VSB TECHNICAL UNIVERSITY OF OSTRAVA | IT4INNOVATIONS NATIONAL SUPERCOMPUTING CENTER

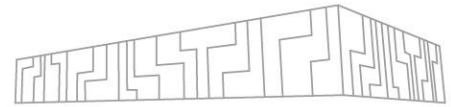
INTRODUCTION TO HPC

Ondřej Vysocký
IT4Innovations

3. 6. 2024

EUROPEAN UNION
European Structural and Investment Funds
Operational Programme Research,
Development and Education

MINISTRY OF EDUCATION, YOUTH AND SPORTS



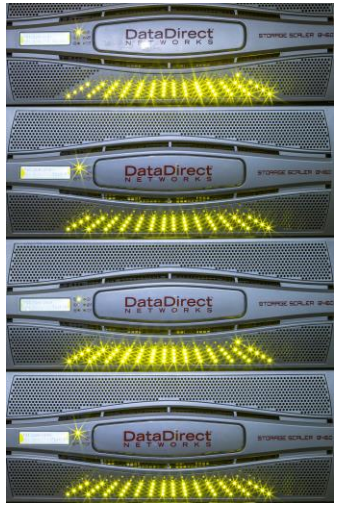
INTRODUCTION

SUPERCOMPUTING

WHAT IS A SUPERCOMPUTER?

Compute nodes

Data storage



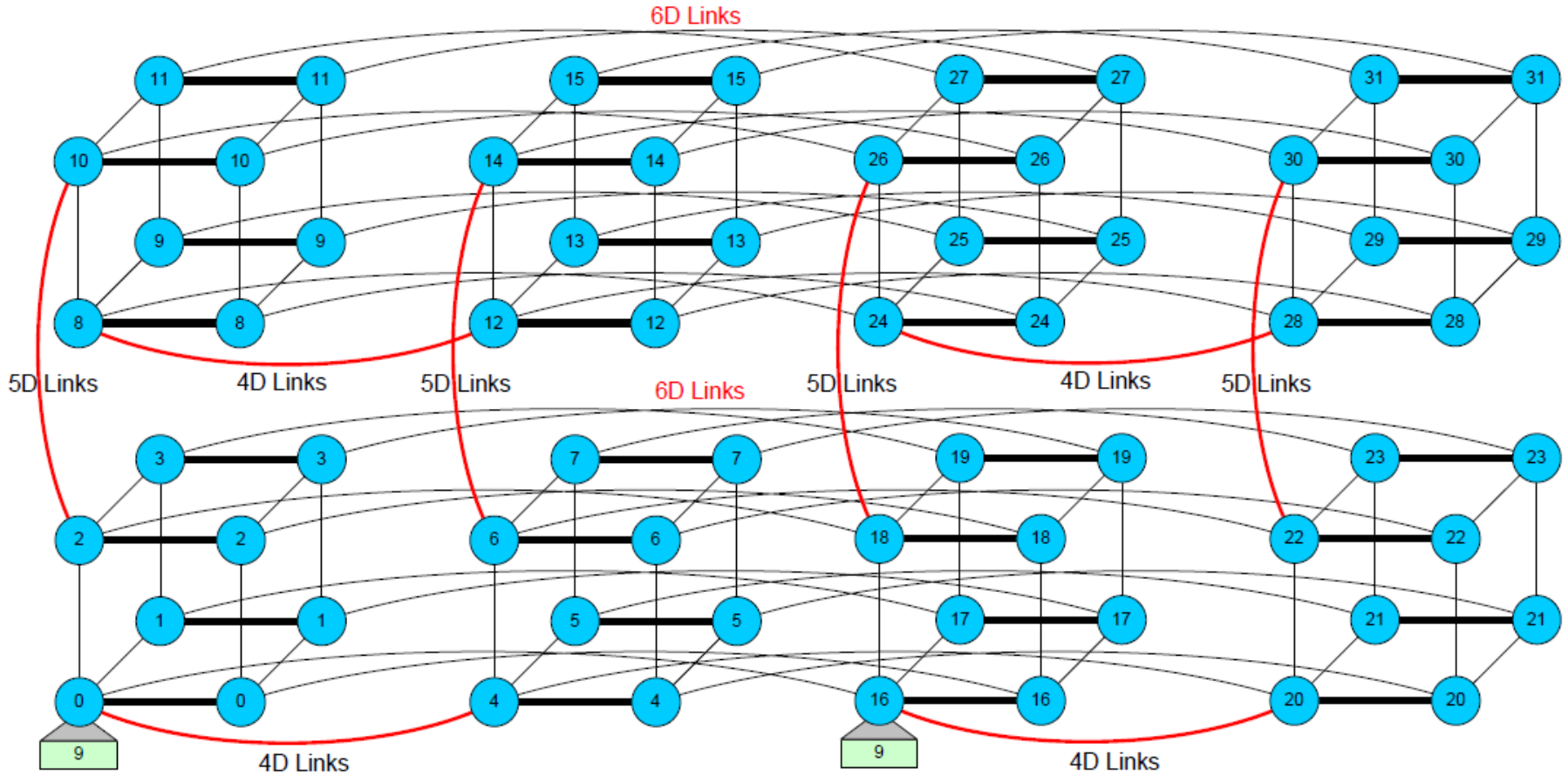
Interconnect

WHAT IS NOT A SUPERCOMPUTER?

WHAT IS NOT A SUPERCOMPUTER?

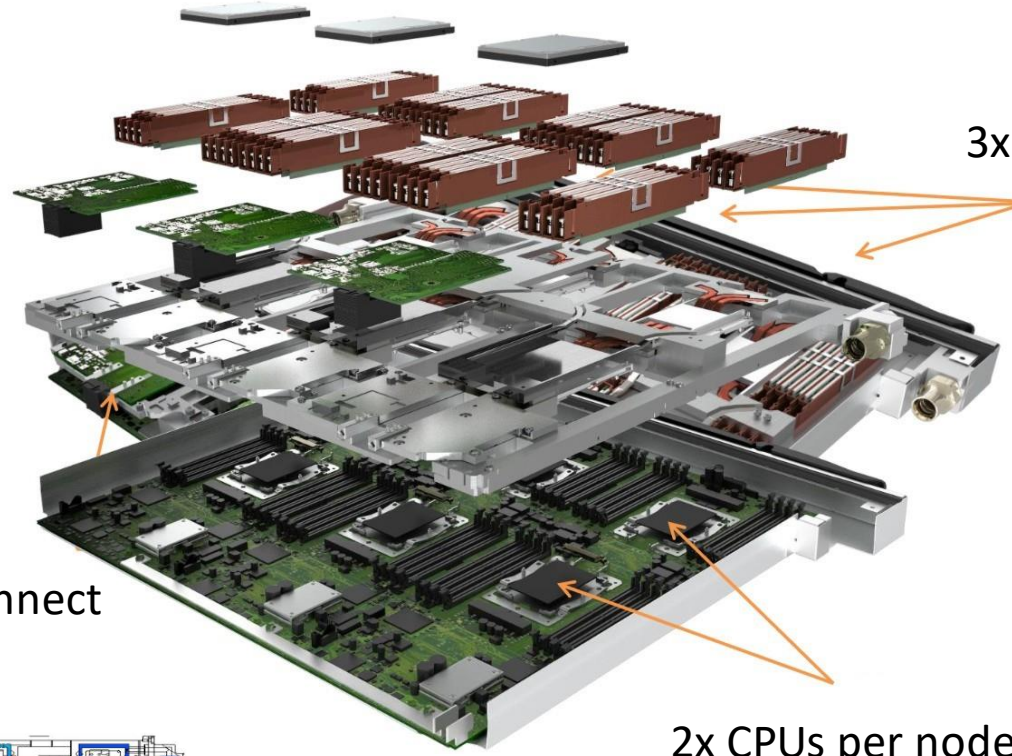
EXAMPLE OF A NETWORK?

- InfiniBand FDR56 / 7D Enhanced hypercube



DATA CENTER

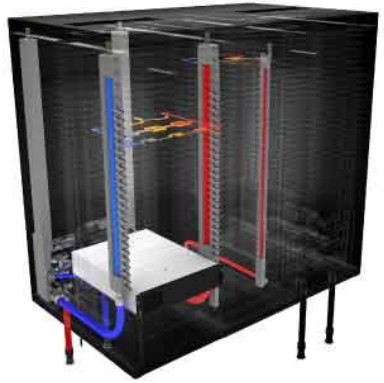
CABINET

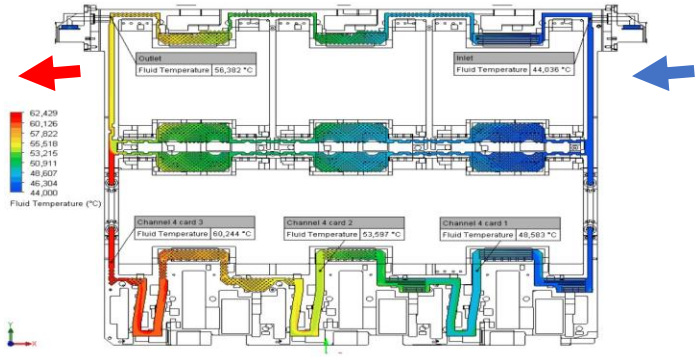


3x compute nodes

interconnect

2x CPUs per node



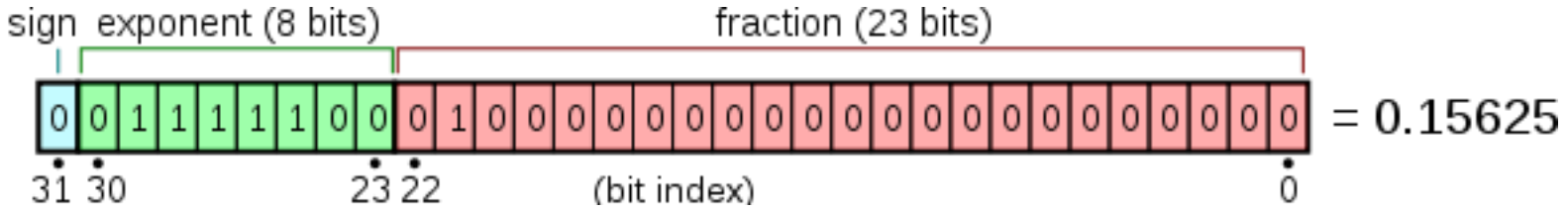


FLOATING POINT COMPUTING

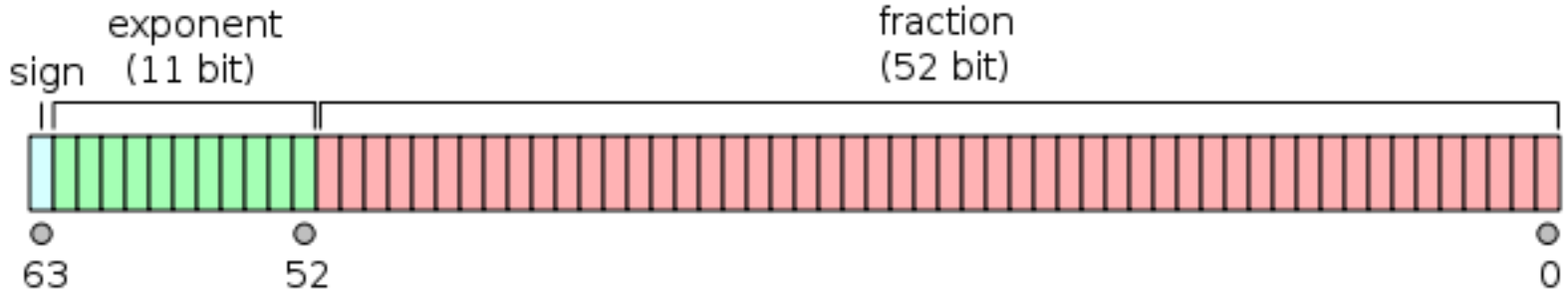
- Floating point number representation

$$25{,}167 = 0{,}25167 \cdot 10^{2} =$$
$$= (-1)^{0} \cdot (2 \cdot 10^{-1} + 5 \cdot 10^{-2} + 1 \cdot 10^{-3} + 6 \cdot 10^{-4} + 7 \cdot 10^{-5}) \cdot 10^{2}$$

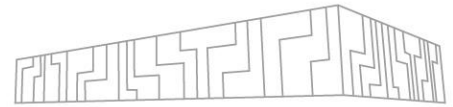
- 25,167 = [0, 2, 2, 5, 1, 6, 7]
- Single precision, 4B = 32bits, fp32



- Double precision, 8B = 64bits, fp64



PEAK PERFORMANCE

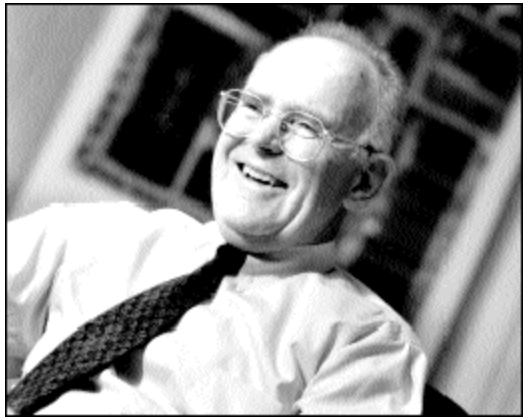


- FLOP = Floating point operation
- **Computer performance** = number of floating-point operations per second FLOPS (Flop/s)st
- Intel® Xeon® Platinum 8280M Processor

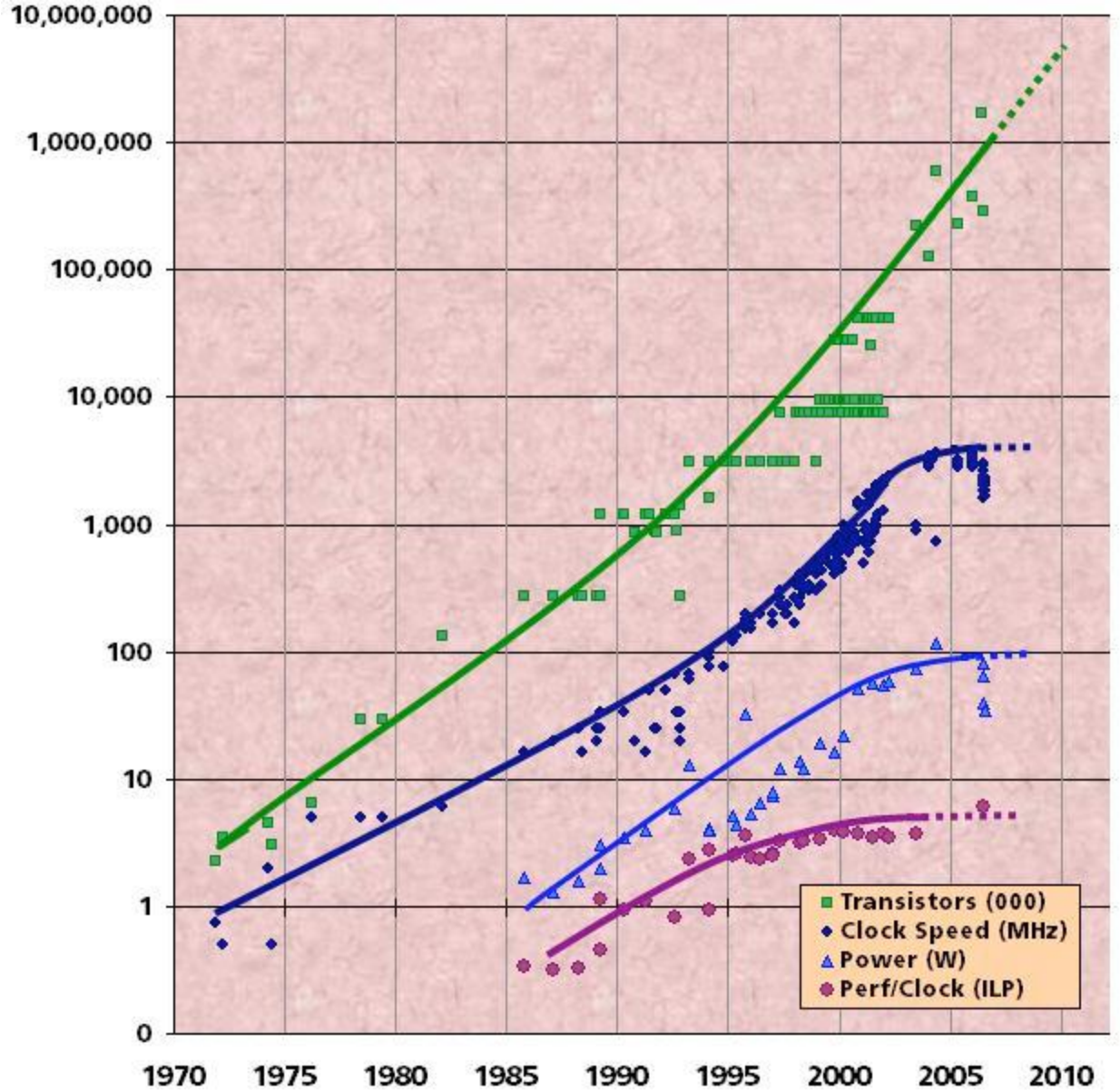
number of compute nodes	**1000**	**1000**
number of CPUs	2	2
frequency	2.7 GHz	2.7
number of cores	28	28
have FMA instruction	yes	2
have 2 FMA units	yes	2
SIMD width	512 bit = 8 double precision	8
		4 838 000 Gflop/s
		4 838 Tflop/s
		4.8 Pflop/s

MOORE'S LAW

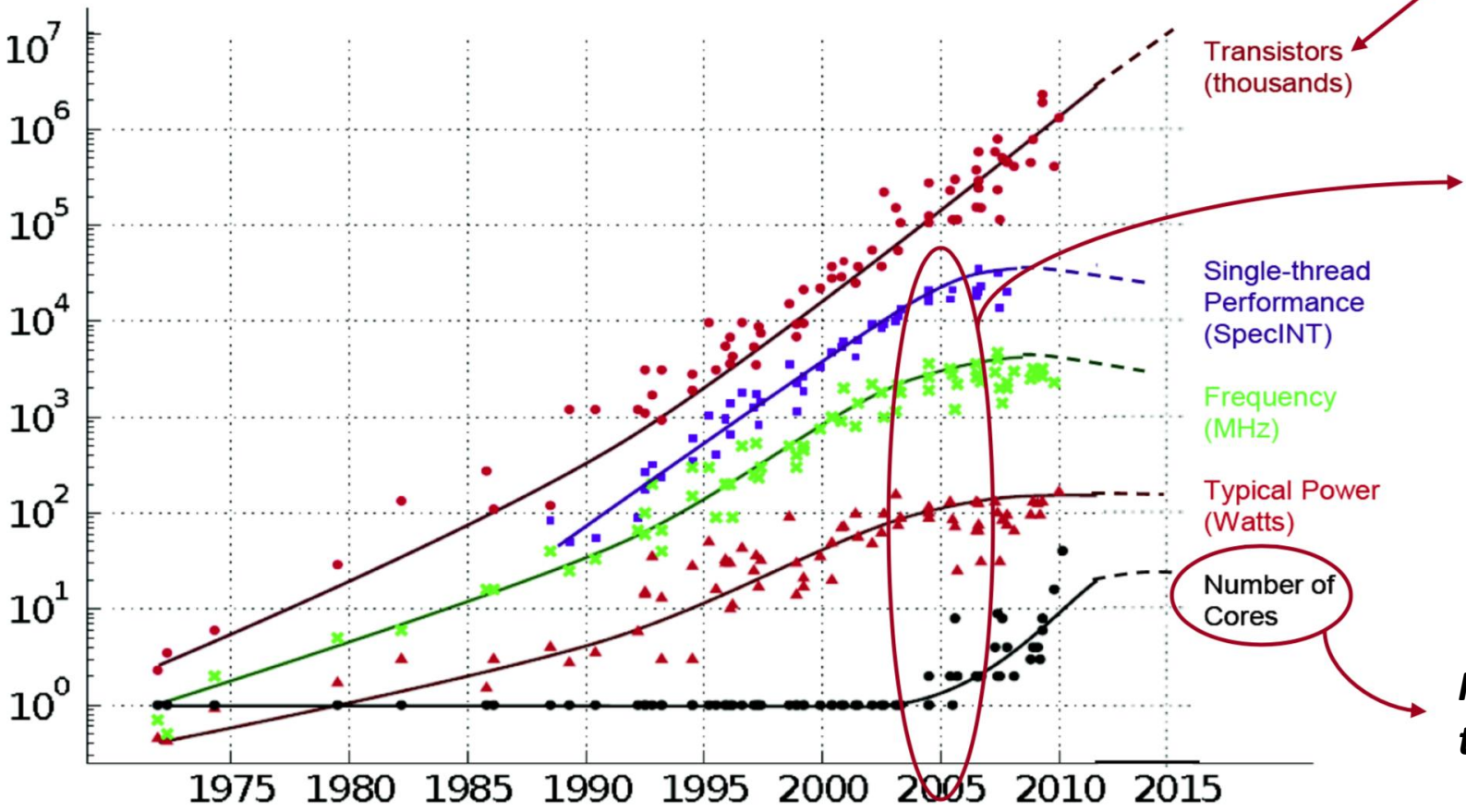
- Chip density is continuing increase ~2x every 2 years
- Clock speed is not
- Number of processor cores has to double instead
- Parallelism must be exposed to and managed by software



Slide source: Jack Dongarra



MOORE'S LAW



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

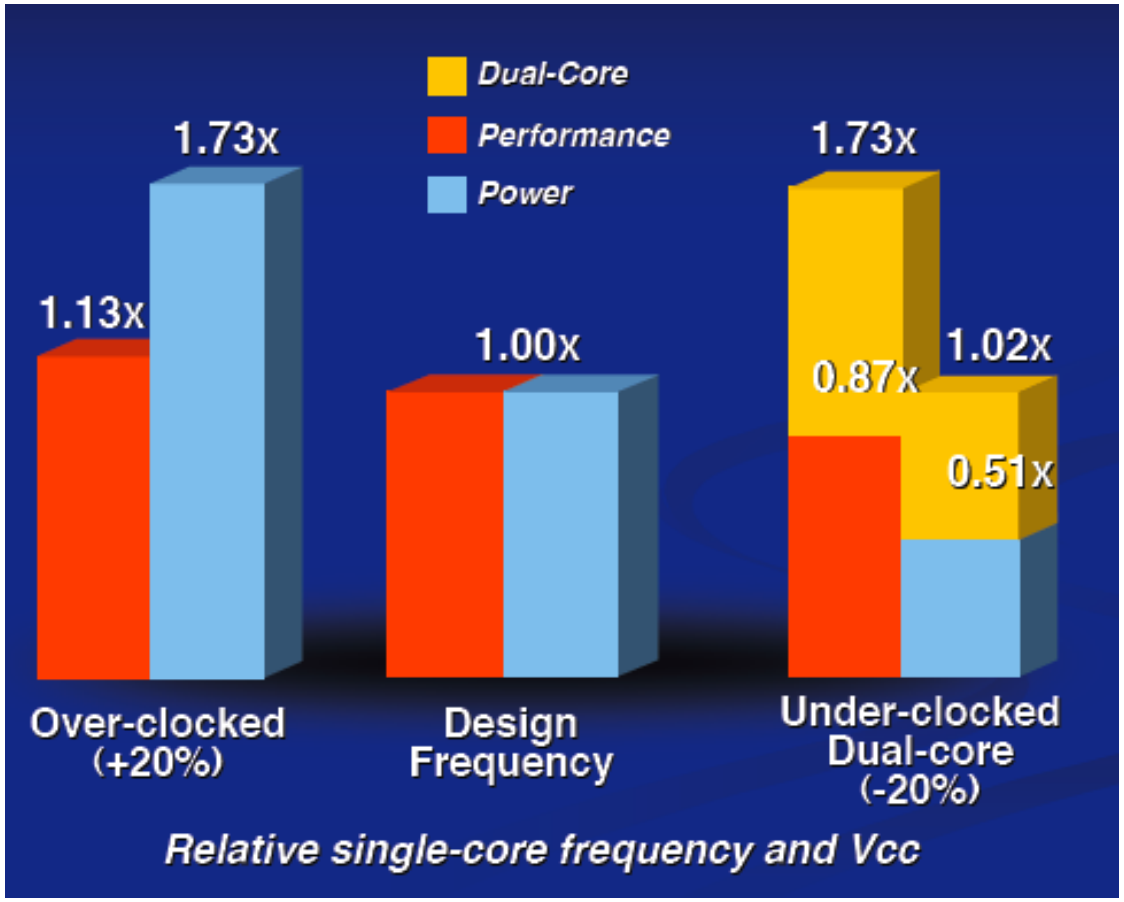
Transistor count doubles every 18 months, Moore's Law

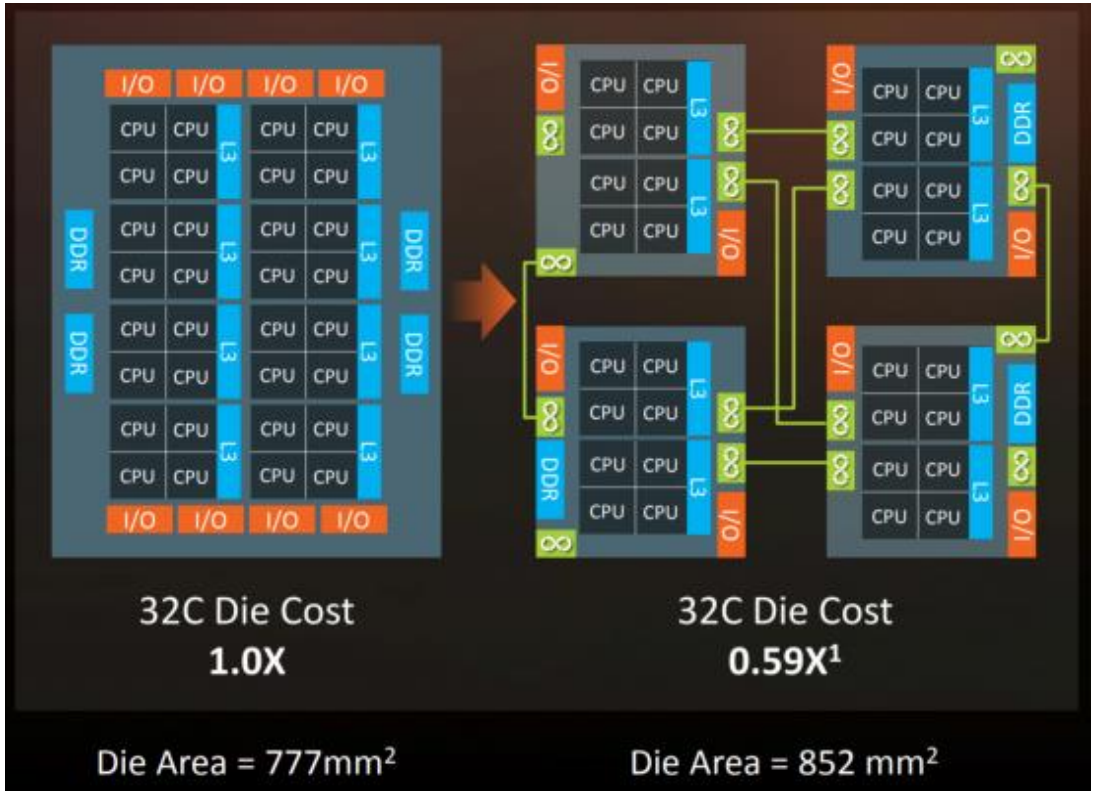
The Power Wall

- Power dissipation of single-core processors becomes prohibitive
- The "Free Performance Lunch" of frequency scaling is over!

Performance can only grow through node-level parallelism!

MODERN CPU DESIGN





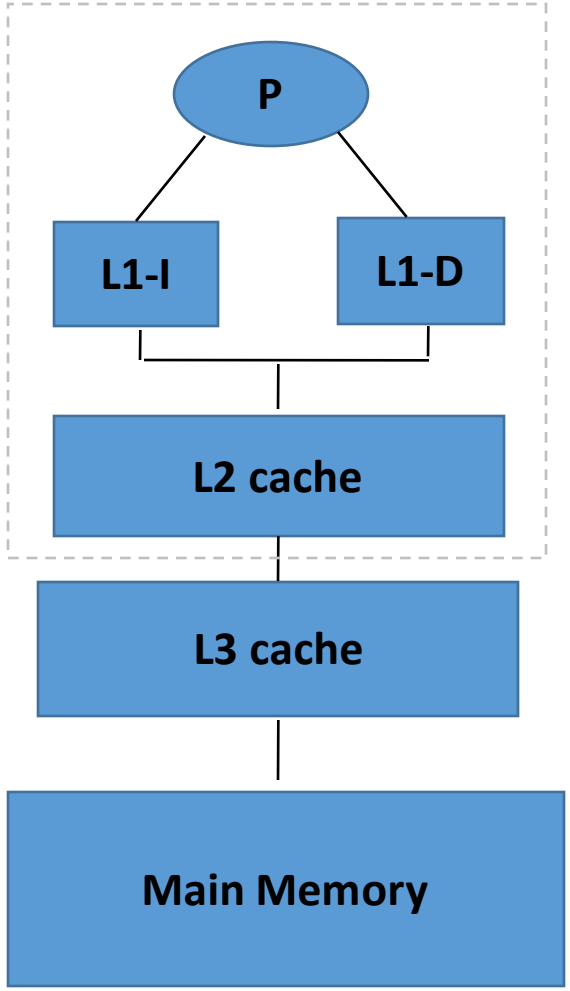
TYPICAL MEMORY HIERARCHY

16-64 KB, 1-4 cycles

512KB-8MB, 6-15 cycles

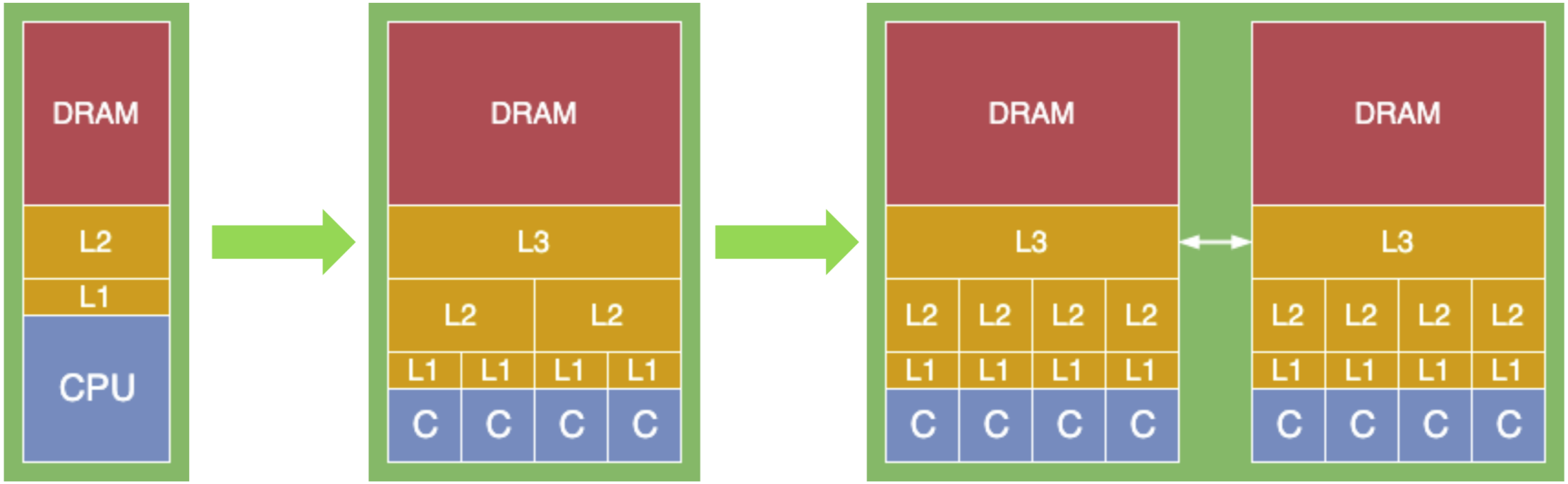
4MB-32MB, 30-50 cycles

>1GB, >300 cycles



- Access time to main memory is 100's of clock cycles
- Use a small but fast storage near processor
- Works due to locality

HPC BUILDING BLOCKS: CPU



Single-core

Multicore SMP

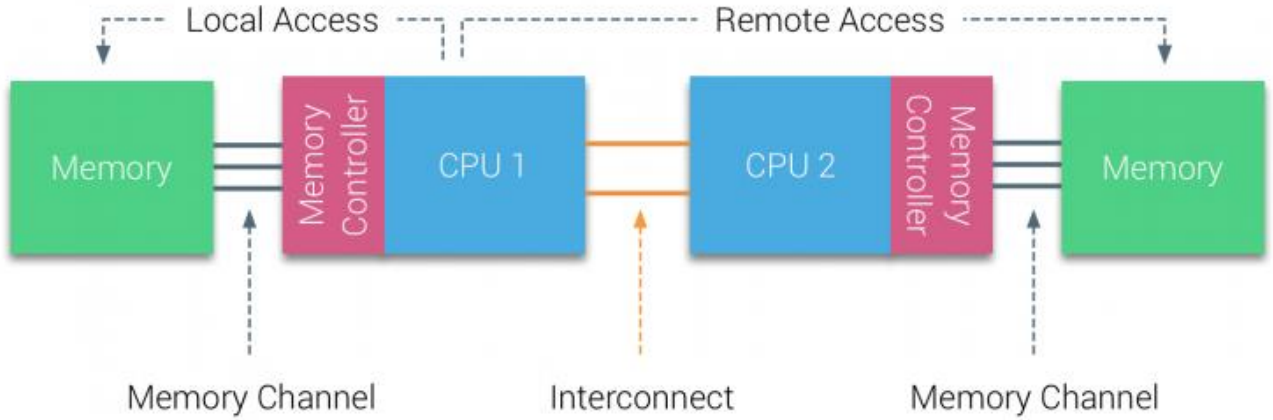
Multicore NUMA

SMP: Symmetric Multi-processor
NUMA: Non-Uniform Memory Access

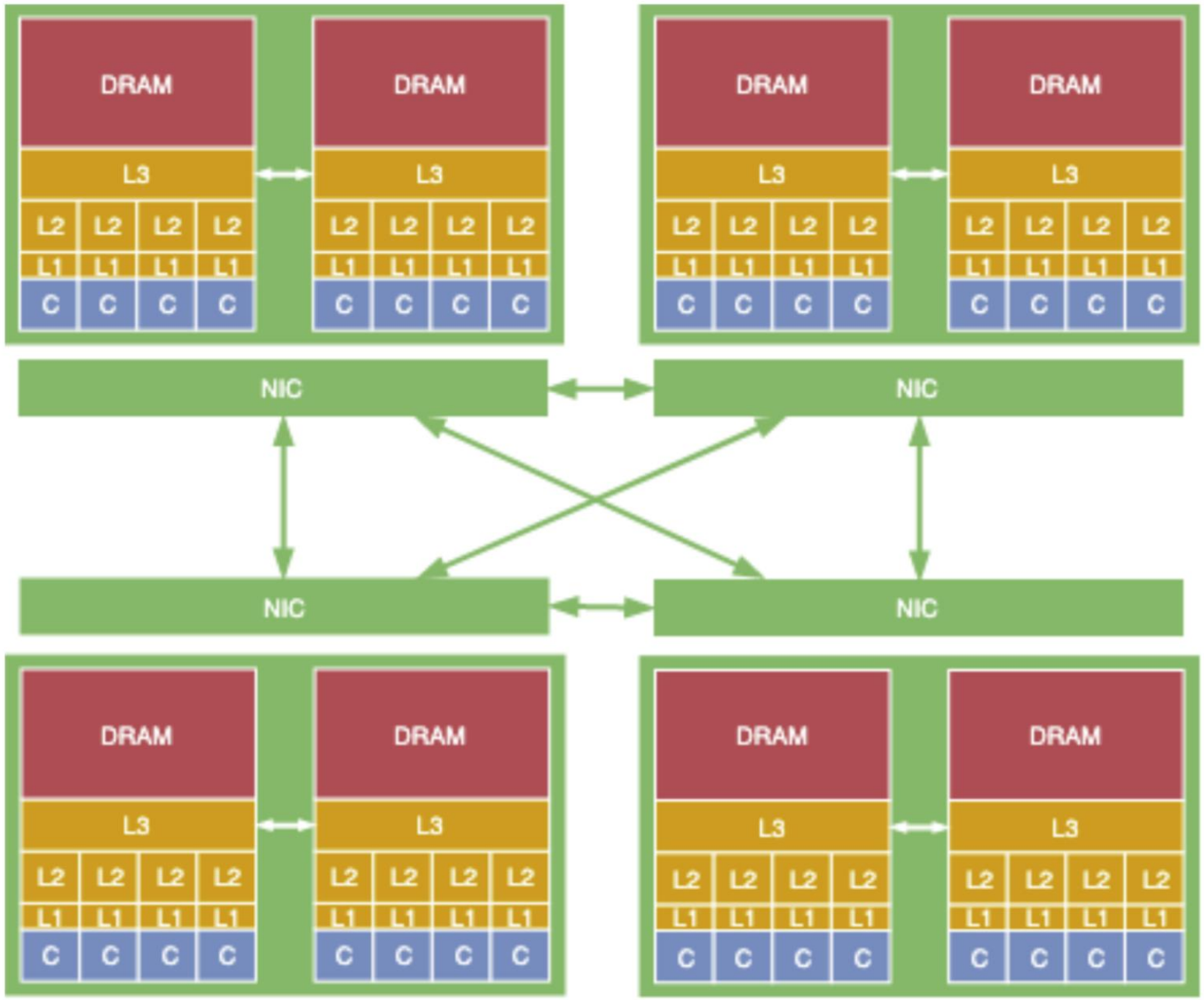
Source: CSCS-USI Summer School 2019

NUMA & CC-NUMA

- **NUMA** – Non-Uniform Memory Access
- Aims at surpassing the scalability limits of the UMA architecture due to **memory bandwidth bottleneck**
- Memory physically shared, but access to different portions of the memory may require **significantly different times**
  - local memory access is the fastest, access across link is slower
- **Caches** used to level access times
  - technically difficult to maintain cache consistency
- **Cache coherency (CC)** accomplished at the **hardware level** (expensive)
  - if one processor updates a location in shared memory, all the other processors learn about the update

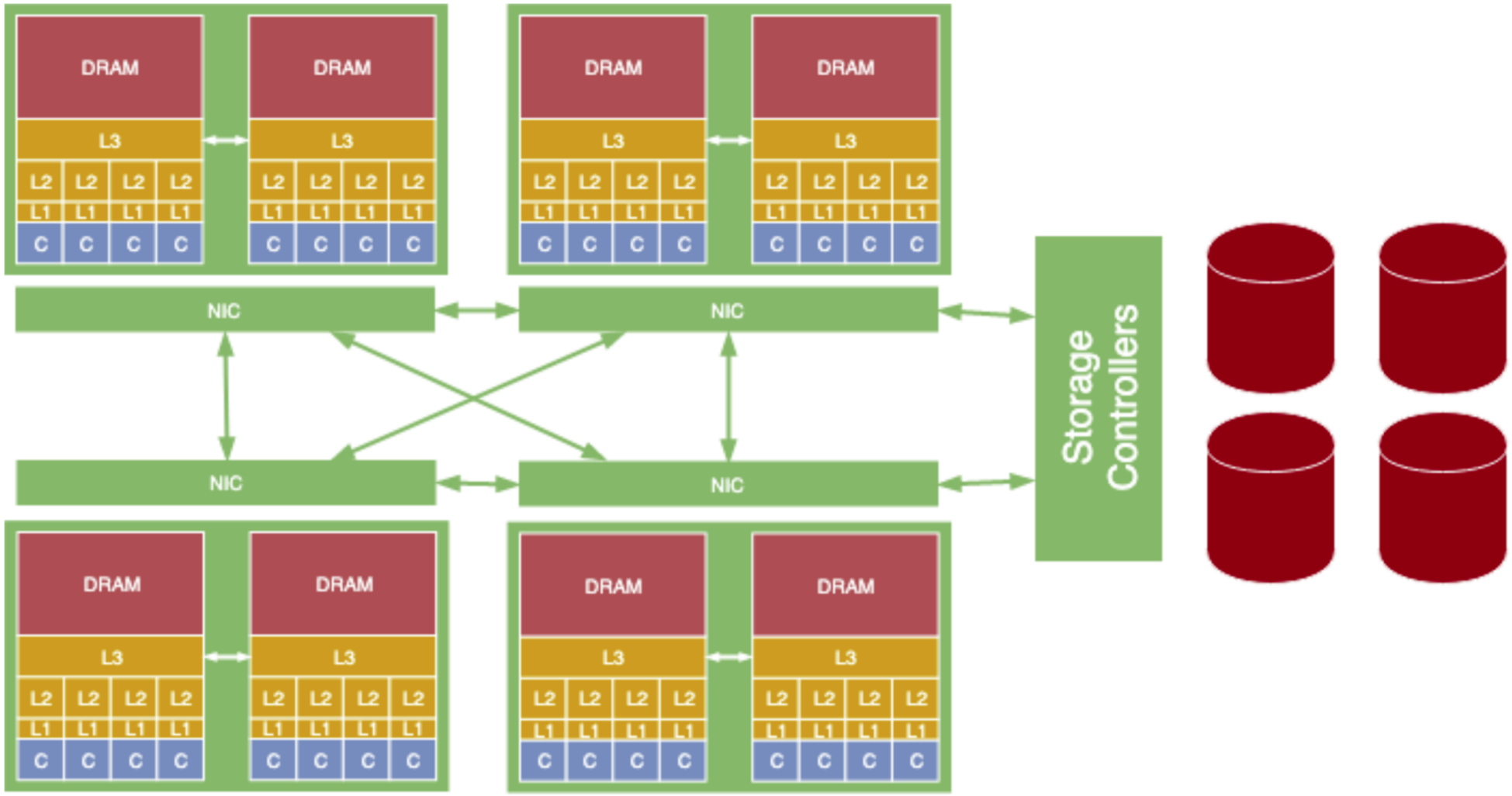


HPC BUILDING BLOCKS: NETWORK



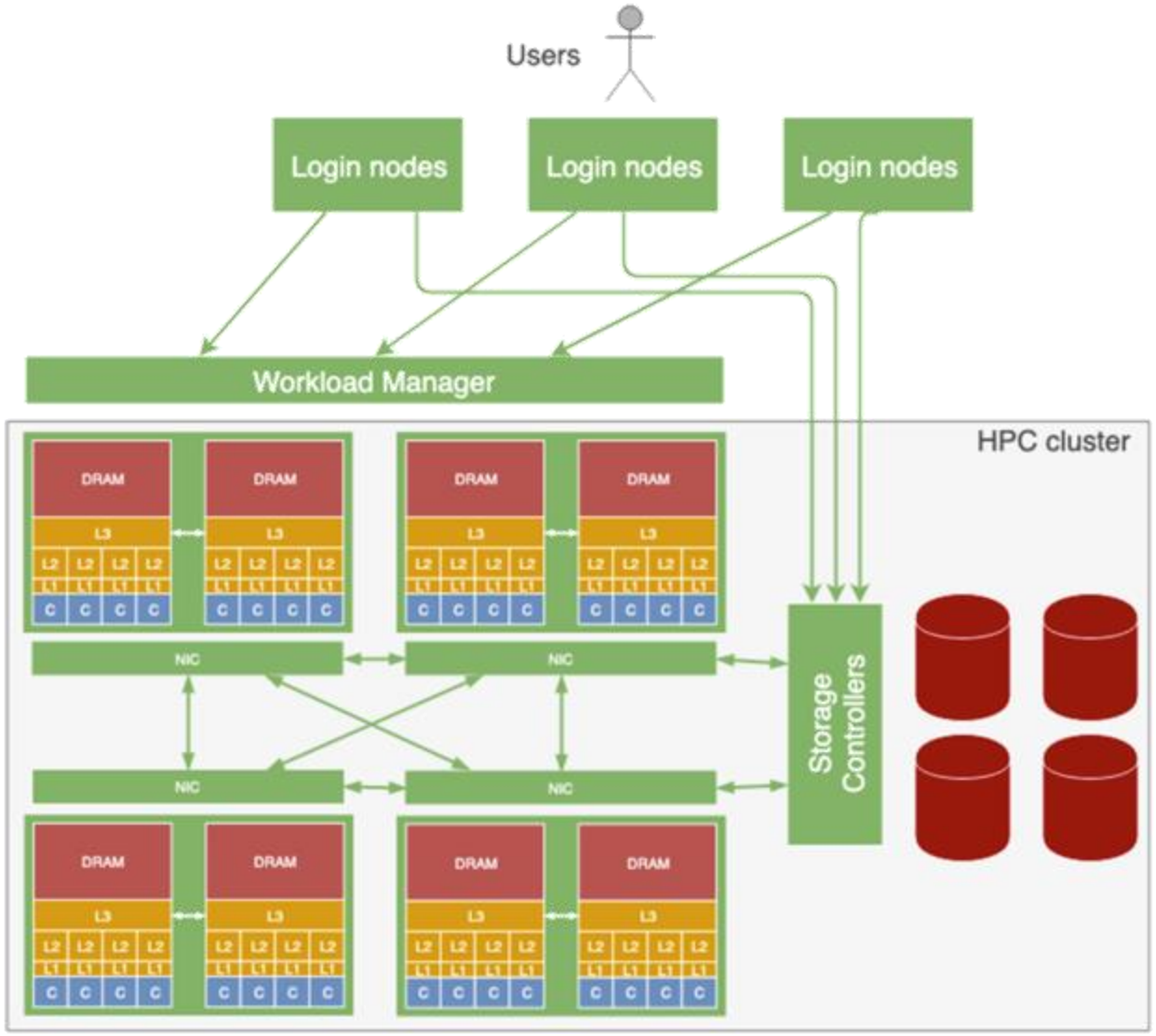
Source: CSCS-USI Summer School 2019

HPC BUILDING BLOCKS: STORAGE



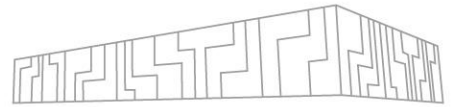
Source: CSCS-USI Summer School 2019

HPC BUILDING BLOCKS: LOGIN+SCHEDULER



Source: CSCS-USI Summer School 2019

BEYOND MULTICORE

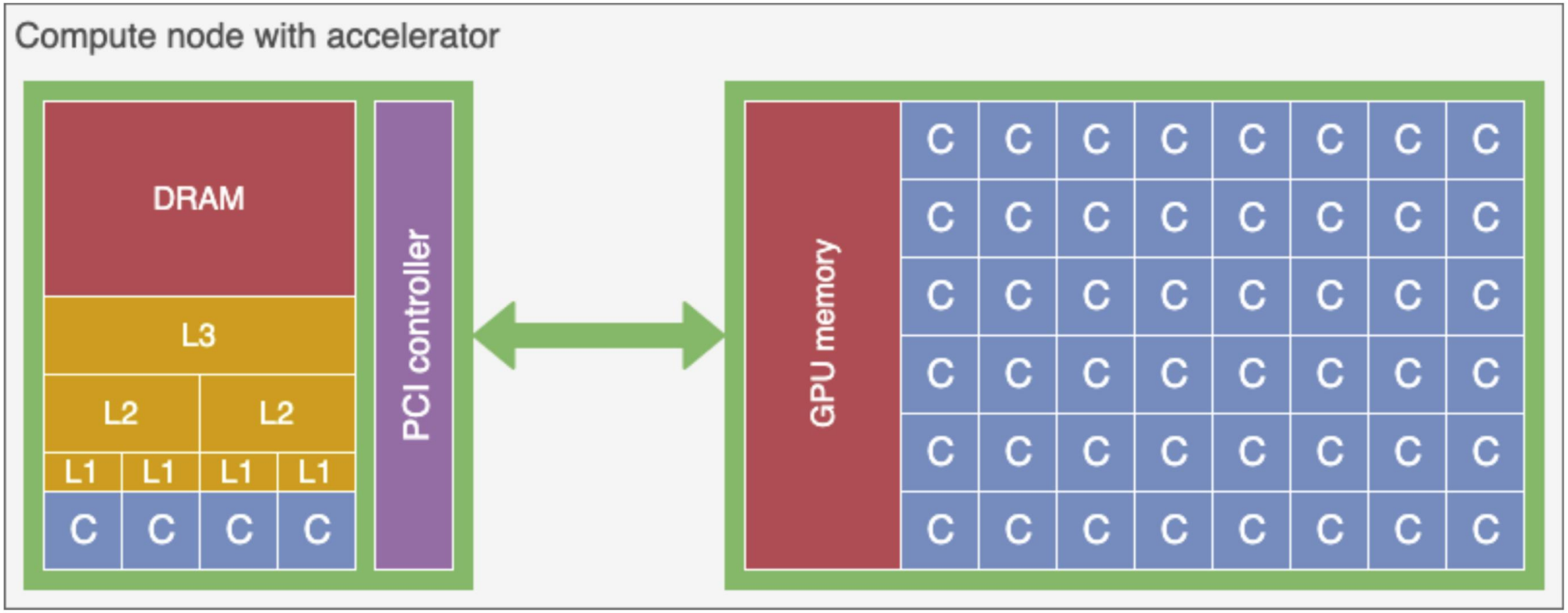


- Multicores have **limitations**
  - Fat cores (branch prediction, out-of-order execution, large caches)
    - Optimized for latency and multiprocessing
  - Still high frequencies
  - Still high-power consumption
  - But programming is easy; matches better our brain's serial way of thinking

- **Accelerators** are taking the opposite direction
  - Low frequencies, thus lower power consumption
  - Die area dedicated to processing units rather than control or caches
  - Suitable for very specific workloads; not for general-purpose tasks
  - Programming not so straightforward; we must think "parallel" now

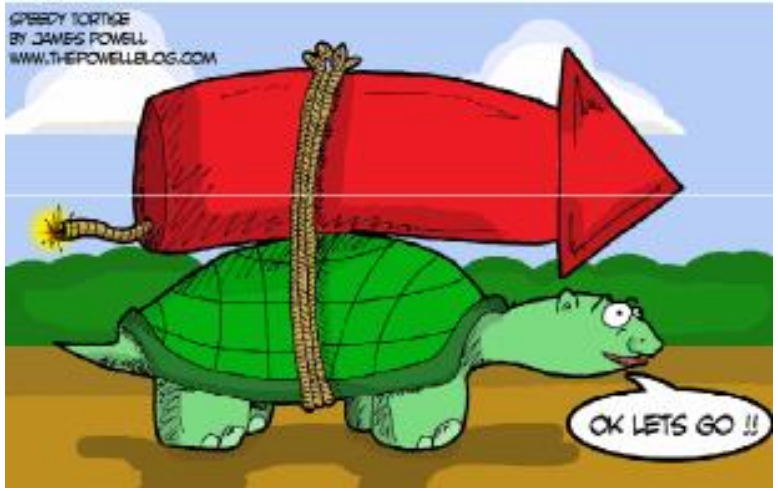
Source: CSCS-USI Summer School 2019

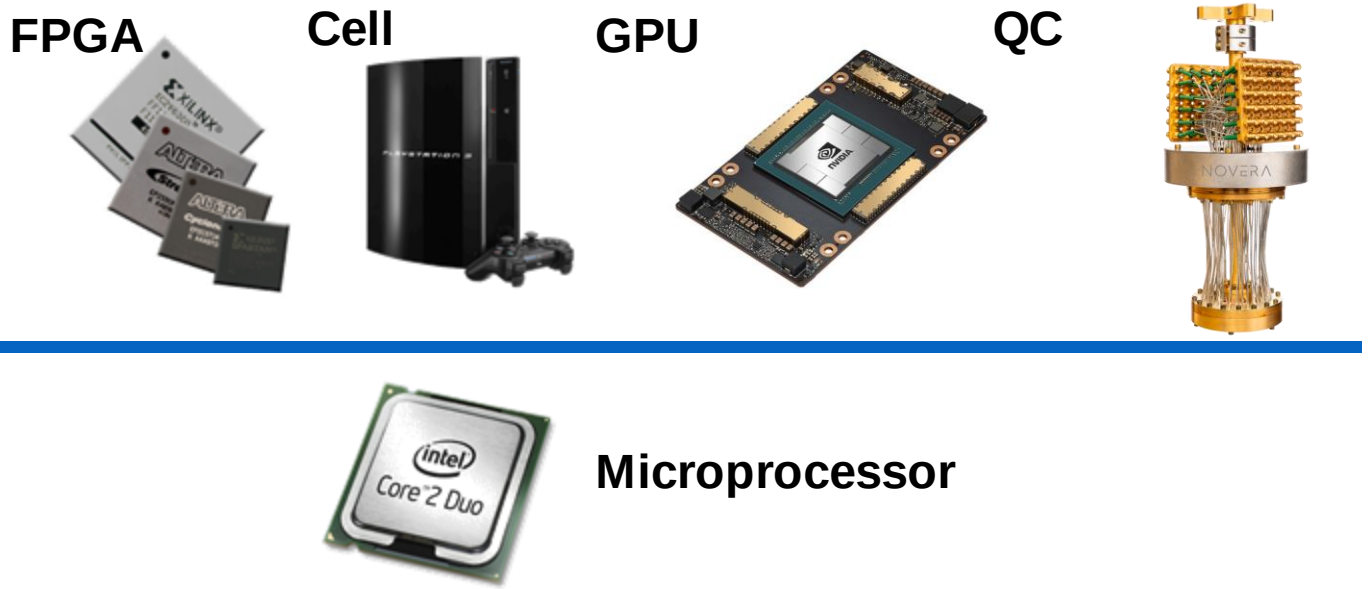
HPC BUILDING BLOCKS: ACCELERATOR



Source: CSCS-USI Summer School 2019

HETEROGENOUS COMPUTING

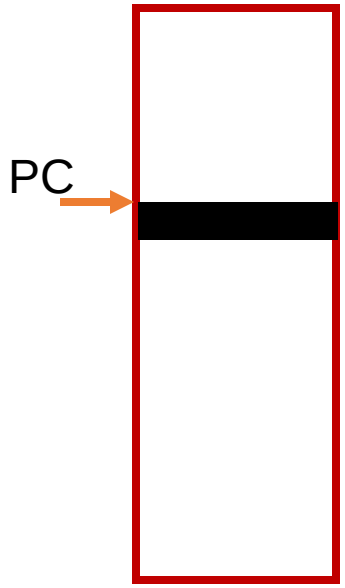


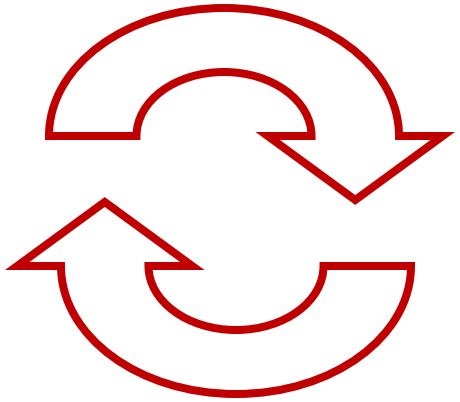


Hardware Accelerators - Speeding up the Slow Part of the Code

- Enable higher performance through fine-grained parallelism
- Offer higher computational density than CPUs
- Accelerators present heterogeneity!

ACCELERATED EXECUTION MODEL

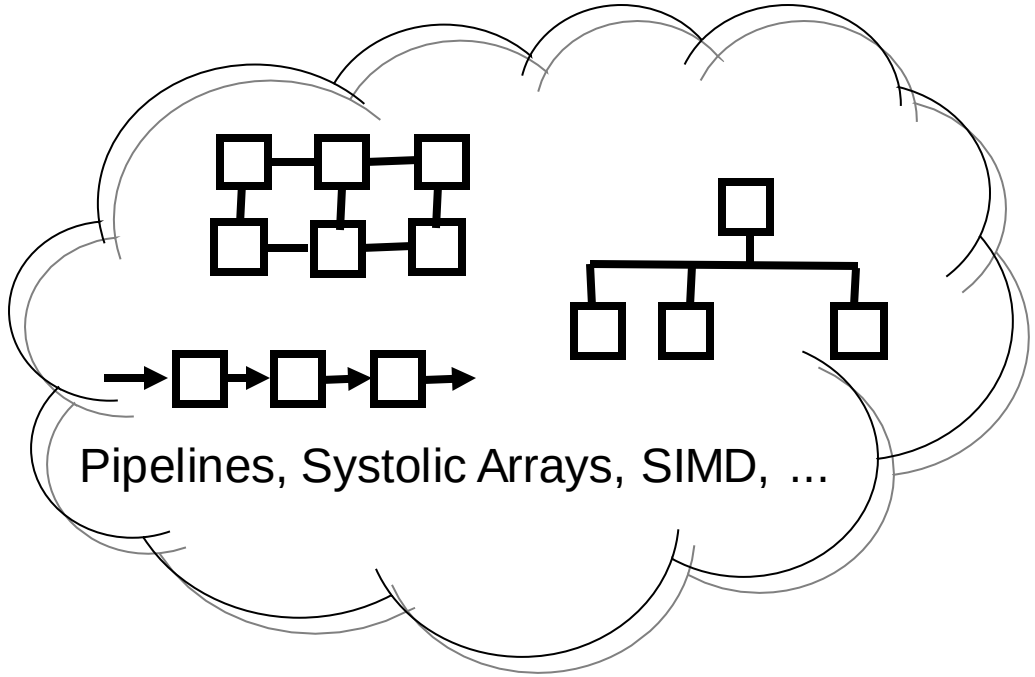




- Transfer of Control
- Input Data

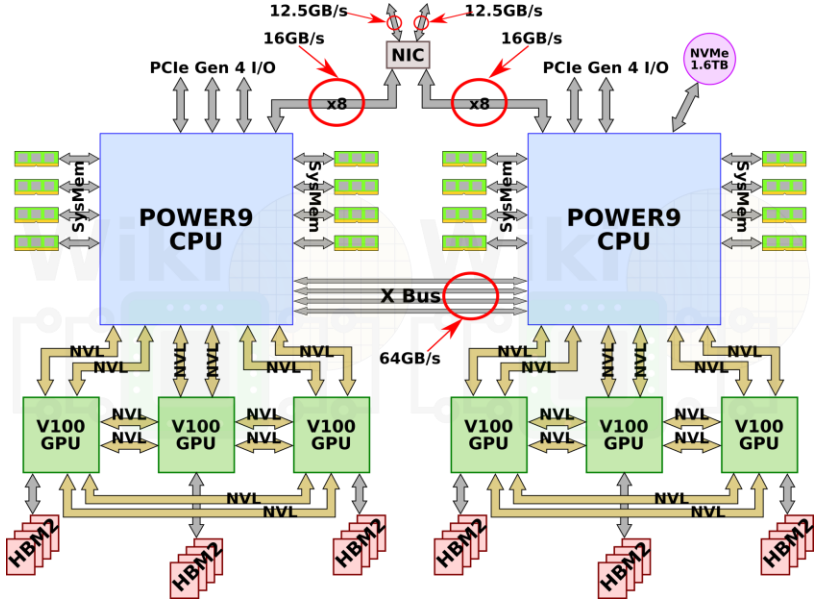
- Output Data
- Transfer of Control

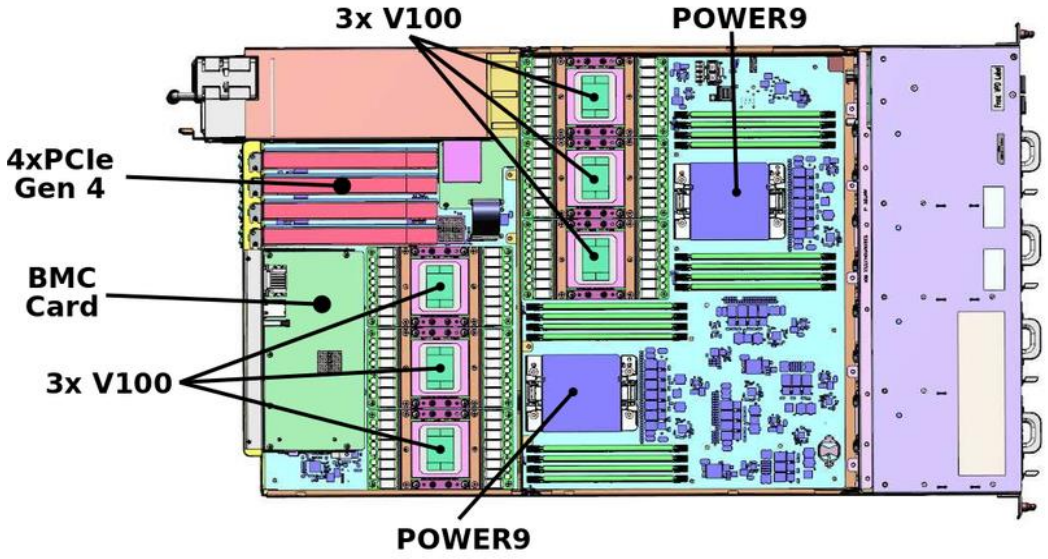
FPGA, GPU, Cell CBE, …

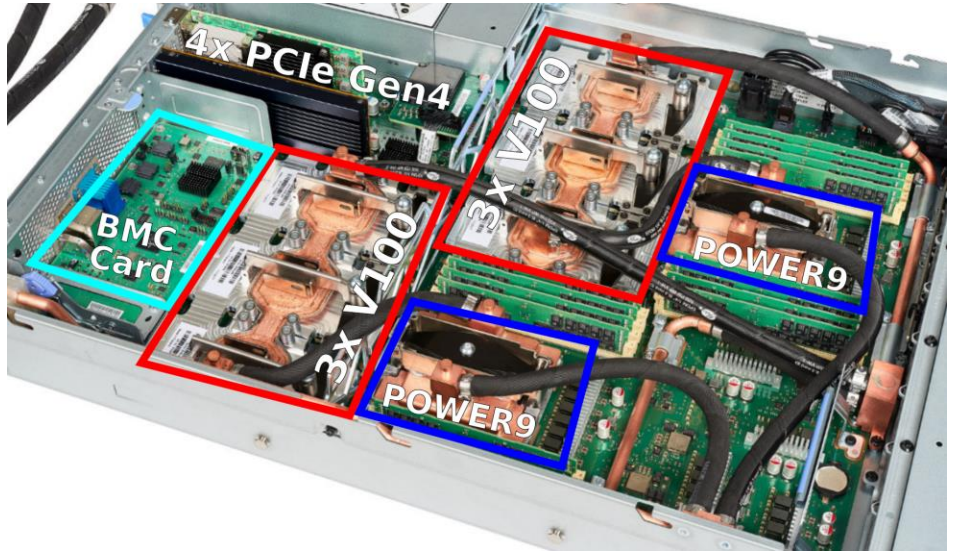


- Fine grain computations with the accelerators, others with the MP
- Interaction between accelerator and MP can be blocking or asynchronous
- This scenario is replicated across the whole system and standard HPC parallel programming paradigms used for interactions

SUMMIT SUPERCOMPUTER (2018)





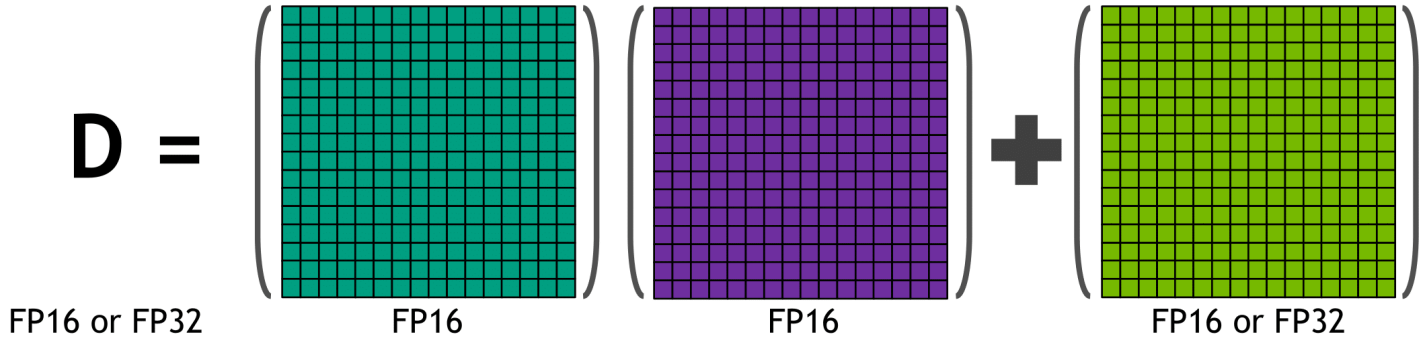


TENSOR CORES

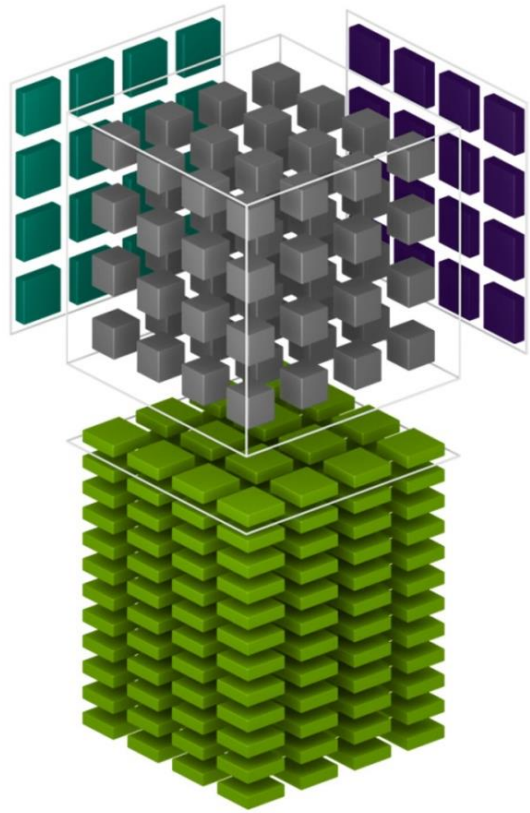
CUDA TENSOR CORE PROGRAMMING

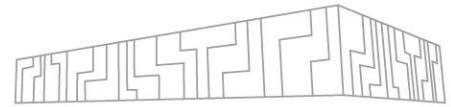
16x16x16 Warp Matrix Multiply and Accumulate (WMMA)

```
wmma::mma_sync(Dmat, Amat, Bmat, Cmat);
```



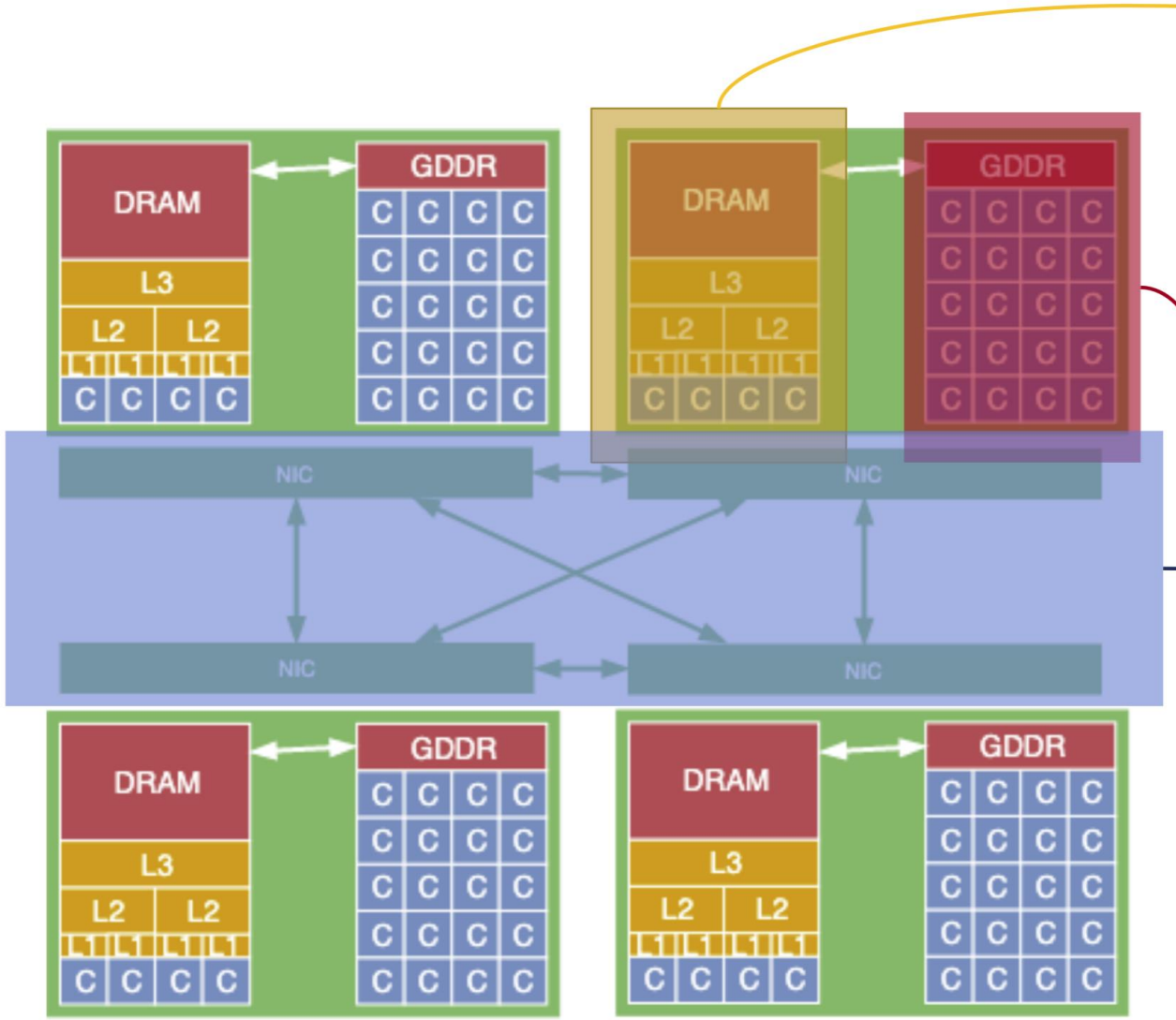
$$D = AB + C$$





SOFTWARE

HOW TO WRITE HPC CODE?



Shared memory: OpenMP, Task-based, POSIX threads, etc.

GPUs: CUDA, OpenACC, OpenMP 4.5, OpenCL, etc.

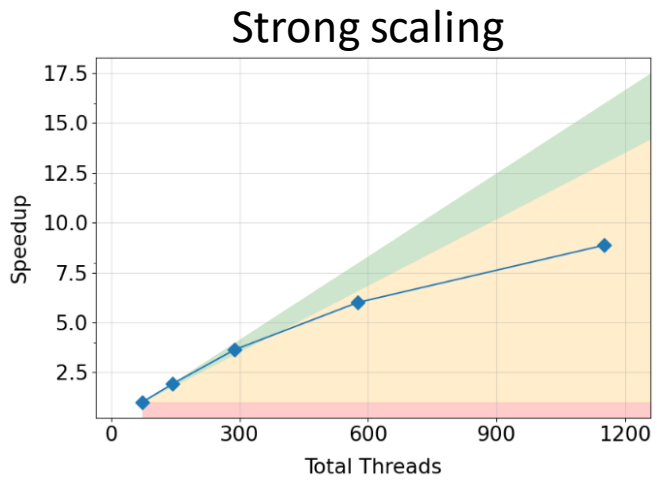
Distributed memory: MPI, Fortran coarrays, UPC, Charm++, etc.

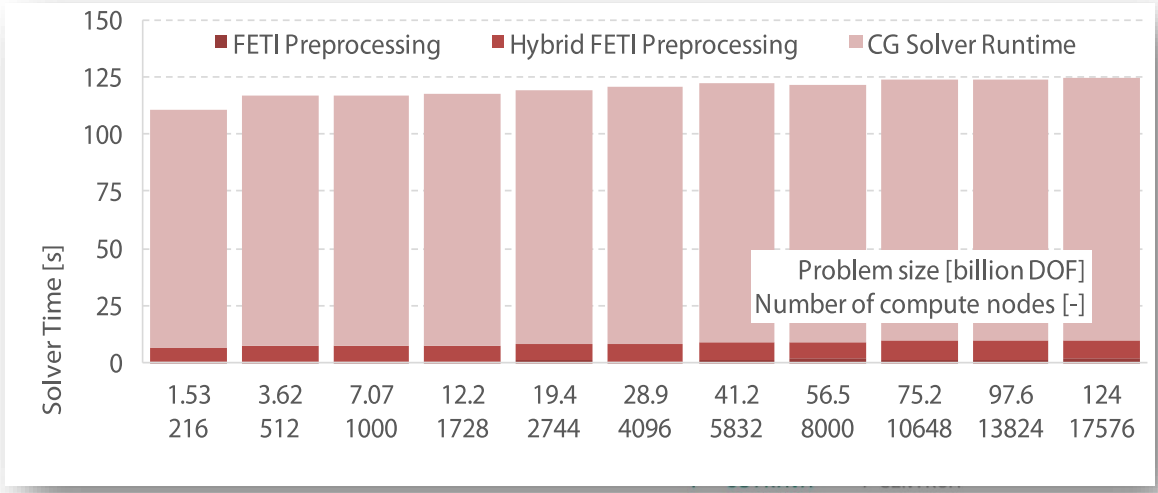
Source: CSCS-USI Summer School 2019

PARALLEL COMPUTING

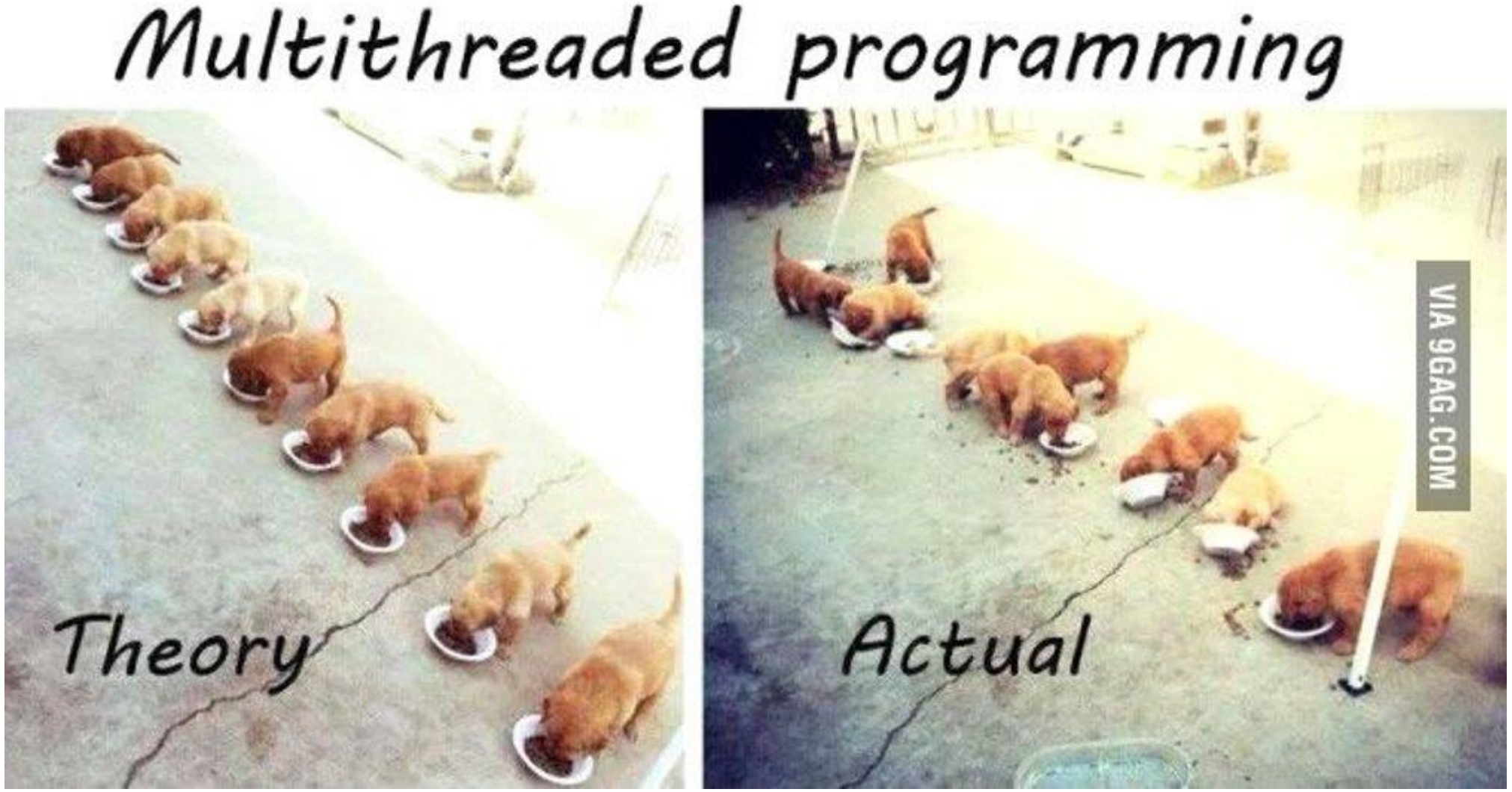
PARALLEL ALGORITHM SCALABILITY

- **Strong scaling**
  - Solve a problem using twice more resources
  - Expected performance – get result in half of time = linear scaling
  - Superlinear scaling
  - Strong scalability has a limitation!
- **Weak scaling**
  - Solving a twice larger problem using twice more resources
  - Expected performance – get result in constant time

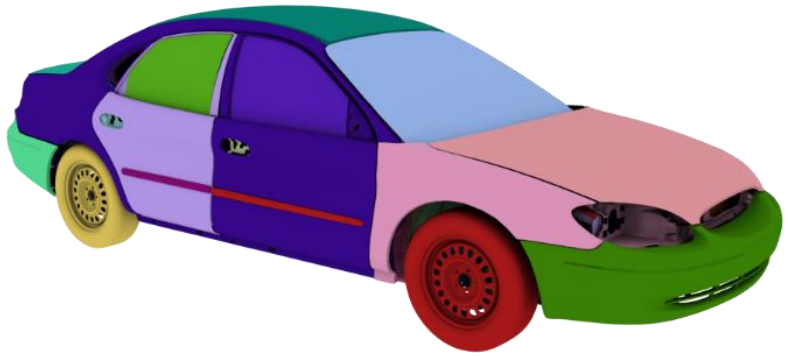


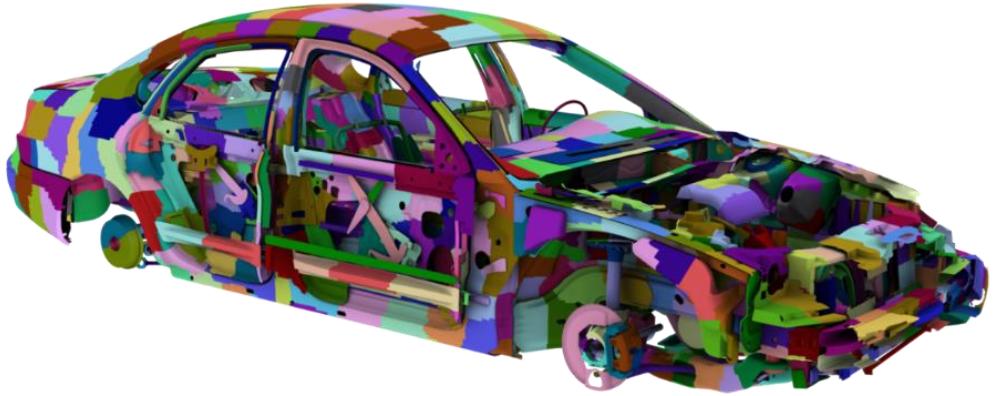


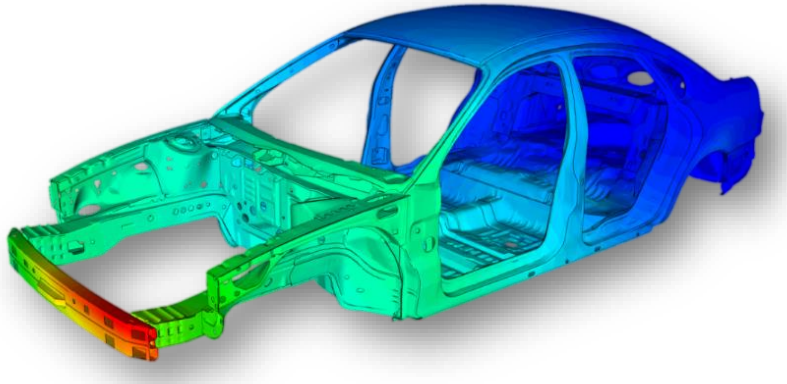
PARALLEL COMPUTING



PARALLEL COMPUTING



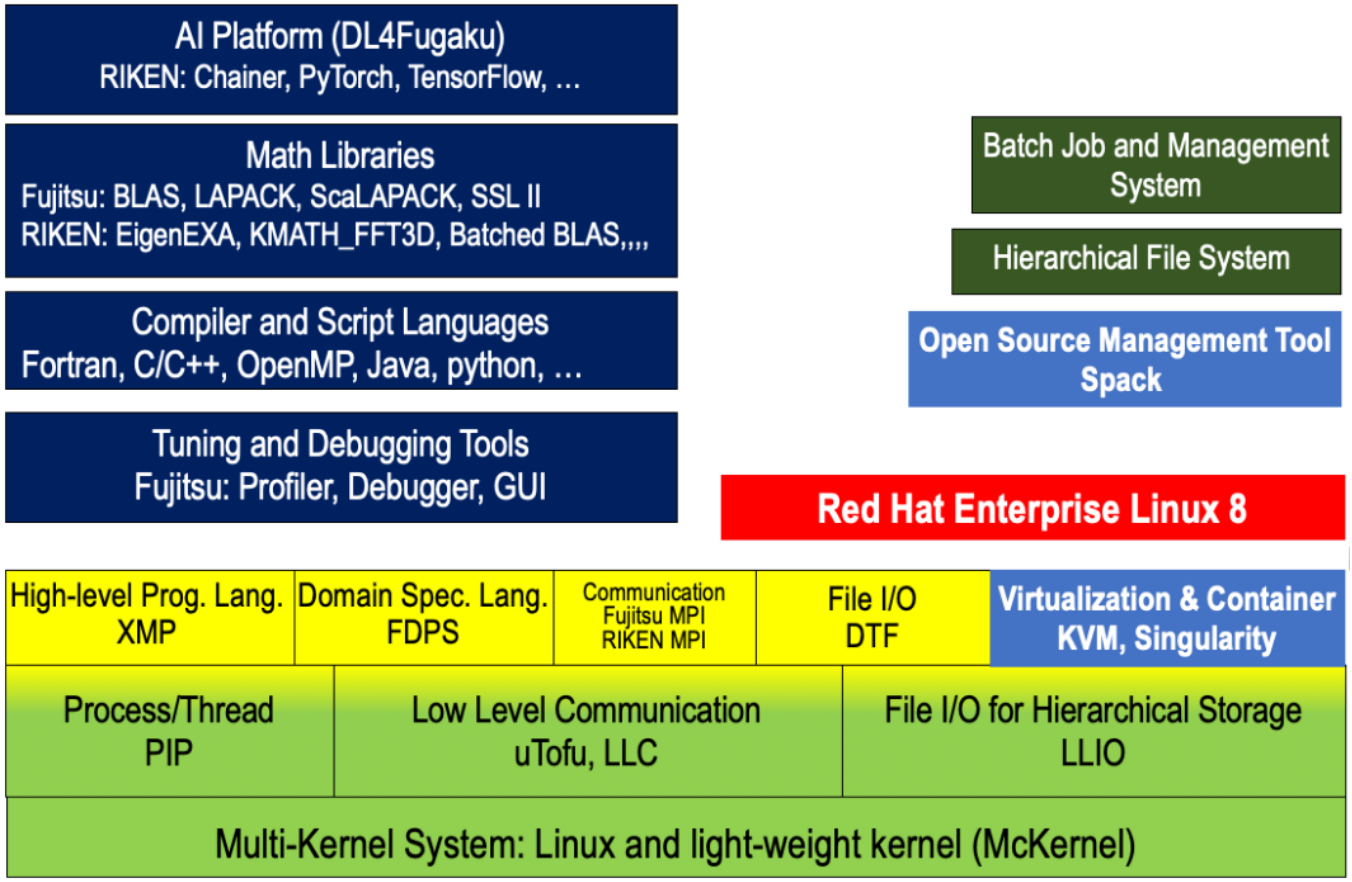




PRE-INSTALLED SOFTWARE

- Environment Module System
  - Modification of the environment paths
  - Software in several versions

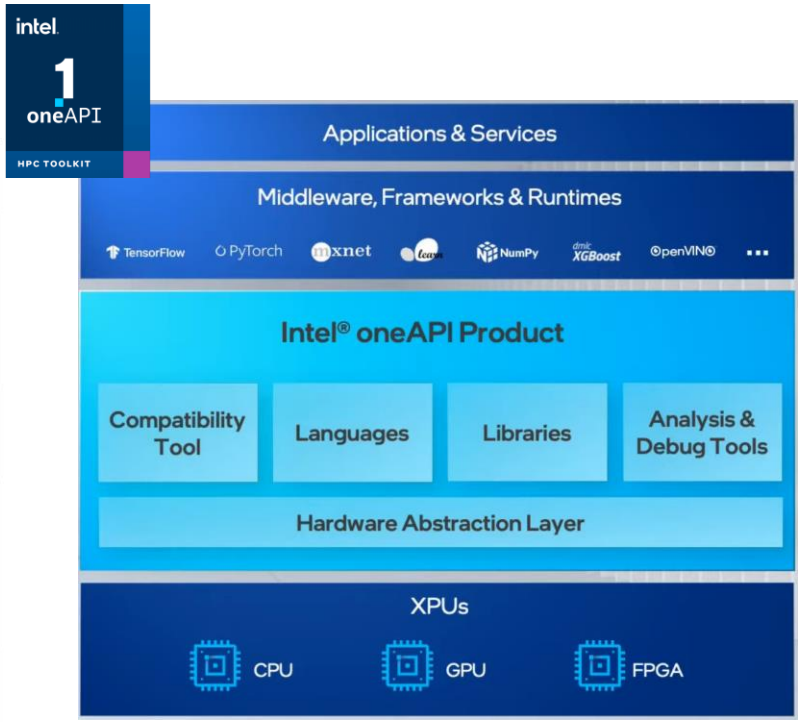
Fugaku software stack

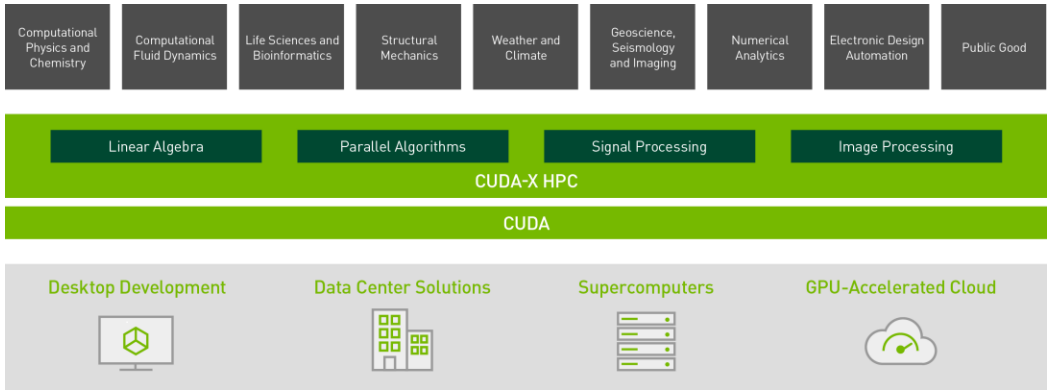


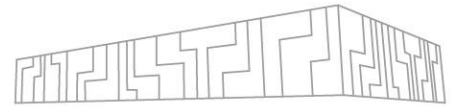
VENDOR'S SOFTWARE STACK

Simplified software development for heterogenous hardware

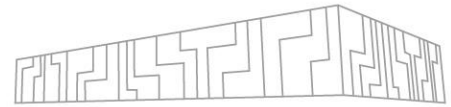
- Intel oneAPI
- AMD ROCm
- CUDA-X HPC & AI software stack







TRENDS



Path to exascale

TOP500 LIST

- List of the most powerful supercomputers
- Updated 2x a year – ISC (June) and SC (November)
- From 1993 High Performance Linpack (HPL) benchmark
- From 2017 also High-Performance Conjugate Gradient (HPCG) Benchmark
- From 2013 Green500 list
- From 2019 HPL-AI – not a list yet - mixed-precision algorithms

TOP500 LIST HPL + HPCG

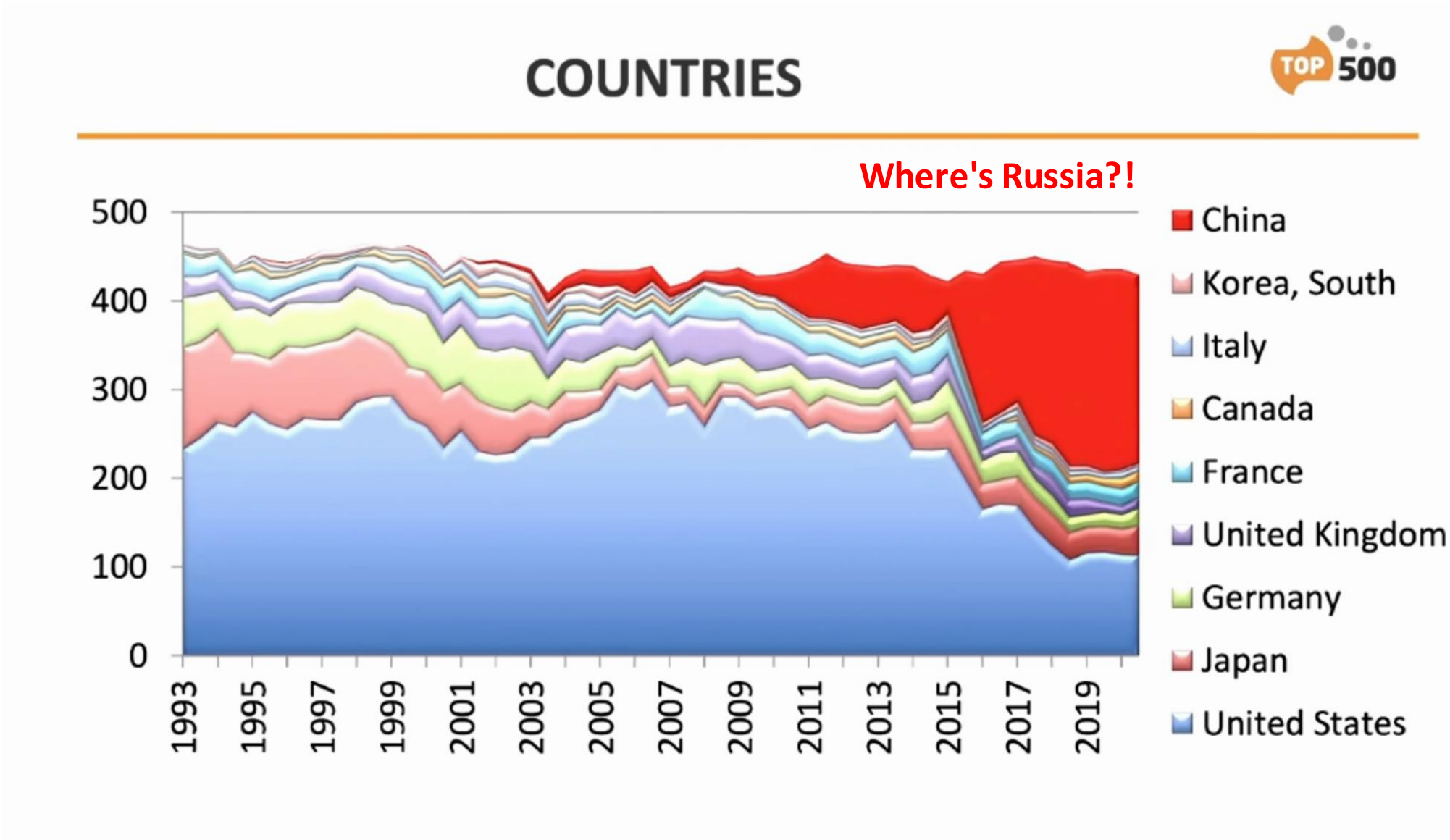
Rank	System	Cores	Rmax (PFlop/s)	Rpeak (PFlop/s)	Power (kW)
1	**Frontier** - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE DOE/SC/Oak Ridge National Laboratory United States	8,699,904	1,206.00	1,714.81	22,786
2	**Aurora** - HPE Cray EX - Intel Exascale 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
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	

Rank	TOP500 Rank	System	Cores	Rmax (PFlop/s)	HPCG (TFlop/s)
1	4	**Supercomputer Fugaku** - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan	7,630,848	442.01	16004.50
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	8,699,904	1,206.00	14054.00
3	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	5612.60
4	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	4586.95
5	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	3671.32
6	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	3113.94
7	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	2925.75
8	14	**Perlmutter** - HPE Cray EX 235n, AMD EPYC 7763 64C 2.45GHz, NVIDIA A100 SXM4 40 GB, Slingshot-11, HPE DOE/SC/LBNL/NERSC United States	888,832	79.23	1905.00
9	12	**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	1795.67
10	15	**Selene** - NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Mellanox HDR Infiniband, Nvidia NVIDIA Corporation United States	555,520	63.46	1622.51

06/2024

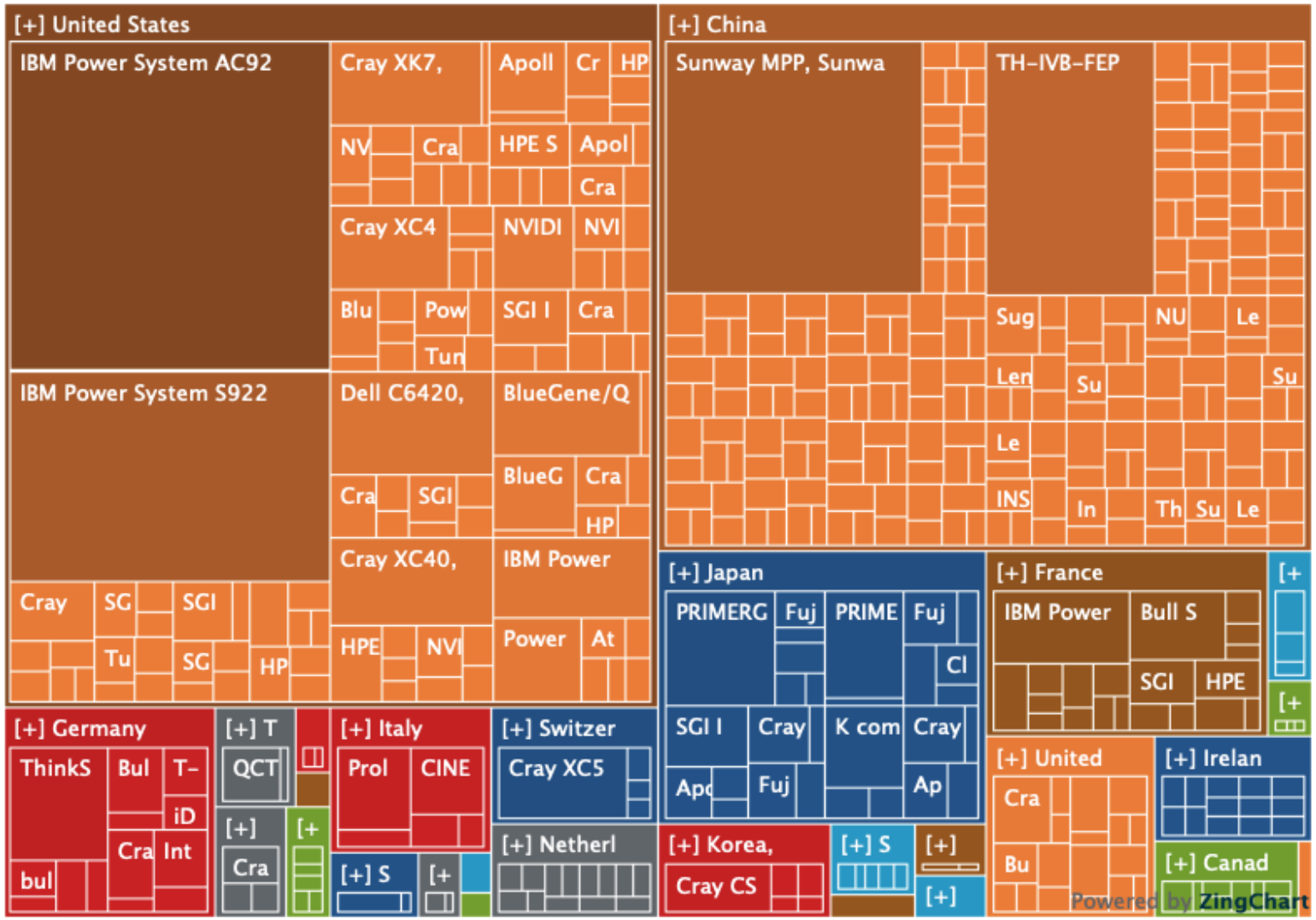
VSB TECHNICAL UNIVERSITY OF OSTRAVA | IT4INNOVATIONS NATIONAL SUPERCOMPUTING CENTER

TOP500 LIST



11/2020

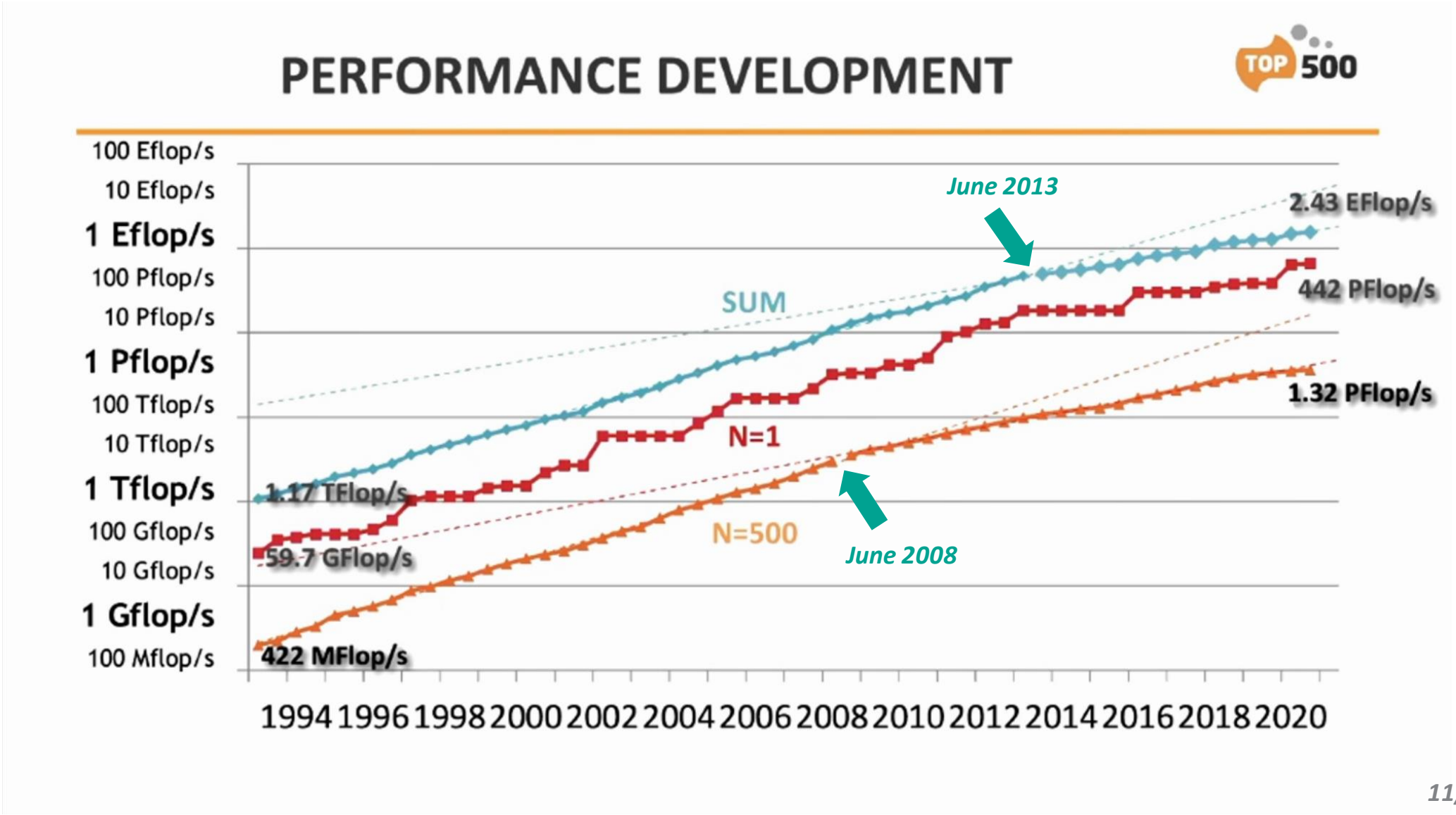
TOP500 LIST



Countries	Count	System Share (%)	Rmax (GFlops)	Rpeak (GFlops)	Cores
China	220	44	466,872,778	887,822,195	26,935,688
United States	116	23.2	600,014,746	851,002,631	17,337,080
Japan	28	5.6	116,184,300	180,998,613	3,355,148
France	20	4	68,205,127	102,530,990	2,212,232
United Kingdom	18	3.6	39,955,369	49,191,669	1,518,312
Ireland	13	2.6	21,438,430	27,555,840	748,800
Netherlands	13	2.6	20,877,830	26,763,264	730,080
Germany	13	2.6	57,856,910	83,721,088	1,442,678
Canada	8	1.6	14,497,480	27,682,534	447,488
Australia	5	1	6,669,188	10,232,963	257,336
Italy	5	1	30,098,790	47,843,836	794,032
Korea, South	5	1	20,966,960	34,322,860	786,020
Singapore	5	1	7,719,590	9,891,840	268,800
Switzerland	4	0.8	25,373,050	32,173,545	529,940
Brazil	3	0.6	4,082,300	7,123,661	125,184
India	3	0.6	7,457,490	8,228,006	241,224
Saudi Arabia	3	0.6	10,109,130	13,858,214	325,940
South Africa	3	0.6	3,275,620	4,193,050	109,656
Finland	2	0.4	2,956,730	4,377,293	80,608
Russia	2	0.4	3,678,350	6,239,795	99,520
Sweden	2	0.4	4,771,700	6,773,346	131,968
Spain	2	0.4	7,615,800	11,699,115	171,576
Taiwan	2	0.4	10,325,150	17,297,190	197,552
Poland	1	0.2	1,670,090	2,348,640	55,728
Austria	1	0.2	2,726,078	3,761,664	37,920
Denmark	1	0.2	1,069,554	2,107,392	31,360
Czech Republic	1	0.2	1,457,730	2,011,641	76,896
Hong Kong	1	0.2	1,649,110	2,119,680	57,600

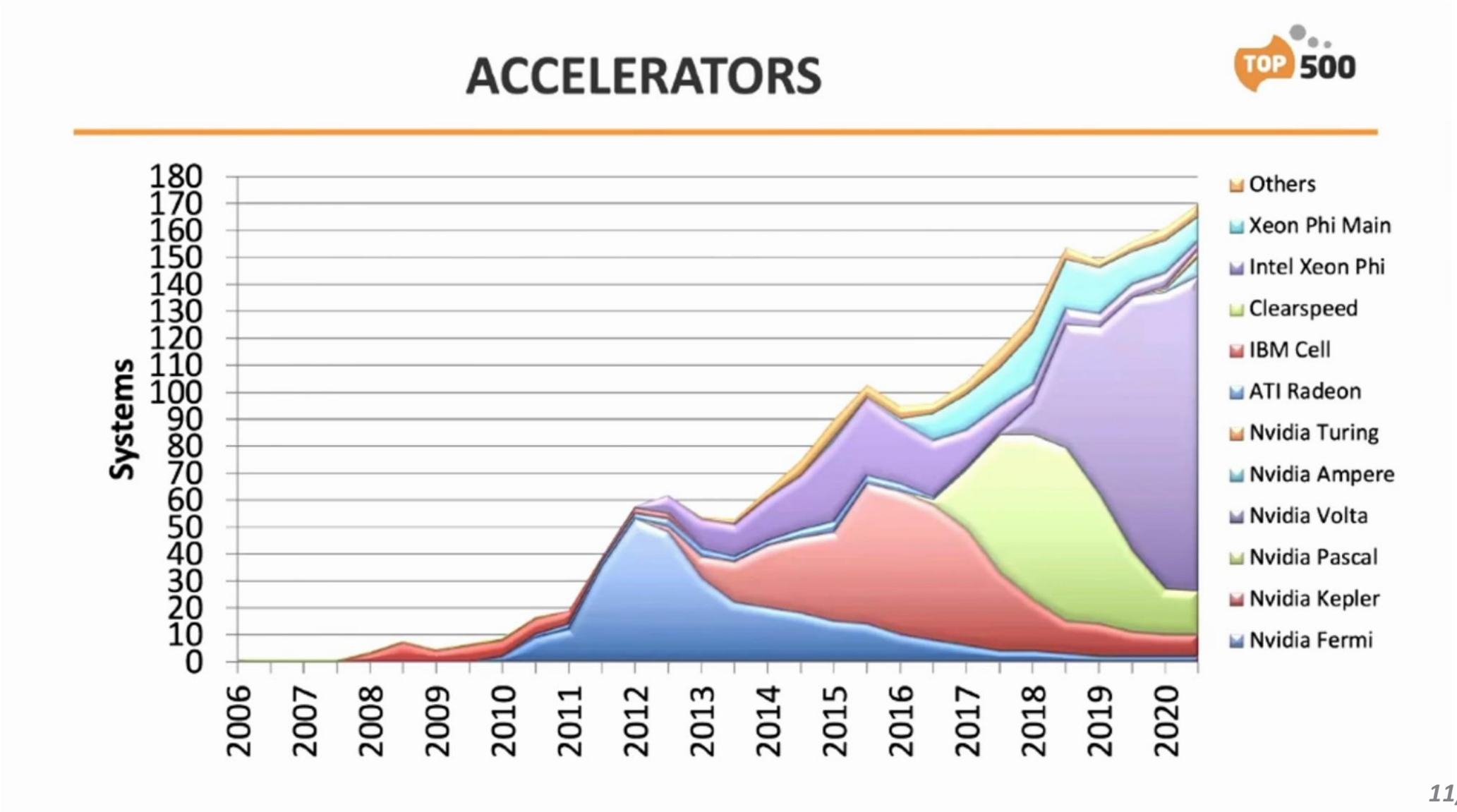
6/2019

TOP500 LIST



11/2020

TOP500 LIST



11/2020

6/2024 - 193 out of 500 systems are accelerated

TOP500 LIST HPL

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
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

06/2024

TOP500 LIST HPL

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
					52.5 GF/W
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
					26.2 GF/W
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
					14.8 GF/W
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
					51.6 GF/W
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
					52.0 GF/W
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
					32.2 GF/W

Exascale goal is
50 GFlops/Watt = 20 MW system

06/2024

TOP500 LIST HPL

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
					52.5 GF/W
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
					26.2 GF/W
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
					14.8 GF/W
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
					51.6 GF/W
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
					52.0 GF/W
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
					32.2 GF/W

The GREEN 500

- Direct Warm-Water Cooling (CPU and GPU cooling separated circles)
- Availability of power controling knobs
- Higher heterogenity of new systems = using accelerators, GPGPUs, FPGAs, single/mixed precission units
- Decarbonization
- AI everywhere
- And many more

06/2024

GREEN500

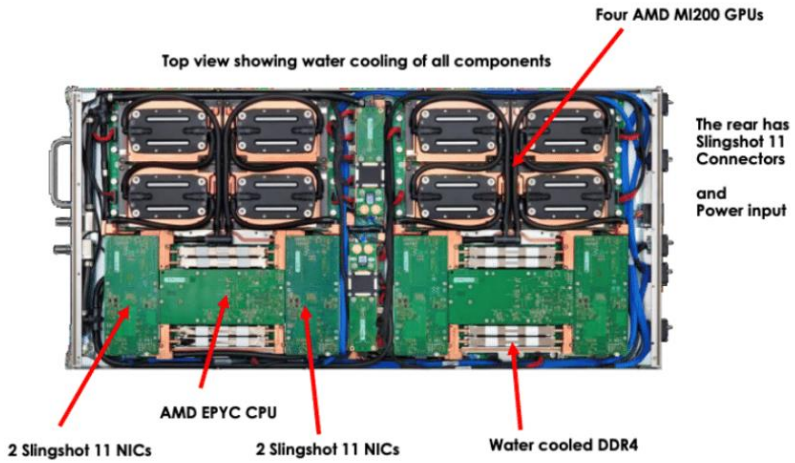
The GREEN 500

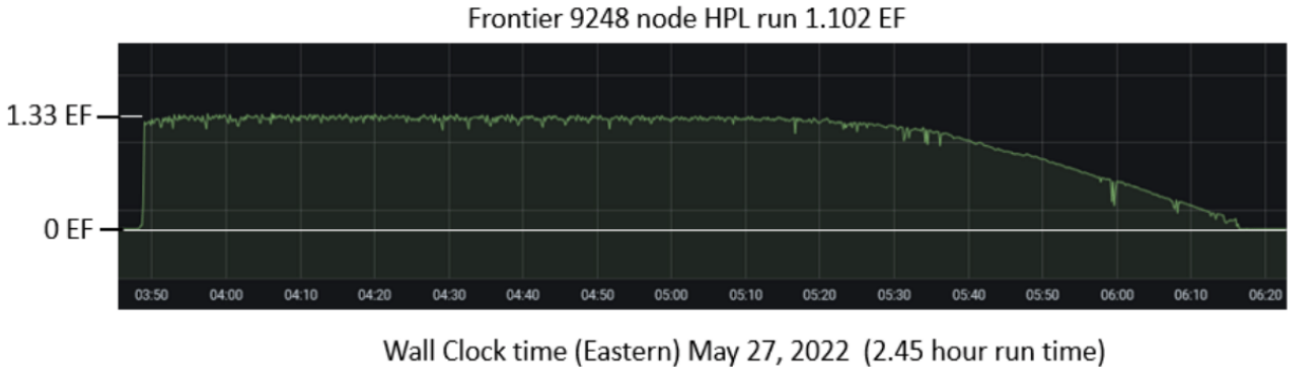
Rank	TOP500 Rank	System	Cores	Rmax (PFlop/s)	Power (kW)	Energy Efficiency (GFlops/watts)	
1	189	**JEDI** - BullSequana XH3000, Grace Hopper Superchip 72C 3GHz, NVIDIA GH200 Superchip, Quad-Rail NVIDIA InfiniBand NDR200, ParTec/EVIDEN EuroHPC/FZJ Germany	19,584	4.50	67	72.733	Nvidia GH200
2	128	**Isambard-AI 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	Nvidia GH200
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	Nvidia GH200
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	Nvidia H100
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	Nvidia GH200
6	299	**HoreKa-Teal** - 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	Nvidia H100
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	AMD MI250X
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	Nvidia GH200
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	AMD MI250X
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	AMD MI250X

06/2024

FRONTIER

- 74 HPE Cray EX cabinets, 9 408 nodes
- 1 AMD Milan “Trento” 7A53 Epyc CPU + 4 AMD Instinct MI250X GPUs
- 512GiB DDR4 + 512GiB HMB2e (128GiB per GPU) coherent memory across node
- HPE Slingshot-11 interconnect (200 Gbit/s)
- 1.102 exaflops of Linpack, 21.1 MW





USA ROADMAP

EXASCALE COMPUTING PROJECT

Pre-Exascale Systems

Future Exascale Systems

2012	2016	2018	2020	2021–2023

ORNL
Cray/AMD/
NVIDIA

LBNL
Cray/Intel

ORNL
IBM/NVIDIA

LBNL
Cray/AMD/
NVIDIA

ORNL
Cray/AMD

1.5 EFlops
AMD CPU + AMD GPU

ANL
IBM BG/Q

ANL
Intel/Cray

ANL
Intel/Cray

~~1 EFlops~~
2 EFlops
Intel CPU + Intel X^{e}

LLNL
IBM BG/Q

LANL/SNL
Cray/Intel

LLNL
IBM/NVIDIA

LANL/SNL
TBD

LLNL
Cray

>2EFlops, ~40 MW
AMD CPU + GPU

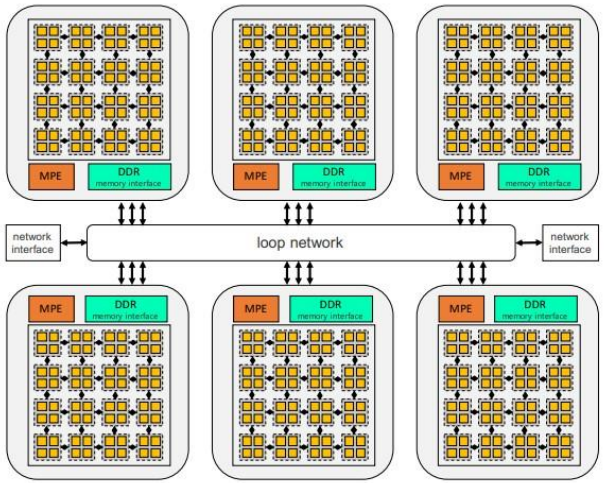
High variability of CPU and GPU vendors

SUPERCOMPUTER #1 ?!

- Frontier (USA) 06/2022 - 1.102 exaflops of Linpack, 21.1 MW

Meanwhile in China:

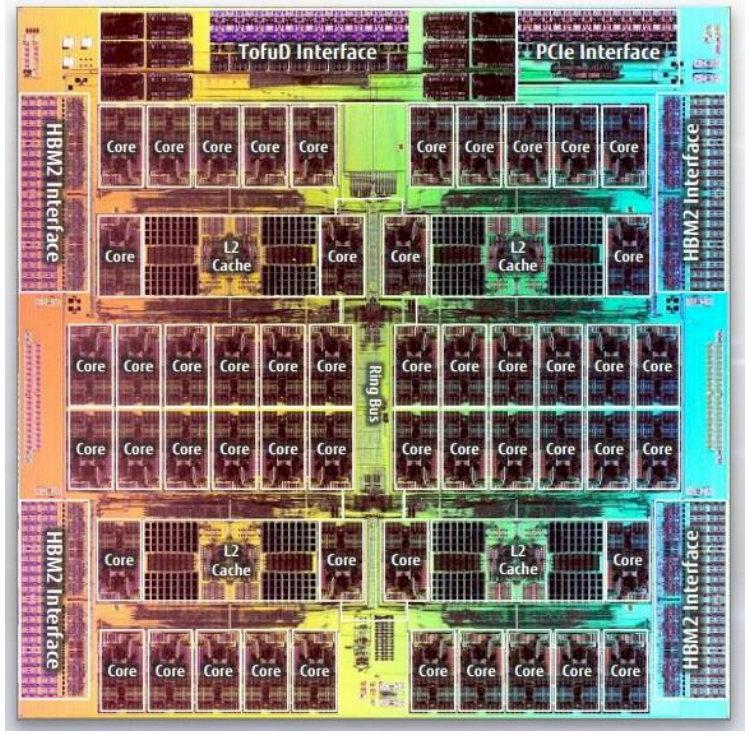
- Sunway Oceanlite (03/2021) - 1.05 exaflops of Linpack, ~35MW
  - ShenWei post-Alpha CPU ISA, 512-bit IS
  - 96 cabinets, 98 304x SW39010 390-core CPU, 14nm
  - Not in the top500.org list
- Tianhe-3 (10/2021) - 1.3 exaflops Linpack
  - 2x Phytium 2000+ FTP ARM CPU (16nm) + Matrix 2000+ MTP accelerator
  - Not in the top500.org list
- Shenzhen Phase 2 - scheduled for 2022
  - 2 exaflops
  - Sugon's Hygon CPU - delayed

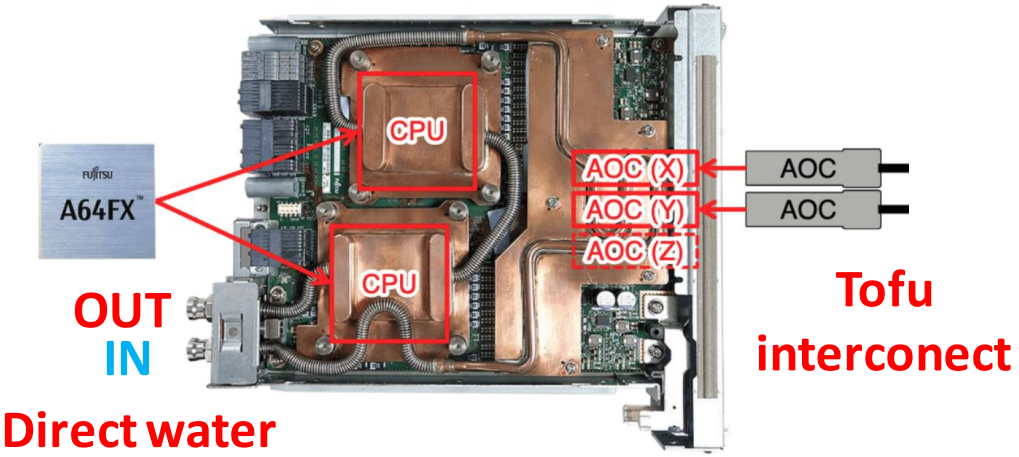


SW26010Pro

FUGAKU SUPERCOMPUTER

- 158 976 nodes, node peak performance 3.4 TFLOP/s
- Fujitsu A64FX ARM v8.2-A, 48(+4) cores, SVE 512 bit instruction
- high bandwidth 3D stacked memory, 4x 8 GB HBM with 1 024 GB/s
- on-die Tofu-D network BW (~400Gbps)
- 29.9 MW





THE EUROHPC JOINT UNDERTAKING

EuroHPC
Joint Undertaking

- A legal and funding agency
- 35 member countries
- **A co-founding programme to build a pan-European supercomputing infrastructure**

Installed medium-to-high range Supercomputers

- **Bulgaria** (6PF, AMD+Nvidia), **Czech Republic** (15PF, AMD+Nvidia), **Luxembourg** (18PF, AMD+Nvidia), **Portugal** (10PF, A64FX, AMD+Nvidia), **Slovenia** (6.8PF, AMD+Nvidia)

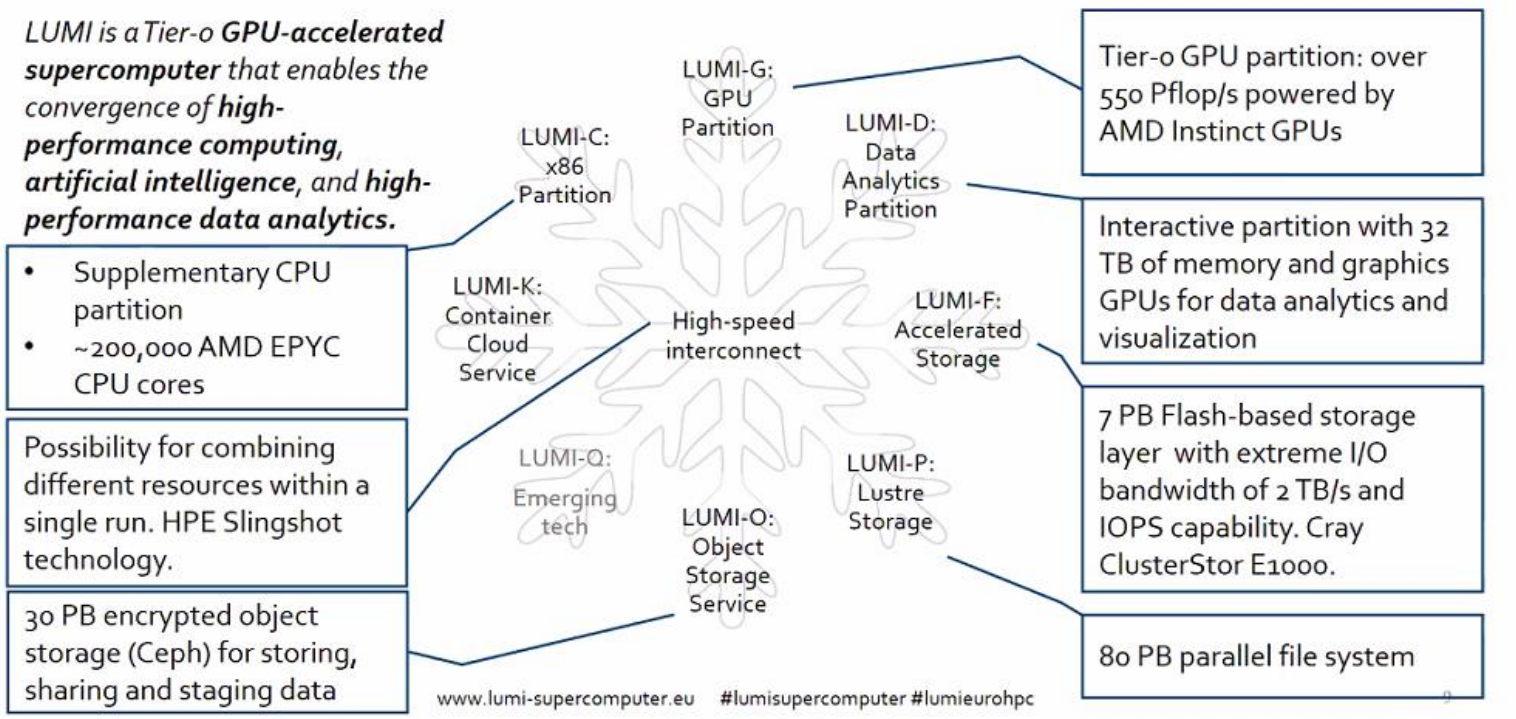
High-range Pre-Exascale Supercomputers

- 150-200 Pflops
- **Finland**, **Spain** and **Italy** consorciums

Next generations of systems planned

EUROPEAN PRE-EXASCALE SYSTEMS

LUMI



- **LUMI-C** - 2xAMD 7763 CPUs
  - 6.3 PFlops linpack
- **LUMI-G** – AMD Trento + 4xAMD MI250X
  - 151.9 PFlops linpack

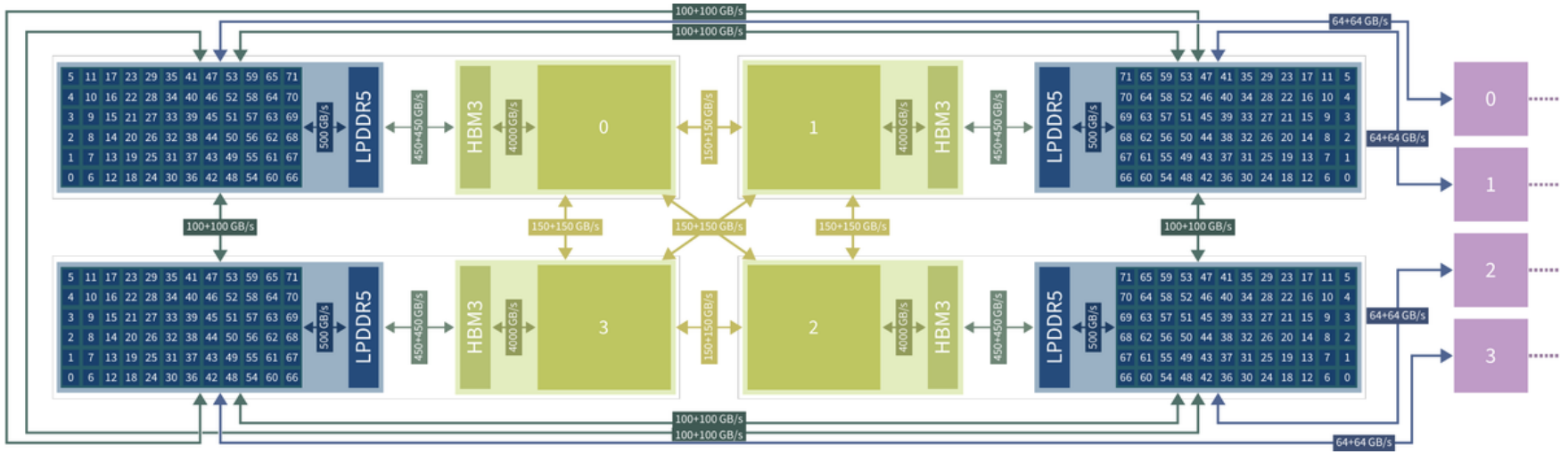
- H2 2022
- 240M €, 248 PFlops
- 3456 accelerated nodes
  2x Intel Xeon Ice Lake CPUs
  + 4 Nvidia A100 GPUs
- 1536 non-accelerated nodes
  2x Intel Xeon Sapphire Rapids

MareNostrum V

- H2 2023
- 223M €, 200 PFlops
- 2x Intel Sapphire Rapids +
  4x Nvidia H100

JUPITER SUPERCOMPUTER

- ~6000 nodes of Nvidia Grace Hopper, 1 ExaFLOP/s
- >1300 nodes of SiPearl Rhea, 5 PetaFLOP/s



EUROPEAN PROCESSOR INITIATIVE (EPI)

European Processor Initiative

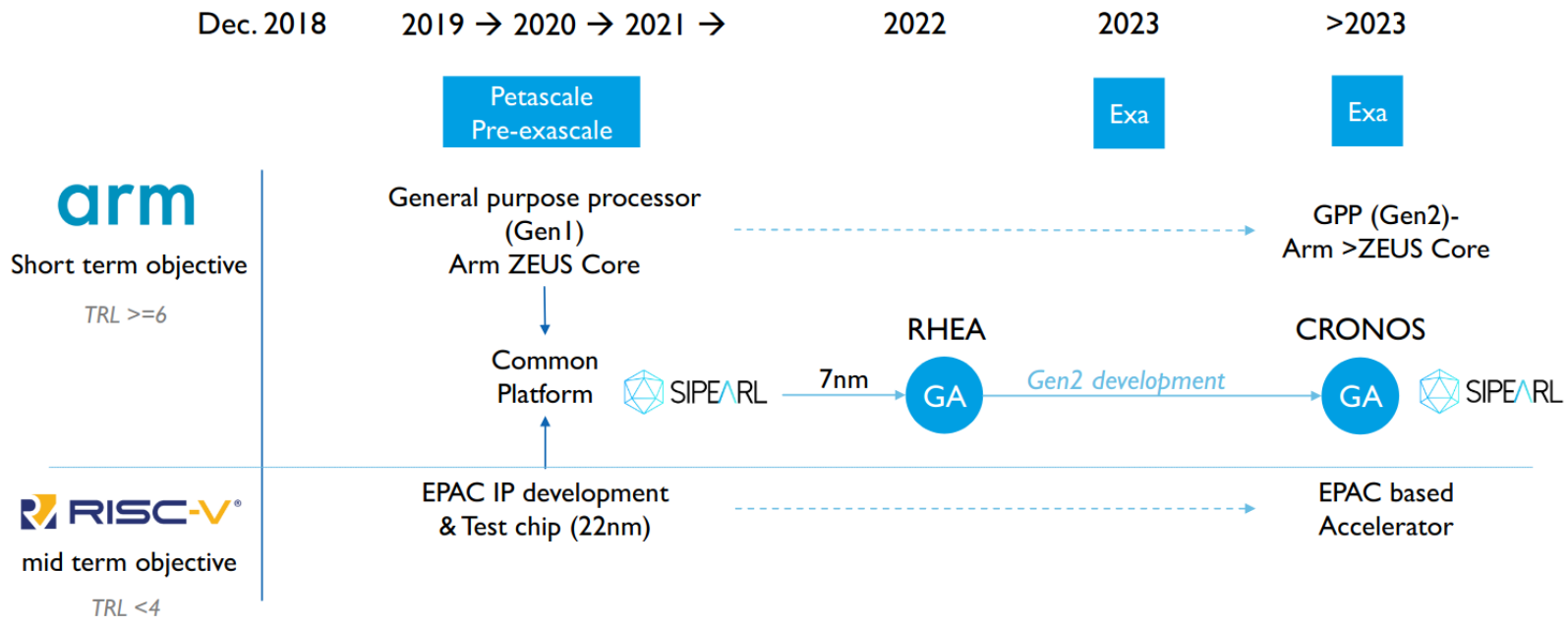
EUPEX
European Pilot for Exascale

Europe invests into development of a new processor

- Security
- Competitiveness

Design a roadmap of future European low power processors

- common platform
- general purpose processor
- accelerator
- automotive

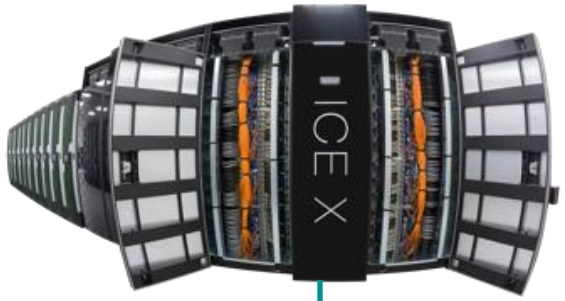


HISTORY OF THE IT4INNOVATIONS

Anselm

5/2011 – 6/2013

Salomon



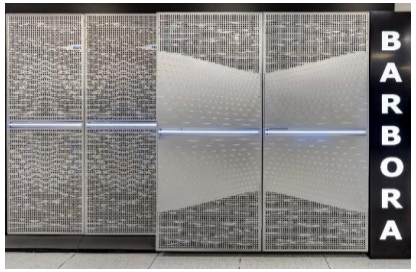
7/2014 – 7/2015

NVIDIA DGX-2

ARTIFICIAL INTELLIGENCE

3/2019

Barbora

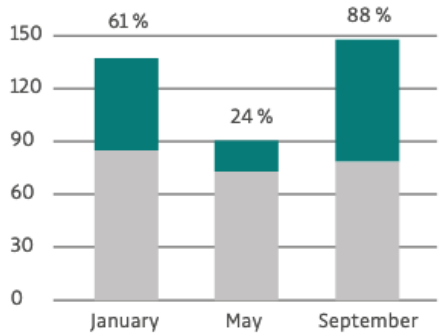


10/2019 – 7/2021

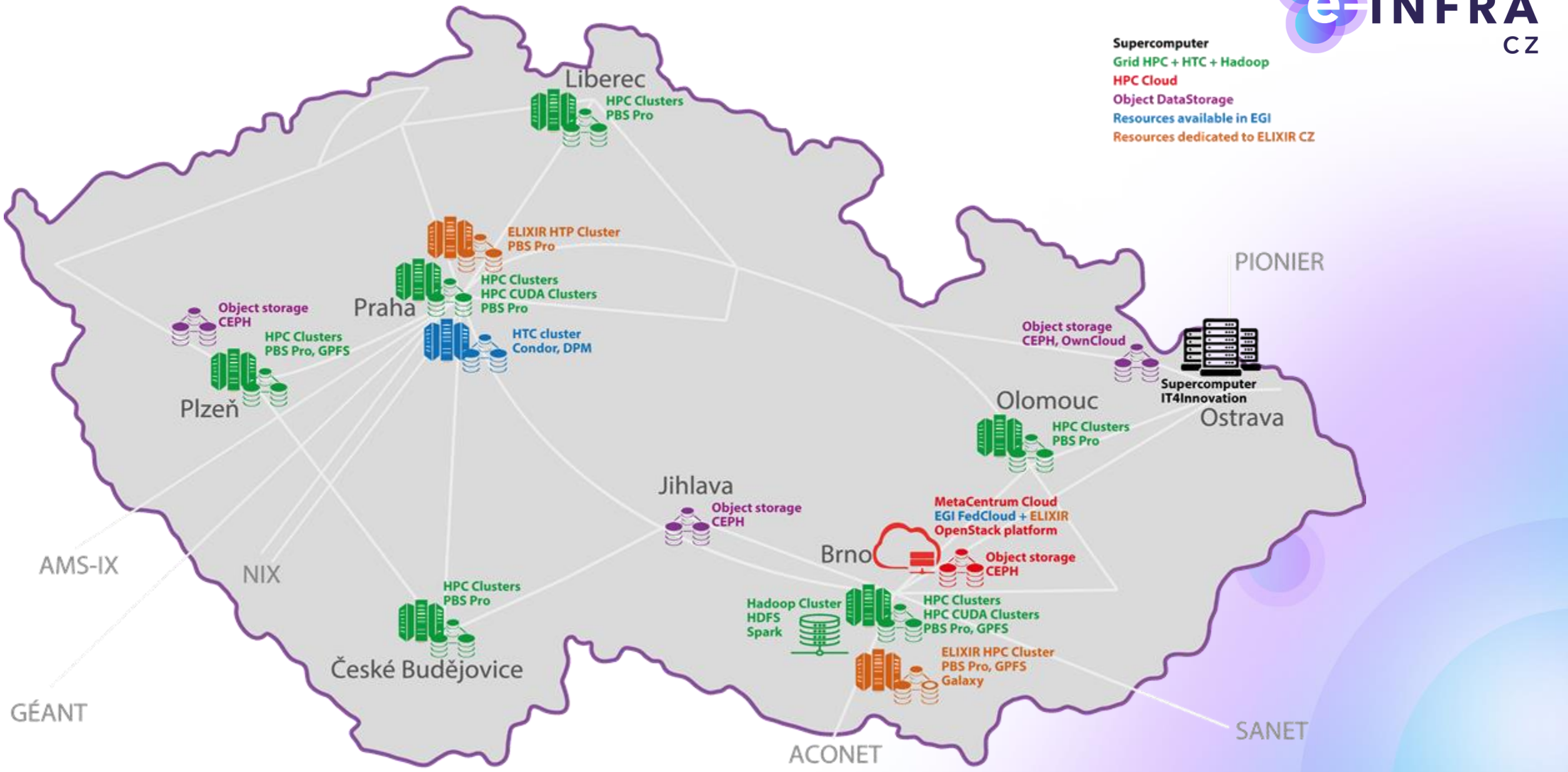
KAROL1NA

Open Access Grant Competitions in 2020

- Granted allocation
- Difference between demand and granted allocation



VSB TECHNICAL UNIVERSITY OF OSTRAVA | IT4INNOVATIONS NATIONAL SUPERCOMPUTING CENTER


e-INFRA CZ
Supercomputer
Grid HPC + HTC + Hadoop
HPC Cloud
Object DataStorage
Resources available in EGI
Resources dedicated to ELIXIR CZ
Liberec
HPC Clusters
PBS Pro
ELIXIR HTP Cluster
PBS Pro
Praha
HPC Clusters
HPC CUDA Clusters
PBS Pro
HTC cluster
Condor, DPM
Object storage
CEPH
HPC Clusters
PBS Pro, GPFS
Plzeň
PIONIER
Object storage
CEPH, OwnCloud
Supercomputer
IT4Innovation
Ostrava
Olomouc
HPC Clusters
PBS Pro
Jihlava
Object storage
CEPH
MetaCentrum Cloud
EGI FedCloud + ELIXIR
OpenStack platform
Brno
Object storage
CEPH
Hadoop Cluster
HDFS
Spark
HPC Clusters
HPC CUDA Clusters
PBS Pro, GPFS
ELIXIR HPC Cluster
PBS Pro, GPFS
Galaxy
HPC Clusters
PBS Pro
České Budějovice
AMS-IX
NIX
GÉANT
ACONET
SANET

IT4I – A MODERN DATA CENTER

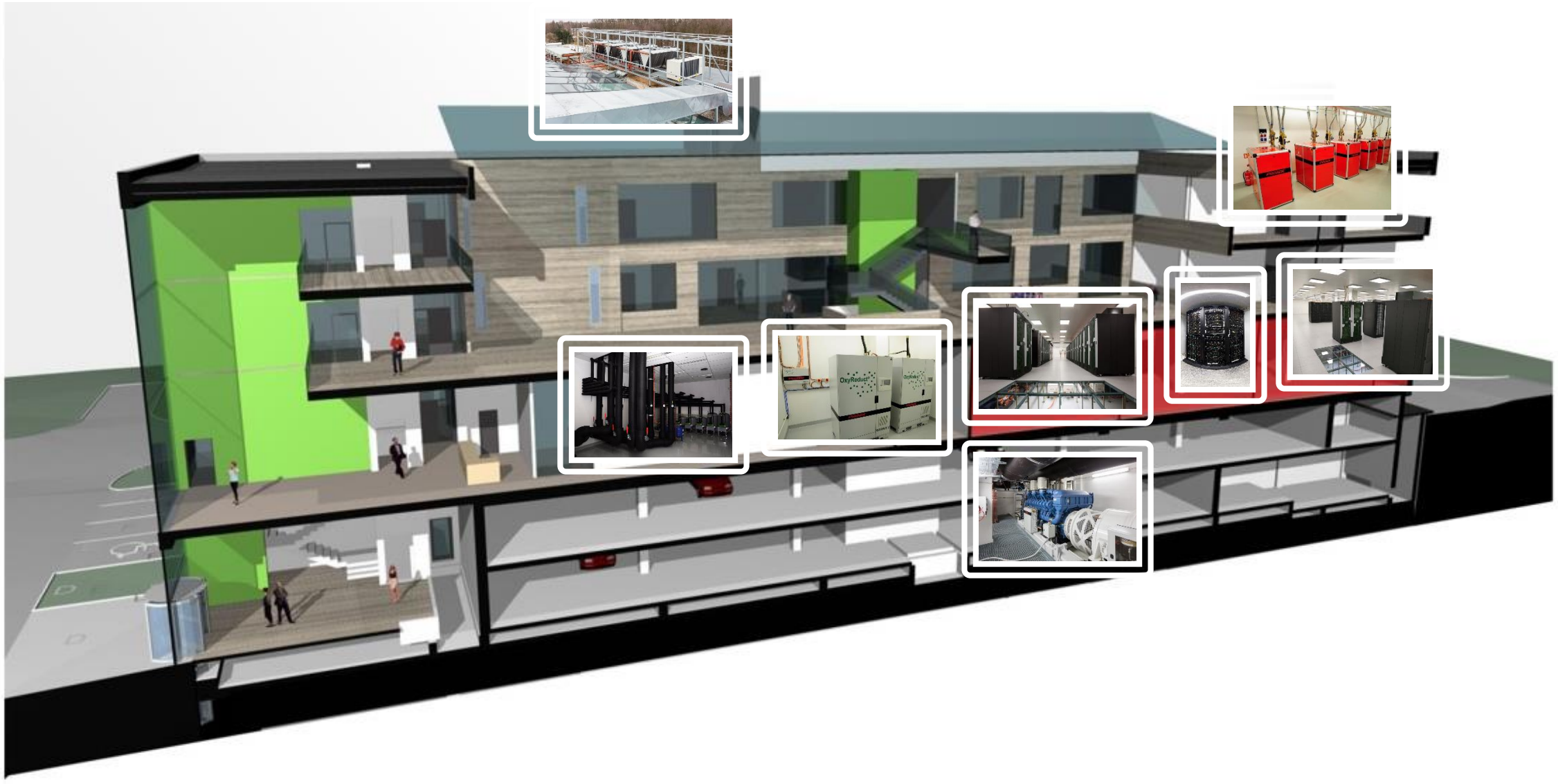
500 sq.m.

OxyReduct fire prevention

Dynamic rotating UPS 2x2,5MVA

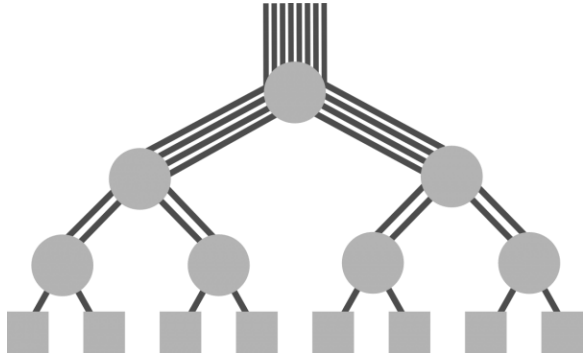
Cold and Hot water cooling

SUPPLEMENTARY INFRASTRUCTURE



KAROLINA SUPERCOMPUTER

- **720x compute nodes, universal partition**
  - 2x AMD EPYC 7H12 (Rome) @2.6GHz, turbo 3.3GHz, 64 jader
  - 256GB RAM
- **72x compute nodes, accelerated partition**
  - 2x AMD EPYC 7763 (Milan) @2.45GHz, turbo 3.5GHz, 64 jader
  - 8x Nvidia A100, 40GB HBM2
  - 1024GB RAM
- 1x fat node, 32x24 cores (Intel Xeon 8268), 24TB RAM
- 36x cloud partition, 2x24 cores (7h12), 256GB RAM
- Network - non-blocking fat tree, 100Gb/s



VSB TECHNICAL UNIVERSITY OF OSTRAVA | IT4INNOVATIONS NATIONAL SUPERCOMPUTING CENTER

KAROLINA SUPERCOMPUTER

- 720x compute nodes, universal partition
  - **3833** TFLOPS Peak performance
- 72x compute nodes, accelerated partition
  - **8645** TFLOPS Peak performance

TOP500 CERTIFICATE

The List.

Karolina, GPU partition - Apollo 6500, AMD EPYC 7452 32C 2.35GHz, NVIDIA A100 SXM4 40 GB, Infiniband HDR200

IT4Innovations National Supercomputing Center, VSB-Technical University of Ostrava, Czechia

is ranked

No. 69

among the World's TOP500 Supercomputers

with 6.05 PFlop/s Linpack Performance

in the 57th TOP500 List published at the ISC Virtual 2021

Conference on June 28, 2021.

Congratulations from the TOP500 Editors

Erich Strohmaier, NERSC/Berkeley Lab
Jack Dongarra, University of Tennessee
Horst Simon, NERSC/Berkeley Lab
Martin Meuer, Prometeus

The GREEN 500 CERTIFICATE

Karolina, GPU partition - Apollo 6500, AMD EPYC 7763 64C 2.45GHz, NVIDIA A100 SXM4 40 GB, Infiniband HDR200

IT4Innovations National Supercomputing Center, VSB-Technical University of Ostrava, Czechia

is ranked

No. 8

among the World's TOP500 Supercomputers

with 27.213 GFlops/watts Performance

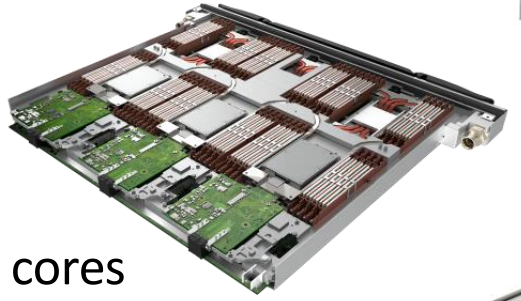
in the Green500 List published at the SC21

Conference on November 16, 2021.

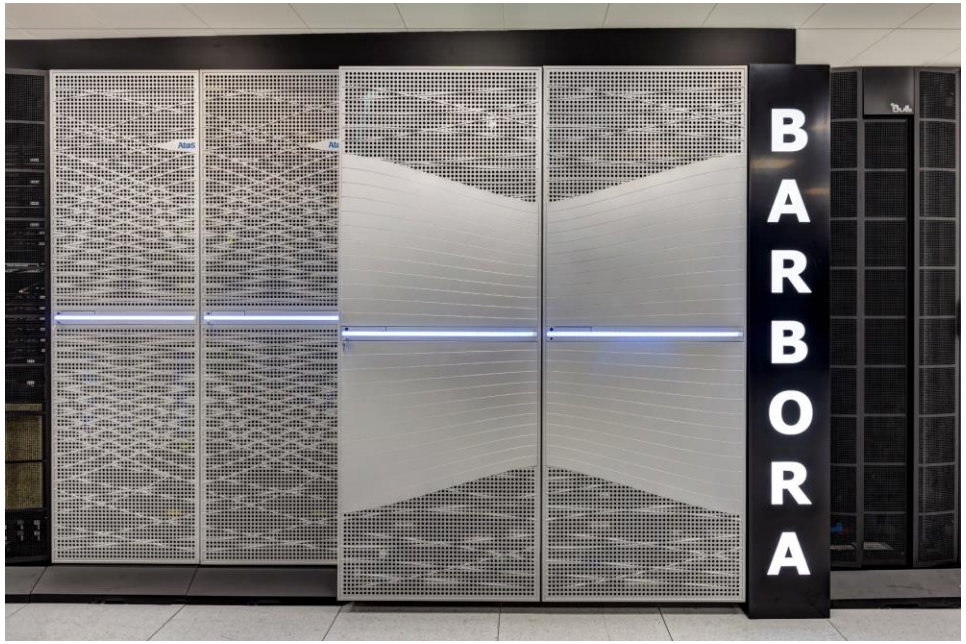
Congratulations from the Green500 Editors

Wu-chun Feng, Virginia Tech
Kirk Cameron, Virginia Tech

BARBORA SUPERCOMPUTER



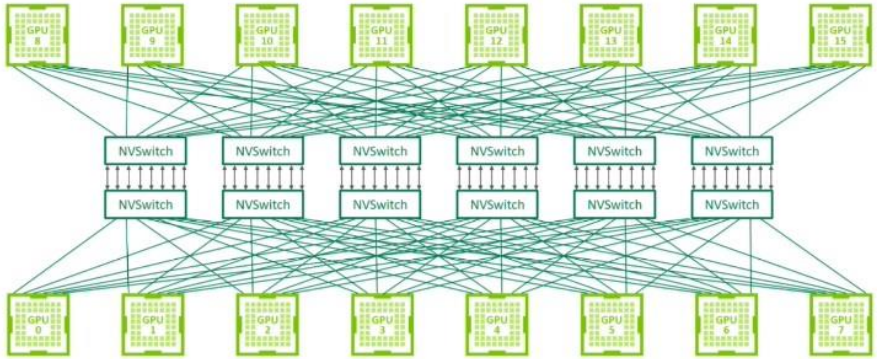
- 189x non-accelerated nodes
  - 2x Intel Xeon Gold 6240 CPU (Cascade Lake) @2.6GHz, 18 cores
- 8x accelerated nodes
  - 2x Intel Skylake Gold 6126 (Skylake) @2.6GHz, 12 cores
  - 4x Nvidia V100-SMX2
- Infiniband HDR, 200Gb/s link
- Fat tree topology
- 840 TFlops peak performance



NVIDIA DGX PLATFORM

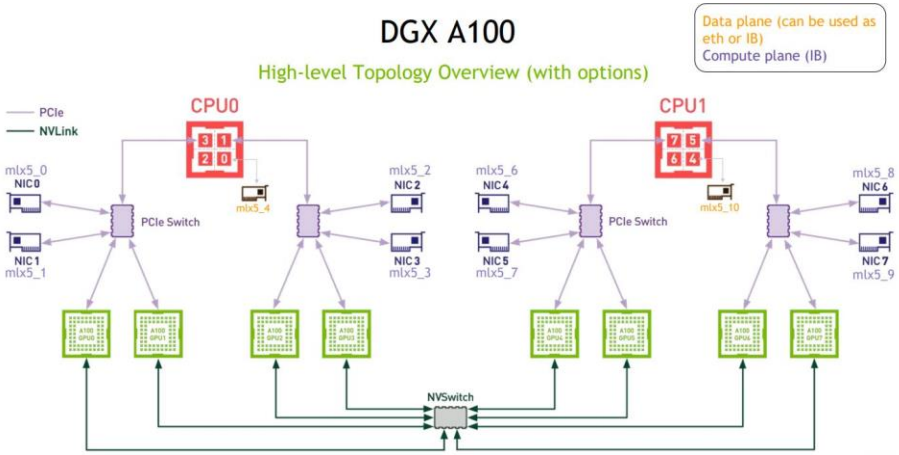
DGX-2

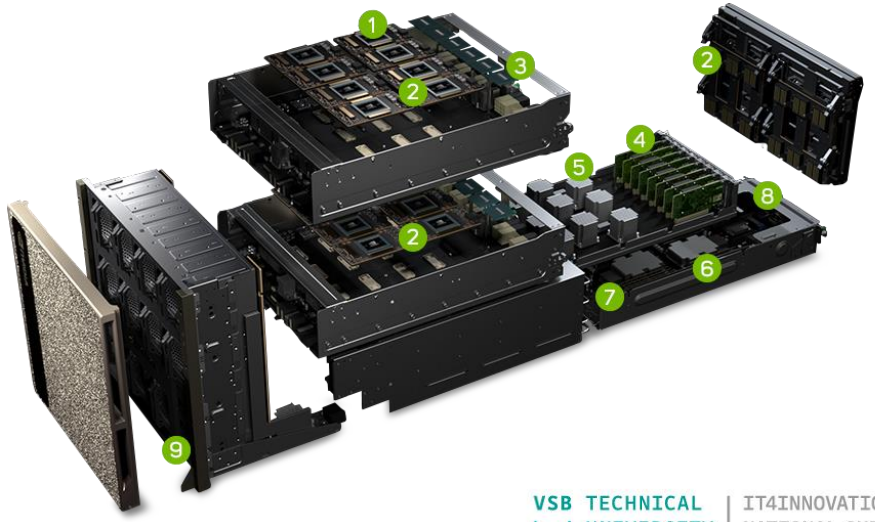
- 16x NVIDIA Tesla V100
- 2x Intel Xeon Platinum
- NVSwitch - 2.4 TB/s of bisection bandwidth



DGX-A100

- Almost the same as one Karolina node
- 8× NVIDIA A100 SXM4
- 2x AMD EPYC 7742





IT4I IN THE TOP500.ORG

Salomon ranking

List	Rank
11/2020	460
06/2020	423
11/2019	375
06/2019	282
11/2018	214
06/2018	139
11/2017	88
06/2017	79
11/2016	68
06/2016	56
11/2015	48
06/2015	40

			CPU cores	Rmax [Flop/s]	Rpeak [Flop/s]	power [kW]
375	IT4Innovations National Supercomputing Center, VSB-Technical University of Ostrava Czech Republic	**Salomon** - SGI ICE X, Xeon E5-2680v3 12C 2.5GHz, Infiniband FDR, Intel Xeon Phi 7120P HPE	76,896	1,457.7	2,011.6	4,806

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	9,080.2	311

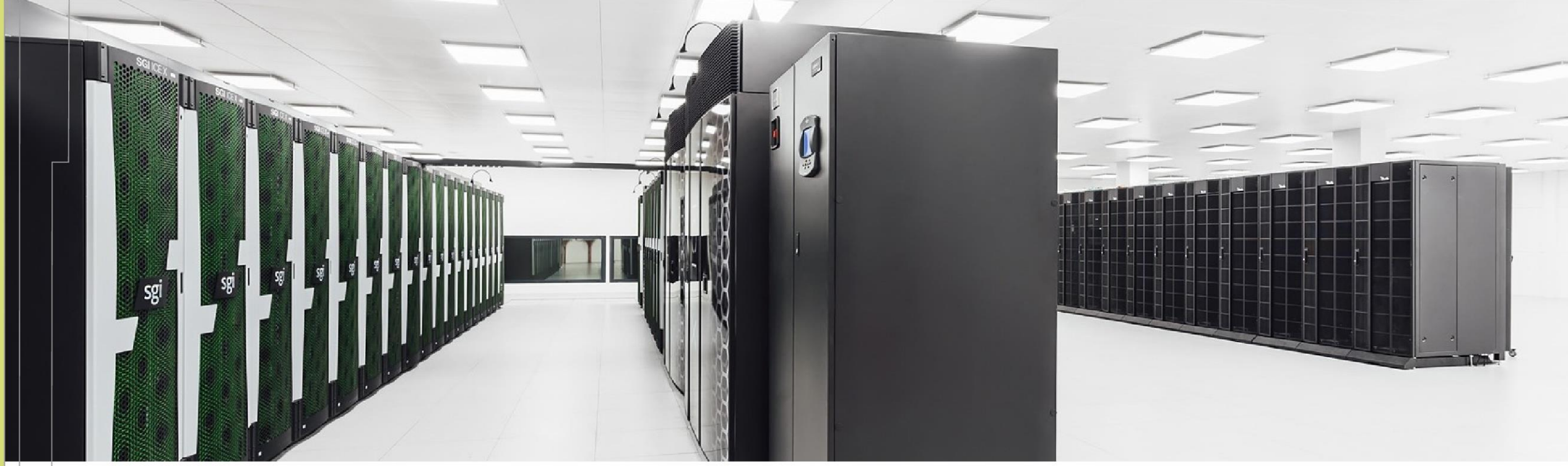
IT4I IN THE TOP500.ORG

Karolina GPU ranking

List	Rank
06/2024	135
11/2023	112
06/2023	95
11/2022	85
06/2022	79
11/2021	71

			CPU cores	Rmax [Flop/s]	Rpeak [Flop/s]	power [kW]
375	IT4Innovations National Supercomputing Center, VSB-Technical University of Ostrava Czech Republic	**Salomon** - SGI ICE X, Xeon E5-2680v3 12C 2.5GHz, Infiniband FDR, Intel Xeon Phi 7120P HPE	76,896	1,457.7	2,011.6	4,806

71	IT4Innovations National Supercomputing Center, VSB-Technical University of Ostrava Czechia	**Karolina, GPU partition** - Apollo 6500, AMD EPYC 7763 64C 2.45GHz, NVIDIA A100 SXM4 40 GB, Infiniband HDR200, HPE	71,424	6,752.0	9,080.2	311



Ondřej Vysocký
Ondrej.vysocky@vsb.cz

IT4Innovations National Supercomputing Center
VSB – Technical University of Ostrava
Studentská 6231/1B
708 00 Ostrava-Poruba, Czech Republic
www.it4i.cz

VSB TECHNICAL UNIVERSITY OF OSTRAVA | IT4INNOVATIONS NATIONAL SUPERCOMPUTING CENTER

EUROPEAN UNION
European Structural and Investment Funds
Operational Programme Research,
Development and Education

MINISTRY OF EDUCATION, YOUTH AND SPORTS