

中国科学院国家天文台

International Astronomical Observatories, CAS



ZENTRUM FÜR ASTRONOMIE

Univ. Heidelberg

UNIVERSITÄT HEIDELBERG
Zukunft. Seit 1386.



Introduction to GPU Accelerated Computing:

1. History of Computer Architecture Many-Core, GPU, and more...

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

Picture: Xishuangbanna, Yunnan, China by R.Sp.



北京大学
PEKING UNIVERSITY

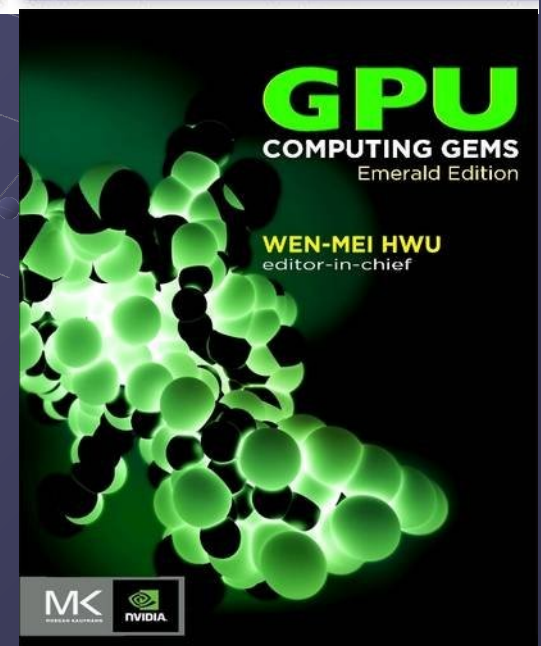
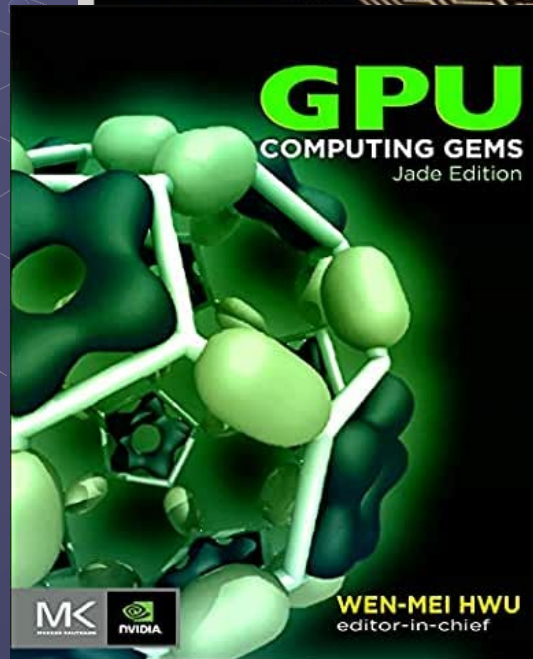
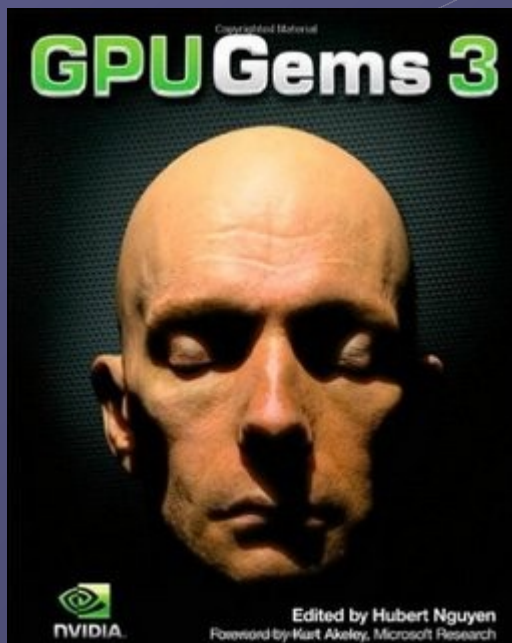
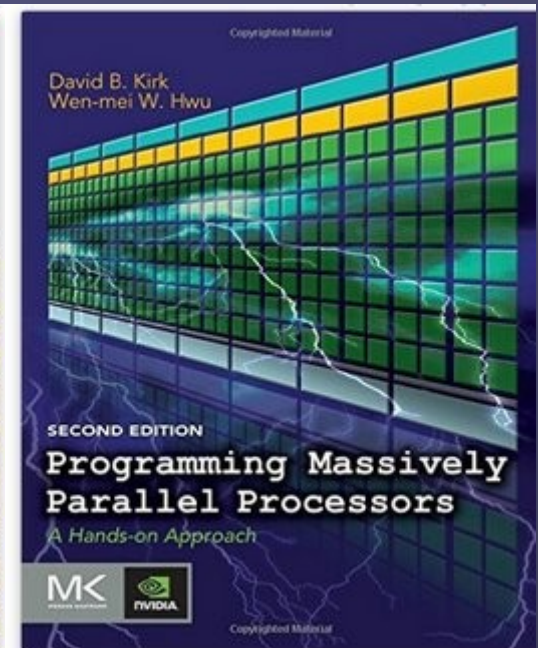
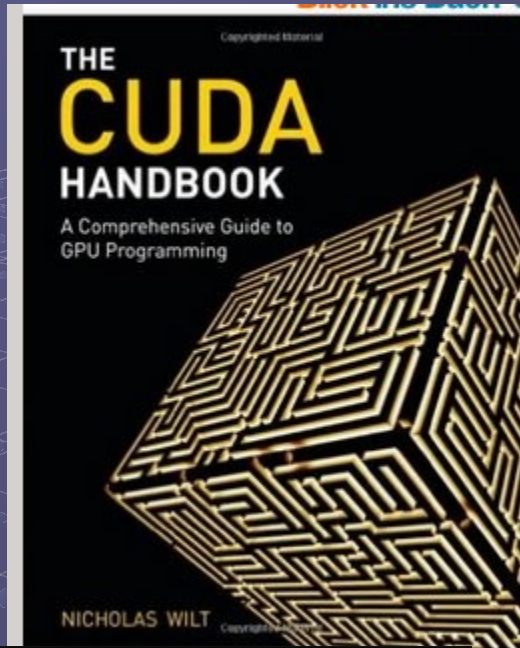
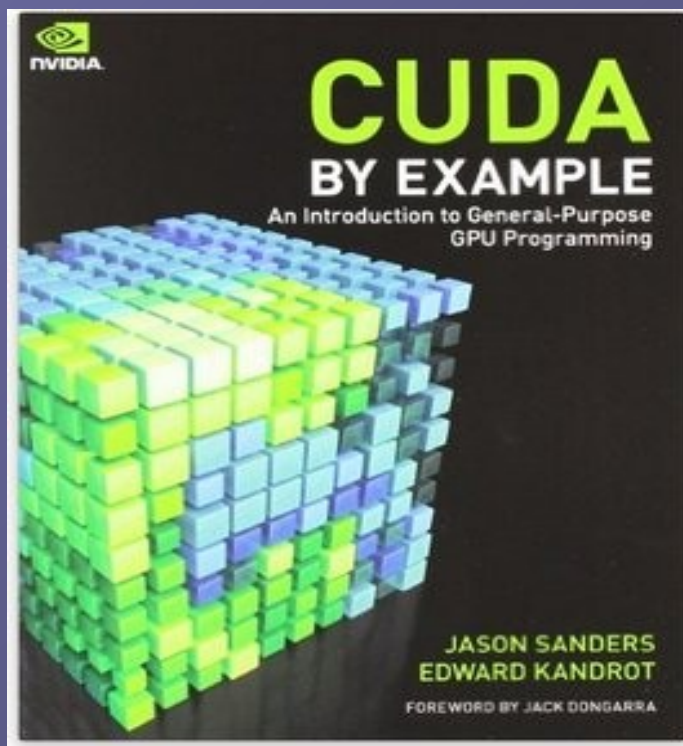
Introduction to GPU Accelerated Computing

Feb. 21-25, 2022

Table of Contents (subject to adjustment/change):

1. Monday morning: General Introduction Computer Architecture, Many-Core, GPU and others..., Access...
2. Monday afternoon: Access to kepler, CUDA Hello, GPU Properties, Simple Add, Vector Add
3. Tuesday morning: More on GPU Software and Hardware
4. Tuesday afternoon: CUDA More Vector Add, Scalar Products, Using Blocks and Threads
5. Wednesday morning: Parallelization and Amdahl's Law, GPU Acceleration, Future Architecture
6. Wednesday Afternoon: Events, Histograms, Matrix Multiplication
7. Thursday Morning: Astrophysical N-Body Code
8. Thursday Afternoon: Astrophysical Parallel N-Body Code Using MPI and GPU
9. Friday Morning: Continue Histograms, Wrap-Up, Q+A, Other Lectures (Wen-Mei Hwu)

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



LUMA METALL ^W ₁₀

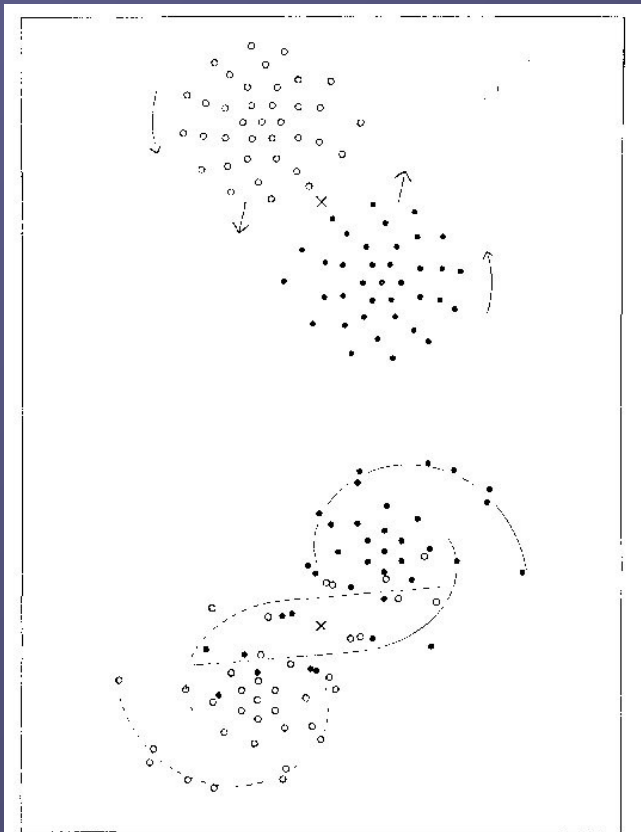


FIG. 4b

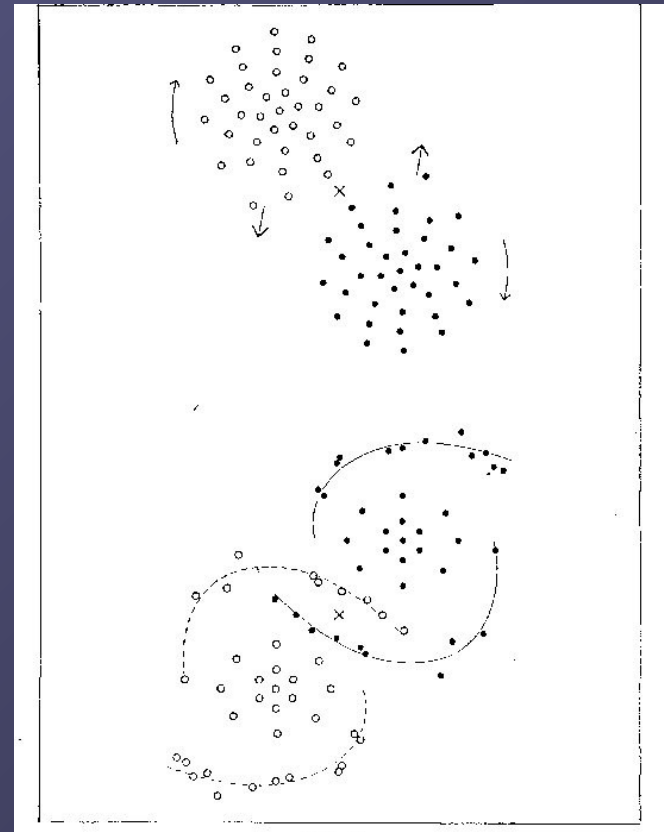


FIG. 4a

HARDWARE

...before von Neumann...

● Konrad Zuse (1910-1995) Berlin



Invented freely programmable Computer



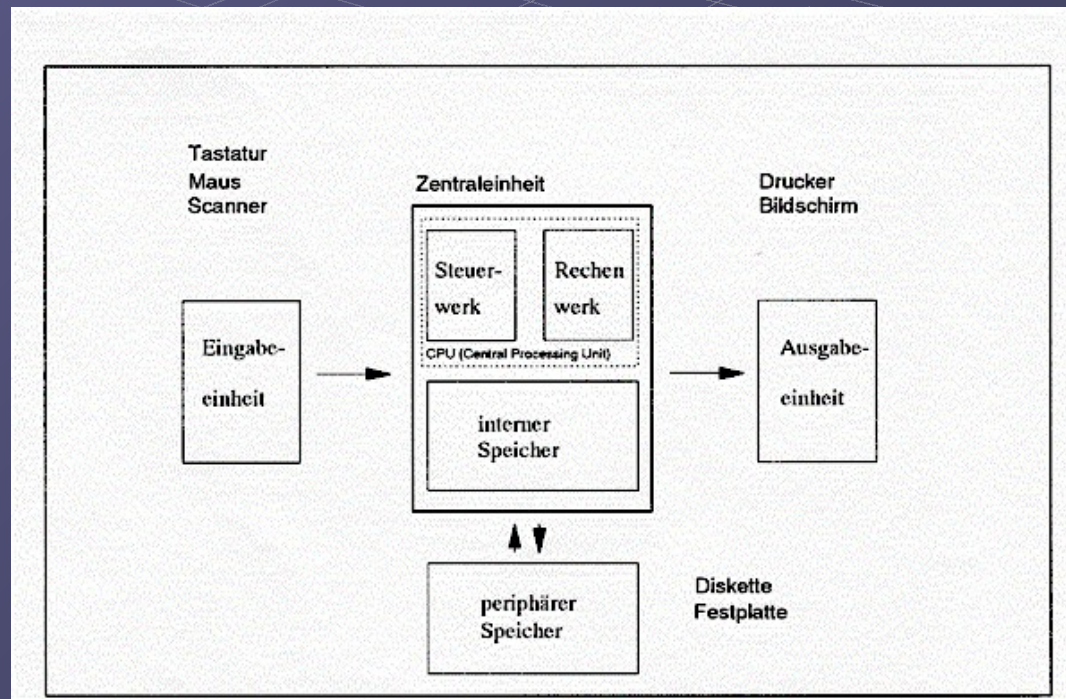
Z1 in parental flat 1936

HARDWARE

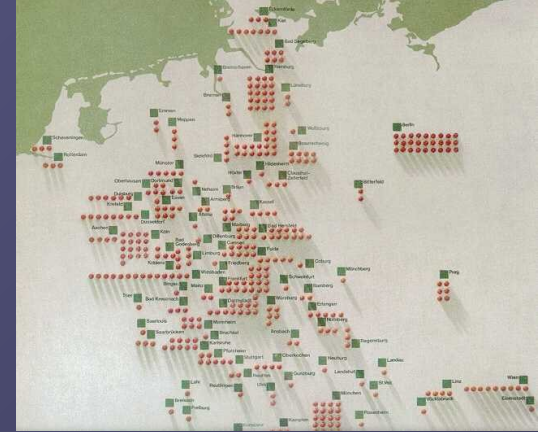
- John von Neumann (1903-1957)

Born Budapest, Lecturer Berlin, since 1930 Princeton Univ.

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



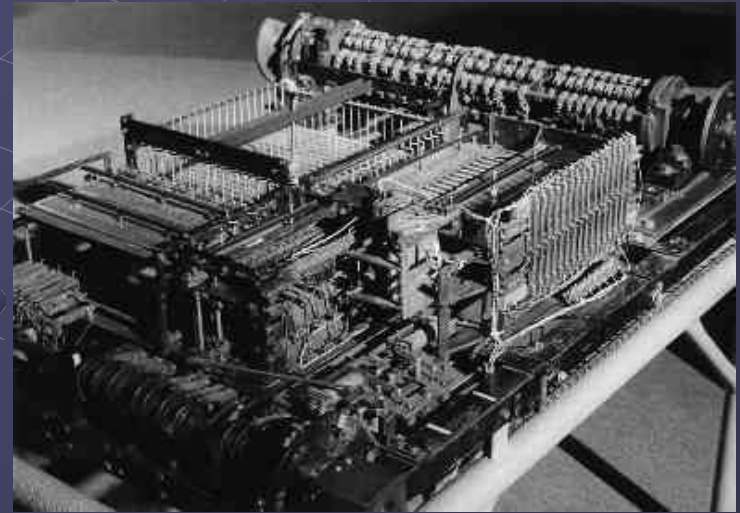
History



**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)

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>

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

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

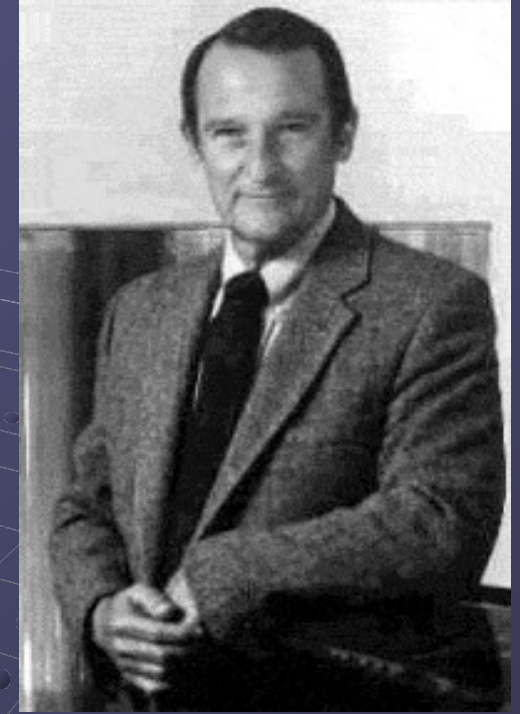
Siemens 2002
N=4,8,12,16 (4 Trx)

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

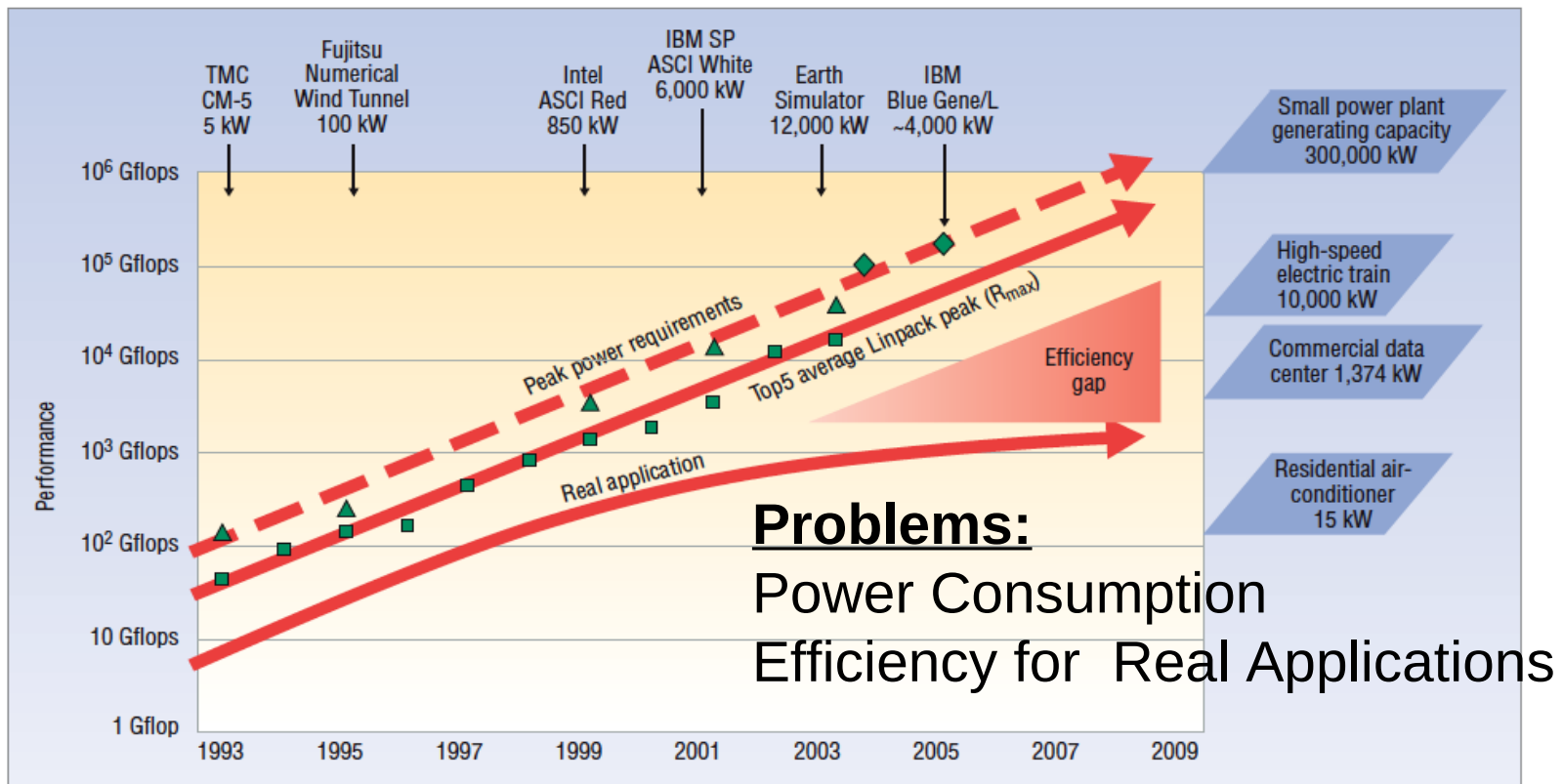


Figure 1. Rising power requirements. Peak power consumption of the top supercomputers has steadily increased over the past 15 years. Thanks to Horst Simon, LBNL/NERSC for this diagram.

GPU Computing

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

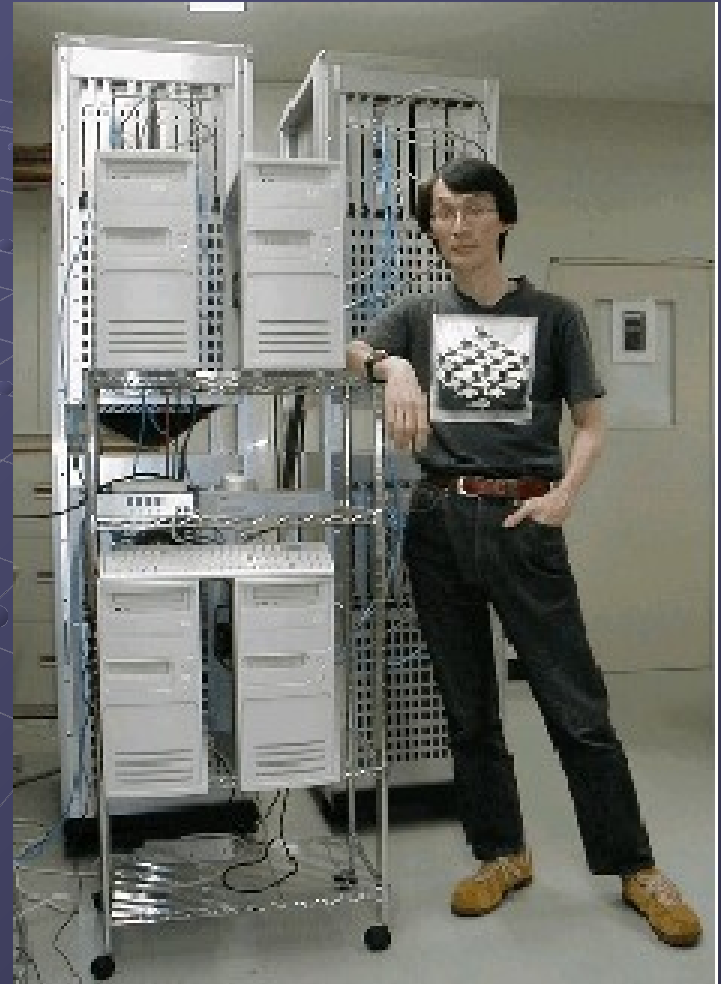
Slide: Guillermo Marcus



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>

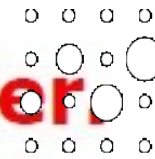
Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)
1	1664.20	IBM Thomas J. Watson Research Center	NNSA/SC Blue Gene/Q Prototype	38.80
2+	1448.03	National Astronomical Observatory of Japan	GRAPE-DR accelerator Cluster, Infiniband	24.59
2	958.35	GSIC Center, Tokyo Institute of Technology	HP ProLiant 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)

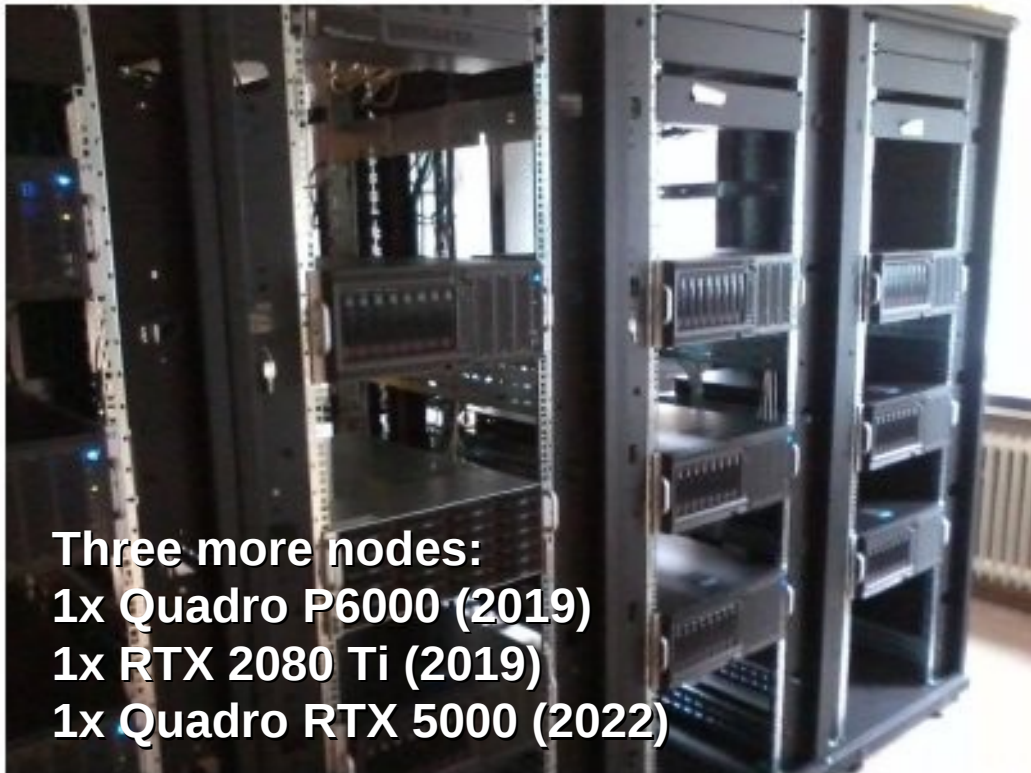
12 x 64 GB = 768 GB RAM CPU memory

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.



Three more nodes:

1x Quadro P6000 (2019)

1x RTX 2080 Ti (2019)

1x Quadro RTX 5000 (2022)

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



With NVLINK

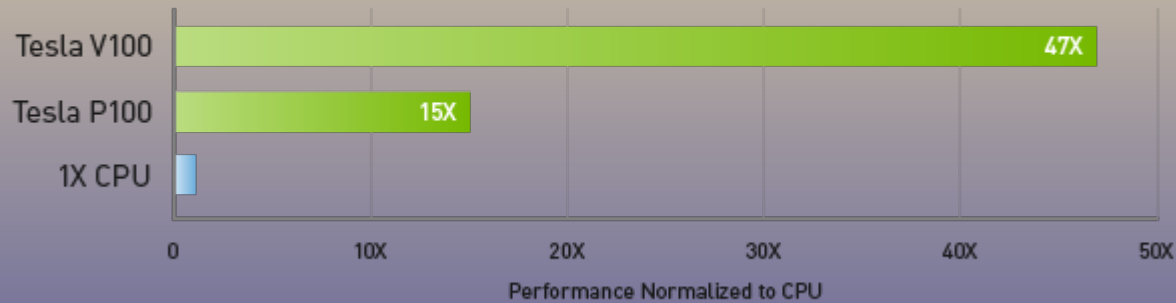
Without NVLINK



PERFORMANCE with NVIDIA GPU Boost*	DOUBLE-PRECISION 7.8 _{tera} FLOPS	DOUBLE-PRECISION 7 _{tera} FLOPS
SINGLE-PRECISION 15.7 _{tera} FLOPS	SINGLE-PRECISION 14 _{tera} FLOPS	
DEEP LEARNING 125 _{tera} FLOPS	DEEP LEARNING 112 _{tera} FLOPS	
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

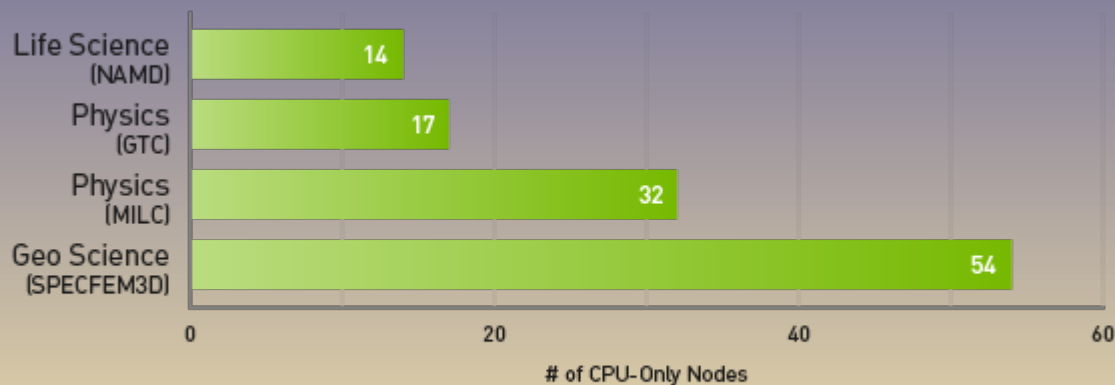
47X Higher Throughput Than CPU Server on Deep Learning Inference



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



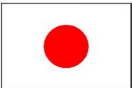

From www.top500.org - list of fastest

supercomputers in the world...
... last year Nov. 2010:

► 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).

1	National Supercomputing Center in Tianjin China		Tianhe-1A - NUDT TH MPP, X5670 2.93Ghz 6C, NVIDIA GPU, FT-1000 8C NUDT	<u>GPU</u>
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	<u>GPU</u>
4	GSIC Center, Tokyo Institute of Technology Japan		TSUBAME 2.0 - HP ProLiant SL390s G7 Xeon 6C X5670, Nvidia GPU, Linux/Windows NEC/HP	<u>GPU</u>
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.	

NCSA director: GPU is future of supercomputing

by Brooke Crothers



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.

"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.**)



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

Intel MIC Hardware

INSPUR, NAOC - 2013.XI.26



icpc ... "-mmic" ... $61 \times 4 = 244 \times 1.1$ GHz omp cores !!!
Full fp64 !!!

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





From <https://www.top500.org/>
Nov. 2021 List



Japan



USA



USA



China



USA

Rank	System	Cores	(TFlop/s)	(TFlop/s)	(kW)
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	29,899
		<u>Fujitsu Arm</u>			
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
		<u>GPU NVIDIA Volta</u>			
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
		<u>GPU NVIDIA Volta</u>			
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
		<u>Chinese Processor</u>			
5	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	70,870.0	93,750.0	2,589
		<u>GPU NVIDIA Ampere</u>			



USA

6	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

7	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
---	---	-----------	----------	-----------	--------

Chinese Matrix-2000 Processor



Germany

8	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

9	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
---	---	---------	----------	----------	-------

GPU NVIDIA Volta



USA

10	Voyager-EUS2 - ND96amsr_A100_v4, AMD EPYC 7V12 48C 2.45GHz, NVIDIA A100 80GB, Mellanox HDR Infiniband, Microsoft Azure Azure East US 2 United States	253,440	30,050.0	39,531.2	
----	---	---------	----------	----------	--

GPU NVIDIA Ampere

RIKEN, Kobe, JAPAN

FUGAKU



Nature's Secrets

富岳

Mt. Fuji

The world's fastest Super Computer 2020 /2021



source :nytimes



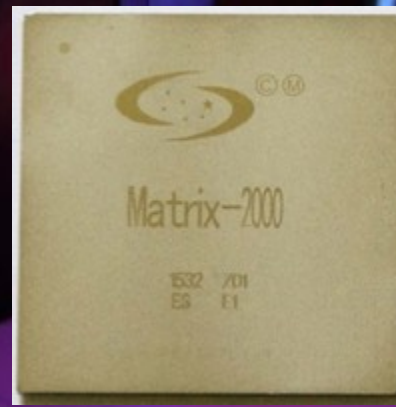
Fugaku extends its reign as champion of supercomputers

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** (Xeon Phi replaced by Chinese Matrix-2000)
Total: ~5 million cores



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



The Matrix-2000 features
128 RISC cores operating
at 1.2 GHz achieving
2.46 / 4.92 TFLOPS
(DP/SP) with a peak power
dissipation of 240 W

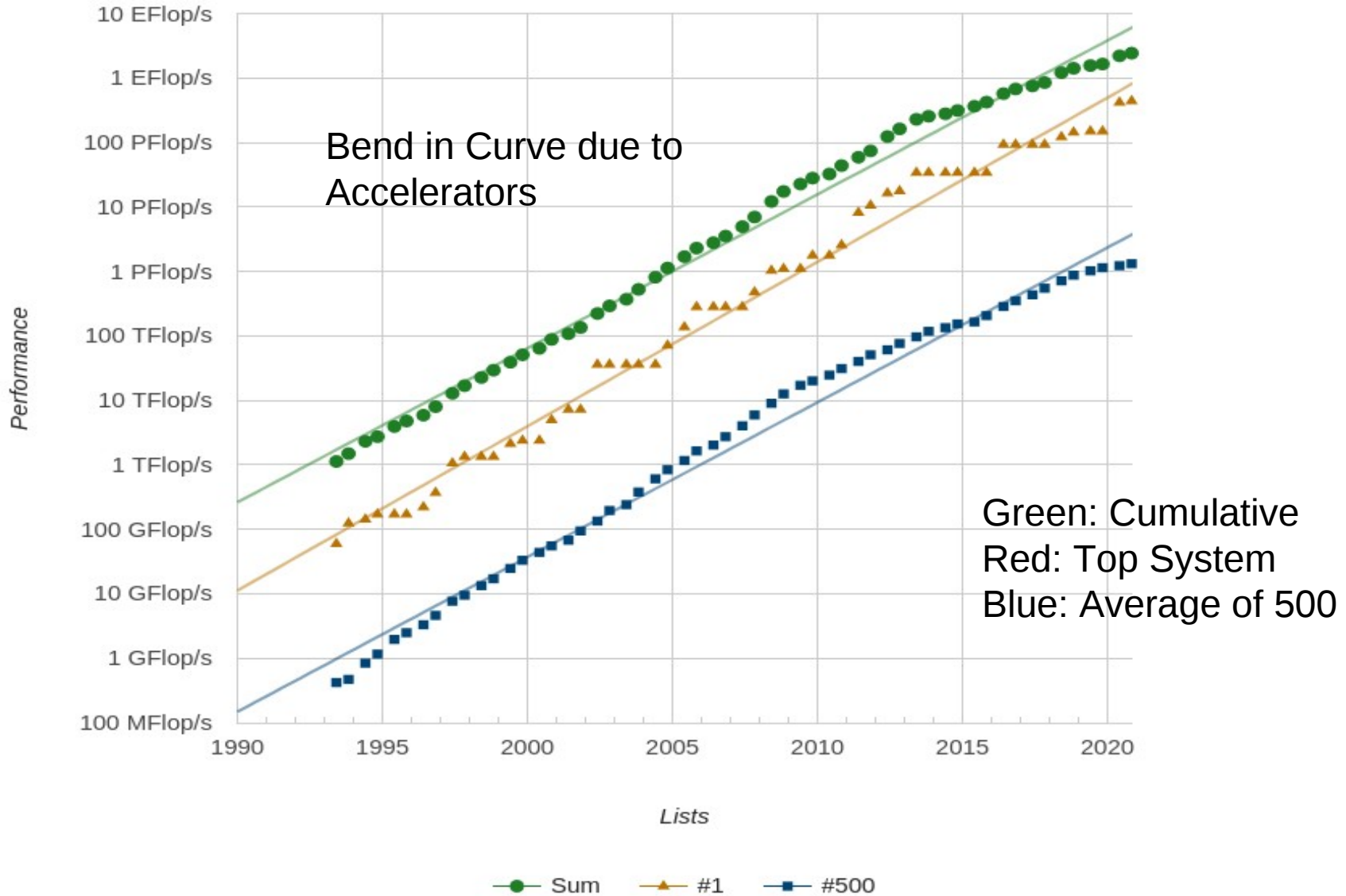


~450.000 cores, NVIDIA Ampere A100 GPU
~ 70 Pflop/s SP ~ 44 Pflop/s DP

Jülich Wizard for European Leadership Science



Projected Performance Development



GREEN 500 list 2020 Power Efficiency (Gflops/Watts), see also top500 webpage

Rank	TOP500 Rank	System	Cores	Rmax [TFlop/s]	Power [kW]	Power Efficiency [GFlops/watts]
1	301	MN-3 - MN-Core Server, Xeon Platinum 8260M 24C 2.4GHz, Preferred Networks MN-Core, MN-Core DirectConnect, Preferred Networks Japan	1,664	2,181.2	55	39.379
						<u>MN-500 special Japan</u>
2	291	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	2,274.1	63	32.882
						<u>GPU Ampere</u>
3	295	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	2,255.0	72	31.538
						<u>GPU Ampere</u>
4	280	Wilkes-3 - PowerEdge XE8545, AMD EPYC 7763 64C 2.45GHz, NVIDIA A100 80GB, Infiniband HDR200 dual rail, DELL EMC University of Cambridge United Kingdom	26,880	2,287.0	74	30.797
						<u>GPU Ampere</u>
5	30	HiPerGator AI - NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Infiniband HDR, Nvidia University of Florida United States	138,880	17,200.0	583	29.521
						<u>GPU Ampere</u>
6	403	Snetillus Phase 1 GPU - ThinkSystem SD650-N V2, Xeon Platinum 8360Y 36C 2.4GHz, NVIDIA A100 SXM4 40 GB, Infiniband HDR, Lenovo SURF Netherlands	6,480	1,818.0	63	29.046
						<u>GPU Ampere</u>
7	5	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	70,870.0	2,589	27.374
						Top 500 #5
						<u>GPU Ampere</u>
8	71	Karolina, GPU partition - Apollo 6500, AMD EPYC 7763 64C 2.45GHz, NVIDIA A100 SXM4 40 GB, Infiniband HDR200, HPE IT4Innovations National Supercomputing Center, VSB-Technical University of Ostrava Czechia	71,424	6,752.0	311	27.213
						<u>GPU Ampere</u>
9	45	Meluxina - Accelerator Module - BullSequana XH2000, AMD EPYC 7452 32C 2.35GHz, NVIDIA A100 40GB, Mellanox HDR InfiniBand/ParTec ParaStation ClusterSuite, Atos LuxProvide Luxembourg	99,200	10,520.0	390	26.957
						<u>GPU Ampere</u>
10	262	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
						<u>GPU Ampere</u>

Juwels Booster: #11 / Fugaku #26