he SILK ROAD PROJECT at NAOC/
丝绸之路计划

ZAH
ARI ITA LSW

ZENTRUM FÜR ASTRONOMIE
Univ. Heidelberg

UNIVERSITÄT HEIDELBERG
Zukunft. Seit 1386.

Introduction to GPU Accelerated Computing

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
https://silkroad.zah.uni-heidelberg.de

VolkswagenStiftung

spurzem@ari.uni-heidelberg.de
spurzem@nao.cas.cn

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

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

北京大学
PEKING UNIVERSITY

Introduction to GPU Accelerated Computing
Feb. 12-16, 2024

Table of Contents (subject to adjustment/change):

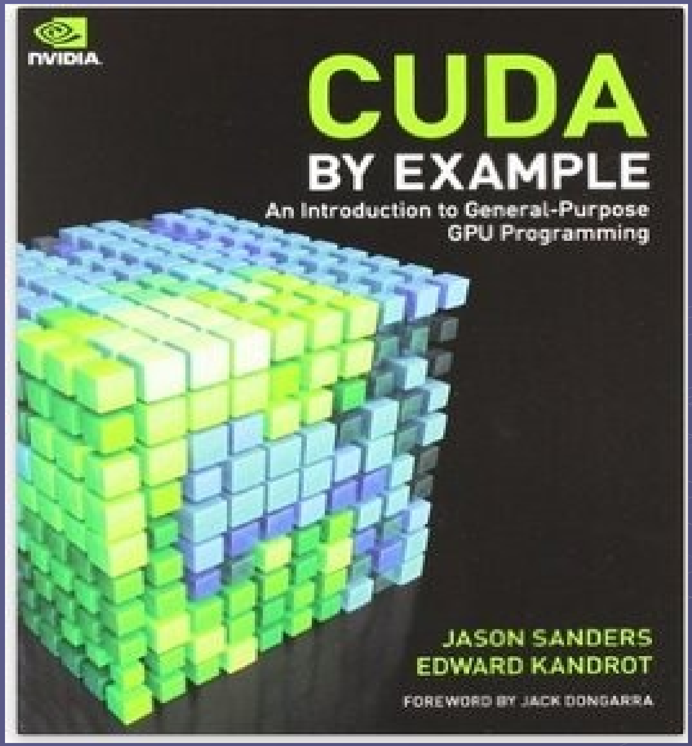
1. Monday morning: General Introduction Computer Architecture, Many-Core, GPU and others…, Access...
2. Monday afternoon: Access to kepler, CUDA Hello, GPU Properties, Simple Add, Vector Add
3. Tuesday morning: More on GPU Software and Hardware
4. Tuesday afternoon: CUDA More Vector Add, Scalar Products, 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)

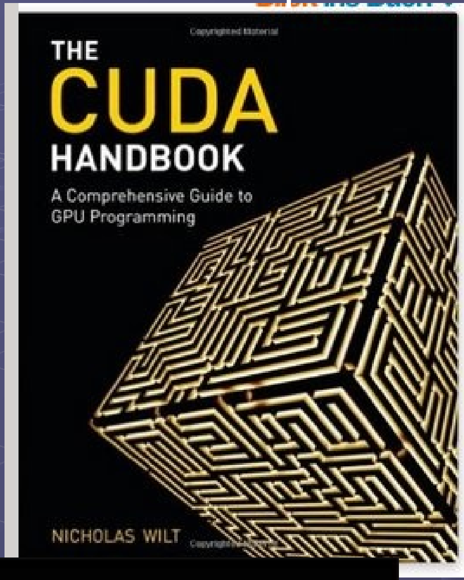
Observations (Experiment)

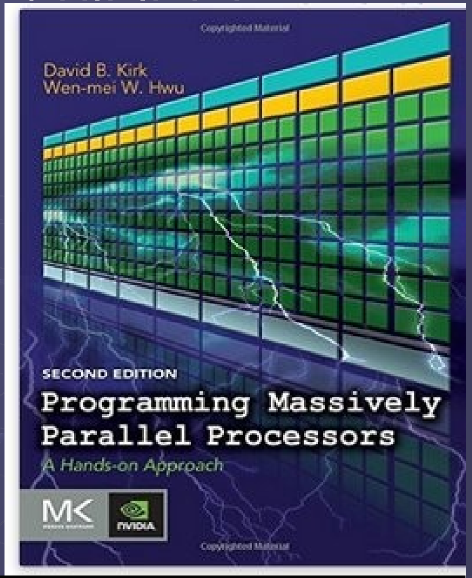
Theory

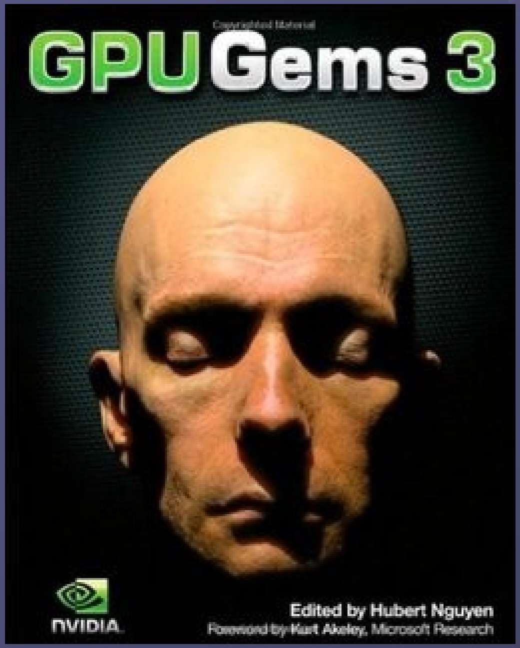
Computational Physics

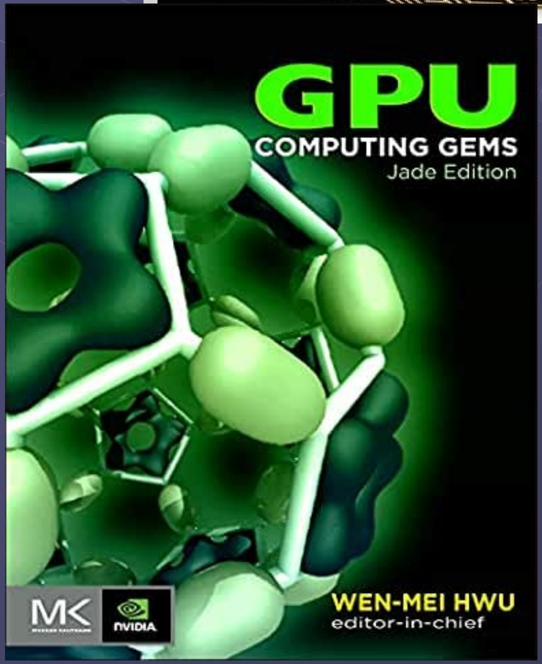
Literature: why NVIDIA? CUDA … ?
easy to learn!! runs on our training system kepler
future? SYCL/openCL? HIP / HIPIFY ?

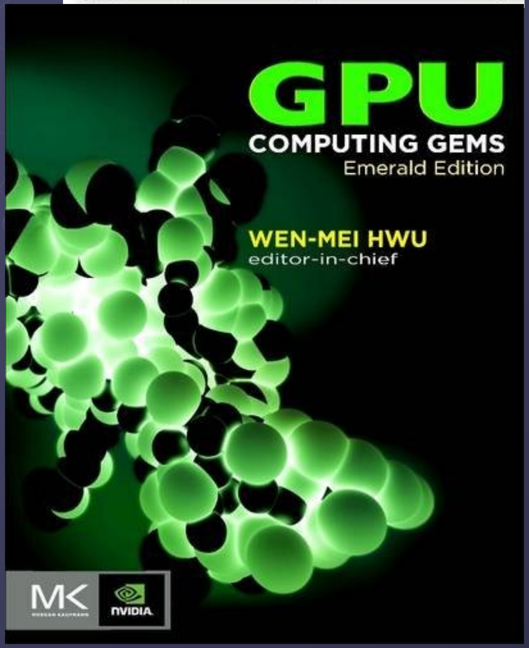












GPU Computing

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

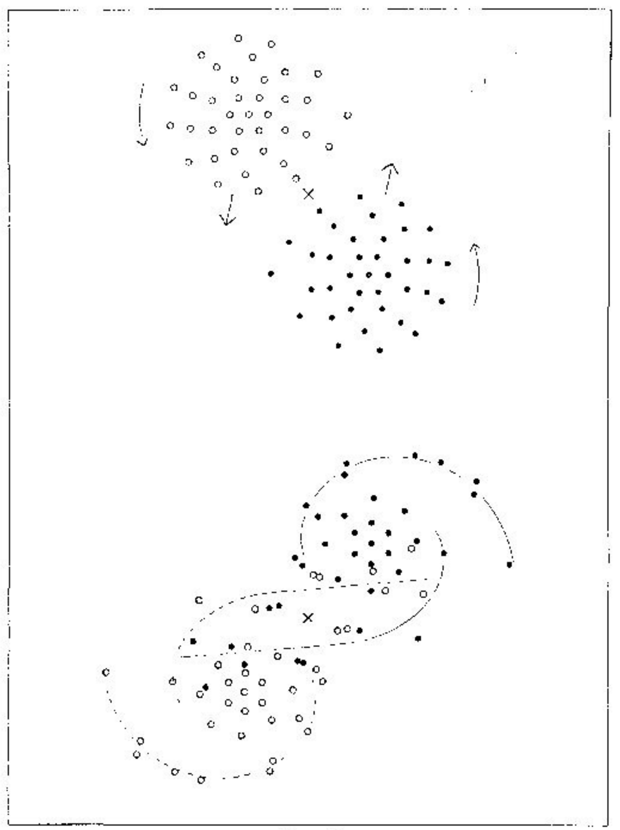


FIG. 4b

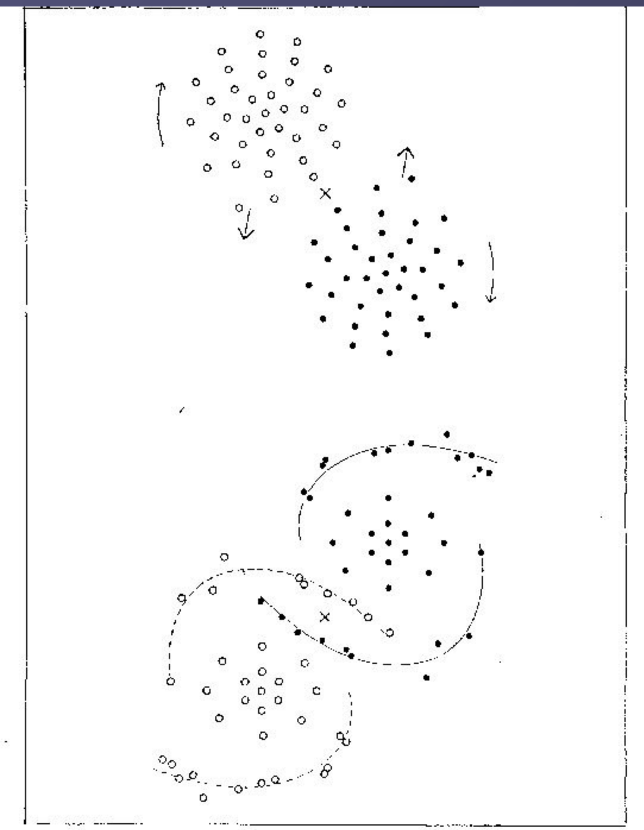


FIG. 4a

HARDWARE

...before von Neumann...

- Konrad Zuse (1910-1995) Berlin

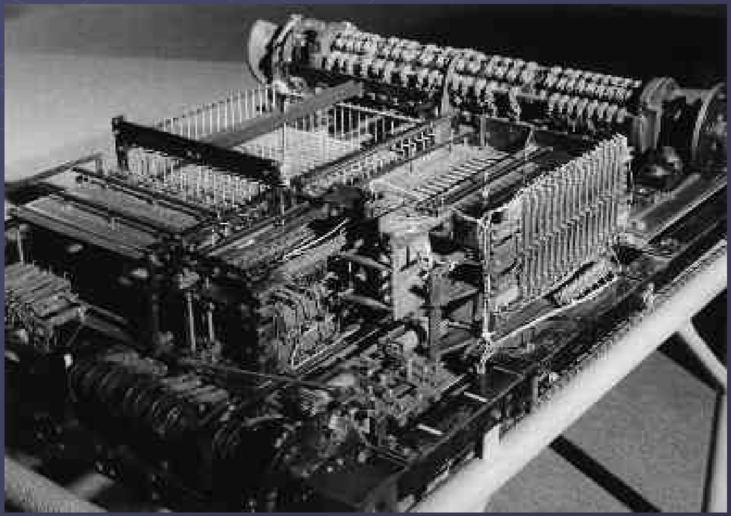
Invented freely programmable Computer

Z1 in parental flat 1936

History

Zuse Z4: 1944 Berlin, 1950 Zürich, 1954 Frankreich

1959 Deutsches Museum München



Computing Speed 0.03 MHz

Memory 256 byte

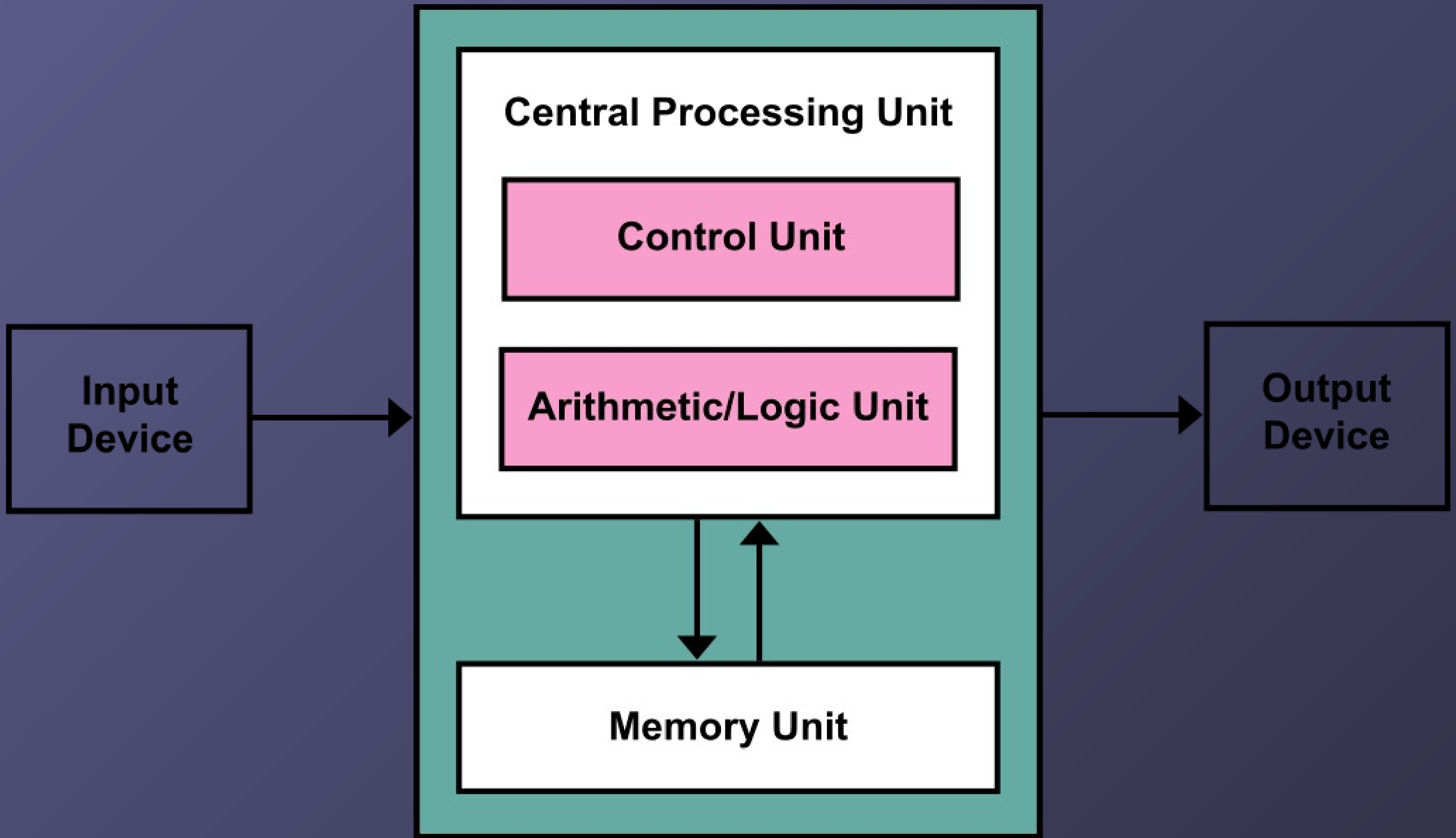
HARDWARE

- John von Neumann (1903-1957)

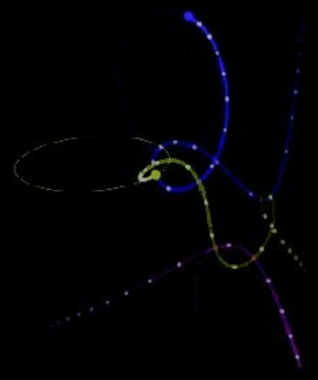
Born Budapest, Lecturer Berlin, since 1930 Princeton Univ.

Fundamental Architecture of an electronic computing device(1946)

Source: https://en.wikipedia.org/wiki/Von_Neumann_architecture#/media/File:Von_Neumann_Architecture.svg



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)

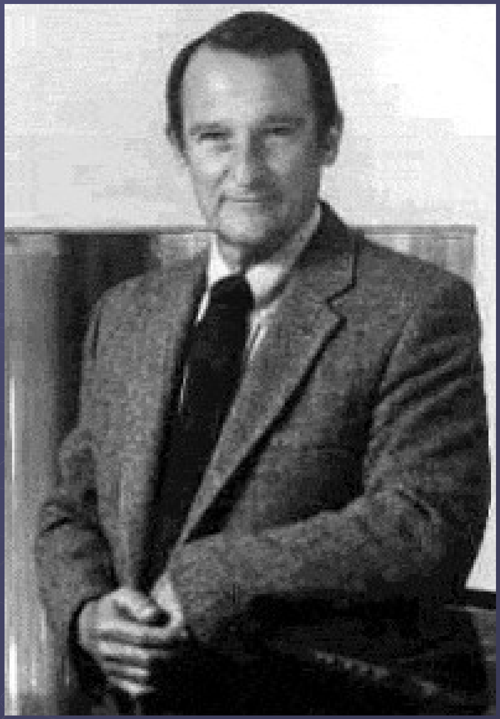
N=16,25 (40 Trx)

History

- Seymour Cray (1925-1996)

"father of supercomputing"

https://en.wikipedia.org/wiki/Women_in_computing



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...

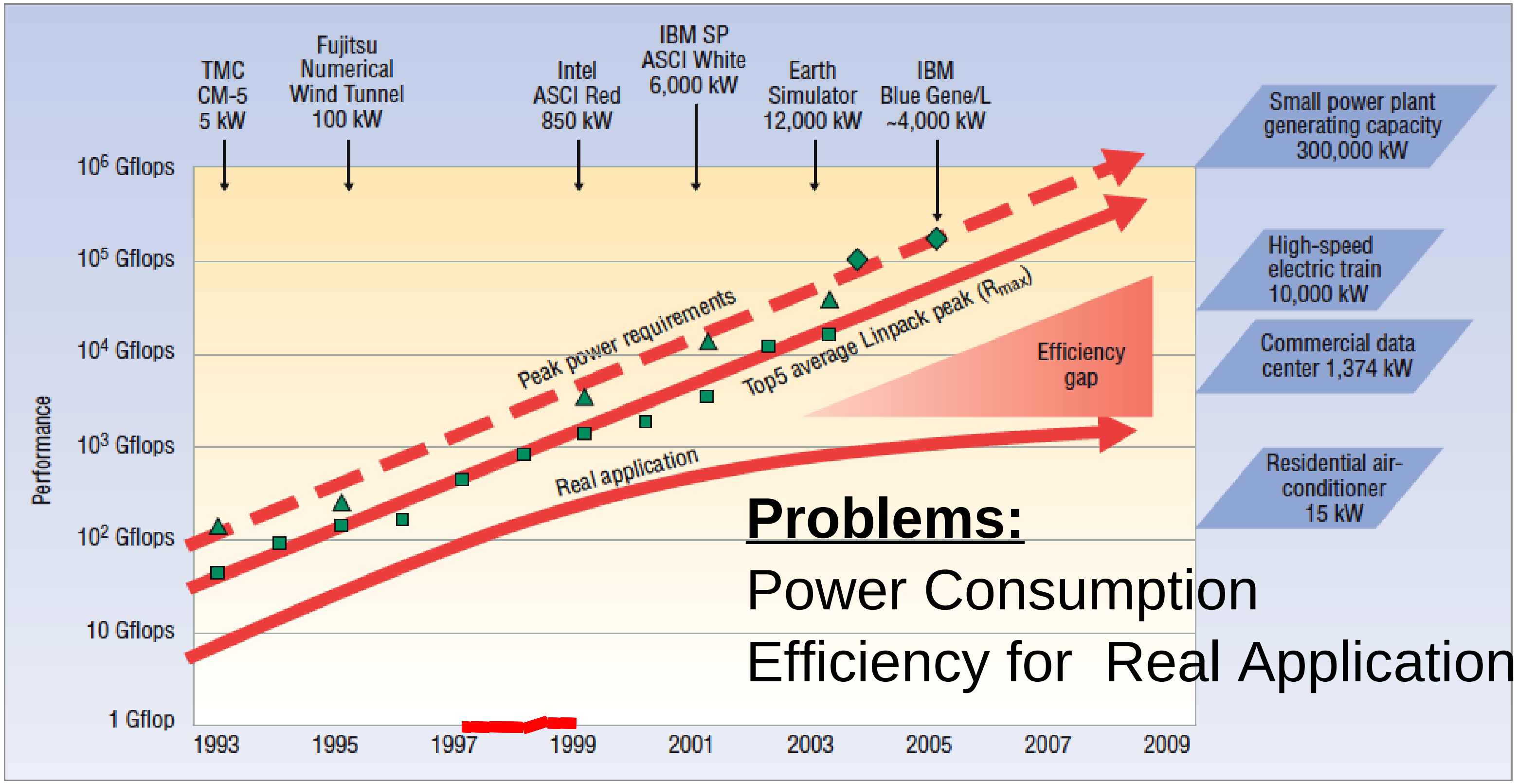
...after von Neumann...

Exaflop/s

Petaflop/s

Teraflop/s

Gigaflop/s



Problems:

Power Consumption

Efficiency for Real Applications

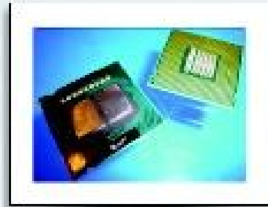
Figure 1. Rising power requirements. Peak power consumption of the top supercomputers has steadily increased over the past 15 years.

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

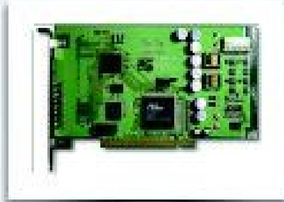
GPUs
Graphic Processing Units

- 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

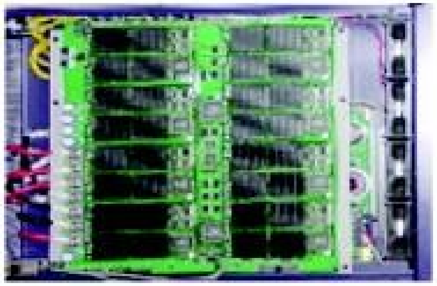
HARDWARE

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)

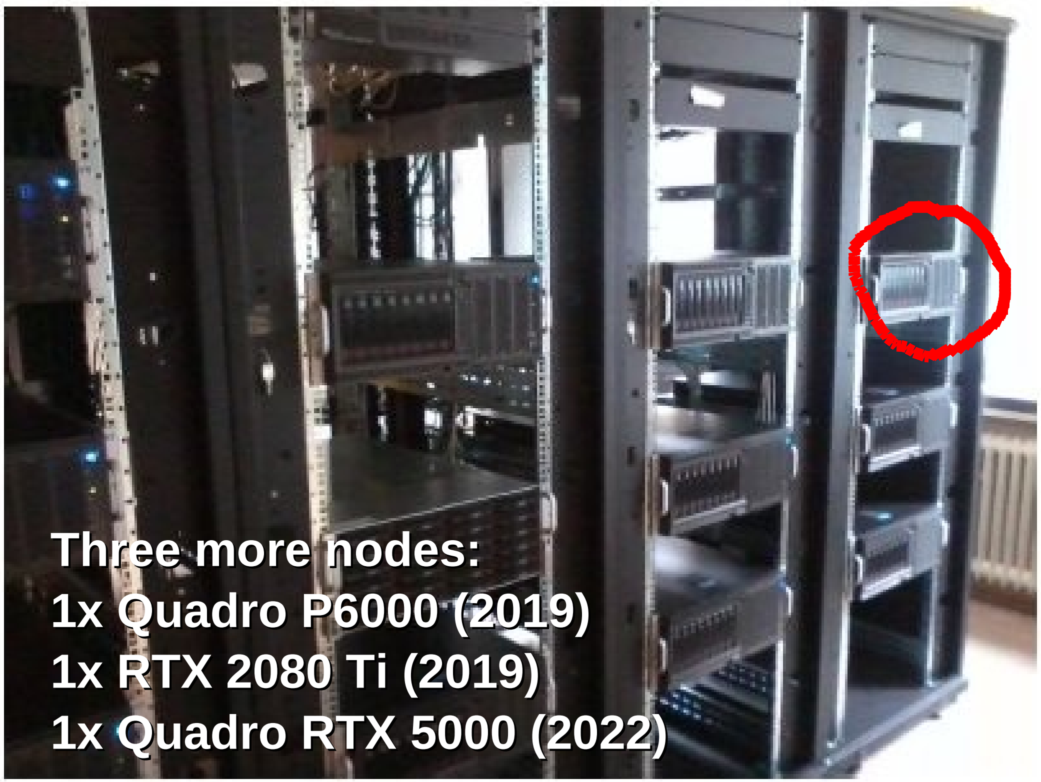
12 x 64 GB = 768 GB RAM CPU memory

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 Ampere A100 GPU, 54 billion transistors, 6920 cores (Hopper H100, ...)

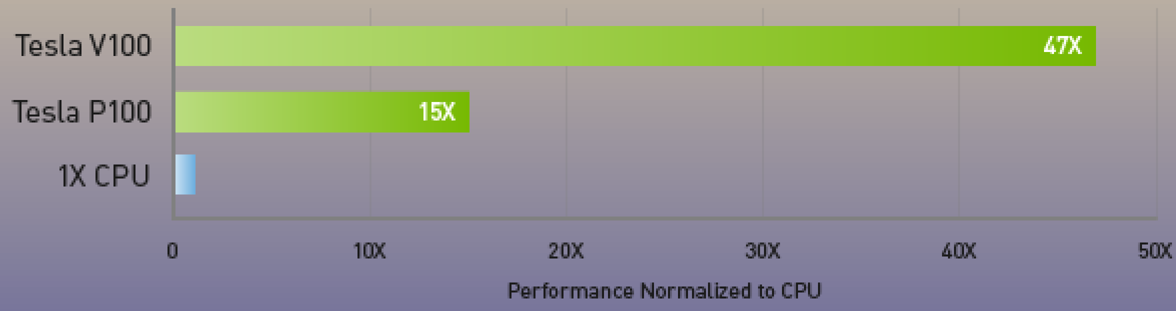
With NVLINK

Without NVLINK

	A100 80GB PCIe	A100 80GB SXM
FP64	9.7 TFLOPS	
FP64 Tensor Core	19.5 TFLOPS	
FP32	19.5 TFLOPS	
Tensor Float 32 (TF32)	156 TFLOPS | 312 TFLOPS*	
BFLOAT16 Tensor Core	312 TFLOPS | 624 TFLOPS*	
FP16 Tensor Core	312 TFLOPS | 624 TFLOPS*	
INT8 Tensor Core	624 TOPS | 1248 TOPS*	
GPU Memory	80GB HBM2e	80GB HBM2e
GPU Memory Bandwidth	1,935 GB/s	2,039 GB/s
Max Thermal Design Power (TDP)	300W	400W ***

NVIDIA Volta V100 GPU, 21 billion transistors, 5120 cores

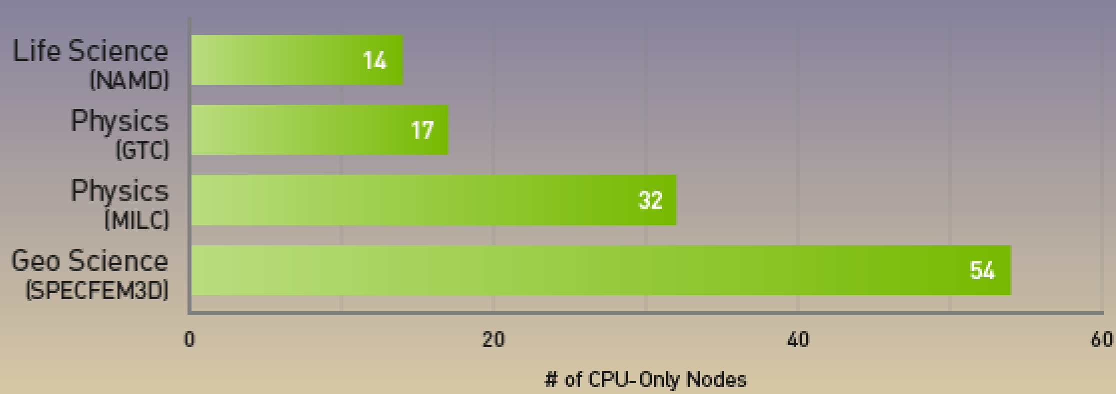
47X Higher Throughput Than CPU Server on Deep Learning Inference



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

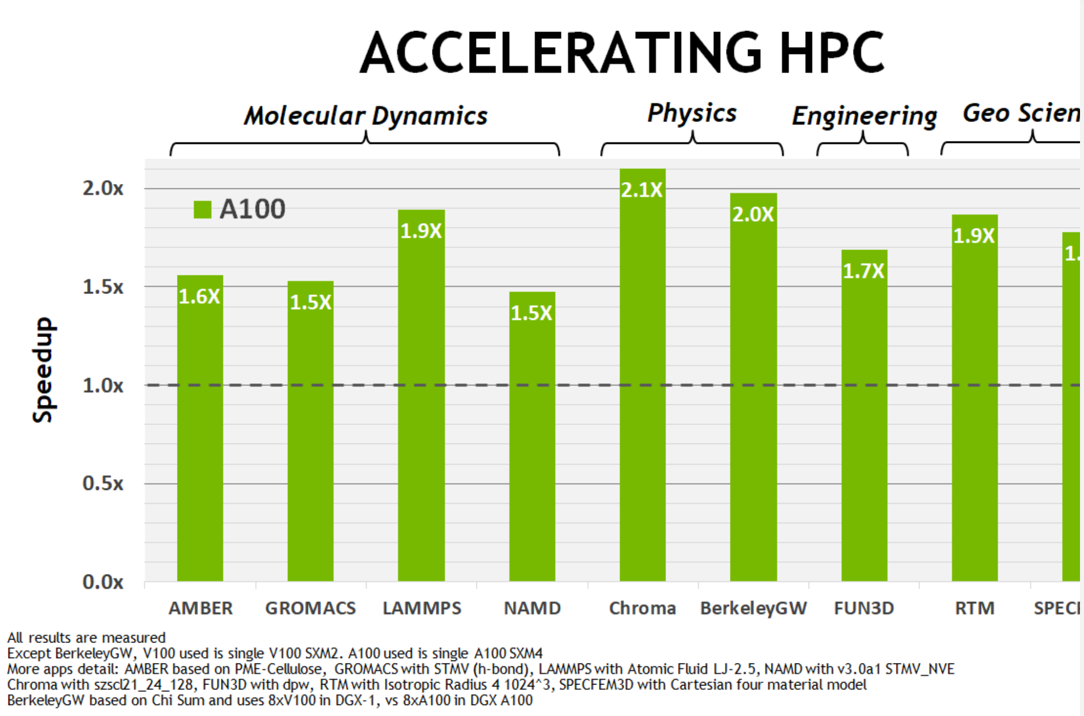
1 GPU Node Replaces Up To 54 CPU Nodes

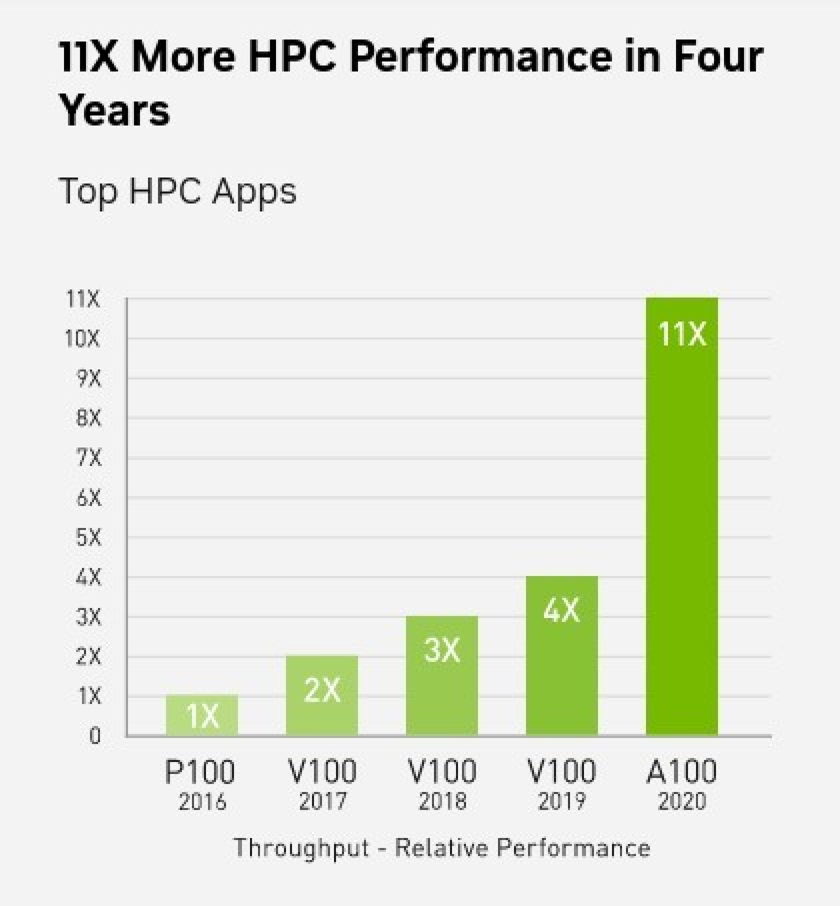
Node Replacement: HPC Mixed Workload



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.

NVIDIA Ampere A100 GPU, 54 billion transistors, 6920 cores



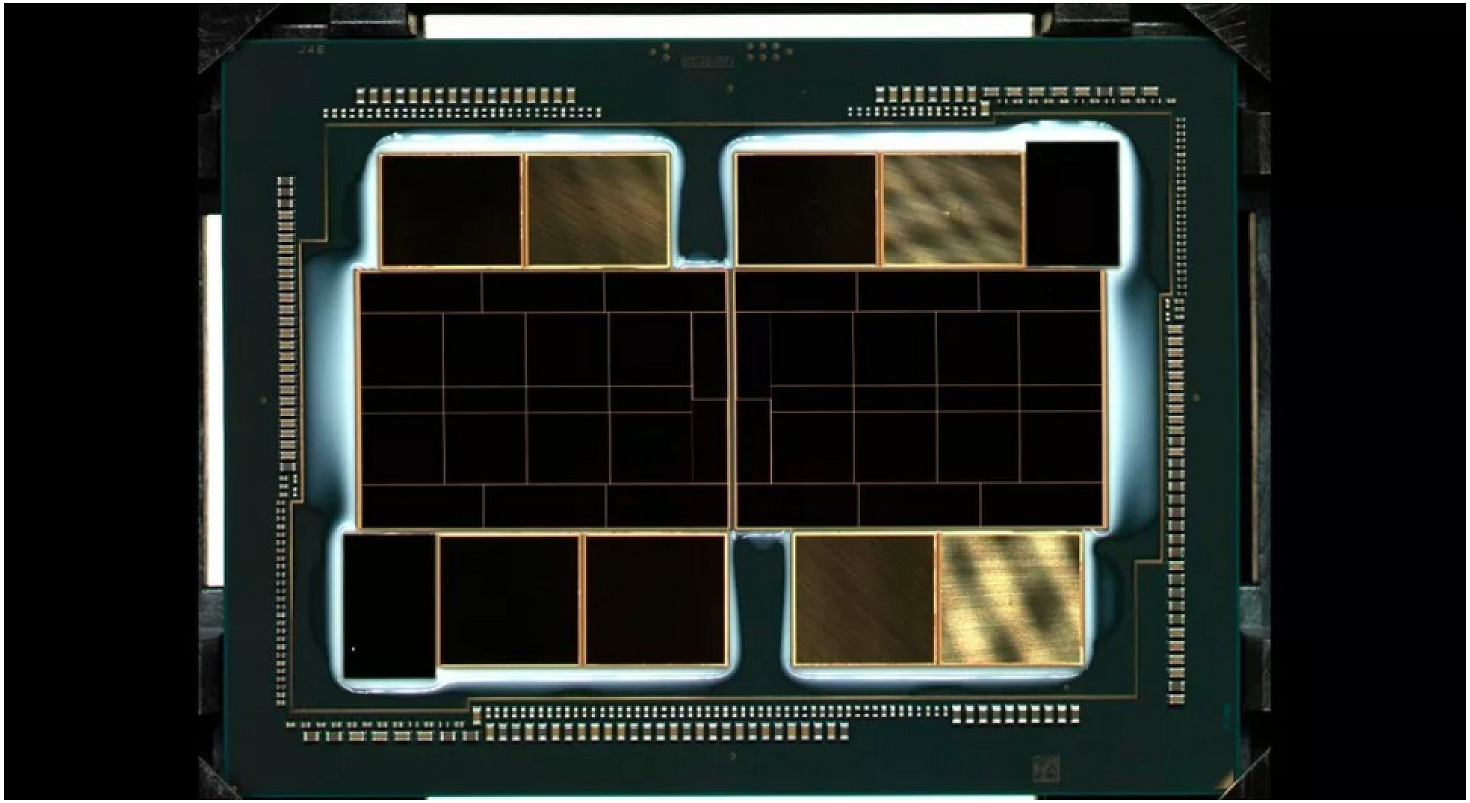


New Intel Ponte Vecchio hardware / Recent Processors

Intel Ponte Vecchio Seemingly Offers 2.5x Higher Performance Than Nvidia's A100

By Zhiye Liu published August 22, 2022

The battle of juggernauts



Intel has detailed the company's Ponte Vecchio Xe-HPC GPU at Hot Chips 34. In the provided benchmarks, the chipmaker claims that Ponte Vecchio delivers up to 2.5x more performance than the Nvidia A100. But, as customary, take vendor-provided benchmarks with a pinch of salt.
From: https://www.tomshardware.com/news/intel-ponte-vecchio-seemingly-offers-25x-higher-performance-than-nvidias-a100

AMD Instinct
MI250X GPU

Nov.2023 Lists:
Used in:

Frontier (#1 US)
And LUMI (#5 FIN)

AMD Instinct™ MI250X

GPU Specifications

GPU Architecture: CDNA2

Lithography: TSMC 6nm FinFET

Stream Processors: 14,080

Compute Units: 220

Peak Half Precision (FP16) Performance: 383 TFLOPs

Peak Engine Clock: 1700 MHz

Peak Single Precision Matrix (FP32) Performance: 95.7 TFLOPs

Peak Double Precision Matrix (FP64) Performance: 95.7 TFLOPs

Peak Single Precision (FP32) Performance: 47.9 TFLOPs

Peak Double Precision (FP64) Performance: 47.9 TFLOPs

Peak INT4 Performance: 383 TOPs

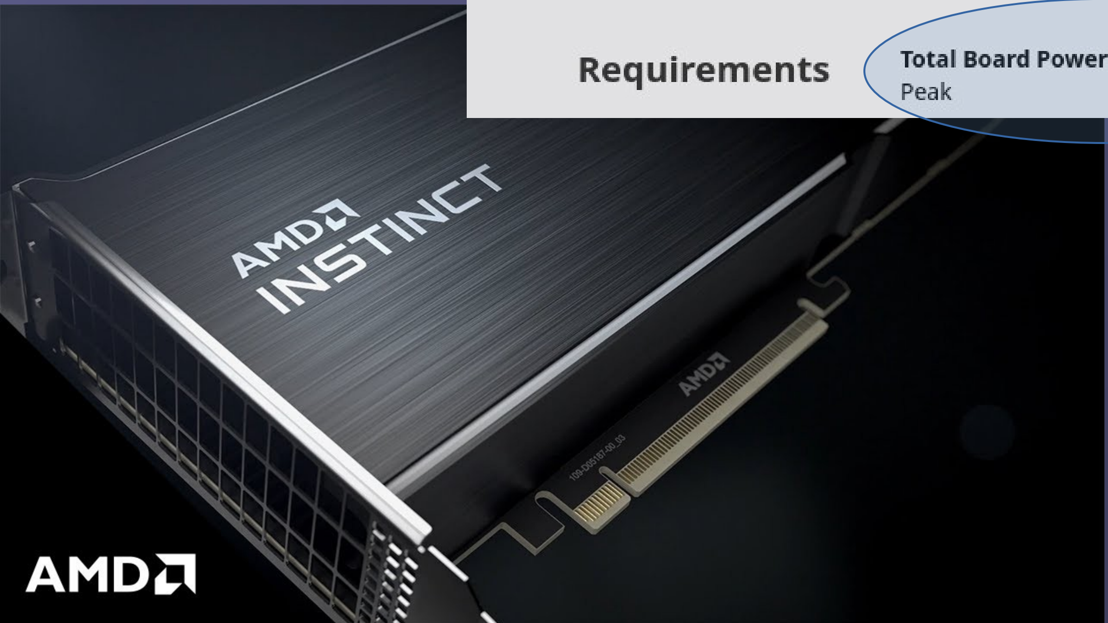
Peak INT8 Performance: 383 TOPs

Peak bfloat16: 383 TFLOPs

OS Support: Linux x86_64

Requirements

Total Board Power (TBP): 500W | 560W Peak

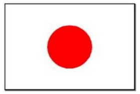


From https://www.top500.org/
Nov. 2023 List


USA


USA


USA


Japan

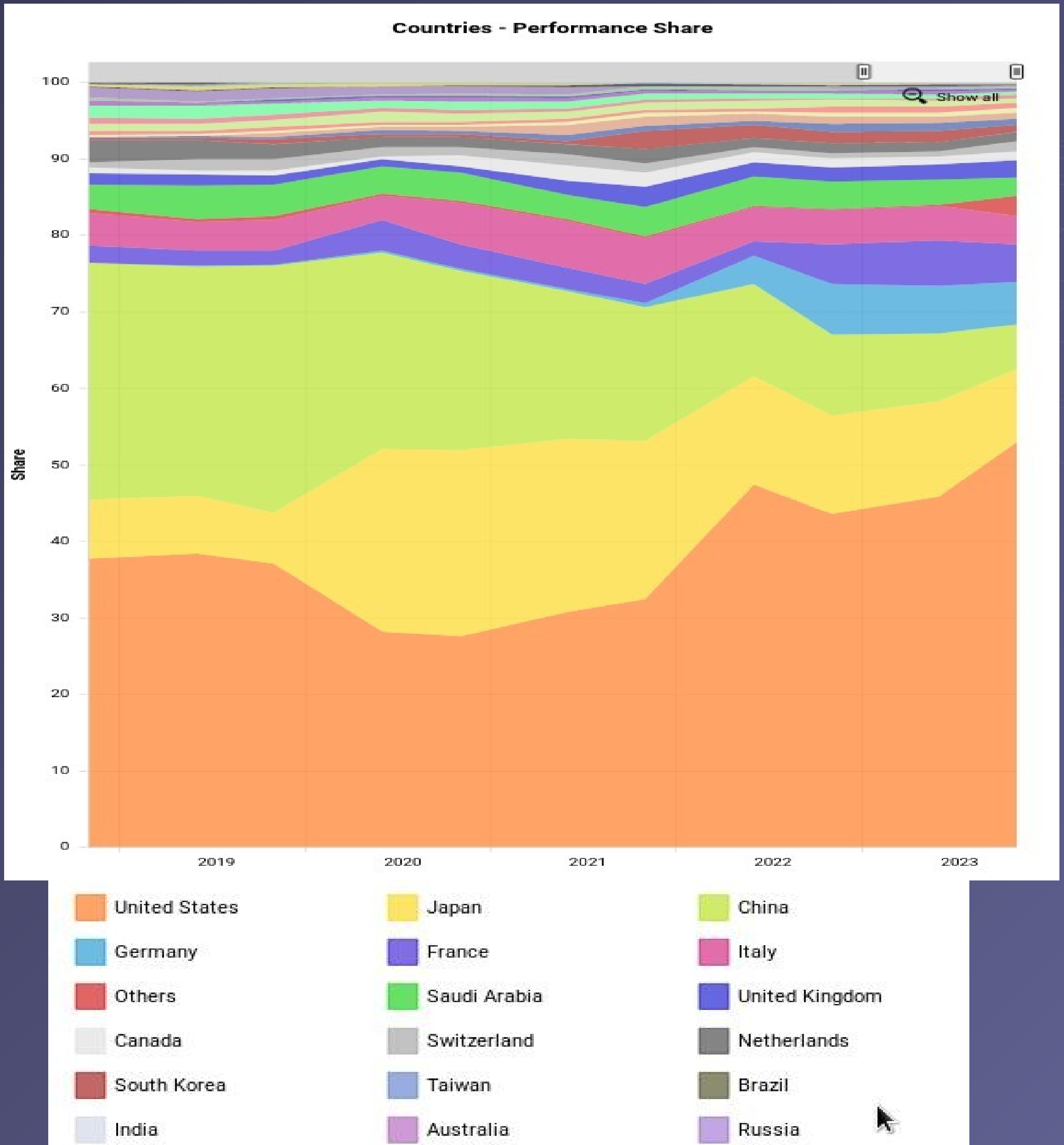

Finland (EUR)

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 ***GPU AMD Instinct***	8,699,904	1,194.00	1,679.82	22,703
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 ***Intel Data Center GPU***	4,742,808	585.34	1,059.33	24,687
3	**Eagle** - Microsoft NDv5, Xeon Platinum 8480C 48C 2GHz, NVIDIA H100, NVIDIA Infiniband NDR, Microsoft Microsoft Azure United States ***GPU NVIDIA Hopper***	1,123,200	561.20	846.84	
4	**Supercomputer Fugaku** - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan ***Fujitsu Arm***	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 ***GPU AMD Instinct***	2,752,704	379.70	531.51	7,107

From https://www.top500.org/
Nov. 2023 List

	Rank	System	Cores	Rmax (PFlop/s)	Rpeak (PFlop/s)	Power (kW)	
USA	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	***GPU NVIDIA Volta***
China	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	***Chinese 26010 Processor*** *(260 cores, 10nm)*
USA	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	2,589	***GPU NVIDIA Ampere***
USA	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	***GPU NVIDIA Ampere***
China	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	***Chinese Matrix-2000 Processor***

Top 500 List November 2023 – Performance Share of Countries From https://www.top500.org



L U M I

Supercomputer, Kajaani, Finland

Using only Hydroelectric Power and its Heat used for heating buildings.

No. 5 in top500
No. 7 in green500

2.2 million cores
~12.000 AMD GPUs

EuroHPC and LUMI consortium:
Finland, Belgium, Czech Republic, Denmark, Estonia, Iceland, Norway, Poland, Sweden, and Switzerland.

RIKEN, Kobe, JAPAN

FUGAKU

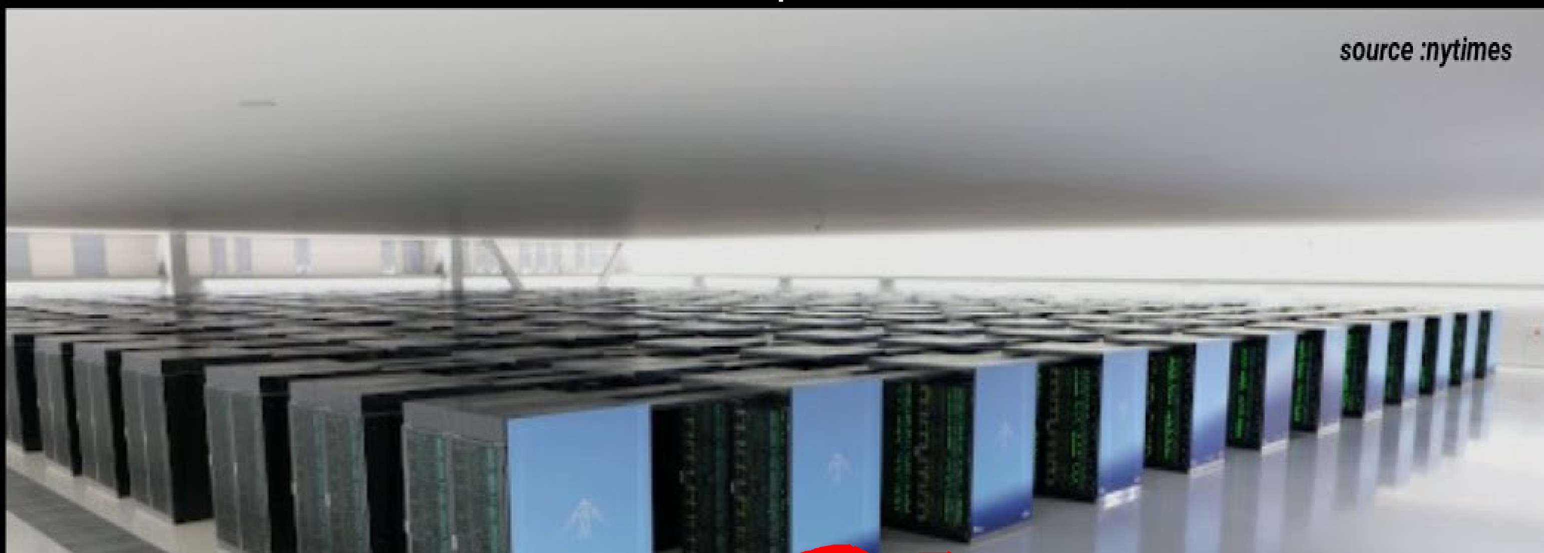
富岳

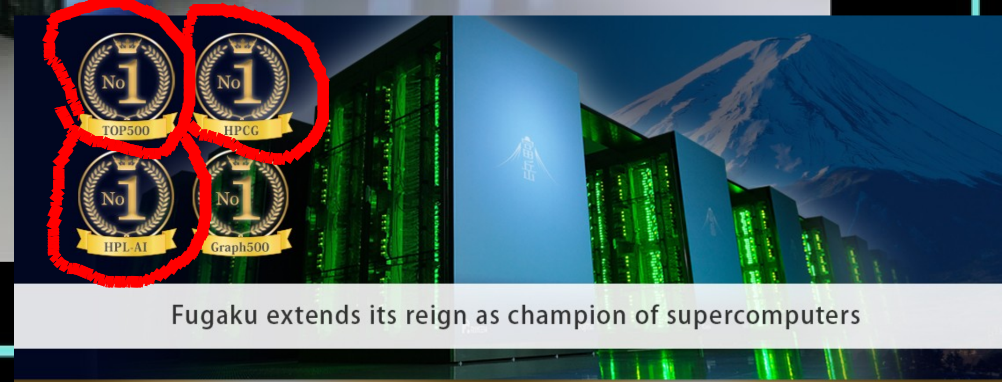
Nature's Secrets

Mt. Fuji

The world's fastest Super Computer 2020 /2021

7.6 million cores, 442 Pflop/s

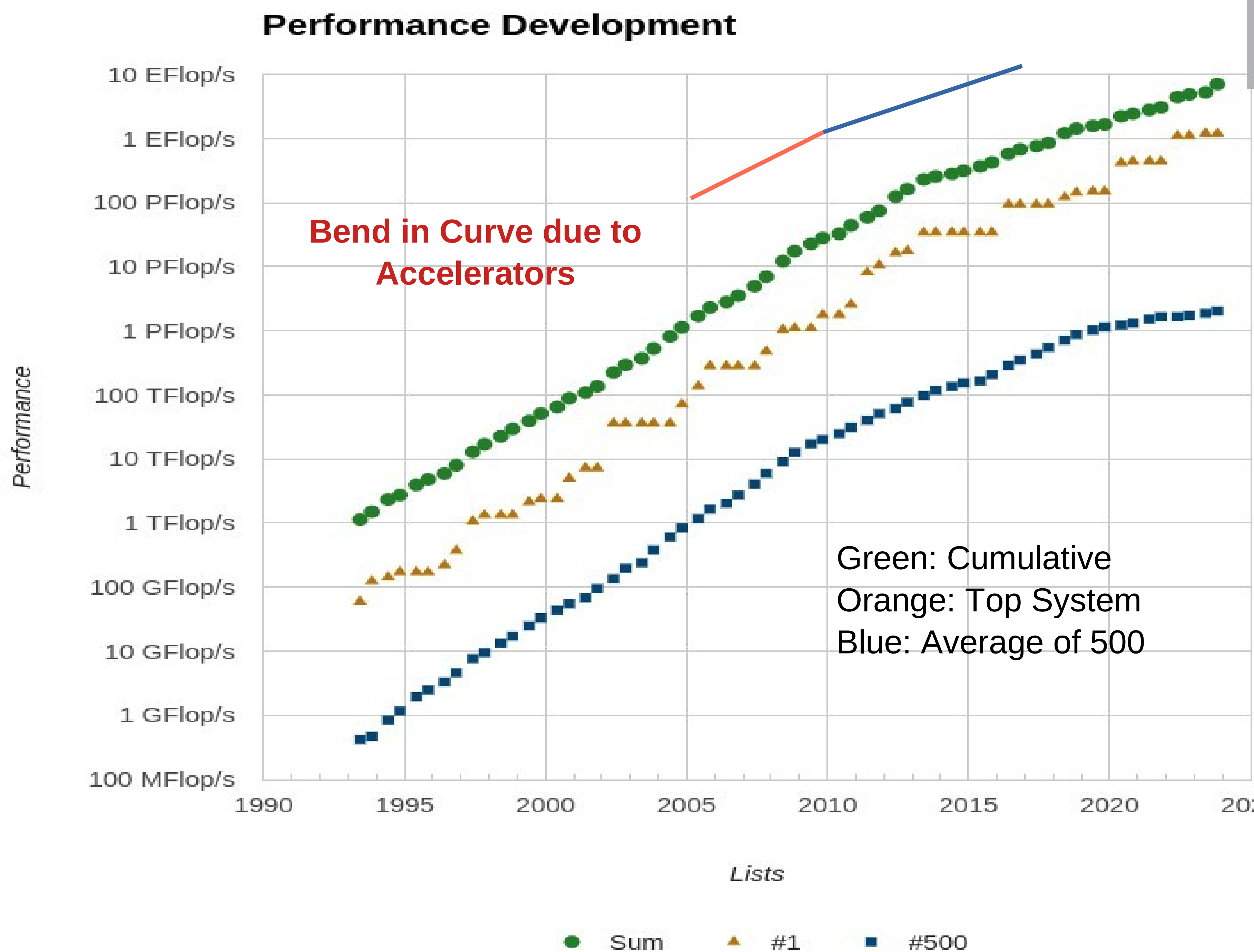




JUWELS Booster 936 nodes (AMD CPU, 4x Ampere GPU)
~450.000 AMD cores, 25 million NVIDIA Ampere GPU cores
~ 70 Pflop/s SP ~ 44 Pflop/s DP
No. 18 in top500 list, No. 3 in green500 list

Jülich Wizard for European Leadership Science

Copyright:
— Forschungszentrum Jülich


Performance Development
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
Bend in Curve due to
Accelerators
Green: Cumulative
Orange: Top System
Blue: Average of 500
Sum
#1
#500

GREEN 500 list Nov. 2023

Power Efficiency (Gflops/Watts),
see also top500 webpage
right: 1-5
below: 6-10

Rank	TOP500 Rank	System	Cores	Rmax (PFlop/s)	Power (kW)	Energy Efficiency (GFlops/watts)	
1	293	**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	*GPU NVIDIA Hopper*
2	44	**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	*GPU AMD Instinct*
3	17	**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	*GPU AMD Instinct*
4	25	**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	*GPU AMD Instinct*
5	92	**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	*GPU AMD Instinct*

6	8	**MareNostrum 5 ACC** - BullSequana XH3000, Xeon Platinum 8460Y+ 40C 2.3GHz, NVIDIA H100 64GB, Infiniband NDR200, EVIDEN EuroHPC/BSC Spain	680,960	138.20	2,560	53.984	*GPU NVIDIA Hopper*
7	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	7,107	53.428	*GPU AMD Instinct*
8	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	22,703	52.592	*GPU AMD Instinct*
9	84	**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	195	46.543	*GPU AMD Instinct*
10	496	**Olaf** - Lenovo ThinkSystem SR675 V3, AMD EPYC 9334 32C 2.7GHz, NVIDIA H100, Infiniband NDR 400, Lenovo Science Institute South Korea	3,936	2.03	45	45.117	*GPU NVIDIA Hopper*