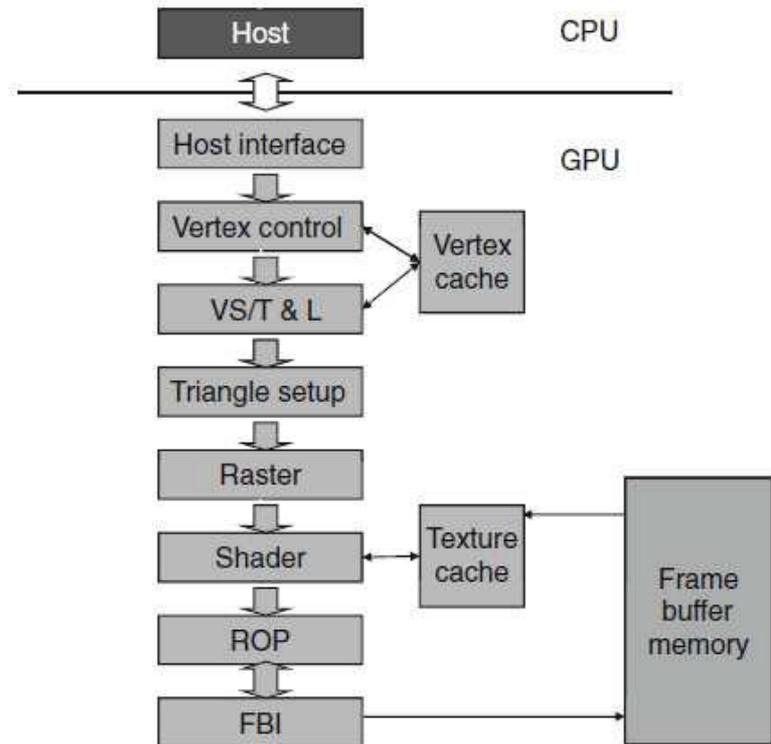
- Las operaciones para visualizar imágenes en pantalla se agrupan en un proceso generalmente conocido como pipeline gráfico.
- Típicamente se deben desplegar imágenes que son proyecciones de escenas tridimensionales en una pantalla bidimensional.
- Durante la década del 90 se diseñan tarjetas gráficas que van incorporando etapas del pipeline gráfico, con el fin de aliviar a la CPU.

El pipeline gráfico

- **En 1999, Nvidia lanza al mercado la GeForce 256 que es considerada la primera GPU (Graphics Processing Unit).**
- **La GeForce 256 es la primera tarjeta de video consumer-level que implementa el pipeline gráfico completo.**
- **El pipeline gráfico ha sido definido con diferente nivel de detalle por diversos autores.**
- **A continuación presentaremos el pipeline gráfico basado en la descripción de Kirk y Hwu (Programming Massively Parallel Processors: A Hands-on Approach).**

El pipeline gráfico

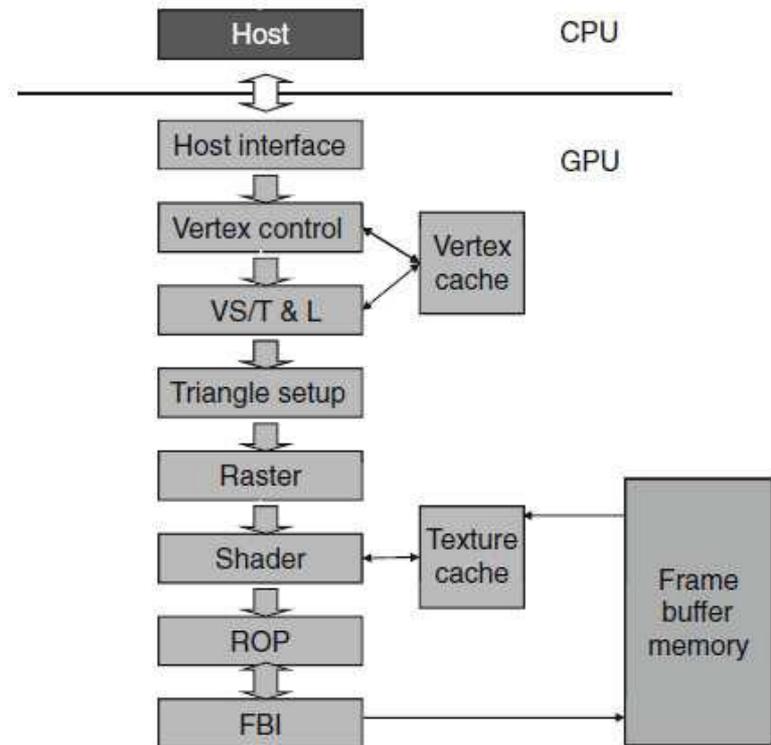
- La primera etapa del pipeline se encarga de recibir las imágenes de la CPU en algún formato manejable.
- Dicho formato es una colección de primitivas.
- Las primitivas suelen ser triángulos ya que son los polígonos más simples y permiten un manejo eficiente.



El pipeline gráfico

- **Vertex control:**

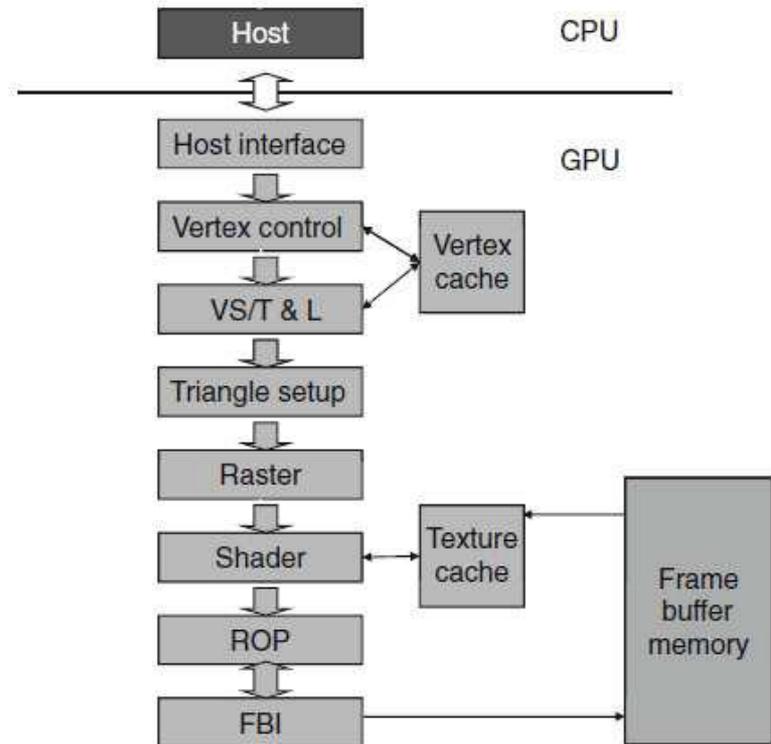
- Cada triángulo se transforma de acuerdo a su posición y orientación respecto al punto de vista.
- Cada uno de los vértices del triángulo se transforma a otra posición.
- Se eliminan los vértices no visibles.



El pipeline gráfico

- **VS/T & L:**

- Sombreado de vértices, transformaciones e iluminación.
- Se transforman los vértices, asignando propiedades a los vértices visibles como son: colores, normales, tangentes, texturas, etc.
- Esta etapa es realizada por los vertex shaders.
- Es una etapa computacionalmente costosa.



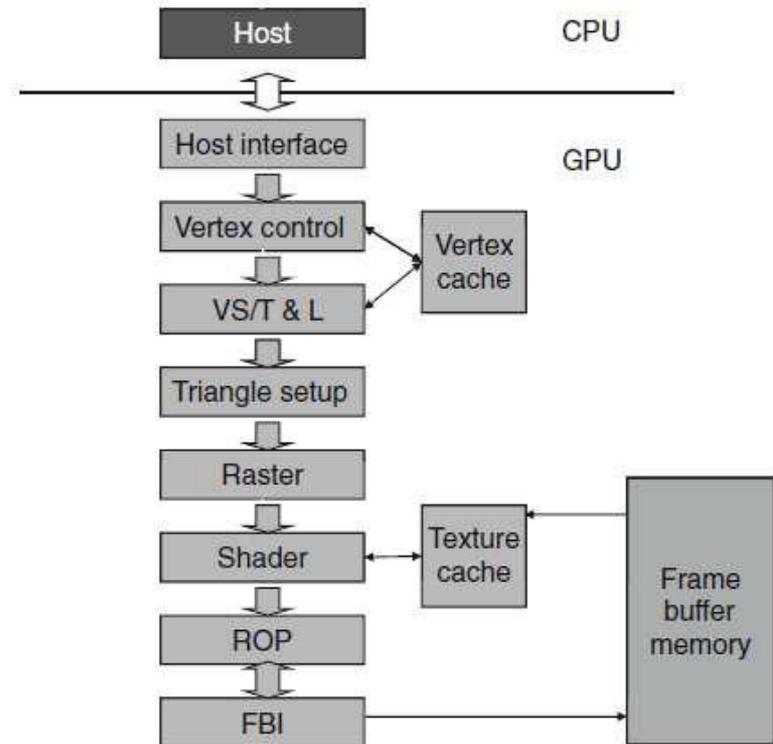
El pipeline gráfico

- **Triangle setup:**

- Realiza cálculos para las aristas, interpolando colores y otros valores de los vértices.

- **Raster:**

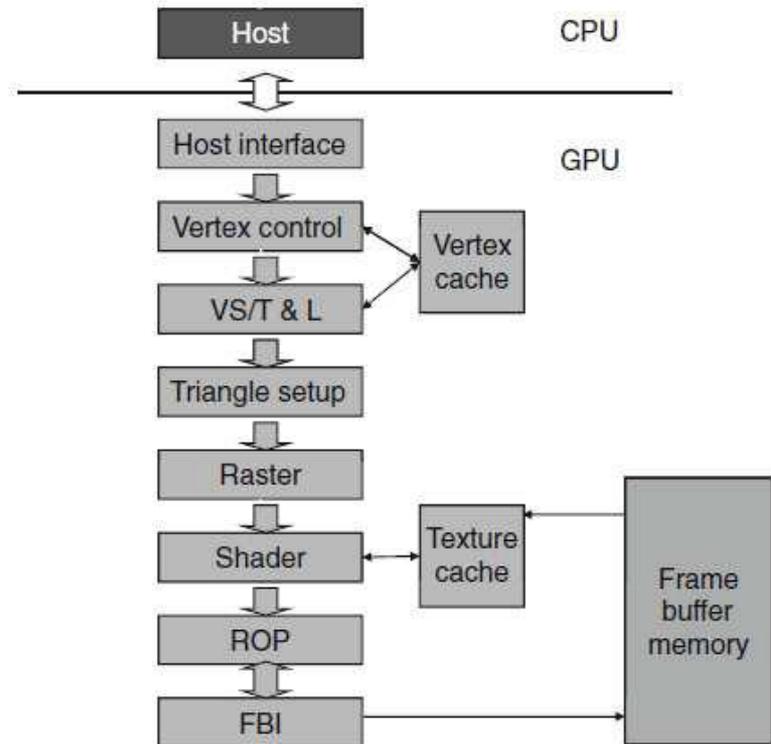
- Determina para cada triángulo que píxeles lo integran.
- Por cada píxel se interpola, a partir de los vértices, los valores necesarios para sombrear los píxeles, evaluar la inclusión de colores y determinar la posición de la textura que será pintada en el píxel.



El pipeline gráfico

- **Shader (sombreado):**

- En esta etapa se ejecutan los efectos que permiten renderizar imágenes en forma más realista.
- Se determina el color final de cada píxel.
- Se utilizan técnicas como la interpolación de colores entre vértices, el mapeo de textura, la iluminación por píxel.
- Esta etapa es computada por los pixel shaders.
- Esta etapa es costosa computacionalmente.



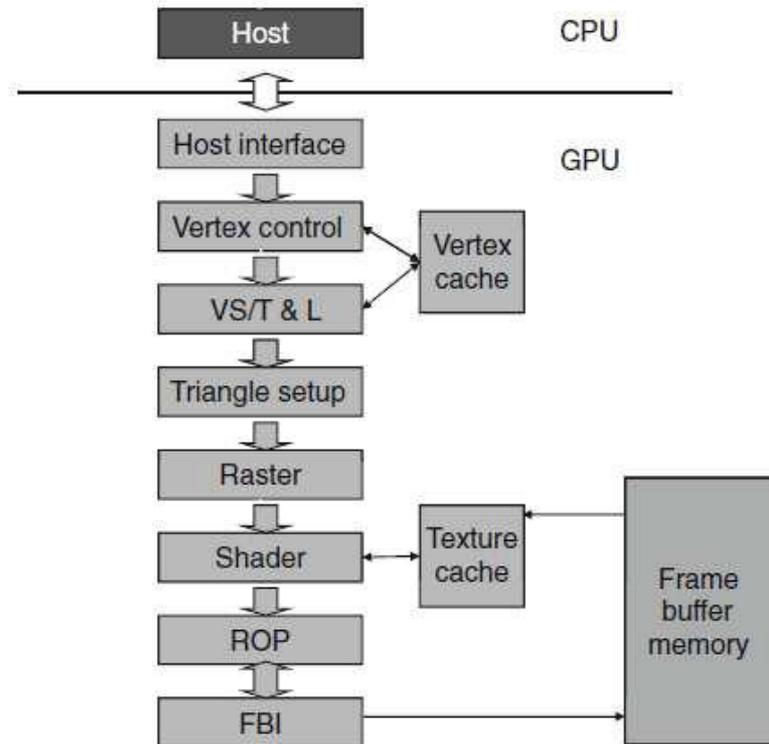
El pipeline gráfico

- **ROP (Raster OPeration):**

- Se aplican restricciones finales sobre cada píxel.
- Por ejemplo: mezcla de colores de triángulos superpuestos para mejorar los efectos de transparencias y antialiasing.
- Se determinan los objetos visibles para un punto de vista dado y se descartan los píxeles ocultos.

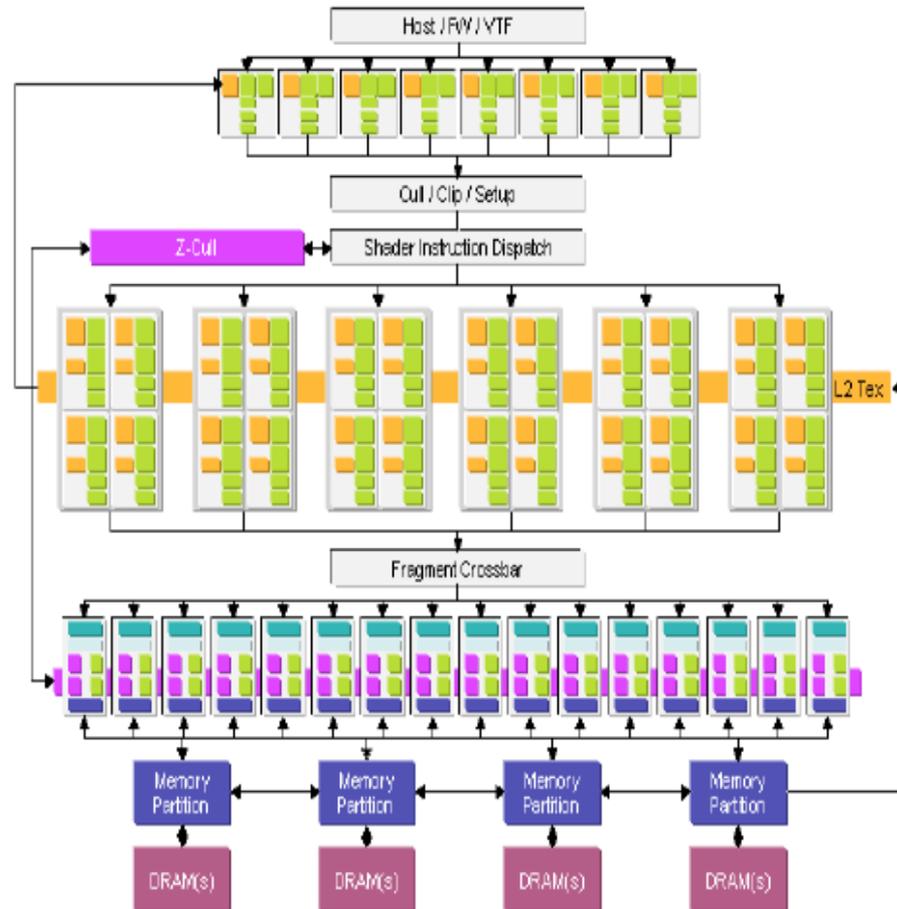
- **FBI (Frame Buffer Interface):**

- Lee y escribe en el buffer de imagen.



El pipeline gráfico

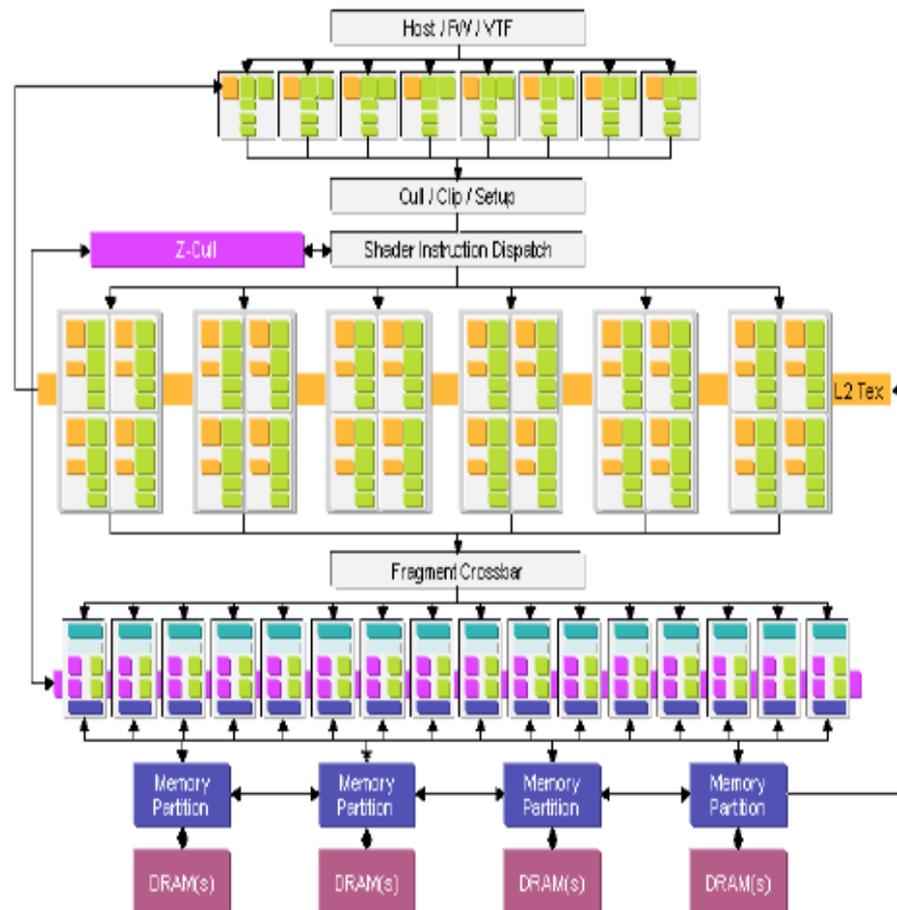
Ejemplo: la arquitectura de la GeForce 7800 (pre-CUDA)



¿Por qué una arquitectura unificada?

¿Por qué una arquitectura unificada?

La arquitectura de la GeForce 7800



¿Por qué una arquitectura unificada?

Why unify?



Heavy Geometry
Workload Perf = 4



Heavy Pixel
Workload Perf = 8

¿Por qué una arquitectura unificada?

Why unify?

Unified Shader



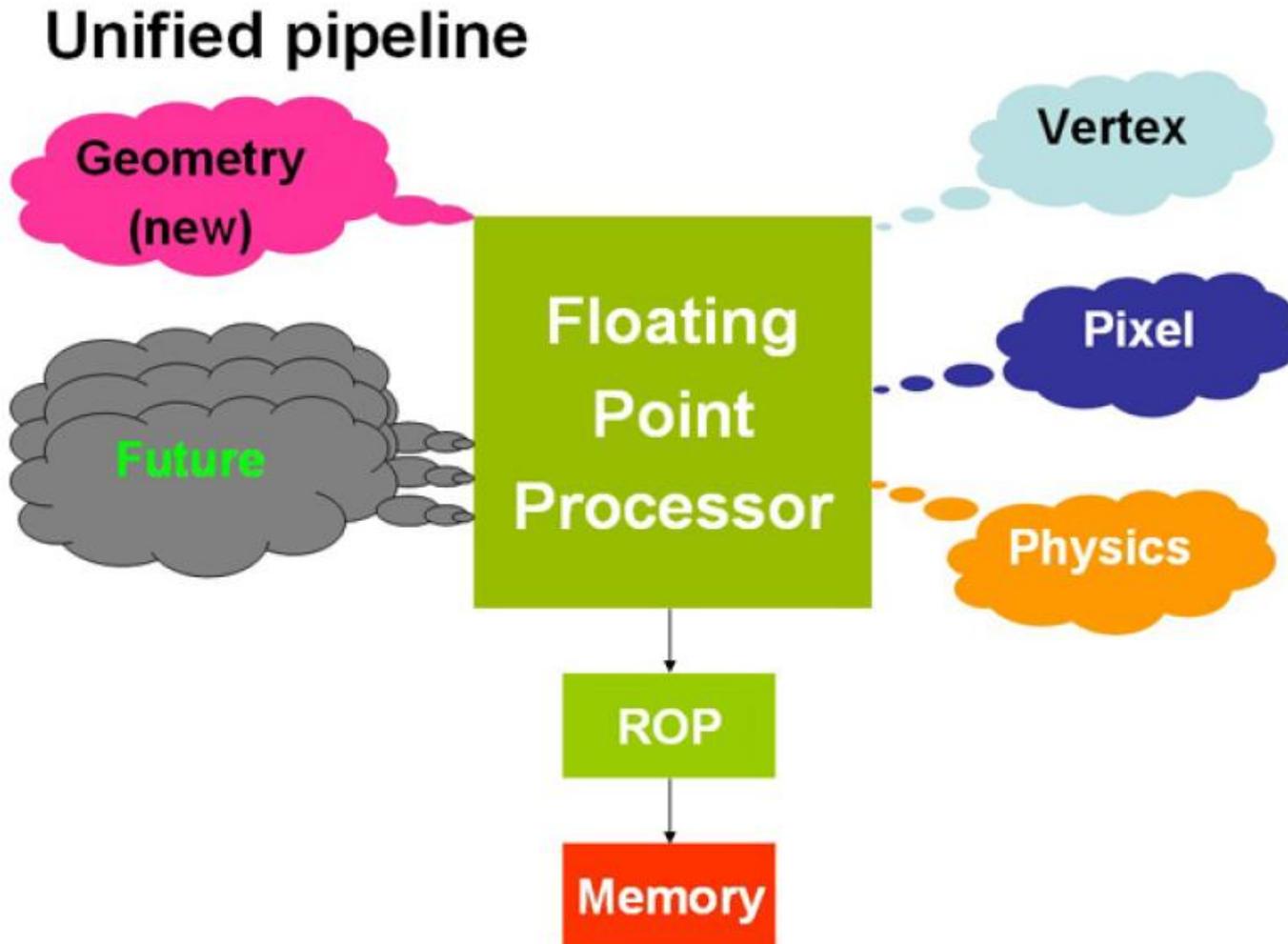
**Heavy Geometry
Workload Perf =12**

Unified Shader

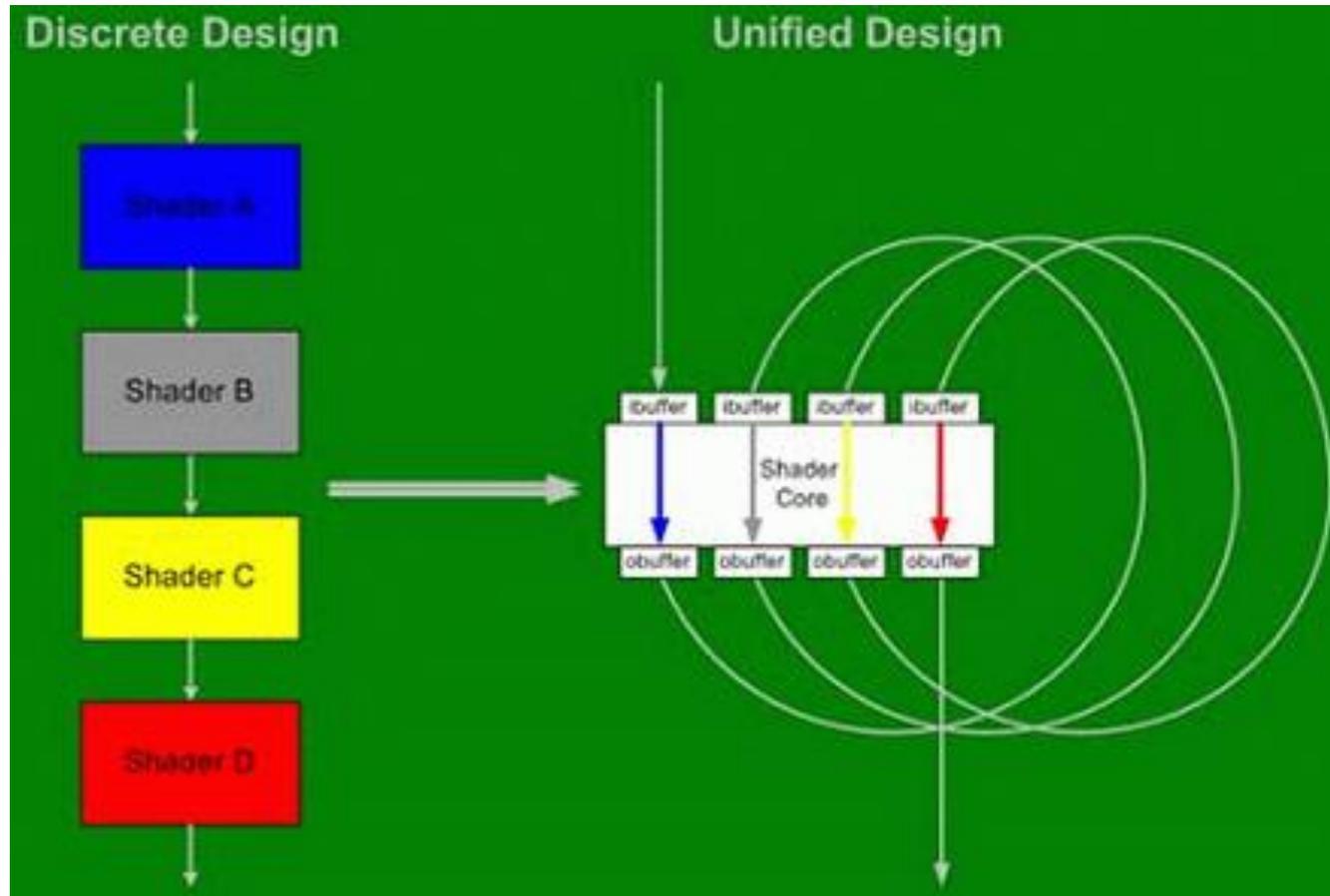


**Heavy Pixel
Workload Perf = 12**

¿Por qué una arquitectura unificada?

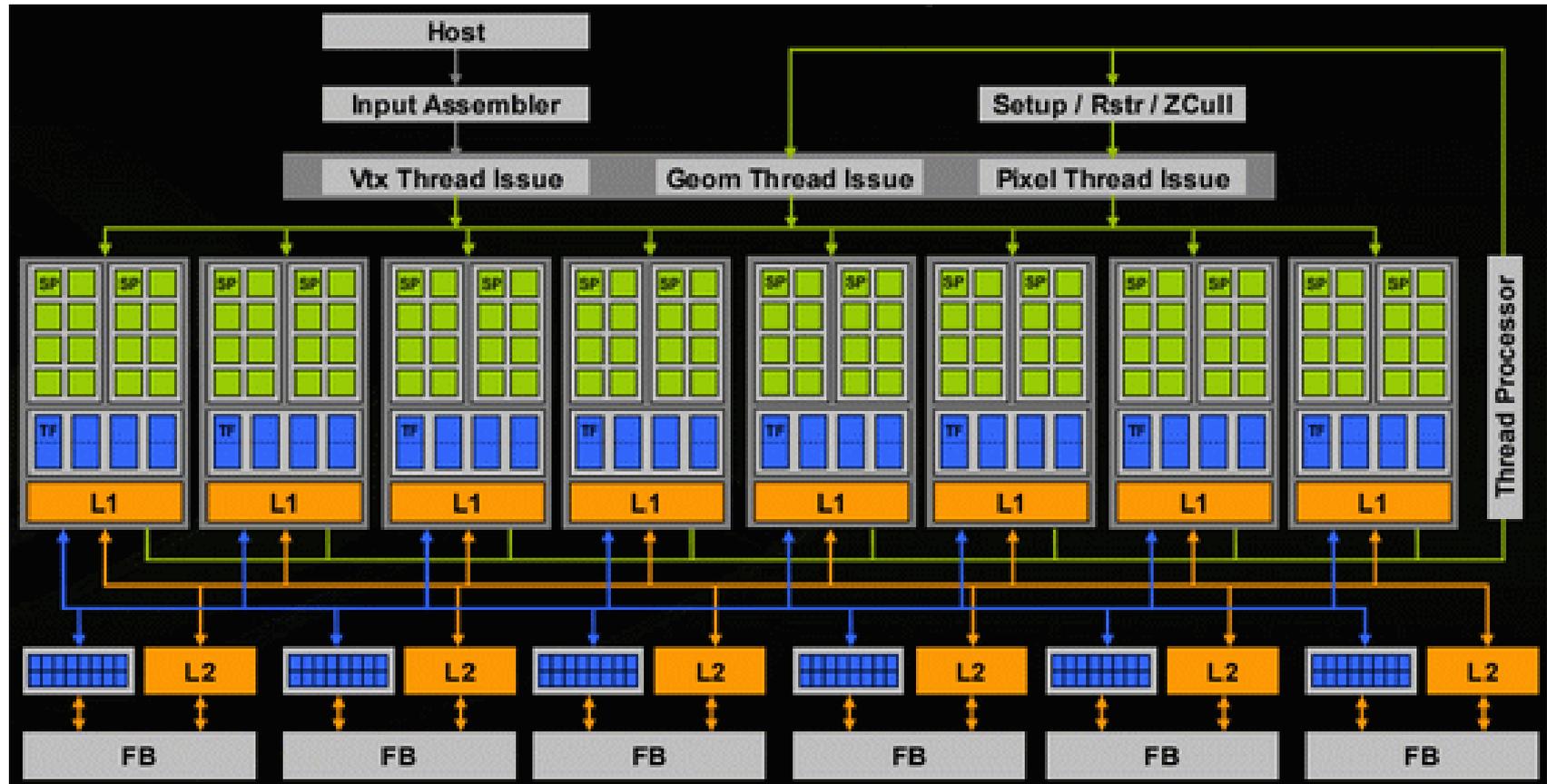


¿Por qué una arquitectura unificada?



¿Por qué una arquitectura unificada?

La arquitectura del chip G80 (GeForce 8800)



¿Por qué una arquitectura unificada?

¿Y por qué siempre terminamos acá?: Arquitectura Ampere (2021)



Tarjetas programables

Tarjetas programables

- **En los últimos años existe un auge de la utilización de GPUs para el cómputo de problemas de propósito general.**
- **Este crecimiento se basa fundamentalmente en:**
 - **La arquitectura es intrínsecamente paralela, en contraste con la arquitectura serial de las CPUs.**
 - **La industria de los videojuegos presiona a los fabricantes de tarjetas gráficas para aumentar las capacidades de procesamiento gráfico para que los juegos sean más realistas y más rápidos.**
 - **El surgimiento de lenguajes de programación de propósito general para GPUs.**

Tarjetas programables

- Las tarjetas de video viejas tenían un pipeline gráfico fijo.
- Las operaciones ejecutadas y el orden en que se aplicaban sobre los datos estaba preconfigurado.
- Originalmente, las GPUs proveían operaciones propias para la transformación e iluminación (vertex shaders) en los vértices, y para determinar el color final de los píxeles (pixel shaders).

Tarjetas programables

- **1999-2006: se produce una mejora significativa en las capacidades de programación de las GPUs.**
- **Algunos hitos son:**
 - **GeForce 3 (2001): primera GPU que ejecutaba vertex shaders programados en DirectX 8 (pudiendo programar la etapa de VS/T & L).**
 - **ATI Radeon 9700 (2002): introdujo la aritmética de punto flotante de 24 bits en los pixel shaders (DirectX 9 y OpenGL).**
 - **GeForce FX (2002-2003): introdujo el trabajo con aritmética de punto flotante de 32 bits.**
 - **Xbox 360 (2005): primera arquitectura unificada de procesadores para el pipeline gráfico. Una sola clase de procesadores computa las distintas secciones del pipeline.**

Tarjetas programables

- **En un principio, el avance en el hardware no fue acompañado por un avance en el software de manejo de las GPUs.**
- **Inicialmente la programación de las GPUs se realizaba mediante llamados a servicios de interrupción de la BIOS.**
- **Posteriormente, se comenzaron a desarrollar los shaders en el lenguaje ensamblador específico de cada modelo.**
- **Esto implicaba la existencia de varios lenguajes y baja portabilidad de los programas.**
- **Otra alternativa era utilizar las APIs gráficas como OpenGL y DirectX.**

Tarjetas programables

- Evidentemente esto representaba una limitación para el desarrollo de aplicaciones.
- Para solucionarlo se desarrollaron diferentes lenguajes de programación de más alto nivel que funcionaran sobre los modelos de GPU existentes, como: High-Level Shading Language (HLSL) y Cg.
- Posteriormente, otros lenguajes de alto nivel surgieron basados en considerar la GPU como un stream processor como: Brook, Sh, PyGPU, Accelerator Language, Close to the Metal (CTM) y ATI Stream.
- Sin embargo, cada herramienta seguía siendo muy dependiente de la arquitectura de la GPU, el modelo, etc.

CUDA

CUDA

- **En el año 2007, Nvidia presenta CUDA (Compute Unified Device Architecture).**
- **Produjo un cambio radical en la arquitectura de las GPUs de Nvidia:**
 - **arquitectura unificada sin distinción entre procesadores de píxeles y vértices.**

CUDA

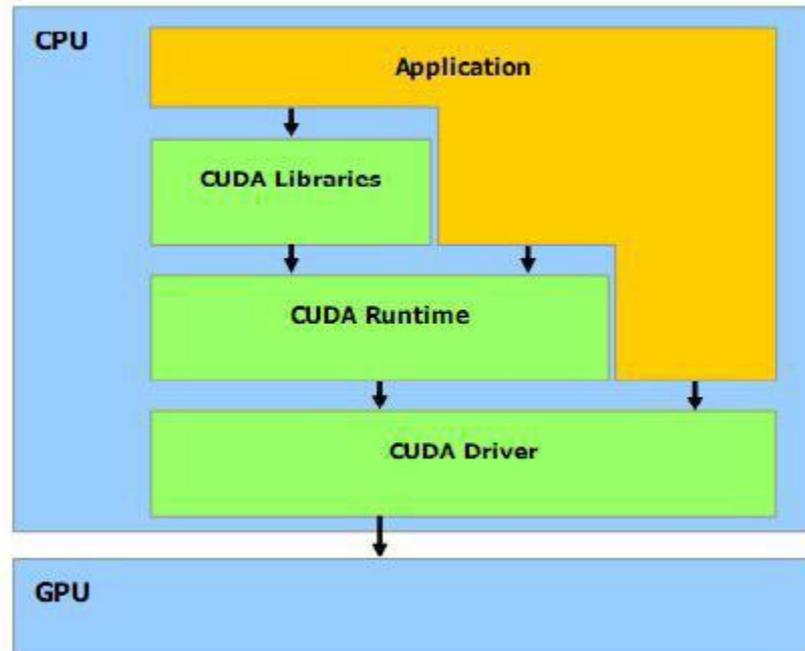
- **Produjo un cambio radical en el software para desarrollo de aplicaciones en las GPUs de Nvidia.**
- **Es el mojón más importante desde el nacimiento de la programación de propósito general en GPUs (GPGPU).**
- **Masificó la GPGPU, ya que dejó de ser un juego para eruditos y se transformó en una alternativa a la alcance de cualquier desarrollador.**

CUDA

- **Pero, ¿qué es CUDA?**
- **CUDA:**
 - Es una arquitectura de computación paralela para el cómputo de problemas de propósito general diseñada por Nvidia.
 - Está enfocada al cálculo masivamente paralelo y las capacidades de procesamiento que brinda las tarjetas gráficas de Nvidia.
 - Permite programar el dispositivo a través de extensiones de lenguajes de programación estándar (C y Fortran)
 - Está disponible para las tarjetas gráficas GeForce de la serie 8 en adelante.
 - Es compatible con Linux de 32/64 bits y Windows de 32/64 bits.

CUDA

- Se compone de una pila de capas de software que incluye bibliotecas, el CUDA Runtime y el CUDA Driver.



CUDA

- **Bibliotecas de CUDA:**
 - Dos bibliotecas matemáticas de alto nivel: CUFFT y CUBLAS.
 - CUFFT para el cálculo de Transformadas de Fourier.
 - CUBLAS es una implementación de BLAS (Basic Linear Algebra Subprograms) en GPU.
- **CUDA Driver:**
 - Es un controlador de hardware.
 - Está dedicado a la transferencia de datos entre la CPU y la GPU.
- **CUDA Runtime:**
 - Provee una interfaz de programación de aplicaciones (API).
 - Provee un conjunto de instrucciones accesible a través de lenguajes de alto nivel como son Fortran y C.

CUDA

- **Al programar en CUDA, la GPU se ve como un dispositivo de cómputo capaz de ejecutar un número muy elevado de hilos en paralelo.**
- **La GPU funciona como un coprocesador de la CPU (host).**
- **A la GPU se le suele llamar dispositivo (device).**

CUDA

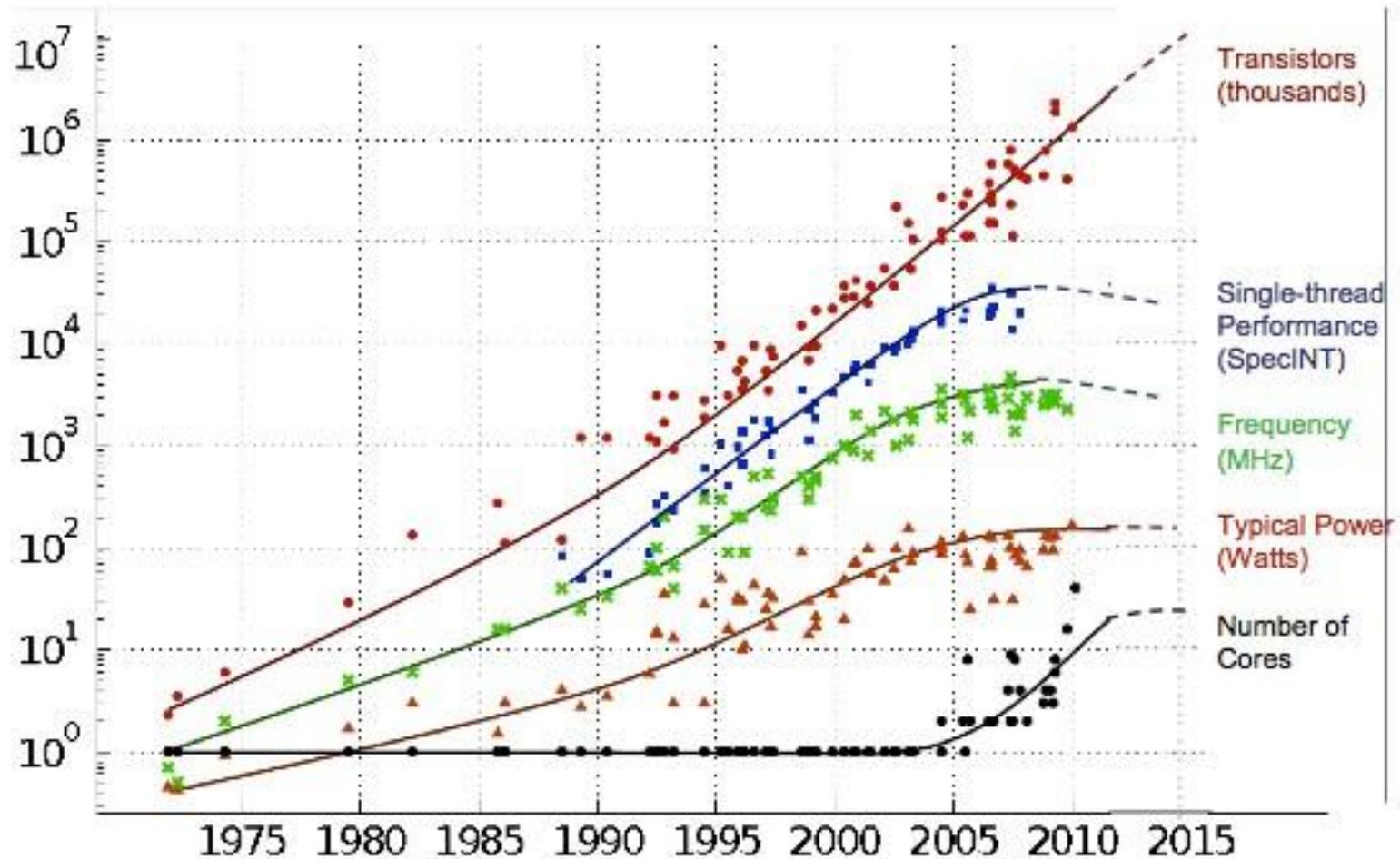
- Si una parte de una aplicación se ejecuta muchas veces pero en forma independiente sobre diferentes datos, puede ser aislada en una función que se ejecutará en el dispositivo mediante muchos hilos de ejecución en forma concurrente.
- La función se compila usando el conjunto de instrucciones del dispositivo.
- El programa resultante, llamado núcleo (kernel), se descarga en el dispositivo para su ejecución.

CUDA

- El curso se centra en CUDA y las arquitecturas de Nvidia.
- Se comentarán algunos aspectos de otros lenguajes que siguen una filosofía similar como OpenCL.
- No se entrará en detalles de las arquitecturas de AMD/ATI ya que presentan características distintas.

Comparación de la performance CPU vs GPU

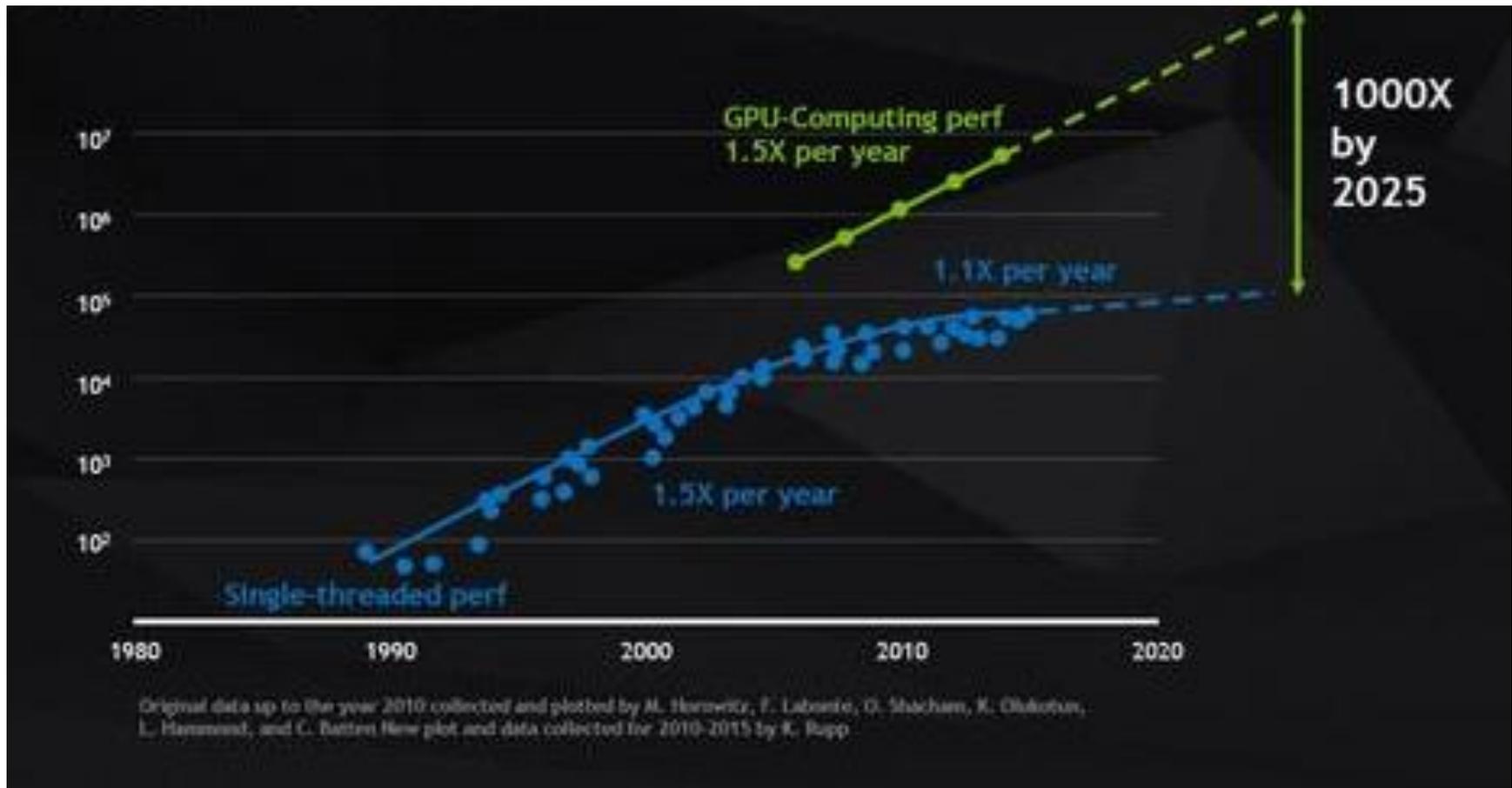
Comparación de la performance CPU vs GPU



Original data collected and plotted by M. Horowitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammond and C. Batten
Dotted line extrapolations by C. Moore

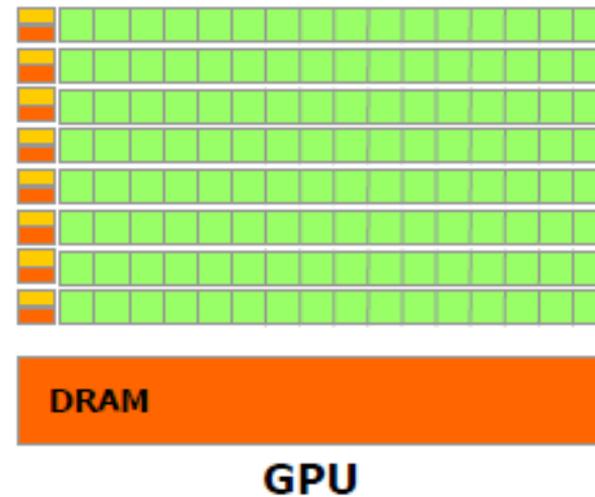
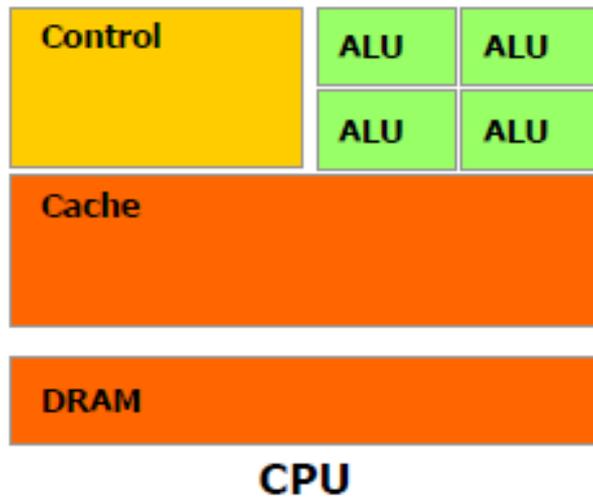
Comparación de la performance CPU vs GPU

Pico teórico de performance en GFLOP/s



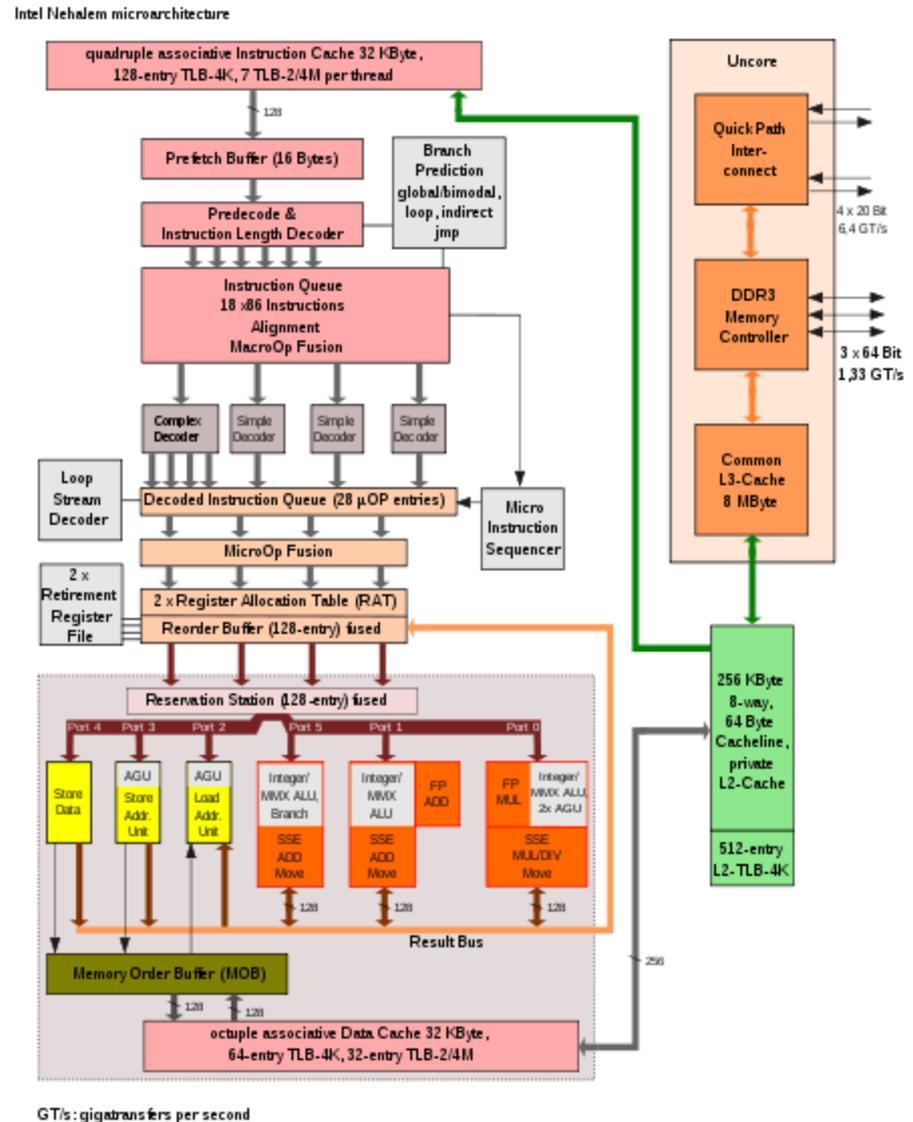
Comparación de la performance CPU vs GPU

- La arquitectura de las GPUs es radicalmente distinta a la de una CPU.



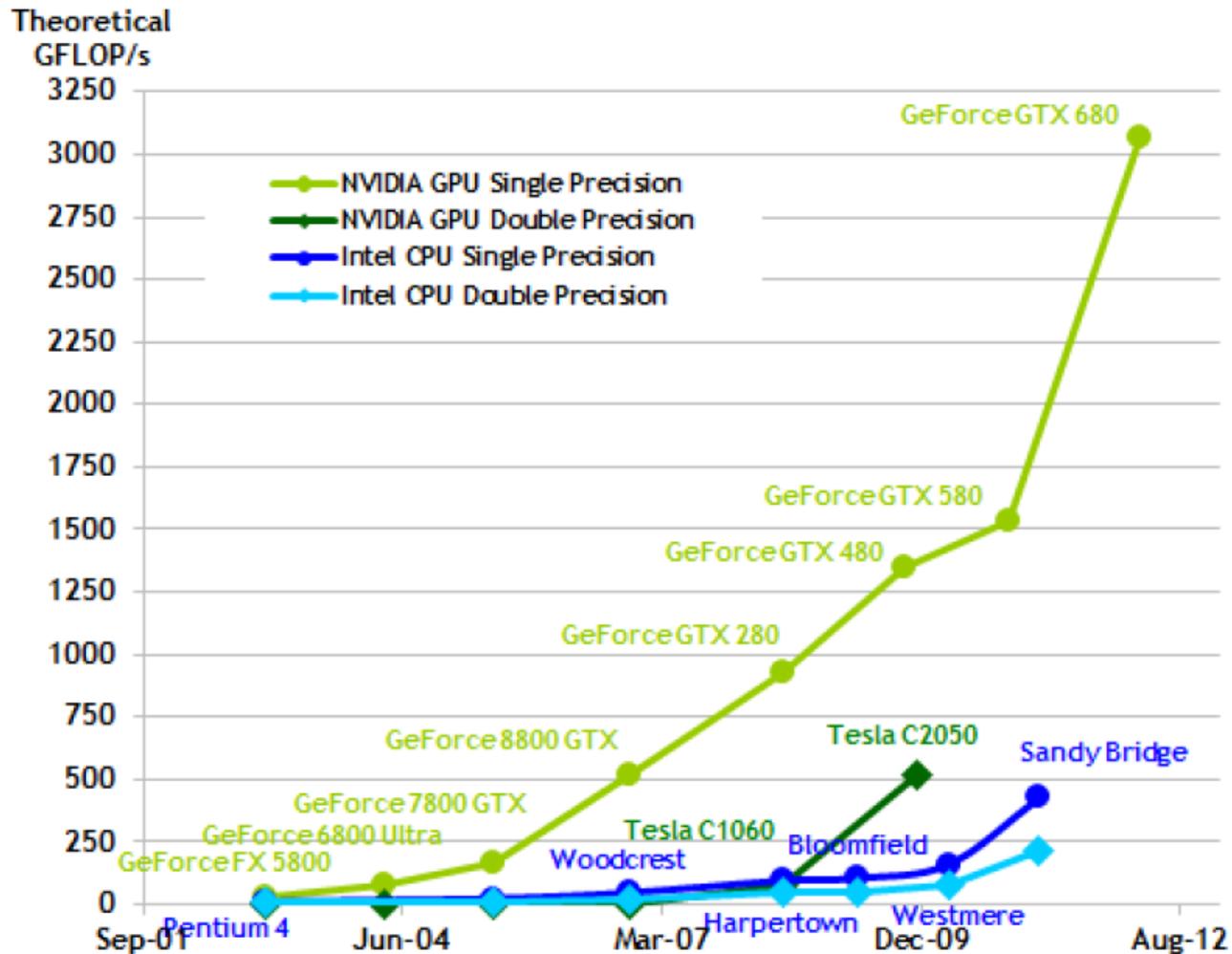
Comparación de la performance CPU vs GPU

- En una CPU tradicional gran parte de los transistores están dedicados a realizar otro tipo de tareas:
 - Predicción de branches.
 - Prefetch de memoria.
 - Ejecución fuera de orden.
 - Caché de datos.
- En las GPUs hay más transistores dedicados al cálculo.



Comparación de la performance CPU vs GPU

Pico teórico de performance en GFLOP/s



Comparación de la performance CPU vs GPU

- **¿Cómo se calcula el pico teórico de performance?**
 - **FLOPS = #cores * Frecuencia * Ops**
 - **FLOPS:** operaciones de punto flotante por segundo
 - **#cores:** número de cores
 - **Frecuencia:** frecuencia del core
 - **Ops:** operaciones de punto flotante por core y por ciclo.
- **La cantidad de operaciones de punto flotante por core depende de:**
 - Si es una CPU o una GPU.
 - Si se trabaja en simple o en doble precisión.

Comparación de la performance CPU vs GPU

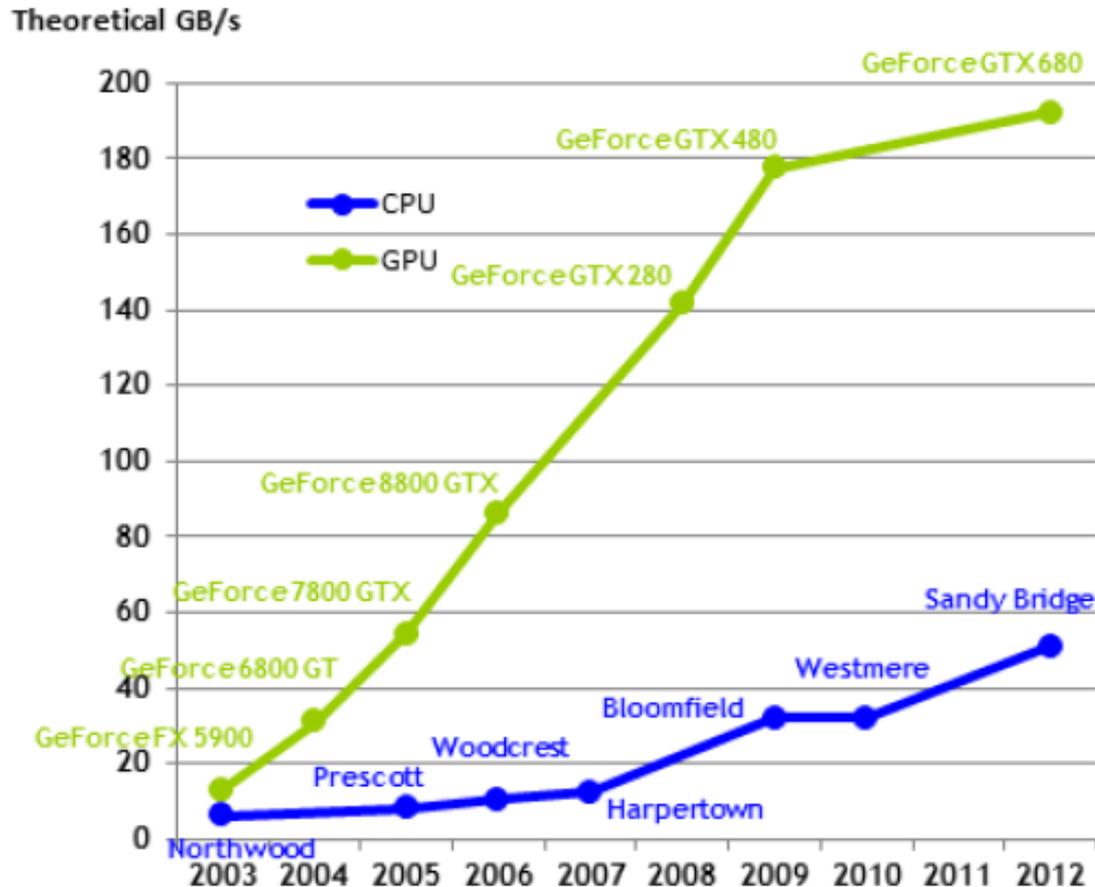
- **Un ejemplo de CPU: Westmere**
 - 8 ops de precisión simple por ciclo por core.
 - 4 ops de precisión doble por ciclo por core.
 - Gulftown/Westmere-EP: 6 Cores a 3.46 GHz (U\$S 1663)
 - Pico en simple = $6 * 3.46 * 8 = 166.08$ GFLOPS
 - Pico en doble = $6 * 3.46 * 4 = 83.04$ GFLOPS
 - El hyperthreading no mejora la performance

Comparación de la performance CPU vs GPU

- **Un ejemplo de GPU: GeForce GTX 480**
 - 2 ops de precisión simple por ciclo por core.
 - 1/2 ops de precisión doble por ciclo por core.
 - GeForce GTX480: 480 CUDA Cores a 1401 MHz (U\$S 500)
 - Pico en simple = $480 * 1401 * 2 = 1344.96$ GFLOPS
 - Pico en doble = $480 * 1401 * 1/2 = 336.24$ GFLOPS
- **Algunos comentarios:**
 - Ojo! Un mega y un giga son 1000 y no 1024.
 - Las GeForce suelen tener peor desempeño en doble precisión (leyenda urbana: deshabilitan parte del funcionamiento). Las Tesla hacen más operaciones por ciclo pero son más caras.

Comparación de la performance CPU vs GPU

Pico teórico de tasa de transferencia de memoria en GB/s



Comparación de la performance CPU vs GPU

- **¿Cómo se calcula el pico teórico de tasa de transferencia de memoria?**
 - Tasa de transferencia = ancho del interfaz de memoria * frecuencia de la memoria.
- **Un ejemplo de CPU: Westmere**
 - ancho del interfaz de memoria = 64 bits = 8 bytes
 - la memoria es 3xDDR3-1333 que puede leer o escribir dos palabras de datos por vez y cuatro veces por ciclo de reloj.
 - frecuencia de la memoria = 166.66 Mhz
 - Tasa de transferencia = $3 * 2 * 4 * 8 * 166.66 = 31.99$ GB/s

Comparación de la performance CPU vs GPU

- **Un ejemplo de GPU: GeForce GTX 480**
 - ancho del interfaz de memoria = 384 bits = 48 bytes
 - frecuencia de la memoria = 1848 Mhz
 - la memoria es GDDR5 que puede leer o escribir dos palabras de datos por ciclo de reloj.
 - Tasa de transferencia = $48 * 1848 * 2 = 177.40$ GB/s

Comparación de la performance CPU vs GPU

- La otra cara de la moneda:
 - Estudio comparativo realizado por Intel entre un i7 de 4 cores y una GTX280 sobre varios problemas.
 - La gráfica muestra la performance relativa entre la GTX280 y el i7.

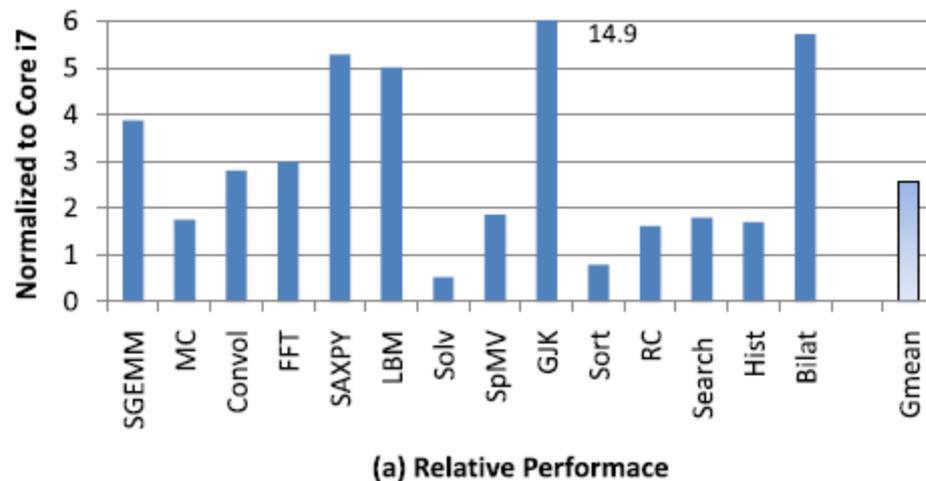


Figure 1: Comparison between Core i7 and GTX280 Performance.