中国科学院国家天文台
National Astronomical Observatories, CAS

千人計劃
RECRUITMENT
PROGRAM OF GLOBAL EXPERTS

UNIVERSITÄT
HEIDELBERG
Zukunft. Seit 1386.

Introduction to GPU Accelerated Computing: 1. History of Computer Architecture Many-Core, GPU, and other ideas...

University

Rainer Spurzem
Astronomisches Rechen-Inst., ZAH, Univ. of Heidelberg, Germany
National Astronomical Observatories (NAOC), Chinese Academy of Sciences
Kavli Institute for Astronomy and Astrophysics (KIAA), Peking University

spurzem@ari.uni-heidelberg.de
http://silkroad.bao.ac.cn

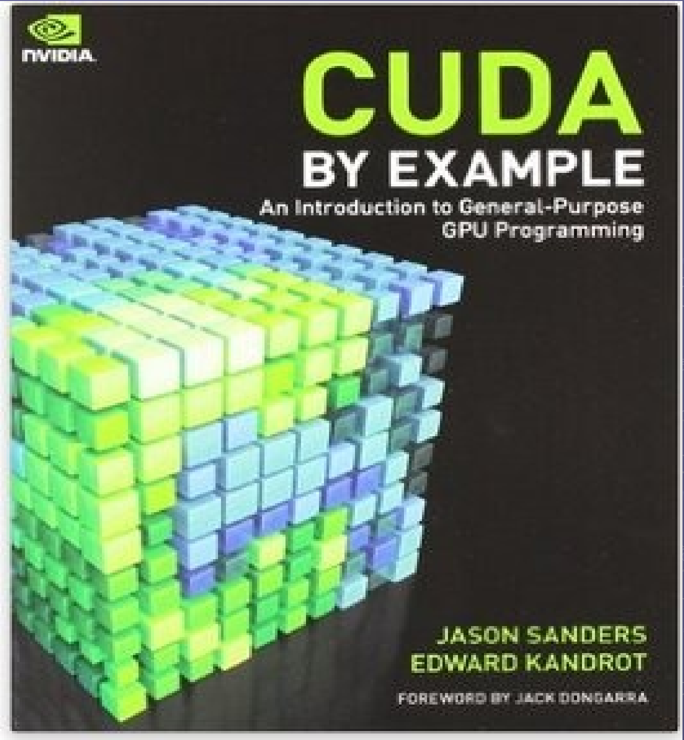
the SILK ROAD PROJECT at NAOC/KIAA
丝绸之路 计划

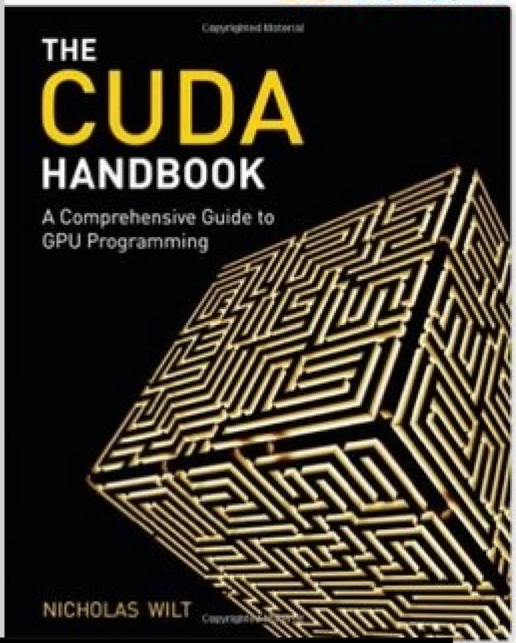
北京大
PEKING UNIVERS

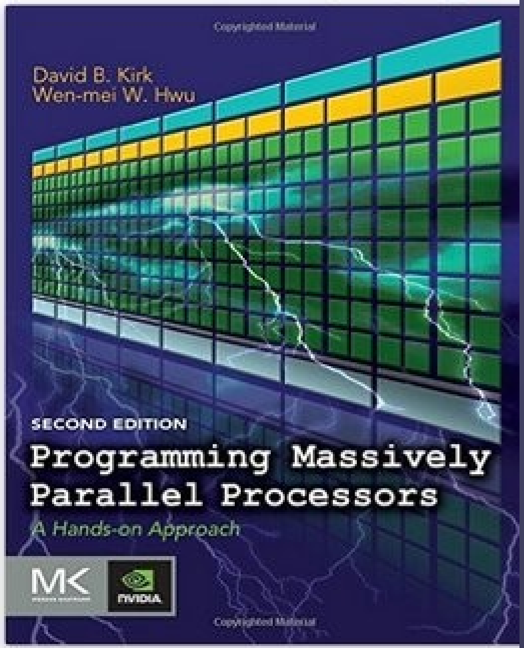
Introduction to GPU Accelerated Computing
March 1 - 5, 2021

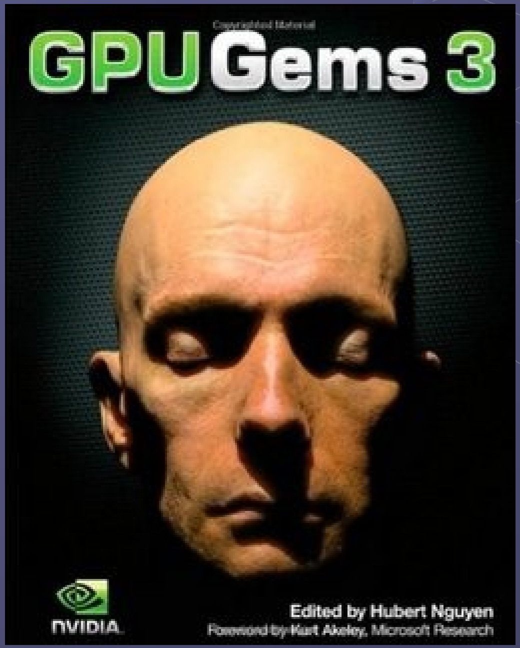
Table of Contents (subject to adjustment/change):

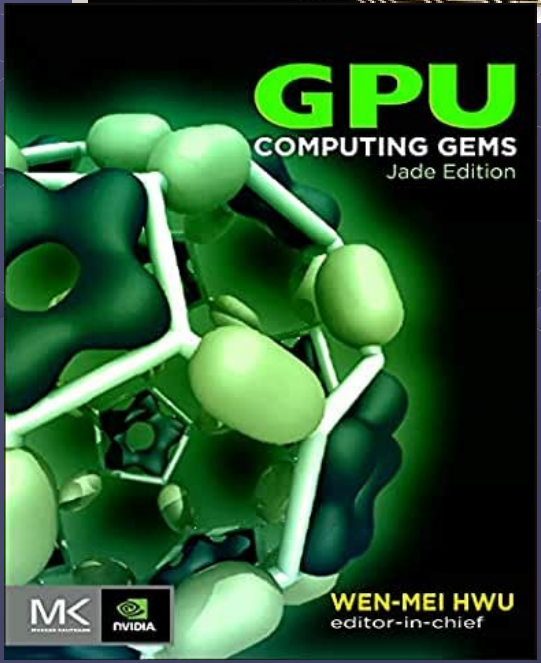
Literature; why NVIDIA? CUDA … openCL?

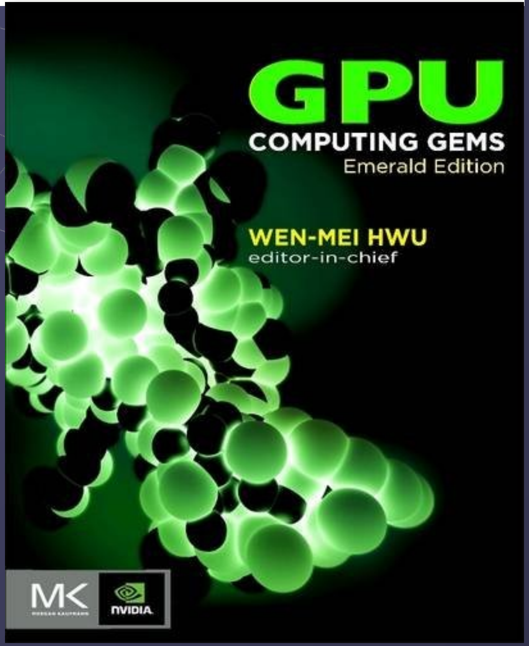












Observations (Experiment)

Theory

Computational Physics

GPU Computing

History

History

Erik Holmberg (1908-2000)

Dissertation Univ. Lund (Schweden) (1937):

``A study of double and multiple galaxies´´

Galaxies often in Groups and Pairs

Irregular Distribution of Satellite Galaxies

(Holmberg-Effect)

Father of numerical astrophysics?

» **...with 200 light bulbs**

History

http://cdsads.u-strasbg.fr/abs/1941ApJ....94..385H

The Astrophysical Journal, Nov. 1941

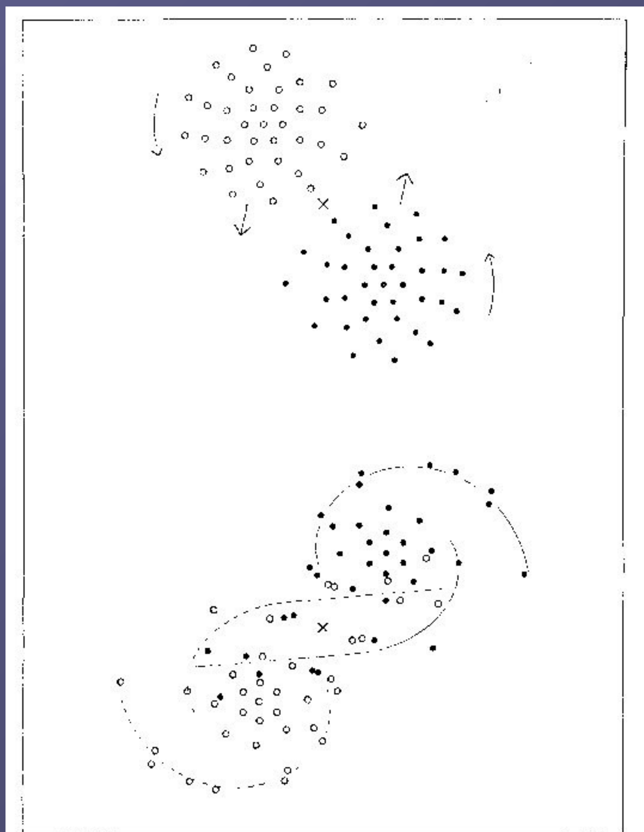


FIG. 4b

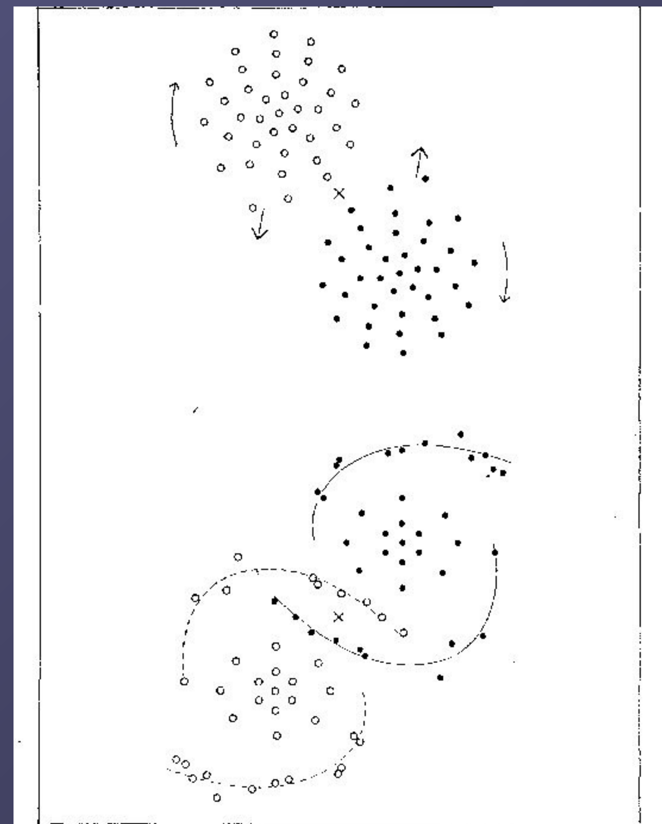


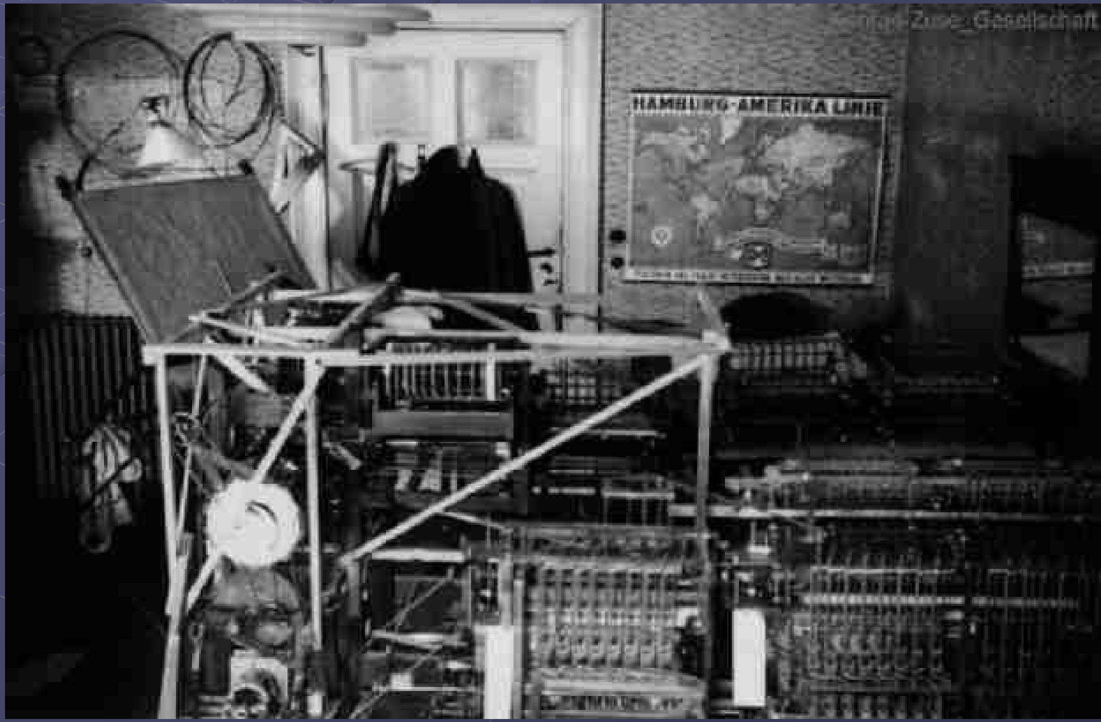
FIG. 4a

HARDWARE

...before von Neumann...

- Konrad Zuse (1910-1995) Berlin

Invented freely programmable Computer



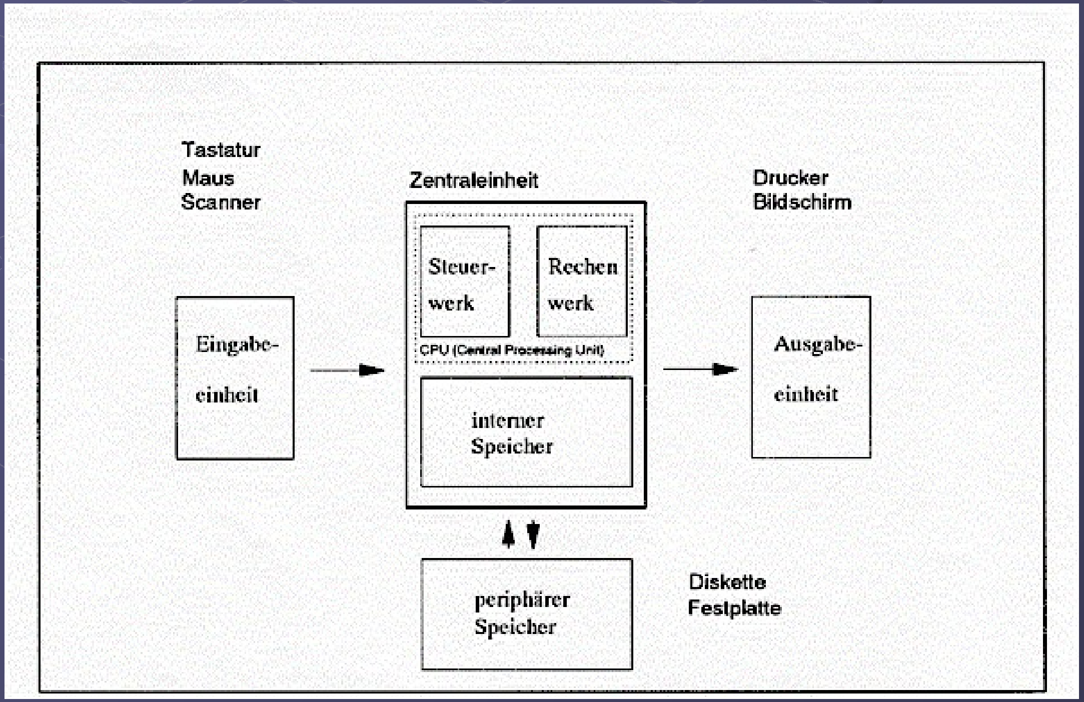
Z1 in parental flat 1936

HARDWARE

- John von Neumann (1903-1957)

Born Budapest, Lecturer Berlin, since 1930 Princeton Univ.

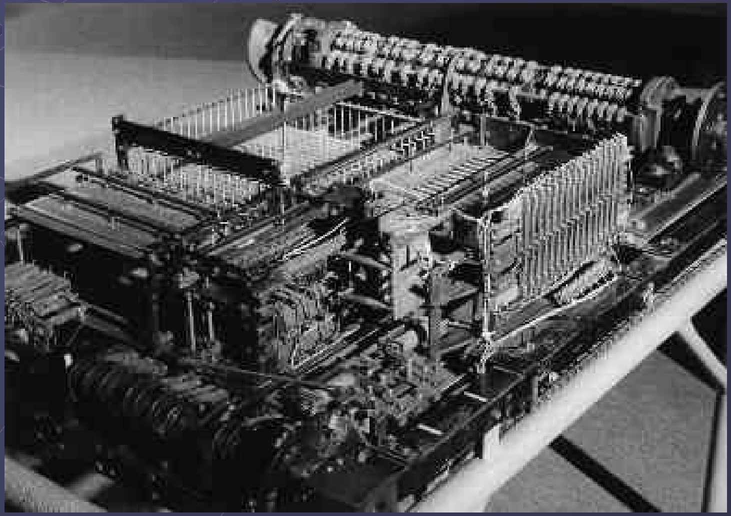
Requirements for the Construction of an electronic computing device(1946)



History

Zuse Z4: 1944 Berlin, 1950 Zürich, 1954 Frankreich

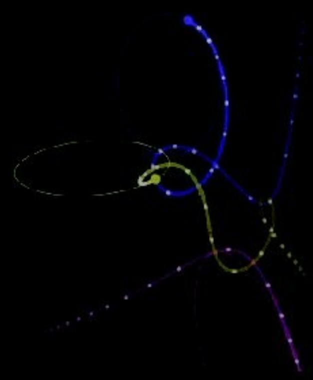
1959 Deutsches Museum München



Computing Speed 0.03 MHz

Memory 256 byte

Astronomisches Rechen-Institut (ARI) at Univ. of Heidelberg, Germany



Siemens 2002 Computer in 1964 At ARI

History

http://cdsads.u-strasbg.fr/abs/1960ZA.....50..184V

Astronomisches Rechen-Institut in Heidelberg
Mitteilungen Serie A Nr. 14

Die numerische Integration des *n*-Körper-Problemes für Sternhaufen I

Von

SEBASTIAN VON HOERNER

Mit 3 Textabbildungen

(Eingegangen am 10. Mai 1960)

Tabelle 5. *Zahl der gegenseitigen Umläufe, Häufigkeit des Auftretens und kleinster gegenseitiger Abstand D_m der engsten Paare.* (Alle engsten Paare mit mehr als zwei vollen Umläufen wurden notiert)

Umläufe	Häufigkeit	D_m
2—3	11	0.0102
3—5	9	0.0177
5—10	5	0.0070
10—20	2	0,0141
20—50	1	0.0007
50—100	1	0.0035
100—200	1	0.0039

Astronomisches Rechen-Institut in Heidelberg
Mitteilungen Serie A Nr. 19

Die numerische Integration des *n*-Körper-Problems für Sternhaufen, II.

Von

SEBASTIAN VON HOERNER

Mit 10 Textabbildungen

(Eingegangen am 19. November 1962)

http://cdsads.u-strasbg.fr/abs/1963ZA.....57...47V

S.v. Hoerner,
Z.f.Astroph. 1960, 63

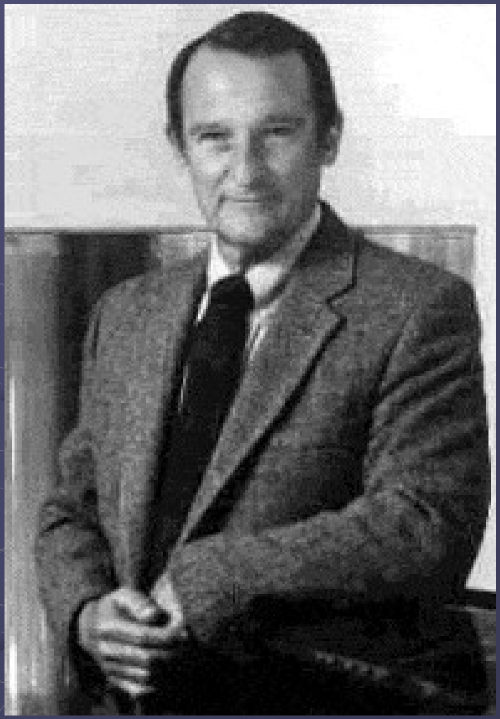
Siemens 2002
N=4,8,12,16 (4 Trx)

N=16,25 (40 Trx)

History

- Seymour Cray (1925-1996)

“father of supercomputing”



CRAY1: Vectorregisters (1976)

160 Mflop, 80 MHz, 8 MByte RAM

CRAY2: (1984)

1Gflop, 120MHz, 2GByte RAM

History

Supercomputer JUGENE IBM Blue Gene At FZ Jülich, Germany

Opening Ceremony June 2008

Computational Science...

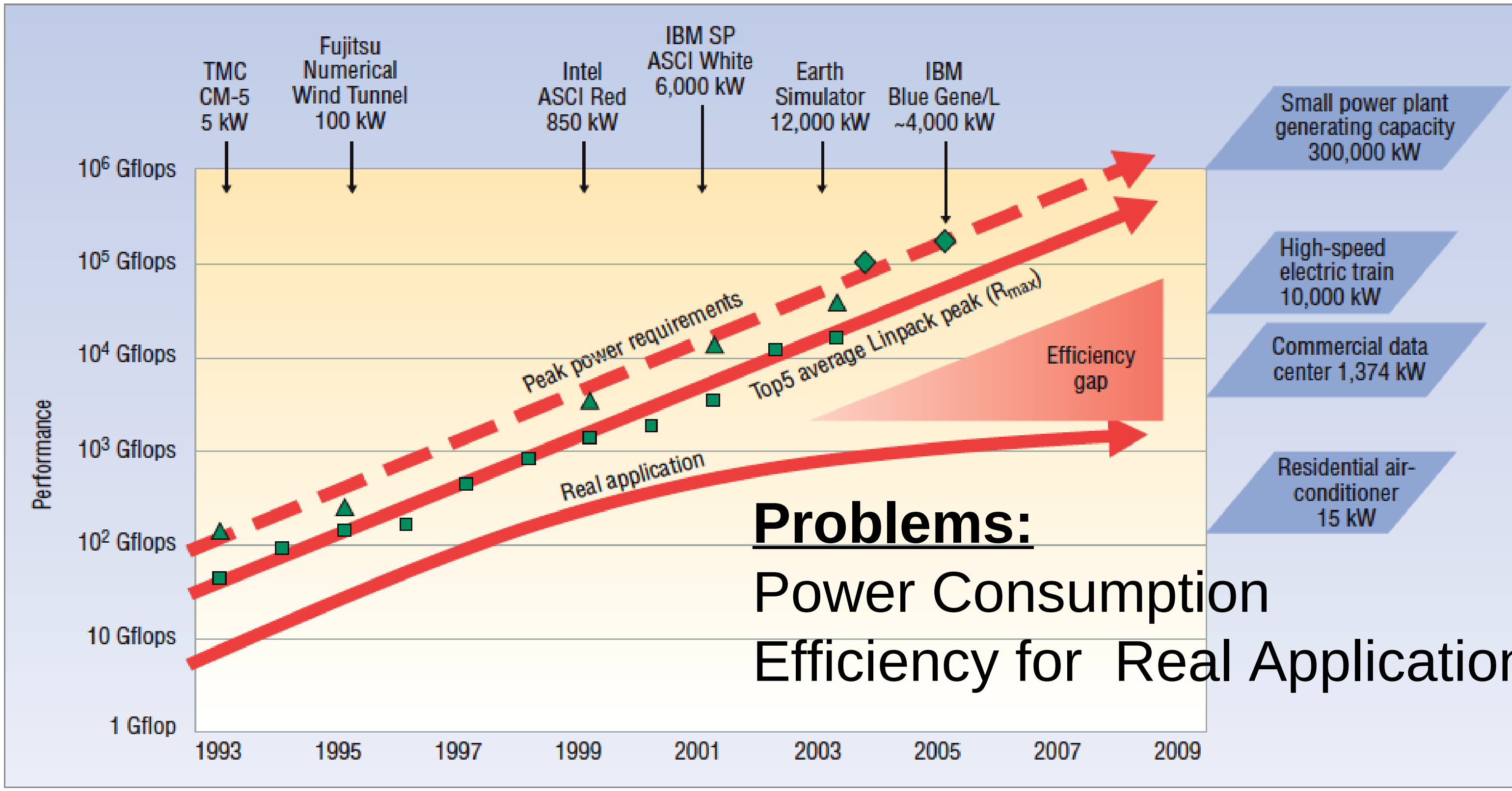
...after von Neumann...

Exaflop/s?

Petaflop/s

Teraflop/s

Gigaflop/s



Problems:

Power Consumption

Efficiency for Real Applications

Figure 1. Rising power requirements. Peak power consumption of the top supercomputers has steadily increased over the past 15 years.

Thanks to Horst Simon, LBNL/NERSC for this diagram.

Special Hardware Accelerators

SPECIAL HARDWARE

CPUs

Central Processing Units

- General Purpose oriented
- 1-12 Cores
- Up to 4 pipes per core using Vector Units
- Fully Programmable, many languages available
- Very well studied
- Max. 125W per processor

GPUs

Graphic Processing Units

- Graphics oriented
- 16-512 Cores
- Massively Parallel Architecture, specialized instructions for parallel processing
- Fully programmable, but limited languages
- Algorithms not fully explored
- Max. 400W per card

FPGAs

Field Programmable Gate Arrays

- Custom designs, best for processing streaming data
- Programmable Logic, Architecture is custom-built for the required application
- Requires extensive knowledge to program, development time is longer than CPUs and GPUs
- Application interface is custom built on each case
- Max. 60W per FPGA

ASICs

Application Specific Integrated Circuits

- Fully custom designs, built for a specific application
- Not flexible, cannot be changed once it is built
- Development is even more specialized than FPGAs
- Power consumption varies with the application, usually best performance per Watt

Slide: Guillermo Marcus

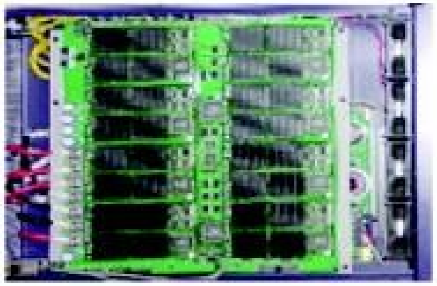
HITS

HARDWARE

GRAPE-6 Gravity/Coulomb Part

- G6 Chip: 0.25μ 2MGate ASIC, 6 Pipelines
- at 90MHz, 31Gflops/chip
- 48Tflops full system (March 2002)
- Plan up to 72Tflops full system (in 2002)
- Installed in Cambridge, Marseille, Drexel, Amsterdam, New York (AMNH), Mitaka (NAO), Tokyo, etc..
  New Jersey, Indiana, Heidelberg

GRAPE-6



1998, 120 Gflops

Developers: Junichiro Makino, Toshiyuki Fukushige, Hiroshi Daisaka, Eiichiro Kokubo, Masaki Koga, Makoto Taiji, Ken Namura

GRAPE-6: Massively-Parallel Special-Purpose Computer for Astrophysical Particle Simulations

Sales information

The Green500 List - November 2010

Listed below are the November 2010 The Green500's energy-efficient supercomputers ranked from 1 to 100.

http://www.green500.org

Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)
1	1684.20	IBM Thomas J. Watson Research Center	NNSA/SC Blue Gene/Q Prototype	38.80
2+	1448.03	National Astronomical Observatory of Japan	GRAPE-DR accelerator Cluster, Infiniband	24.59
2	958.35	GSIC Center, Tokyo Institute of Technology	HP ProLiant SL390s G7 Xeon 6C X5670, Nvidia GPU, Linux/Windows	1243.80
3	933.06	NCSA	Hybrid Cluster Core i3 2.93Ghz Dual Core, NVIDIA C2050, Infiniband	36.00

GPU: NAOC laohu cluster Beijing, China

Heidelberg

Kepler GPU cluster

VolkswagenStiftung

Kepler GPU cluster

12 nodes = 12 x 16 = 192 CPU cores (@ 2 GHz)

12 x 64 GB = 768 GB RAM CPU memory

12 GPUs K20m = 12 x 2496 ~ 30k GPU threads

12 x 4.8 GB ~ 57 GB GPU device memory

4 x Xilinx Virtex-6 FPGA (ML 605)

since beg. 2013 operated.

NVIDIA Volta V100 GPU, 21 billion transistors, 5120 cores (now Ampere A100, 6920 cores, 9.7 Tflops DP, 19.5 SP)

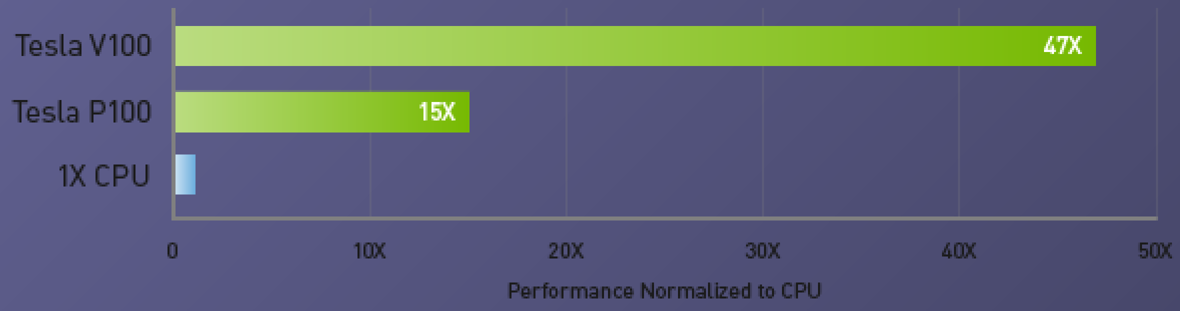
With NVLINK

Without NVLINK

PERFORMANCE with NVIDIA GPU Boost™	DOUBLE-PRECISION 7.8 teraFLOPS	DOUBLE-PRECISION 7 teraFLOPS
	SINGLE-PRECISION 15.7 teraFLOPS	SINGLE-PRECISION 14 teraFLOPS
	DEEP LEARNING 125 teraFLOPS	DEEP LEARNING 112 teraFLOPS
INTERCONNECT BANDWIDTH Bi-Directional	NVLINK 300 GB/s	PCIE 32 GB/s
MEMORY CoWoS Stacked HBM2	CAPACITY 32/16 GB HBM2	
	BANDWIDTH 900 GB/s	
POWER Max Consumption	300 WATTS	250 WATTS

NVIDIA Volta V100 GPU, 21 billion transistors, 5120 cores

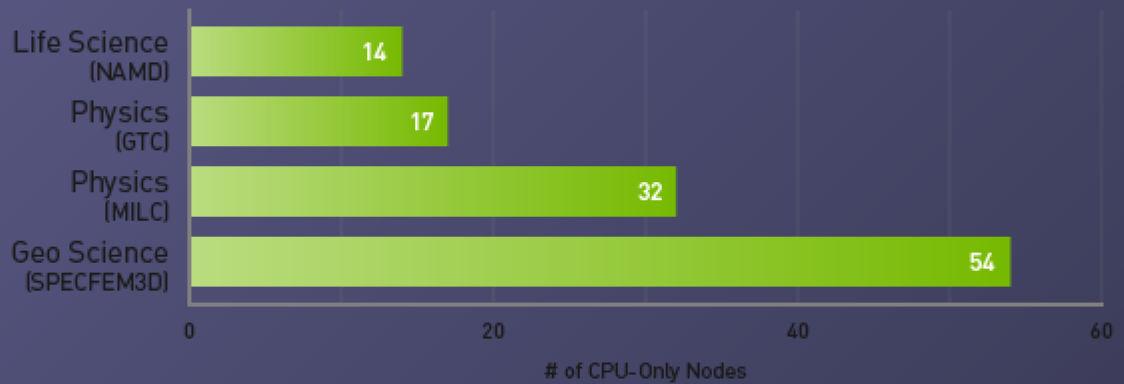
47X Higher Throughput Than CPU Server on Deep Learning Inference



Workload: ResNet-50 | CPU: 1X Xeon E5-2690v4 @ 2.6 GHz | GPU: Add 1X Tesla P100 or V100

1 GPU Node Replaces Up To 54 CPU Nodes

Node Replacement: HPC Mixed Workload



CPU Server: Dual Xeon Gold 6140@2.30GHz, GPU Servers: same CPU server w/ 4x V100 PCIe | CUDA Version: CUDA 9.x| Dataset: NAMD (STMV), GTC (mpi#proc.in), MILC (APEX Medium), SPECFEM3D (four_material_simple_model) | To arrive at CPU node equivalence, we use measured benchmark with up to 8 CPU nodes. Then we use linear scaling to scale beyond 8 nodes.

Top 10 List November 2010

From www.top500.org - list of fastest supercomputers in the world... ... last year Nov. 2010:

Rank	Site	System	
1	National Supercomputing Center in Tianjin China	**Tianhe-1A** - NUDT TH MPP, X5670 2.93Ghz 6C, NVIDIA GPU, FT-1000 8C NUDT	***GPU***
2	DOE/SC/Oak Ridge National Laboratory United States	**Jaguar** - Cray XT5-HE Opteron 6-core 2.6 GHz Cray Inc.	
3	National Supercomputing Centre in Shenzhen (NSCS) China	**Nebulae** - Dawning TC3600 Blade, Intel X5650, NVidia Tesla C2050 GPU Dawning	***GPU***
4	GSIC Center, Tokyo Institute of Technology Japan	**TSUBAME 2.0** - HP ProLiant SL390s G7 Xeon 6C X5670, Nvidia GPU, Linux/Windows NEC/HP	***GPU***
5	DOE/SC/LBNL/NERSC United States	**Hopper** - Cray XE6 12-core 2.1 GHz Cray Inc.	
6	Commissariat a l'Energie Atomique (CEA) France FR	**Tera-100** - Bull bullx super-node S6010/S6030 Bull SA	
7	DOE/NNSA/LANL United States	**Roadrunner** - BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 GHz, Voltaire Infiniband IBM	
8	National Institute for Computational Sciences/University of Tennessee United States	**Kraken XT5** - Cray XT5-HE Opteron 6-core 2.6 GHz Cray Inc.	
9	Forschungszentrum Juelich (FZJ) Germany	**JUGENE** - Blue Gene/P Solution IBM	
10	DOE/NNSA/LANL/SNL United States	**Cielo** - Cray XE6 8-core 2.4 GHz Cray Inc.	

China Grabs Supercomputing Leadership Spot in Latest Ranking of World's Top 500 Supercomputers

Thu, 2010-11-11 22:42

MANNHEIM, Germany; BERKELEY, Calif.; and KNOXVILLE, Tenn.—The 36th edition of the closely watched TOP500 list of the world's most powerful supercomputers confirms the rumored takeover of the top spot by the Chinese Tianhe-1A system at the National Supercomputer Center in Tianjin, achieving a performance level of 2.57 petaflop/s (quadrillions of calculations per second).

NCSA director: GPU is future of supercomputing

by Brooke Crothers

A A Font size | Print | E-mail | Share | 6 comments

Tweet 99 | Share 25 | 2 Digg

The director of the National Center for Supercomputing Applications has seen the future of supercomputing and it can be summed up in three letters: GPU.

Thom Dunning, who directs the NCSA and the Institute for Advanced Computing Applications and Technologies at the famed supercomputing facilities on the campus of University of Illinois at Urbana-Champaign, says high-performance computing will begin to move toward graphics processing units or GPUs. Not coincidentally, **this is exactly what China has done to achieve the world's fastest speeds with its "Tianhe-1A"** supercomputer. That computer combines about 7,000 Nvidia GPUs with 14,000 Intel CPUs: the only hybrid CPU-GPU system in the world of that scale.

Thom Dunning directs the Institute for Advanced Computing Applications and Technologies and the NCSA.

"What we're really seeing in the efforts in China as well as the ones we have in the U.S. is that GPUs are what the future will look like," said Dunning in a phone interview Thursday. "What we're seeing is the beginning of something that's going to be happening all over the world."

NCSA already has a small CPU-GPU hybrid system. "It's something we have been working on for a number of years. We have a CPU-GPU cluster for the NCSA academic community. Made up of Intel CPUs and Nvidia GPUs. A 50 teraflop machine," he said. (Note that **Oak Ridge National Laboratories is also installing a hybrid system now**.)

Intel MIC Hardware

INSPUR, NAOC - 2013.XI.26

icpc ... "-mmic" ... 61 x 4 = 244 x 1.1 GHz omp cores !!!

Full fp64 !!!

Intel MIC Hardware

Intel® Xeon Phi™ Coprocessor Family Reference Table

SKU #	Form Factor, Thermal	Peak Double Precision	Max # of Cores	Clock Speed (GHz)	GDDR5 Memory Speeds (GT/s)	Peak Memory BW	Memory Capacity (GB)	Total Cache (MB)	Board TDP (Watts)	Process
SE10P (special edition)	PCIe Card, Passively Cooled	1073 GF	61	1.1	5.5	352	8	30.5	300	22nm
SE10X (special edition)	PCIe Card, No Thermal Solution	1073 GF	61	1.1	5.5	352	8	30.5	300	
5110P	PCIe Card, Passively Cooled	1011 GF	60	1.053	5.0	320	8	30	225	
3100 Series	PCIe Card, Actively Cooled	>1 TF	Disclosed at 3100 series launch (H1'13)		5.0	240	6	28.5	300	
	PCIe Card, Passively Cooled	> 1 TF			5.0	240	6	28.5	300	

Current Generation:
Knights Landing
14nm

PCIe Card, Actively Cooled

PCIe Card, Passively Cooled

Intel MIC hardware / Recent Processors

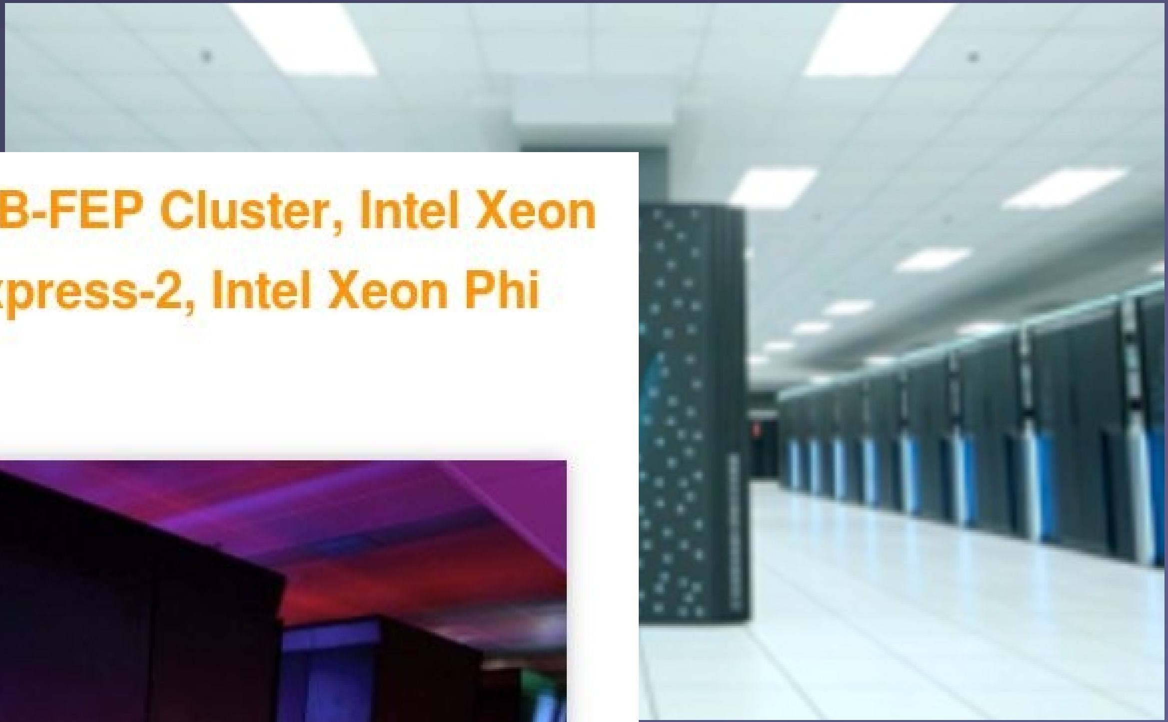
Intel® Xeon Phi™ Processor 7290

- 36 MB L2 Cache
- 72 Cores
- 72 Threads
- 1.70 GHz Max Turbo Frequency

Intel® Xeon Phi™ Processor 7290F

- 36 MB L2 Cache
- 72 Cores
- 72 Threads
- 1.70 GHz Max Turbo Frequency

Supercomputer from China: 96/33 Pflop/s Linpack
Wuxi/Guangzhou/Tianjin National Supercomputing Center
Taihu 10 mill. cores

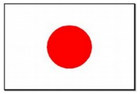
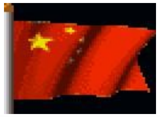


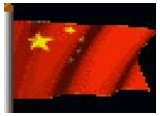
Tianhe-2 (MilkyWay-2) - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P

32000 Intel Xeon 12 core,
48000 Intel Phi Accelerators 57 Core,
now Chinese processor

Test of Taihu planned;
But:
Local cluster with new GPUs at NAOC gives much more resources.

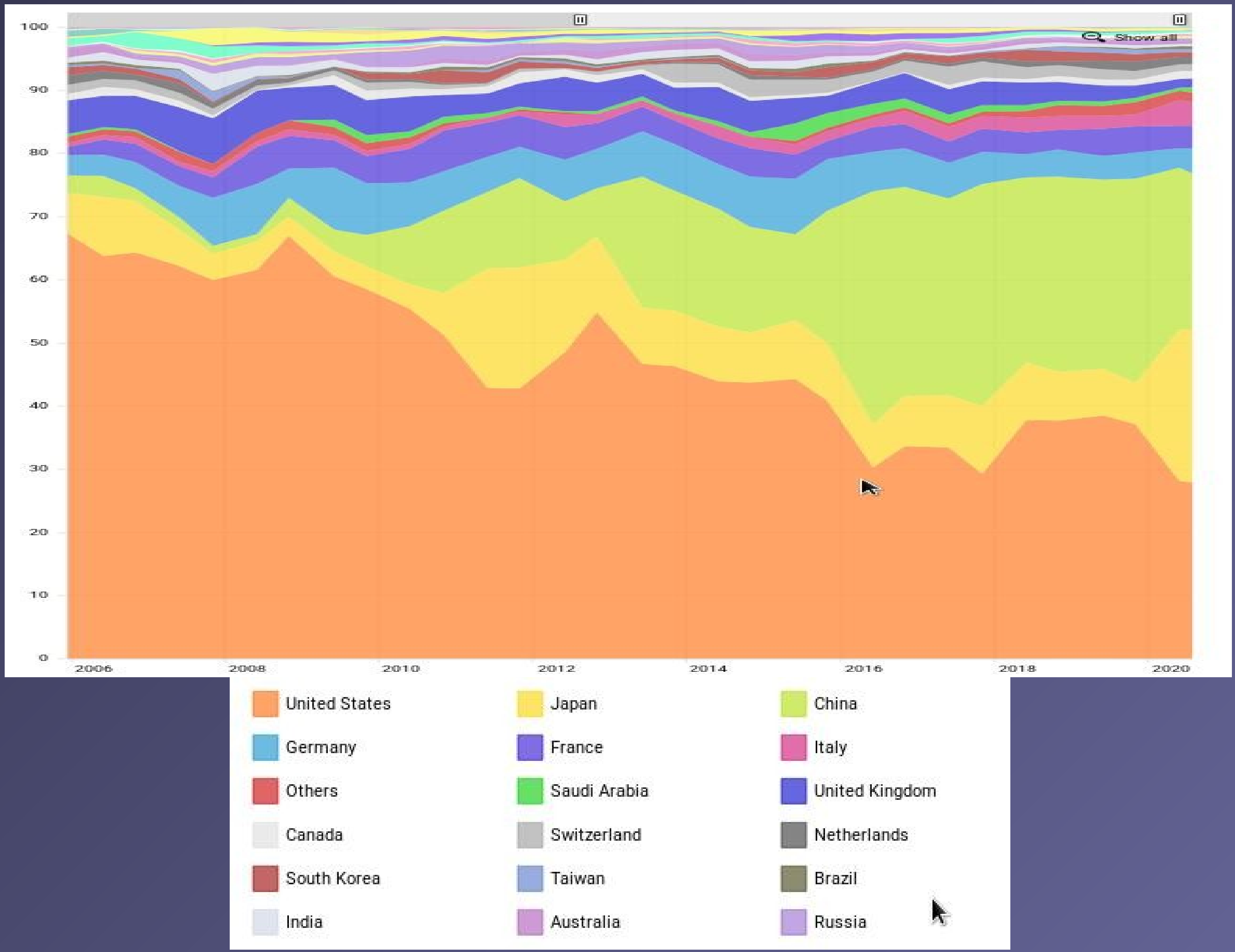
From https://www.top500.org
Nov. 2020 List

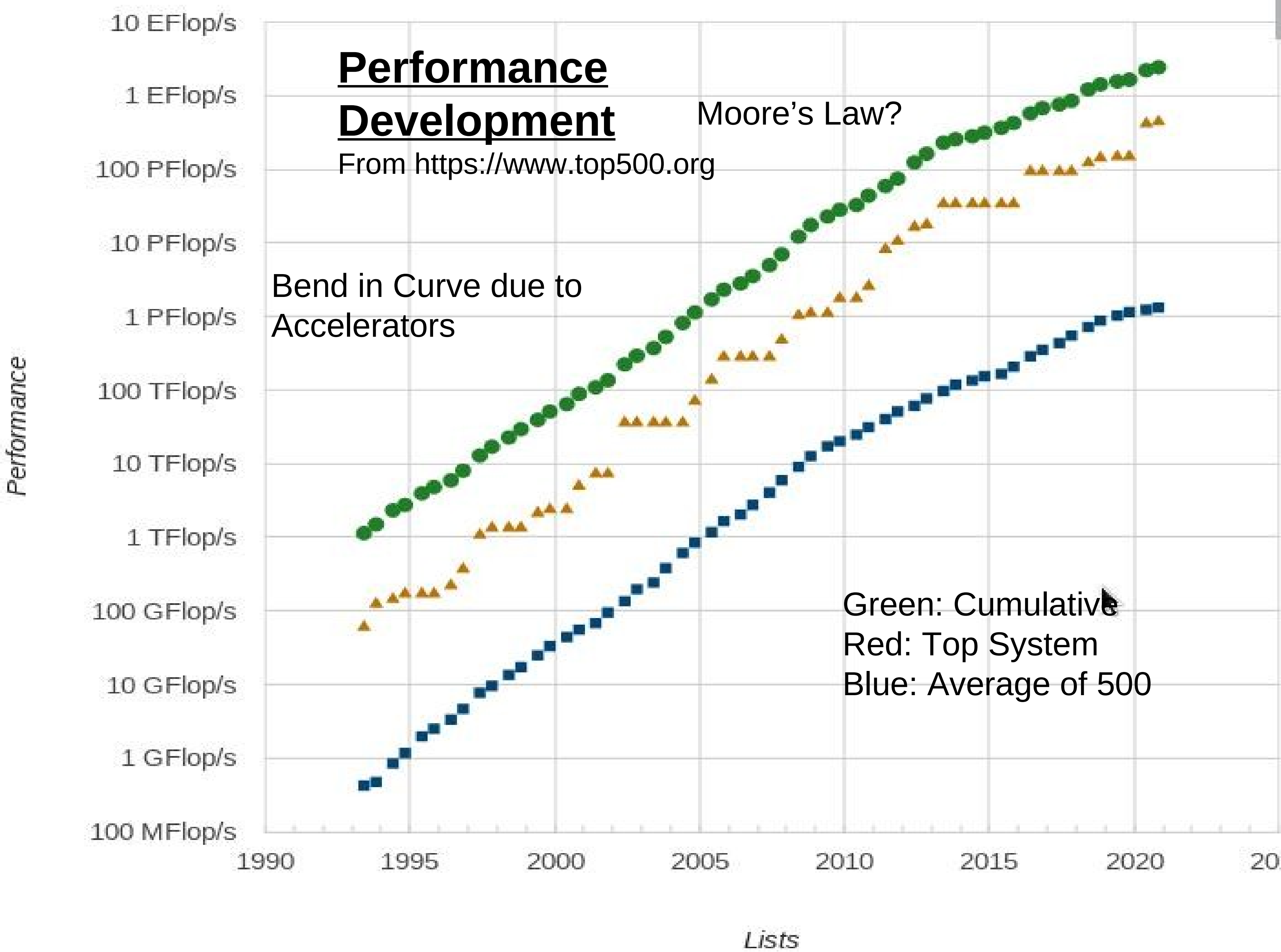
	Rank	System	Cores	(TFlop/s)	(TFlop/s)	(kW)	
 **Japan**	1	**Supercomputer Fugaku** - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan	7,630,848	442,010.0	537,212.0	29,89	***Fujitsu Arm***
USA	2	**Summit** - IBM Power System AC922, IBM POWER9 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	10,09	***GPU NVIDIA Volta***
USA	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	***GPU NVIDIA Volta***
 **China**	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	15,37	***Chinese Processor***
USA	5	**Selene** - NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Mellanox HDR Infiniband, Nvidia NVIDIA Corporation United States	555,520	63,460.0	79,215.0	2,646	***GPU NVIDIA Ampere***

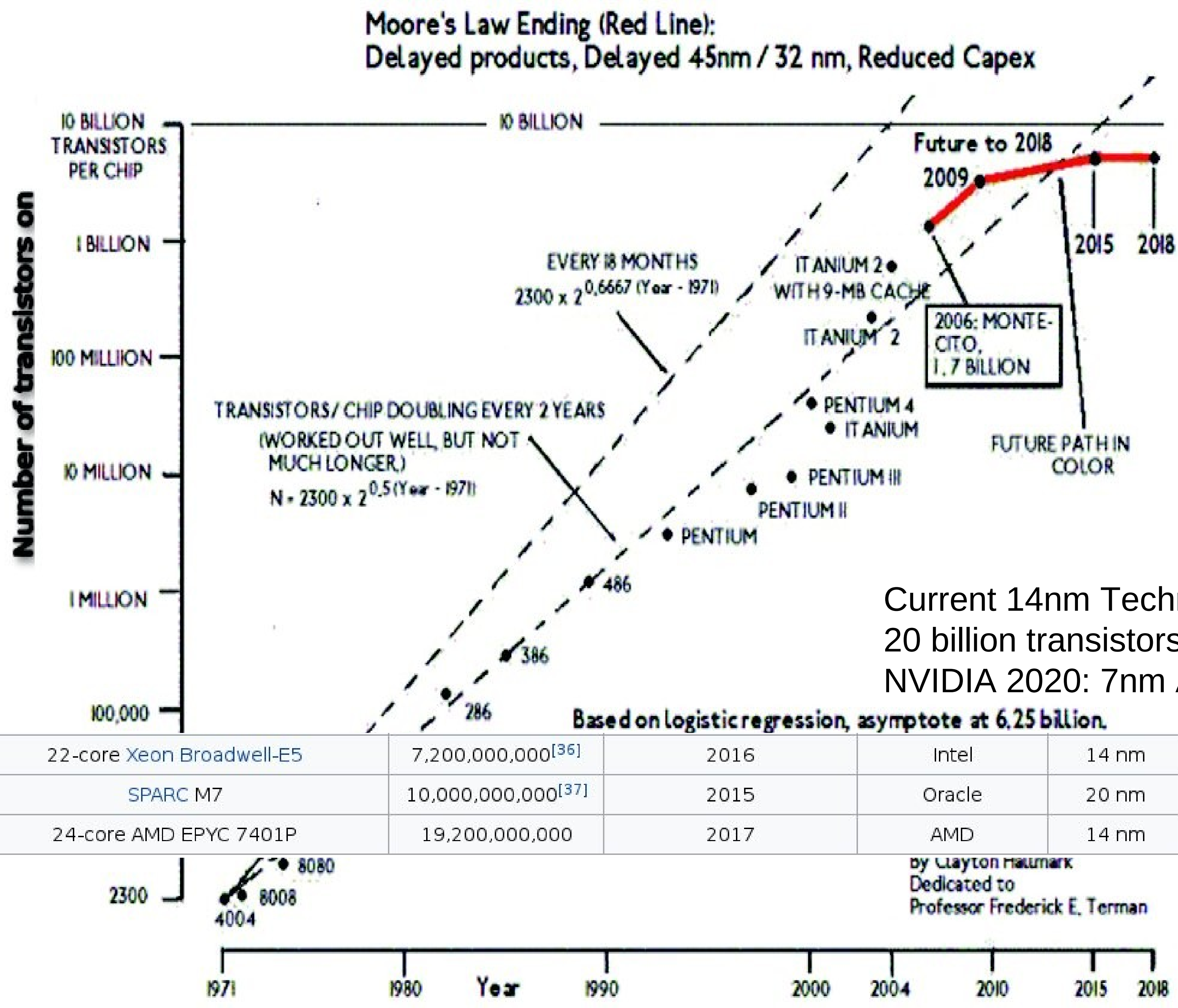
 **China**	6	**Tianhe-2A** - TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000, NUDT National Super Computer Center in Guangzhou China	4,981,760	61,444.5	100,678.7	18,482	***Intel Xeon***
 **Germany**	7	**JUWELS Booster Module** - Bull Sequana XH2000 , AMD EPYC 7402 24C 2.8GHz, NVIDIA A100, Mellanox HDR InfiniBand/ParTec ParaStation ClusterSuite, Atos Forschungszentrum Juelich (FZJ) Germany	449,280	44,120.0	70,980.0	1,764	***GPU NVIDIA Ampere***
Italy	8	**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	***Intel Xeon***
USA	9	**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		***Xeon Platinum***
Saudi-Arabia	10	**Dammam-7** - Cray CS-Storm, Xeon Gold 6248 20C 2.5GHz, NVIDIA Tesla V100 SXM2, InfiniBand HDR 100, HPE Saudi Aramco Saudi Arabia	672,520	22,400.0	55,423.6		***GPU Volta***

Top 500 List November 2020 – Performance Share of Countries

From https://www.top500.org




Performance Development
From https://www.top500.org
Moore's Law?
Bend in Curve due to Accelerators
Green: Cumulative
Red: Top System
Blue: Average of 500
10 EFlop/s
1 EFlop/s
100 PFlop/s
10 PFlop/s
1 PFlop/s
100 TFlop/s
10 TFlop/s
1 TFlop/s
100 GFlop/s
10 GFlop/s
1 GFlop/s
100 MFlop/s
Performance
1990
1995
2000
2005
2010
2015
2020
Lists



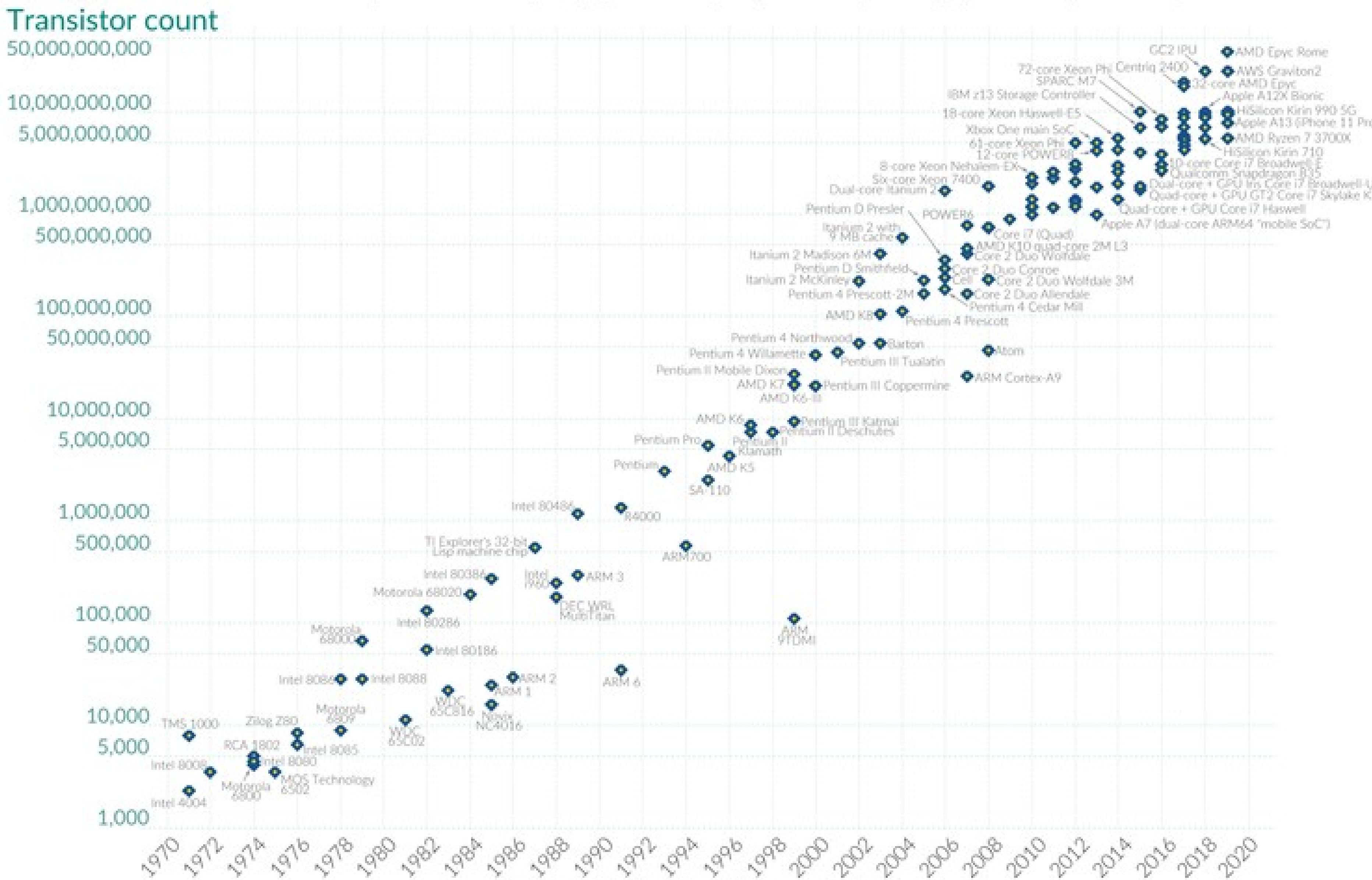
Current 14nm Technology
20 billion transistors
NVIDIA 2020: 7nm Ampere

22-core Xeon Broadwell-E5	7,200,000,000[36]	2016	Intel	14 nm	456 mm²
SPARC M7	10,000,000,000[37]	2015	Oracle	20 nm	
24-core AMD EPYC 7401P	19,200,000,000	2017	AMD	14 nm	195 mm²

Moore's Law: The number of transistors on microchips doubles every two years

Our World in Data

Moore's law describes the empirical regularity that the number of transistors on integrated circuits doubles approximately every two years. This advancement is important for other aspects of technological progress in computing – such as processing speed or the price of computers.



Year in which the microchip was first introduced

GREEN 500 list 2020 Power Efficiency (Gflops/Watts), see also http://www.top500.org/green500

Rank	TOP500 Rank	System	Cores	Rmax (TFlop/s)	Power (kW)	Power Efficiency (GFlops/watts)	
1	170	**NVIDIA DGX SuperPOD** – NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Mellanox HDR Infiniband, Nvidia NVIDIA Corporation United States	19,840	2,356.0	90	26.195	***GPU Ampere***
2	330	**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	1,652.9	65	26.039	***Xeon Platinum***
3	7	**JUWELS Booster Module** – Bull Sequana XH2000 , AMD EPYC 7402 24C 2.8GHz, NVIDIA A100, Mellanox HDR InfiniBand/ParTec ParaStation ClusterSuite, Atos Forschungszentrum Juelich (FZJ) Germany	449,280	44,120.0	1,764	25.008	***GPU Ampere***
4	146	**Spartan2** – Bull Sequana XH2000 , AMD EPYC 7402 24C 2.8GHz, NVIDIA A100, Mellanox HDR Infiniband, Atos Atos France	23,040	2,566.0	106	24.262	***GPU Ampere***
5	5	**Selene** – NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Mellanox HDR Infiniband, Nvidia NVIDIA Corporation United States	555,520	63,460.0	2,646	23.983	***GPU Ampere***
6	239	**A64FX prototype** – Fujitsu A64FX, Fujitsu A64FX 48C 2GHz, Tofu interconnect D, Fujitsu Fujitsu Numazu Plant Japan	36,864	1,999.5	[illegible]	[illegible]	***Fujitsu Arm***
7	29	**AiMOS** – IBM Power System AC922, IBM POWER9 20C 3.45GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM Rensselaer Polytechnic Institute Center for Computational Innovations (CCI) United States	130,000	8,339.0	512	16.285	***GPU Volta***
8	8	**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	2,252	15.740	***GPU Volta***
9	458	**Satori** – IBM Power System AC922, IBM POWER9 20C 2.4GHz, Infiniband EDR, NVIDIA Tesla V100 SXM2, IBM MIT/MGHPCC Holyoke, MA United States	23,040	1,464.0	94	15.574	***GPU Volta***
10	1	**Supercomputer Fugaku** – Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan	7,630,848	442,010.0	29,899	15.418	***Fujitsu Arm***