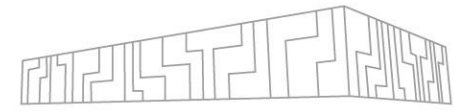




INTRODUCTION TO HPC

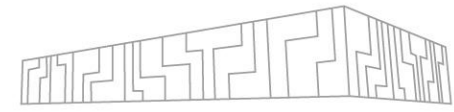
Ondřej Vysocký, Filip Vaverka
IT4Innovations



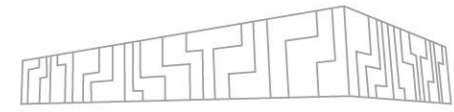


INTRODUCTION

SUPERCOMPUTING



WHAT IS A SUPERCOMPUTER?



Compute nodes



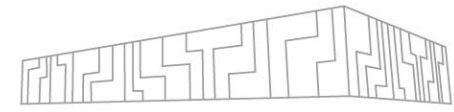
Data storage



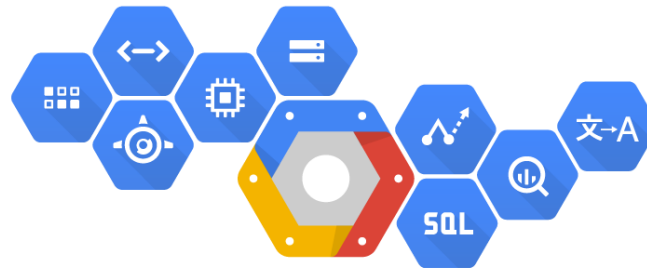
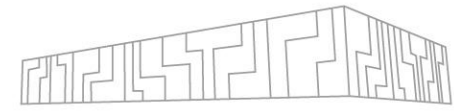
Interconnect



WHAT IS NOT A SUPERCOMPUTER?



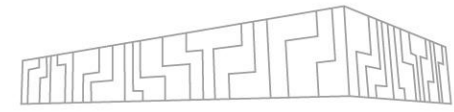
WHAT IS NOT A SUPERCOMPUTER?



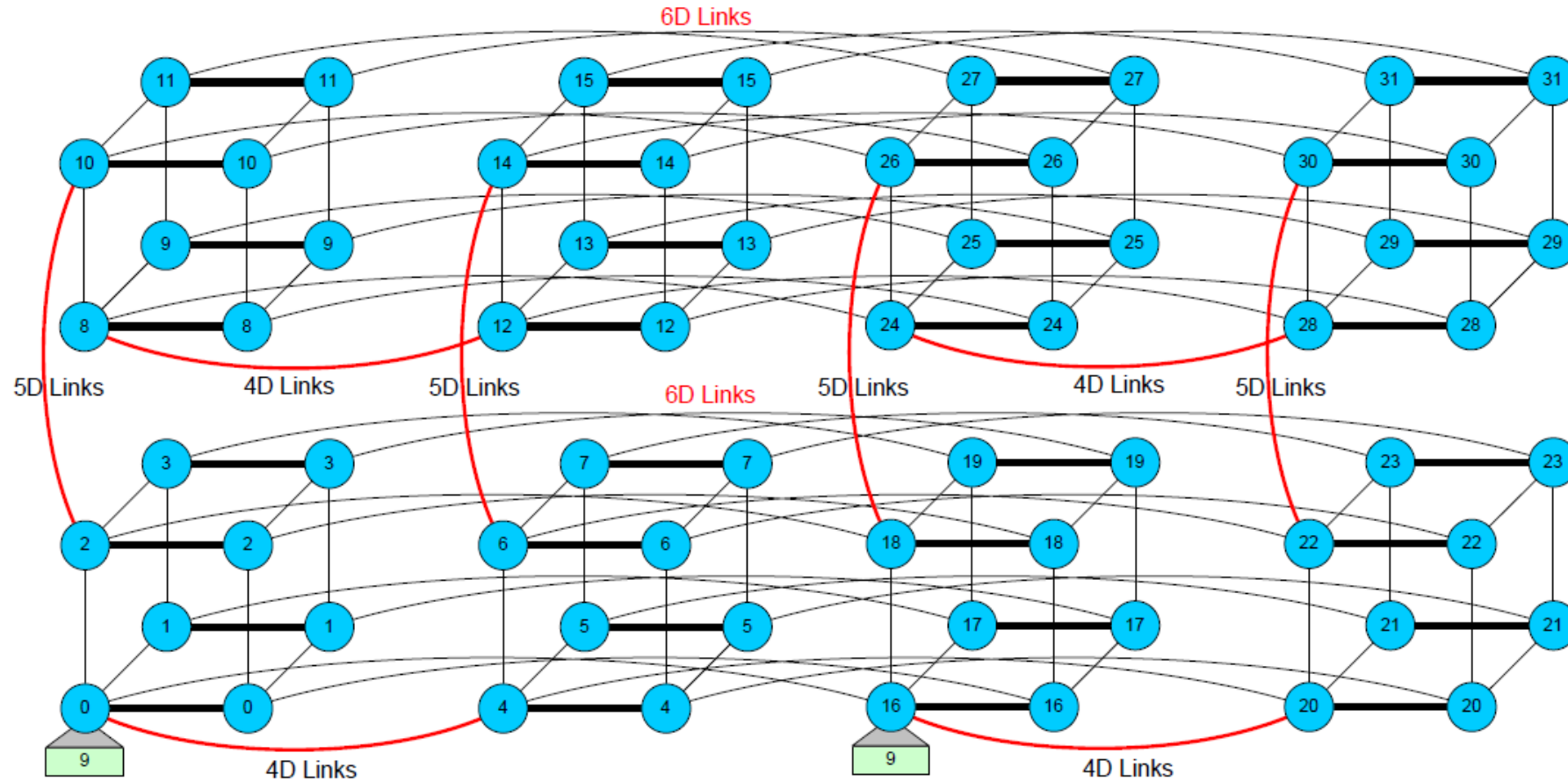
Google Cloud Platform



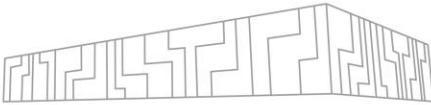
EXAMPLE OF A NETWORK?



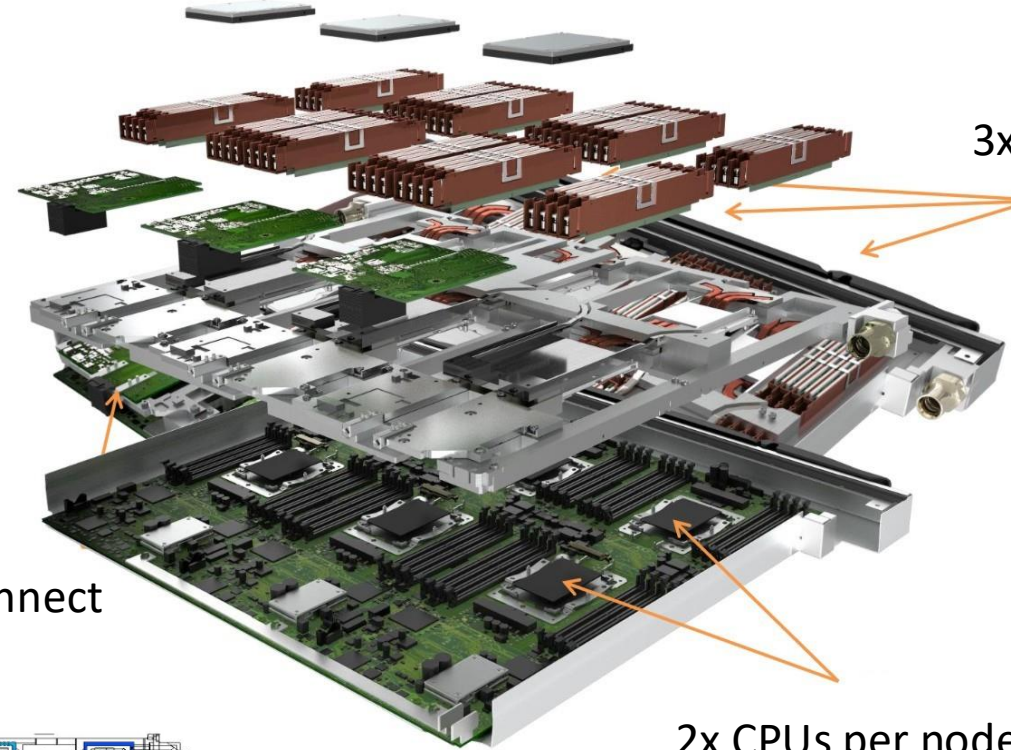
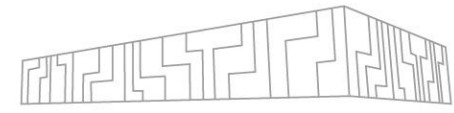
- InfiniBand FDR56 / 7D Enhanced hypercube



DATA CENTER



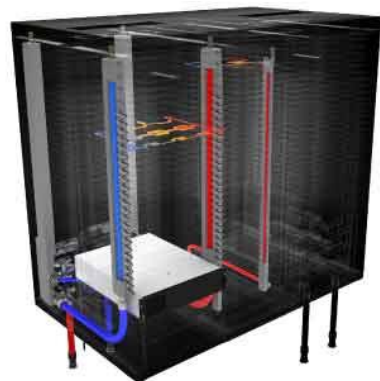
CABINET



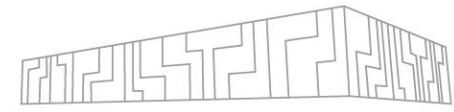
3x compute nodes

interconnect

2x CPUs per node



PEAK PERFORMANCE



- FLOP = Floating point operation
- **Computer performance** = number of floating-point operations per second
FLOPS (Flop/s)

- Intel® Xeon® Platinum 8280M Processor

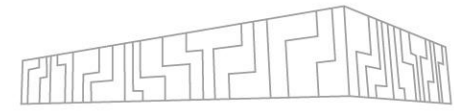
▪ number of compute nodes	1000	1000
▪ number of CPUs	2	2
▪ frequency	2.7 GHz	2.7
▪ number of cores	28	28
▪ have FMA instruction	yes	2
▪ have 2 FMA units	yes	2
▪ SIMD width	512 bit = 8 double precision	8

4 838 000 Gflop/s

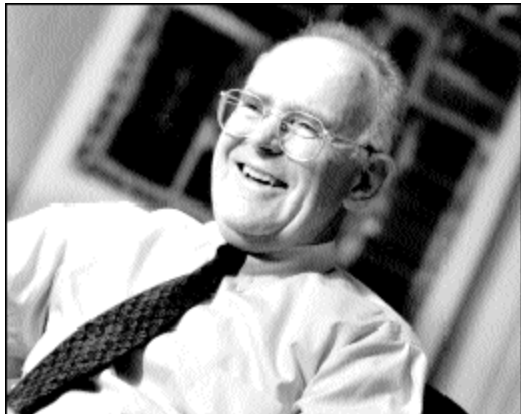
4 838 Tflop/s

4.8 Pflop/s

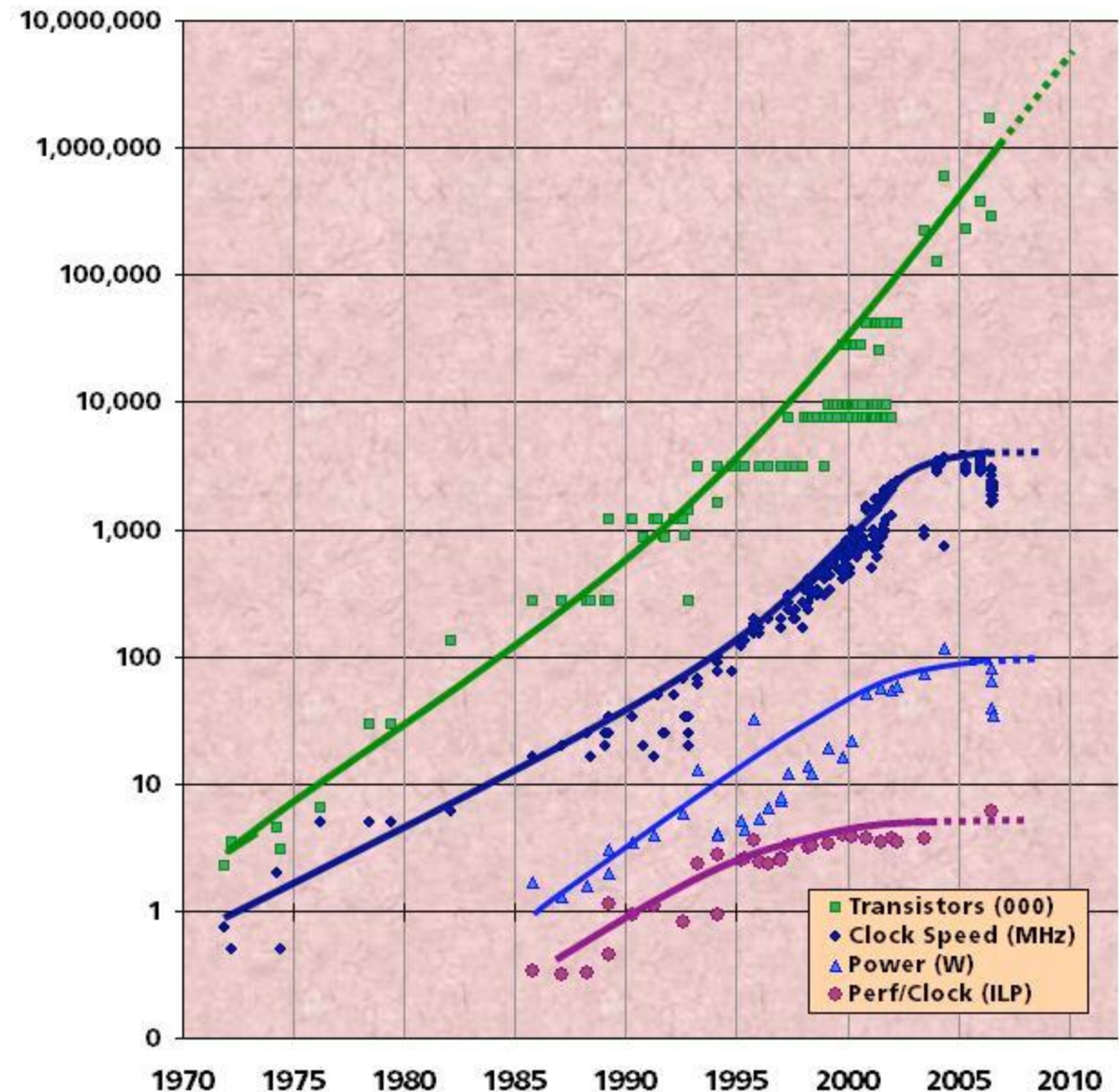
MOORE'S LAW



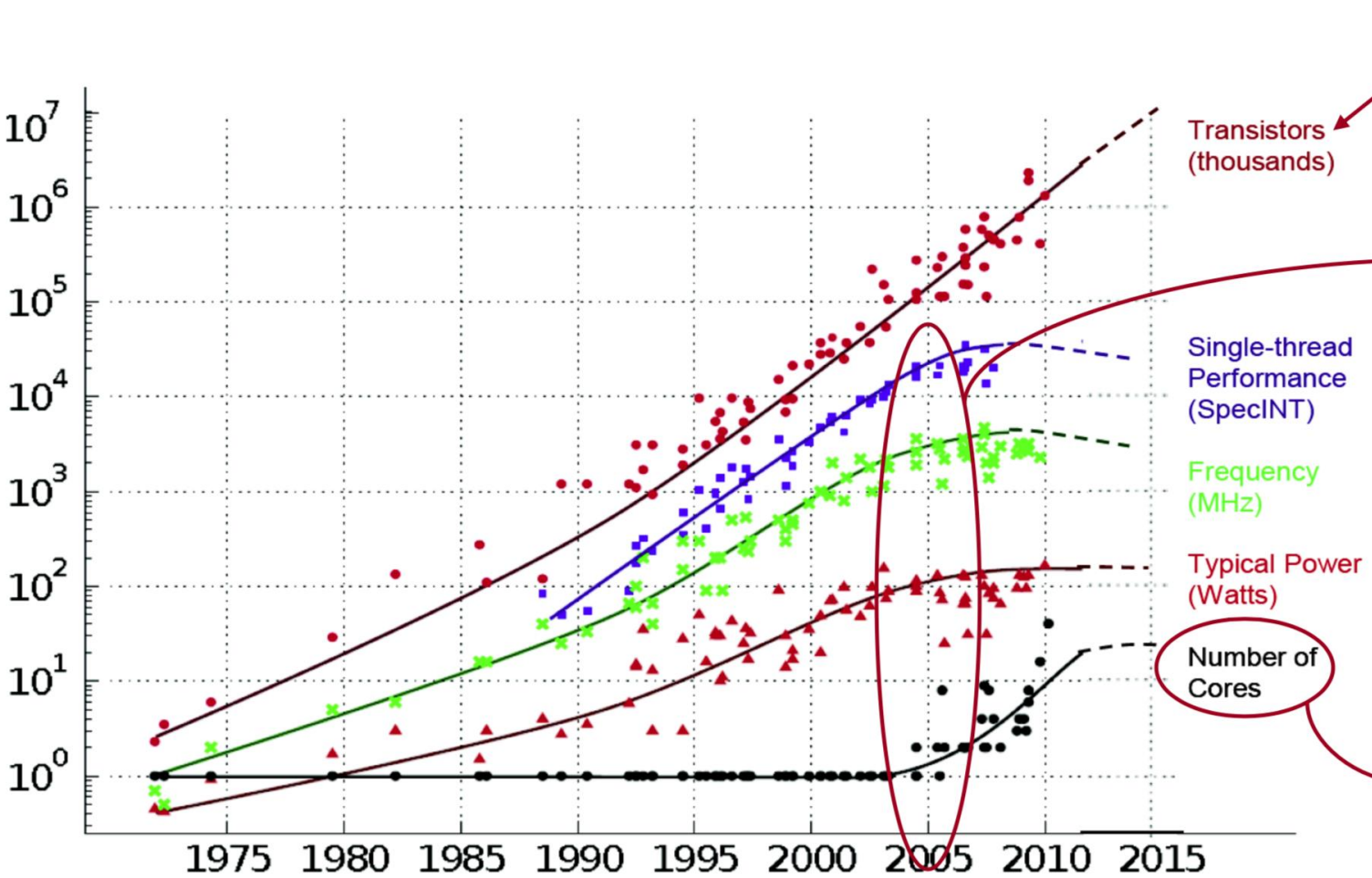
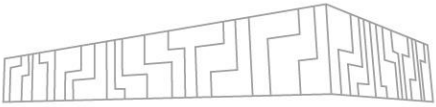
- Chip density is continuing increase $\sim 2x$ every 2 years
- Clock speed is not
- Number of processor cores has to double instead
- Parallelism must be exposed to and managed by software



Slide source: Jack Dongarra



MOORE'S LAW



Transistor count doubles every 18 months, Moore's Law

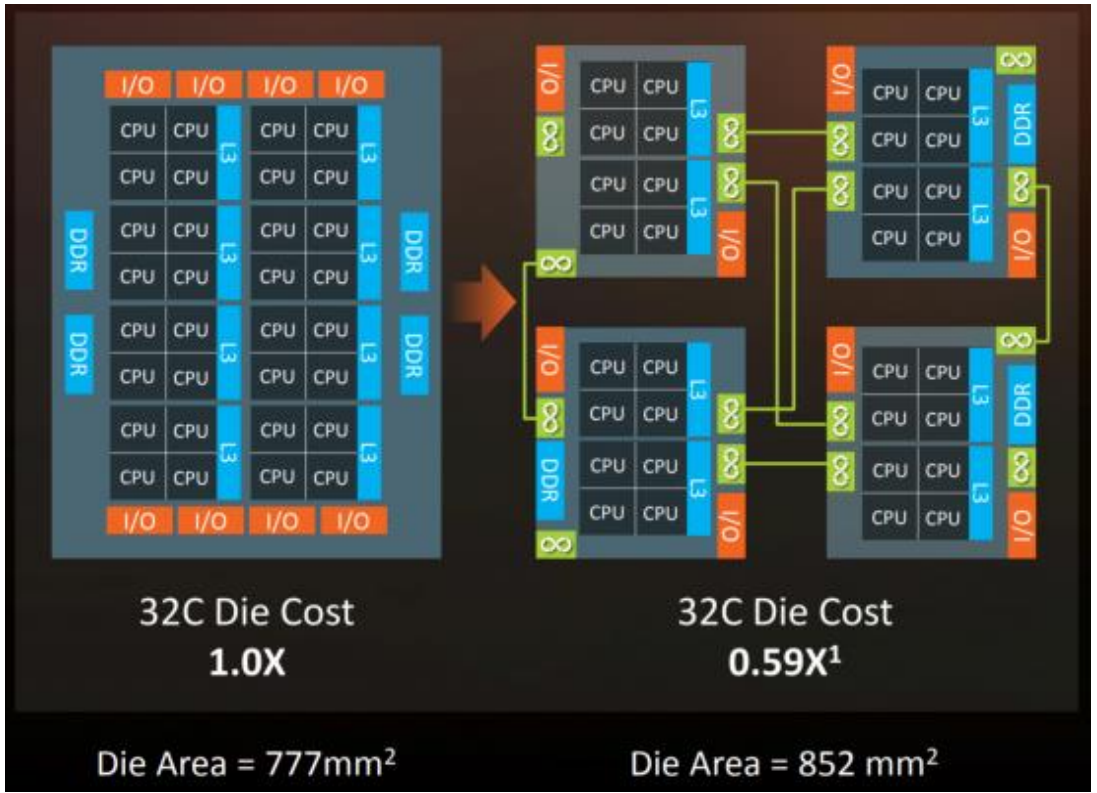
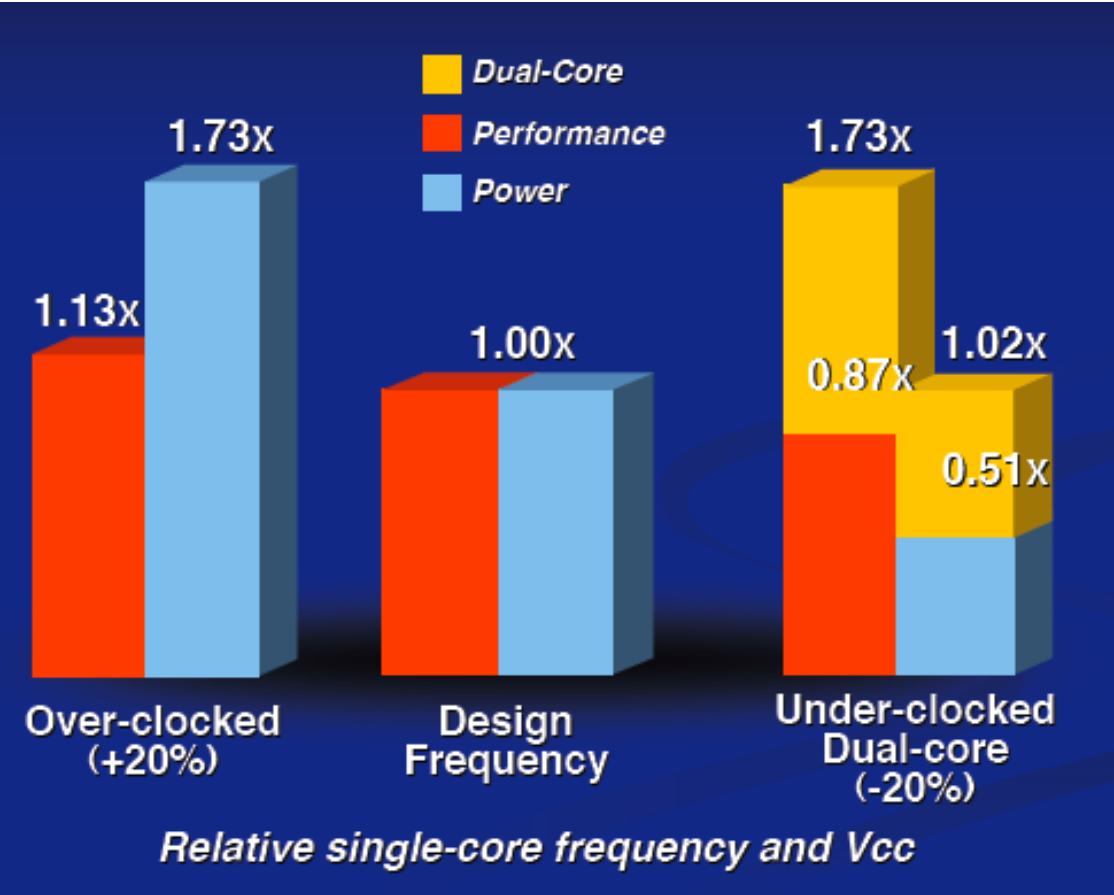
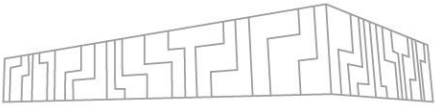
The Power Wall

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

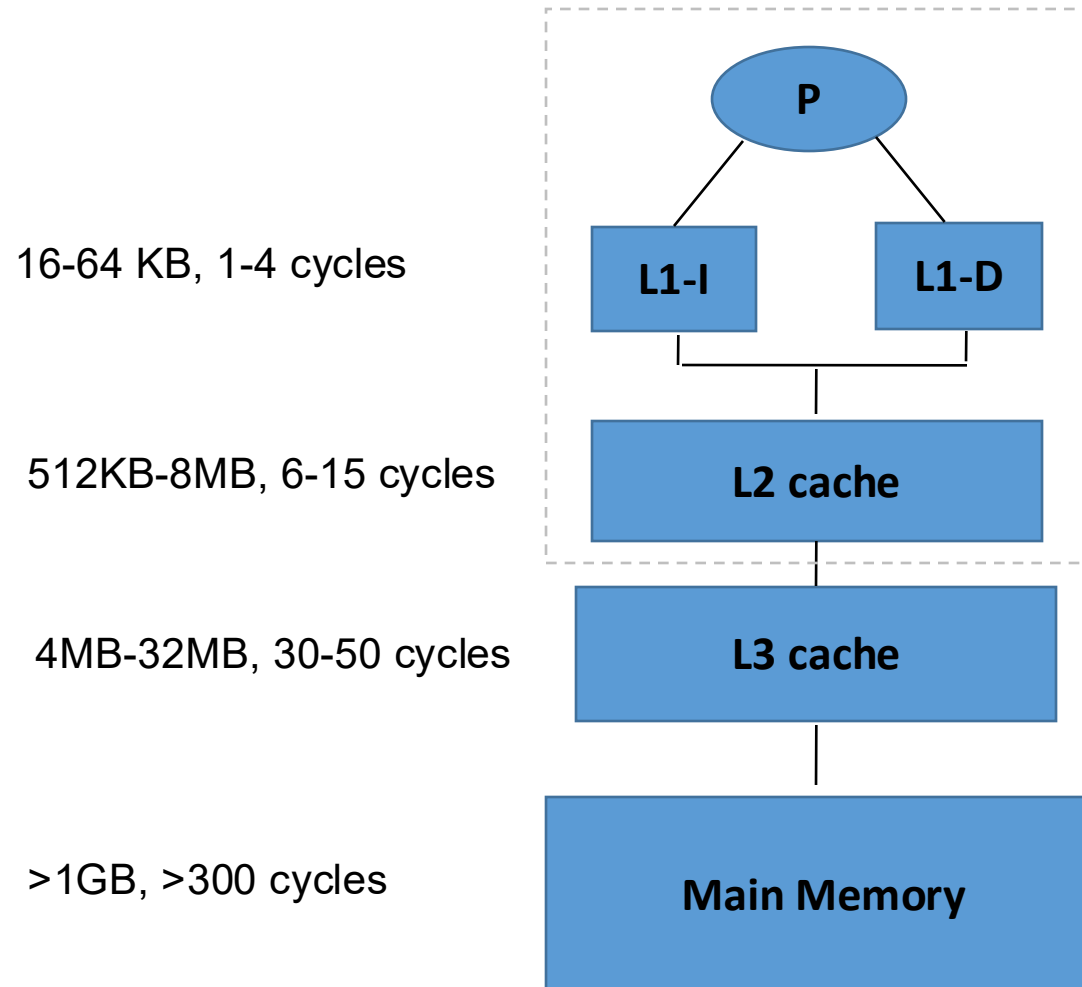
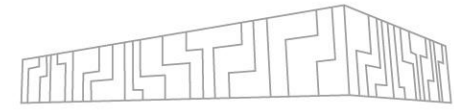
Performance can only grow through node-level parallelism!

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

MODERN CPU DESIGN

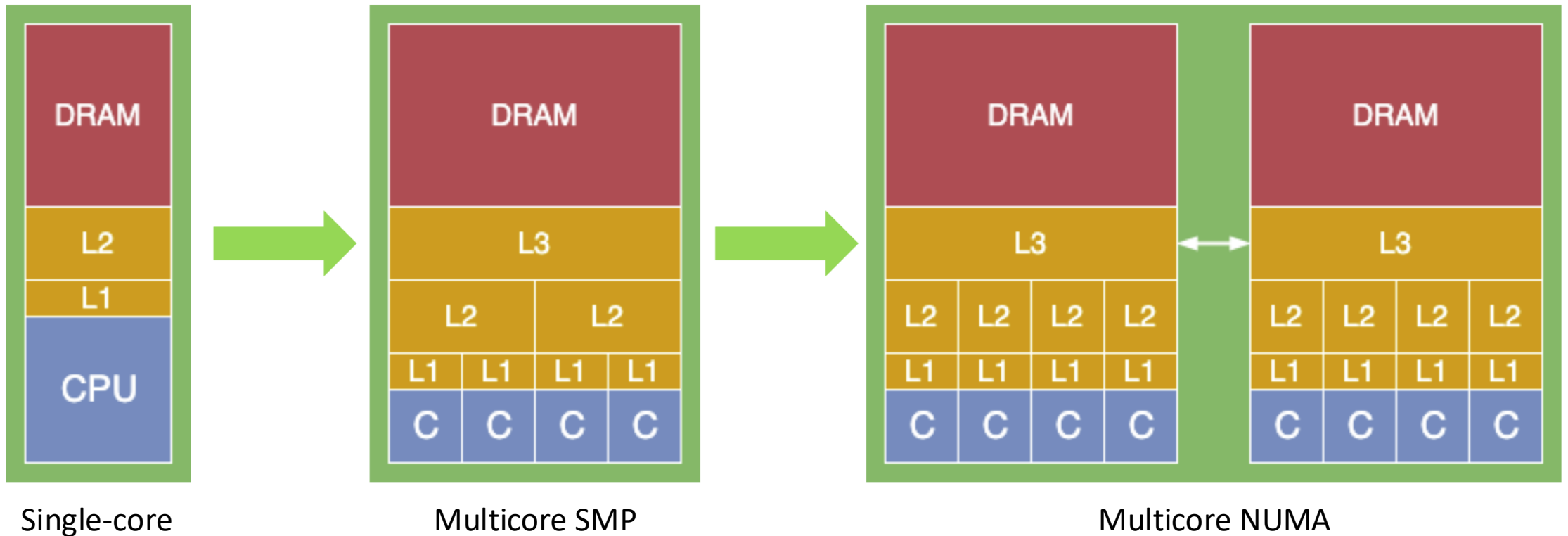
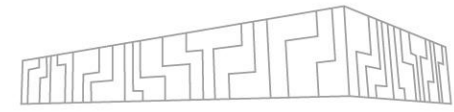


TYPICAL MEMORY HIERARCHY



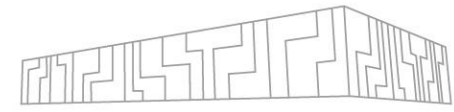
- Access time to main memory is 100's of clock cycles
- Use a small but fast storage near processor
- Works due to locality

HPC BUILDING BLOCKS: CPU

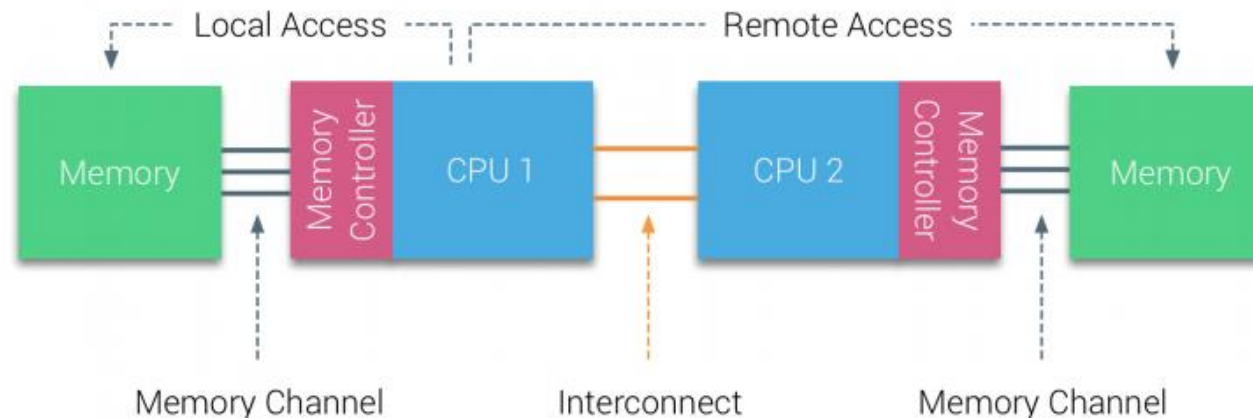


SMP: Symmetric Multi-processor
NUMA: Non-Uniform Memory Access

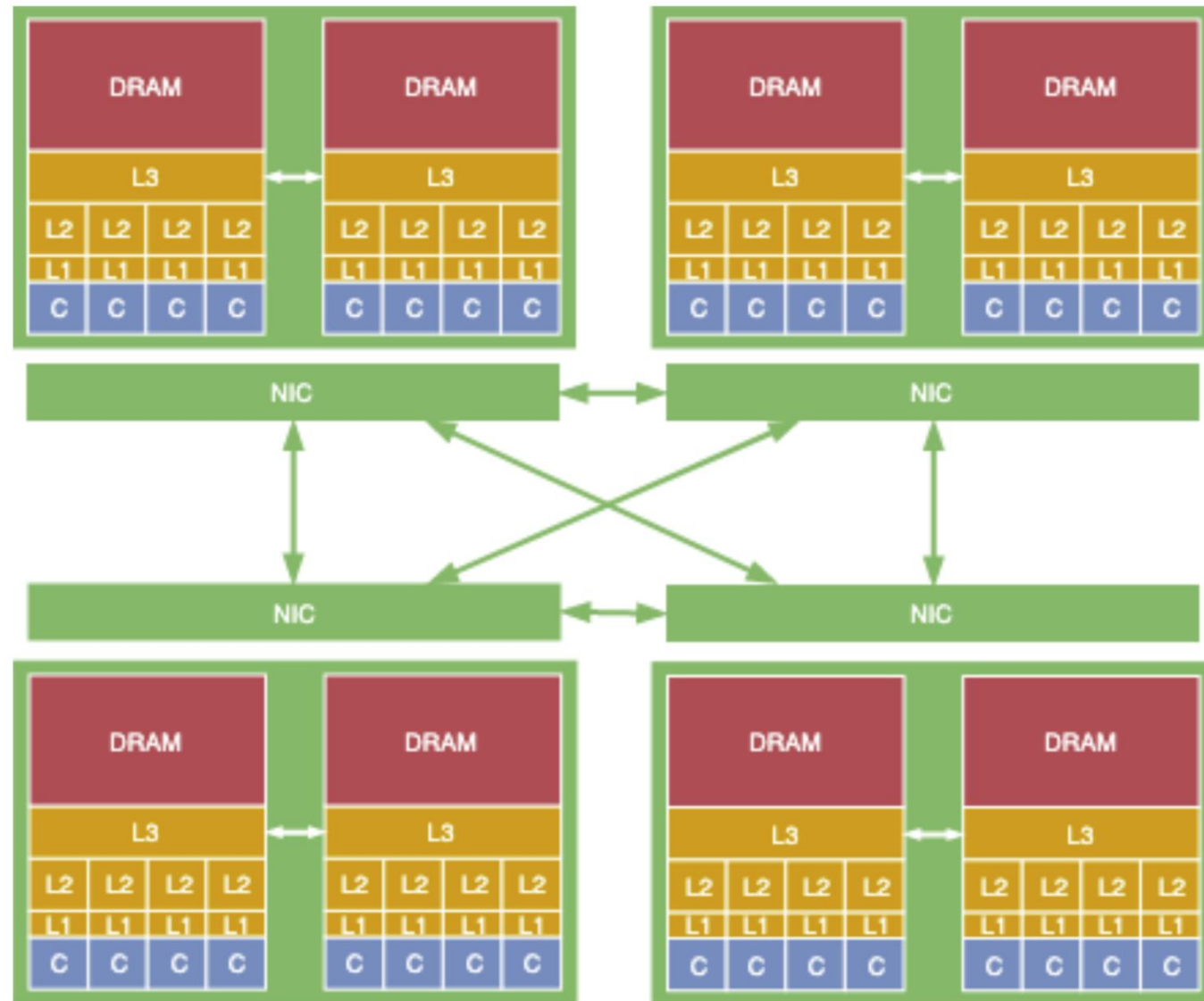
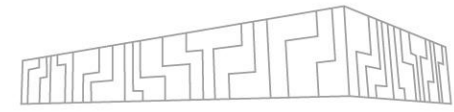
NUMA & CC-NUMA



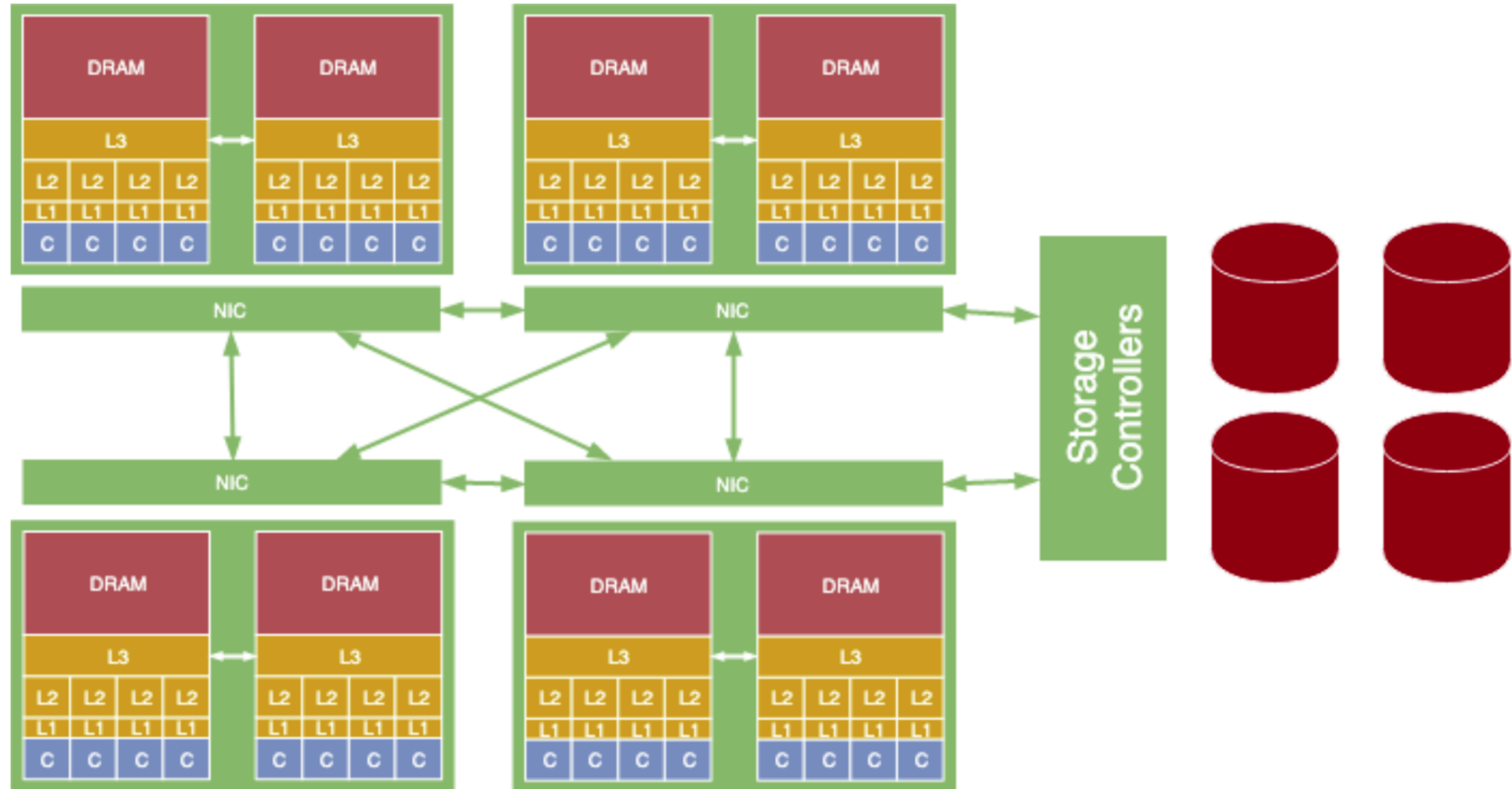
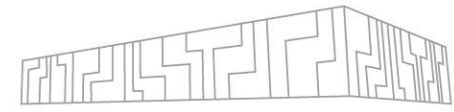
- **NUMA** – Non-Uniform Memory Access
- Aims at surpassing the scalability limits of the UMA architecture due to **memory bandwidth bottleneck**
- Memory physically shared, but access to different portions of the memory may require **significantly different times**
 - local memory access is the fastest, 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



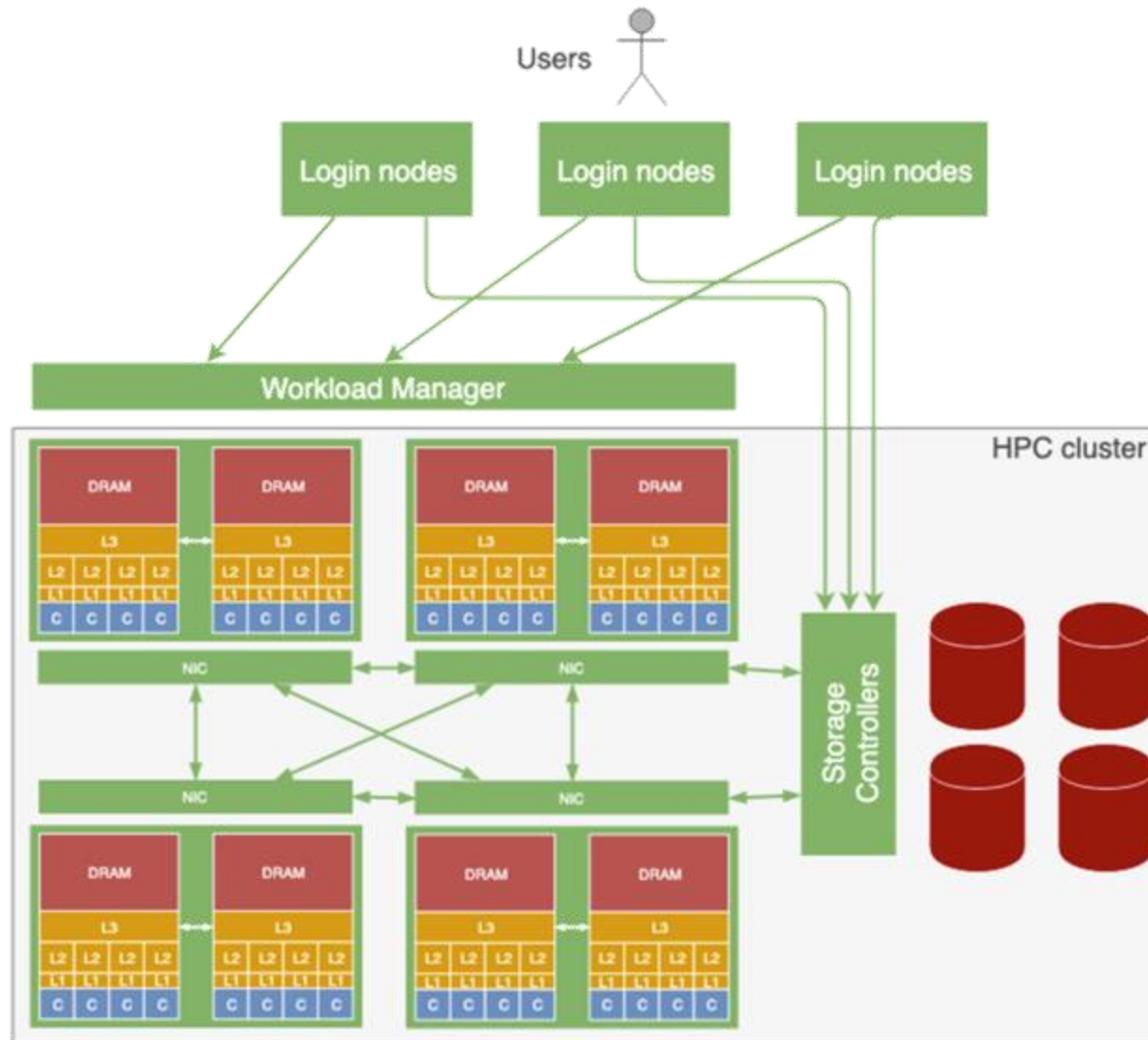
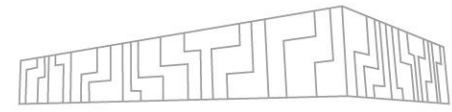
HPC BUILDING BLOCKS: NETWORK



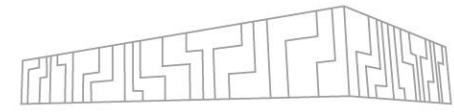
HPC BUILDING BLOCKS: STORAGE



HPC BUILDING BLOCKS: LOGIN+SCHEDULER



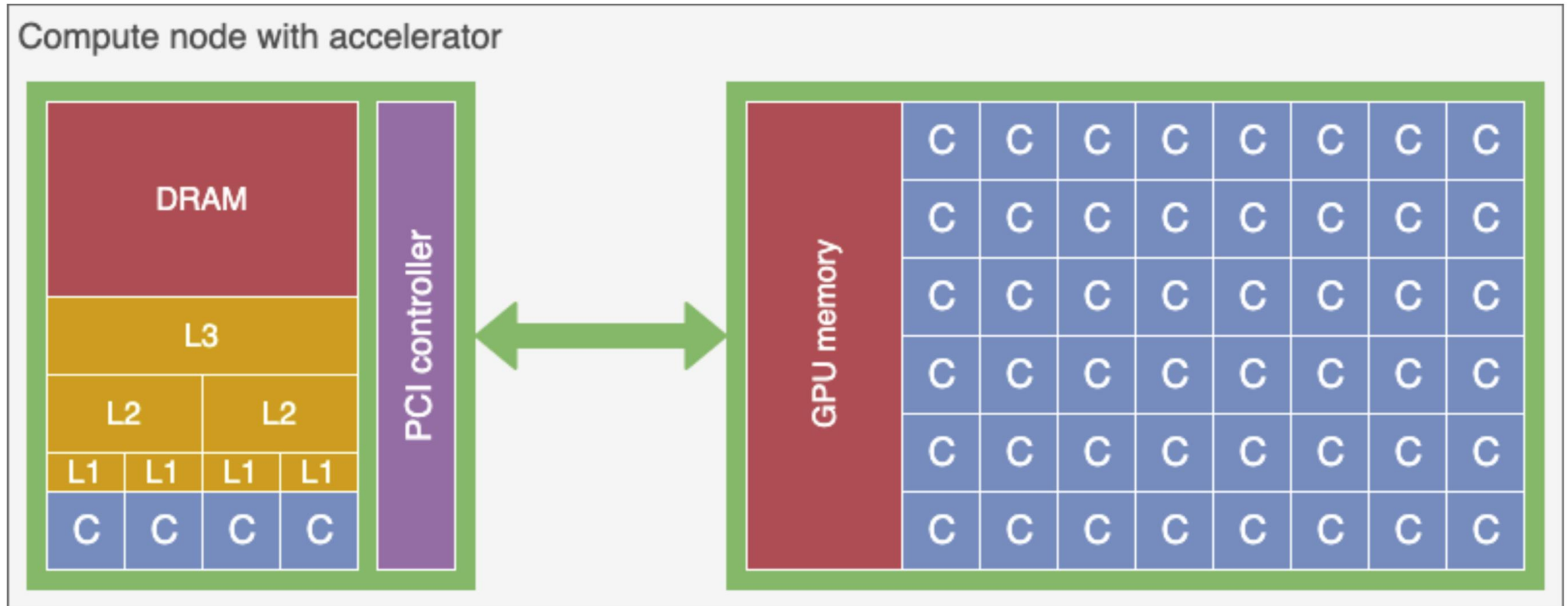
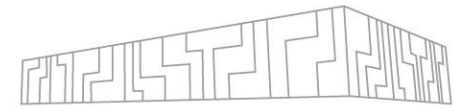
BEYOND MULTICORE



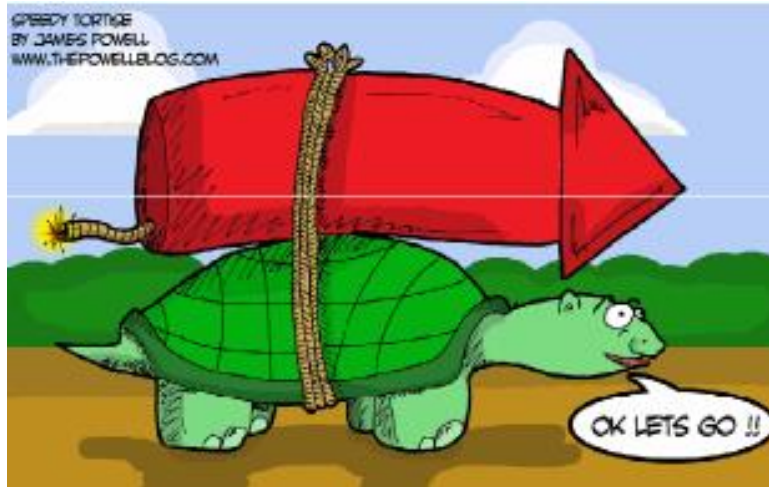
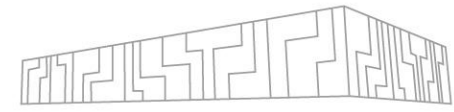
- Multicores have **limitations**
 - Fat cores (branch prediction, out-of-order execution, large caches)
 - Optimized for latency and multiprocessing
 - Still high frequencies
 - Still high-power consumption
 - Strong memory consistency guarantees
 - But programming is easy; matches better our brain's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

HPC BUILDING BLOCKS: ACCELERATOR



HETEROGENOUS COMPUTING



FPGA



Cell



GPU



QC

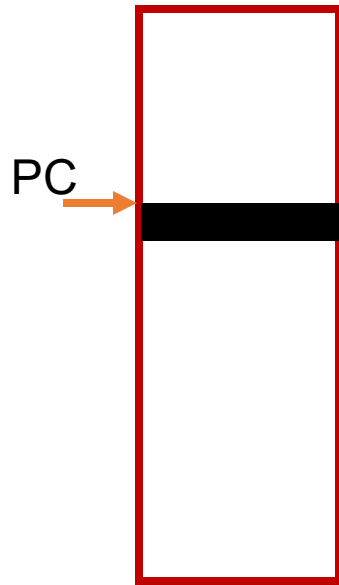
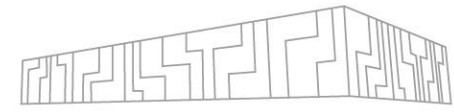


Microprocessor

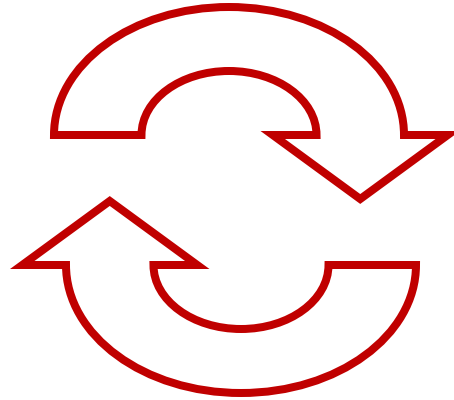
Hardware Accelerators - Speeding up the Slow Part of the Code

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

ACCELERATED EXECUTION MODEL

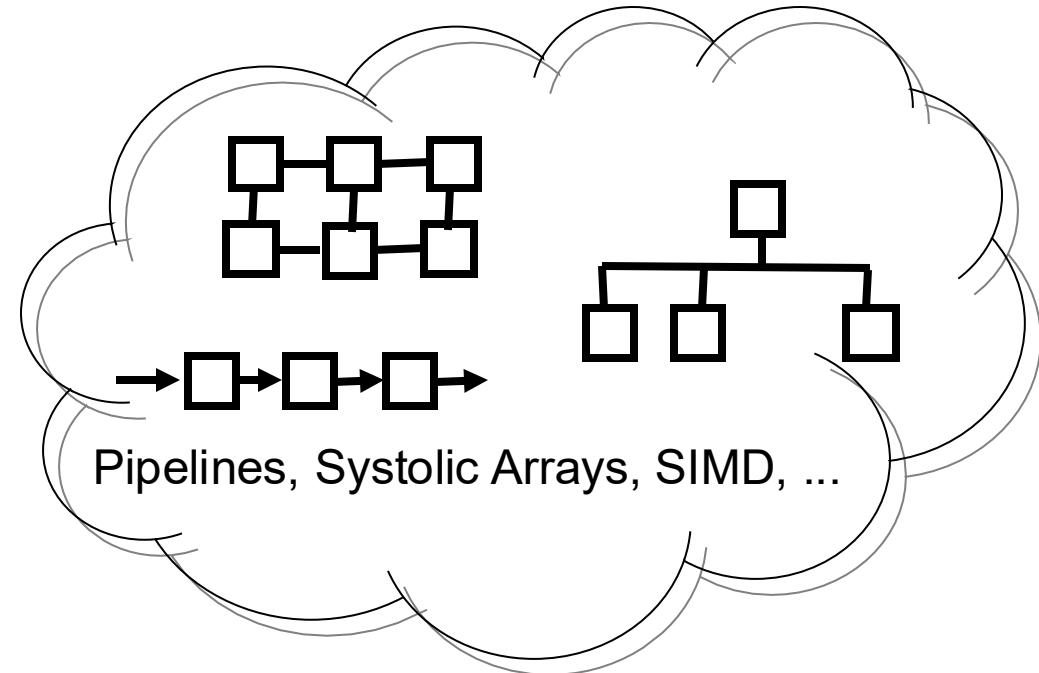


- Transfer of Control
- Input Data



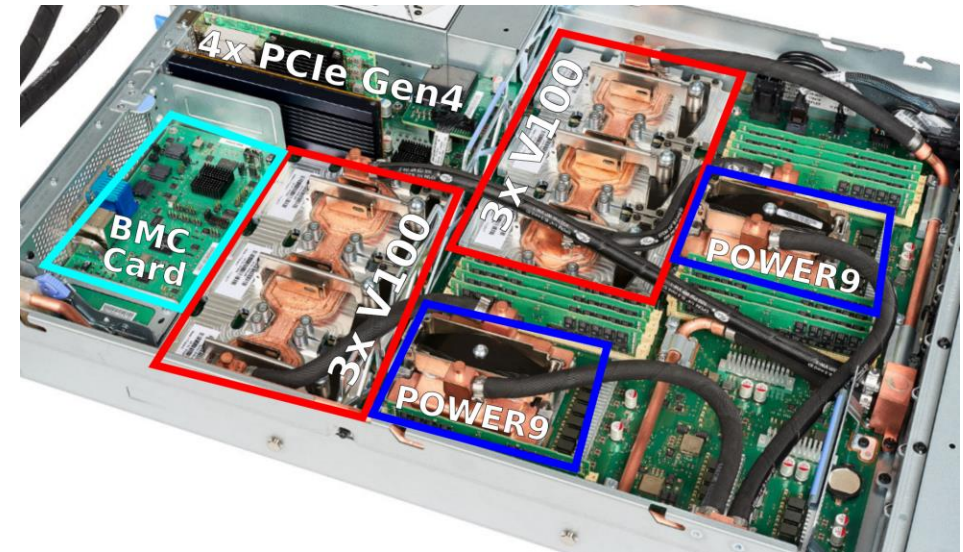
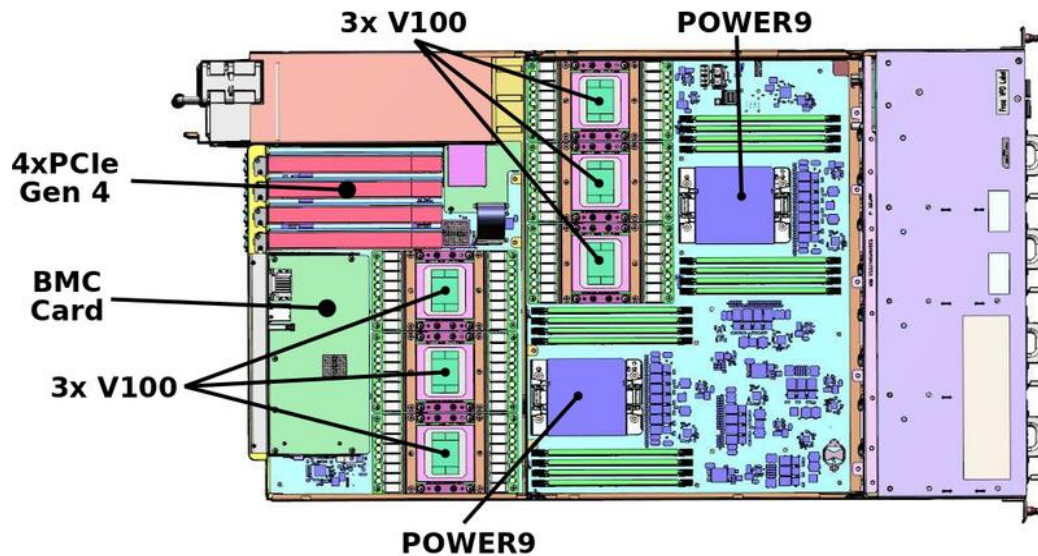
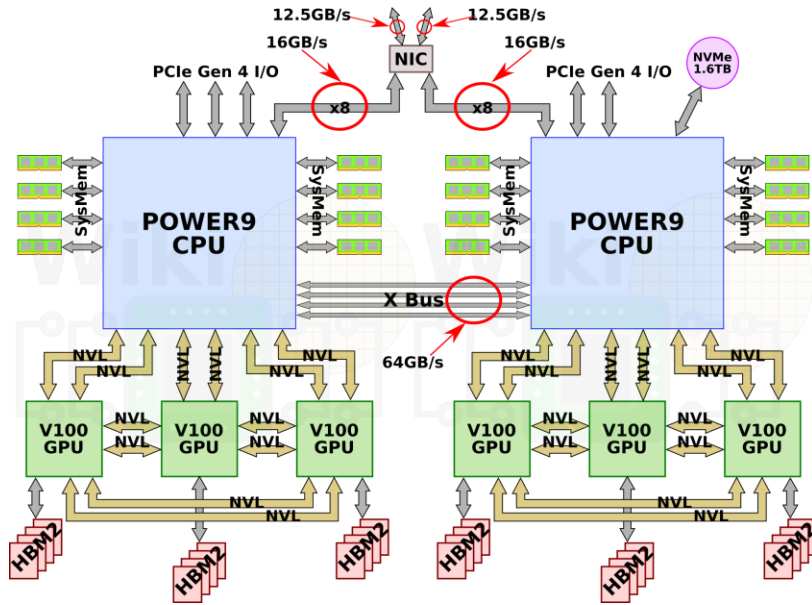
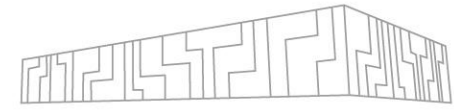
- Output Data
- Transfer of Control

FPGA, GPU, Cell CBE, ...

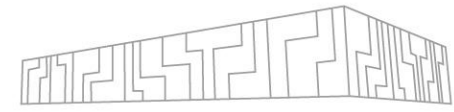


- Fine grained computations with the accelerators, others with the MP
- Interaction between accelerator and MP can be blocking or asynchronous
- This scenario is replicated across the whole system and standard HPC parallel programming paradigms used for interactions

SUMMIT SUPERCOMPUTER (2018)



TENSOR CORES – ACCELERATOR?



CUDA TENSOR CORE PROGRAMMING

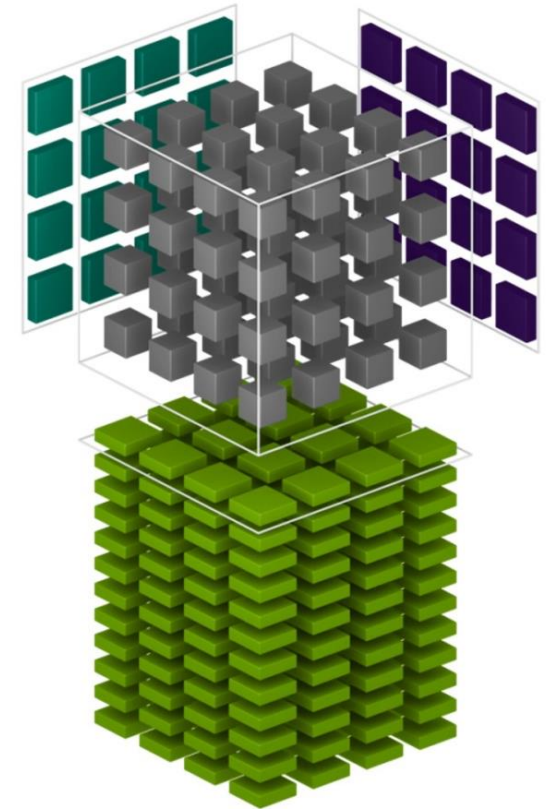
16x16x16 Warp Matrix Multiply and Accumulate (WMMA)

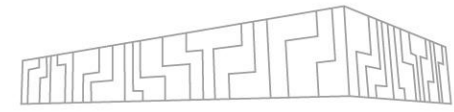
```
wmma::mma_sync(Dmat, Amat, Bmat, Cmat);
```

$$D = \begin{pmatrix} \text{FP16 or FP32} & \text{FP16} & \text{FP16} & \text{FP16 or FP32} \end{pmatrix} + \begin{pmatrix} \text{FP16 or FP32} \end{pmatrix}$$

FP16 or FP32 FP16 FP16 FP16 or FP32

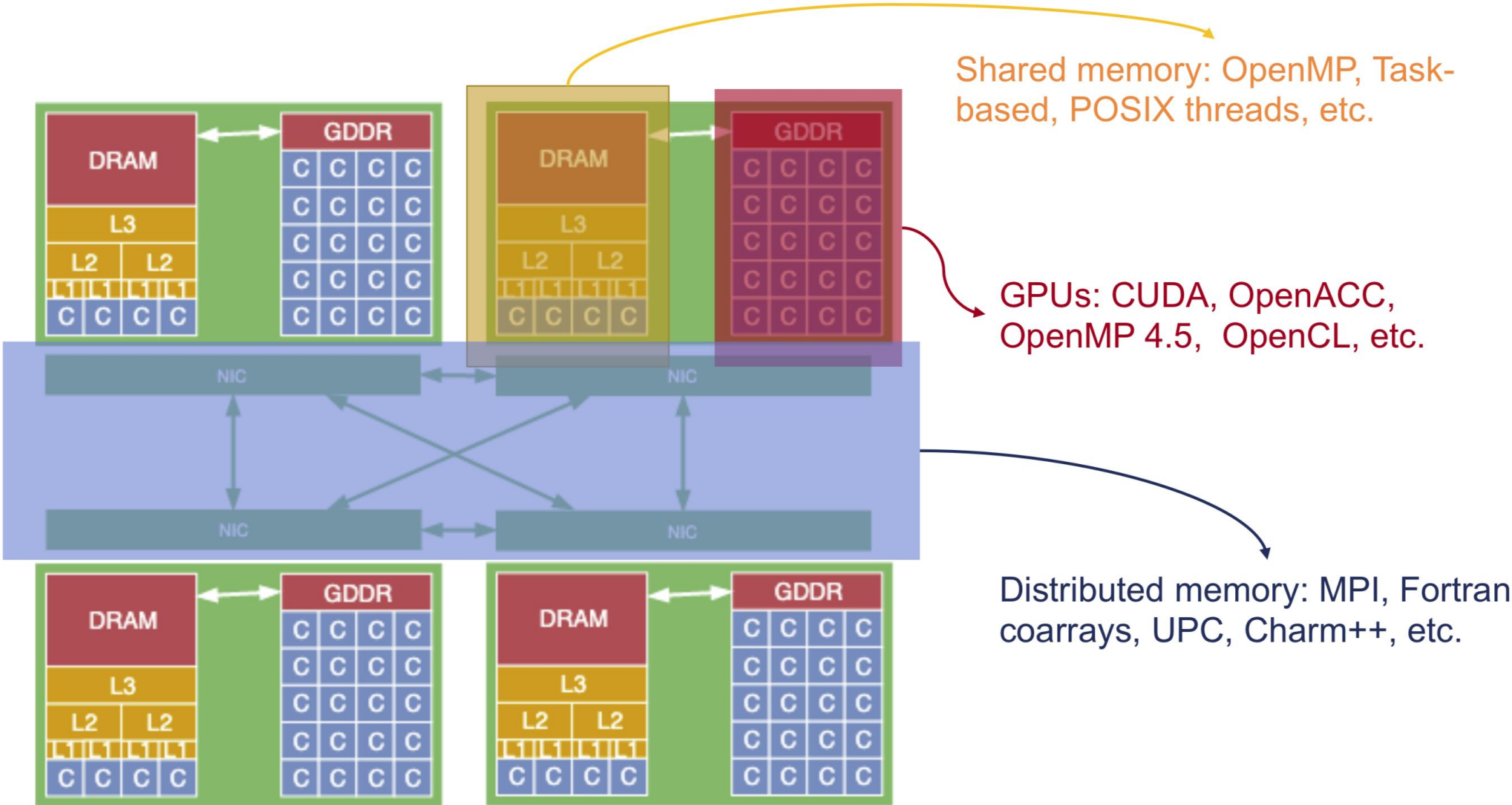
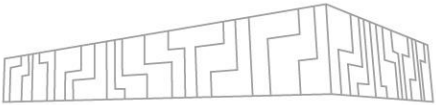
$$D = AB + C$$





SOFTWARE

HOW TO WRITE HPC CODE?

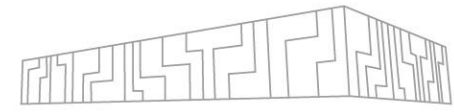


Shared memory: OpenMP, Task-based, POSIX threads, etc.

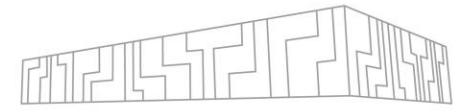
GPUs: CUDA, OpenACC, OpenMP 4.5, OpenCL, etc.

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

PARALLEL COMPUTING



PARALLEL ALGORITHM SCALABILITY

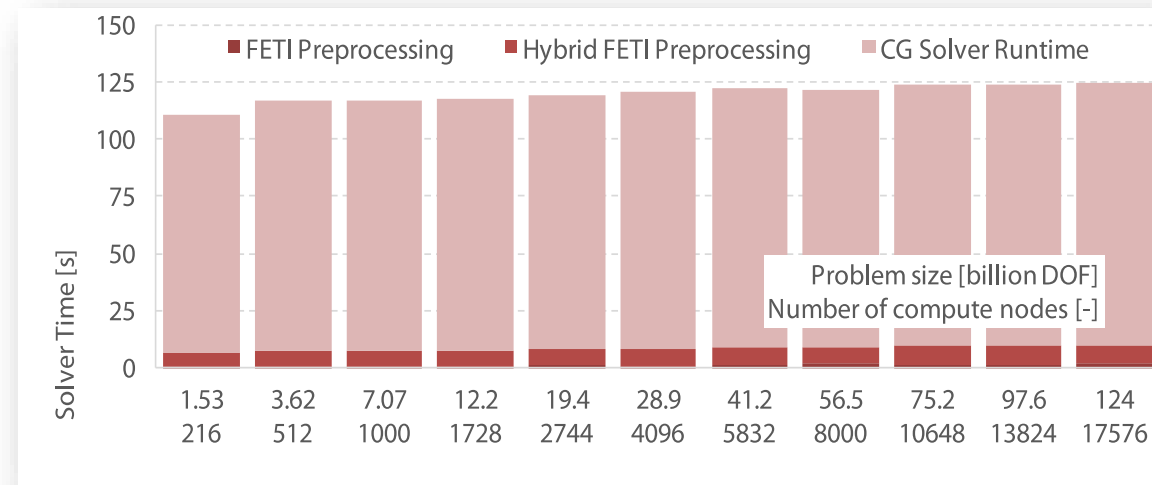
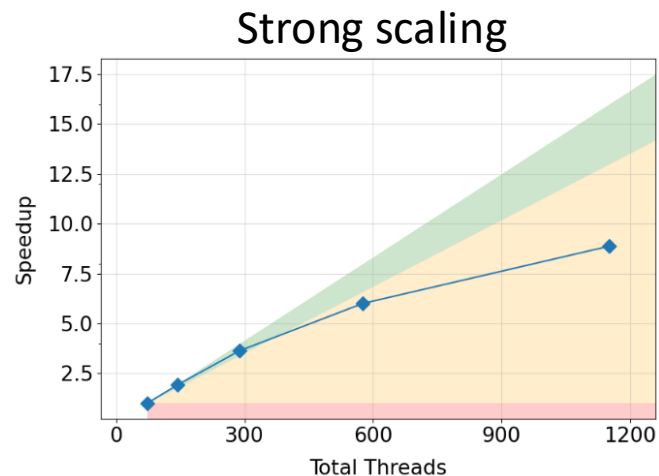


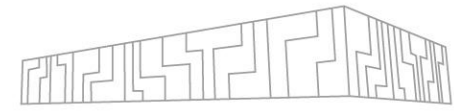
| Strong scaling

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

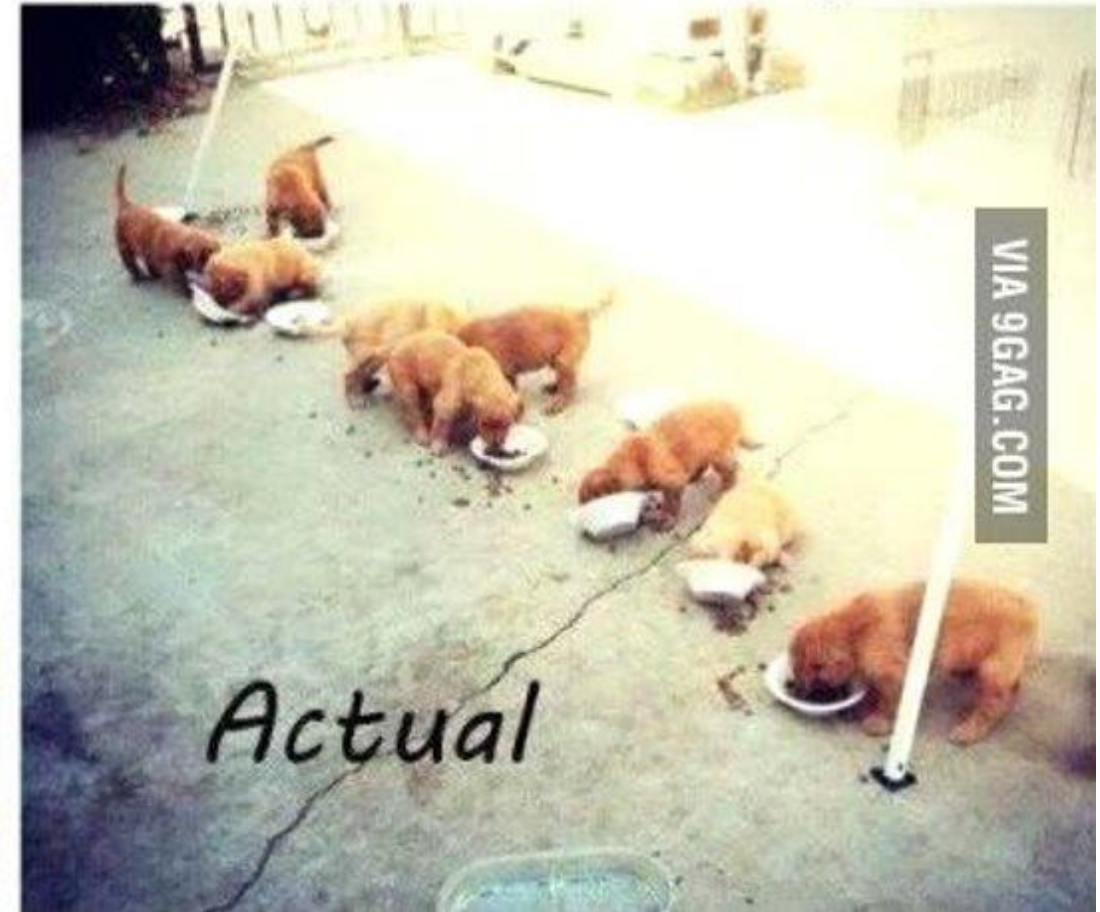
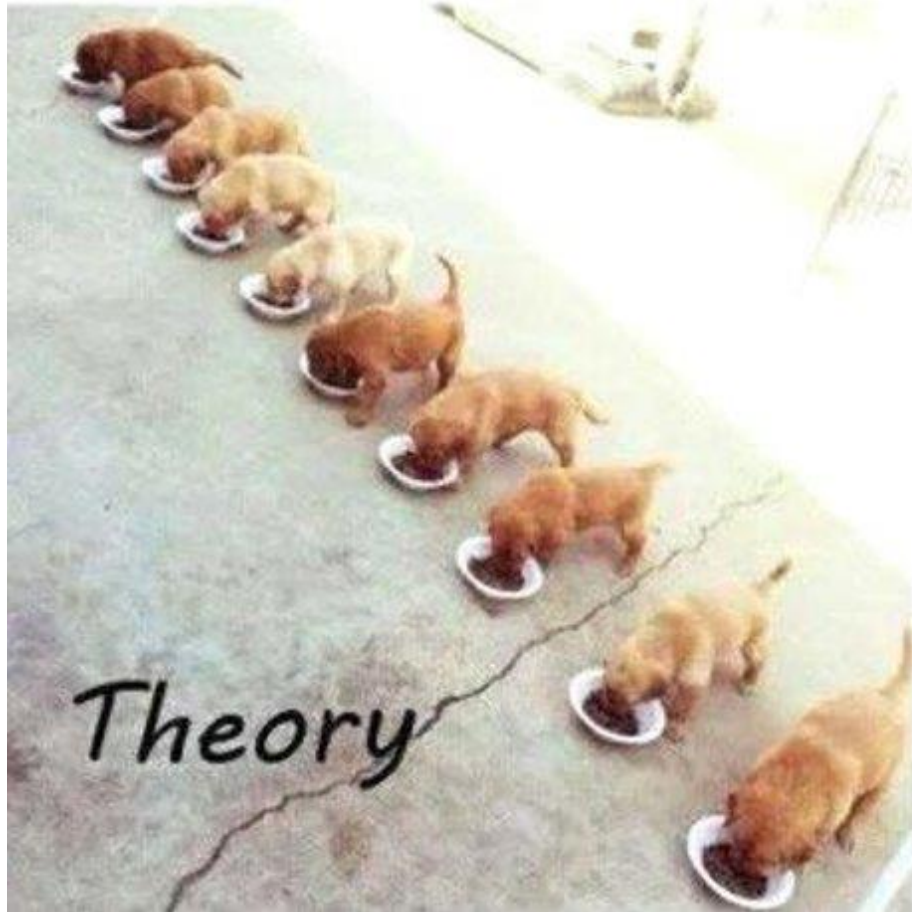
| Weak scaling

- | Solving a twice larger problem using twice more resources
- | Expected performance – get result in constant time

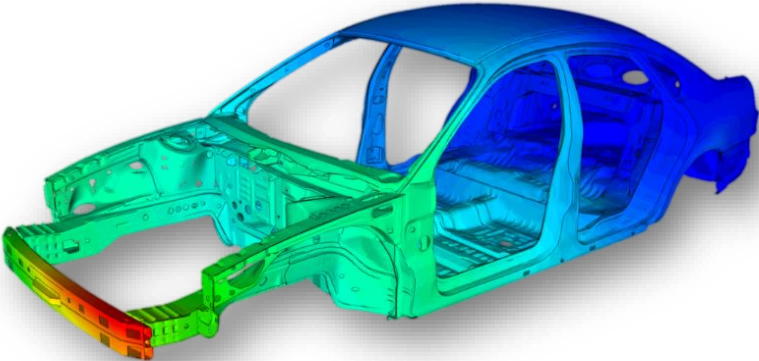
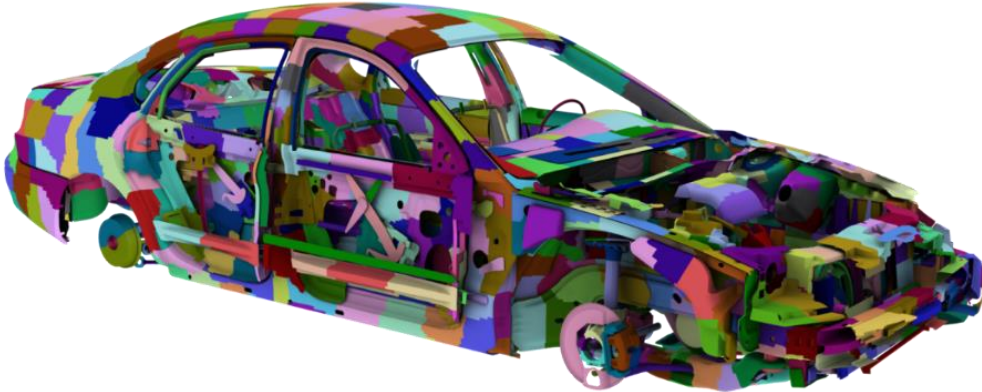
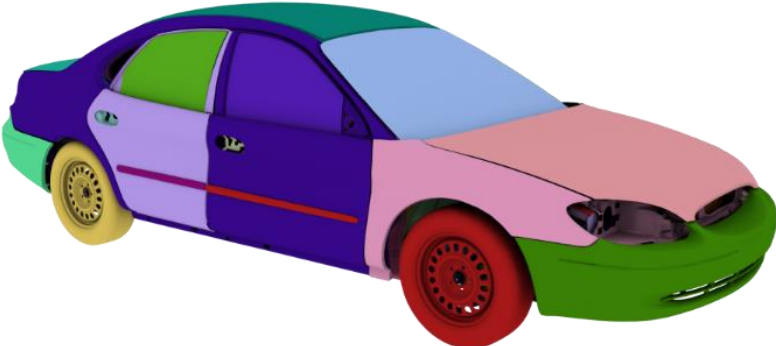
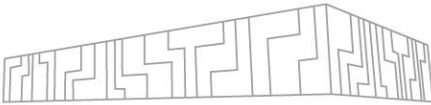




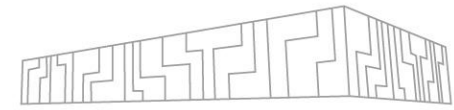
Multithreaded programming



PARALLEL COMPUTING

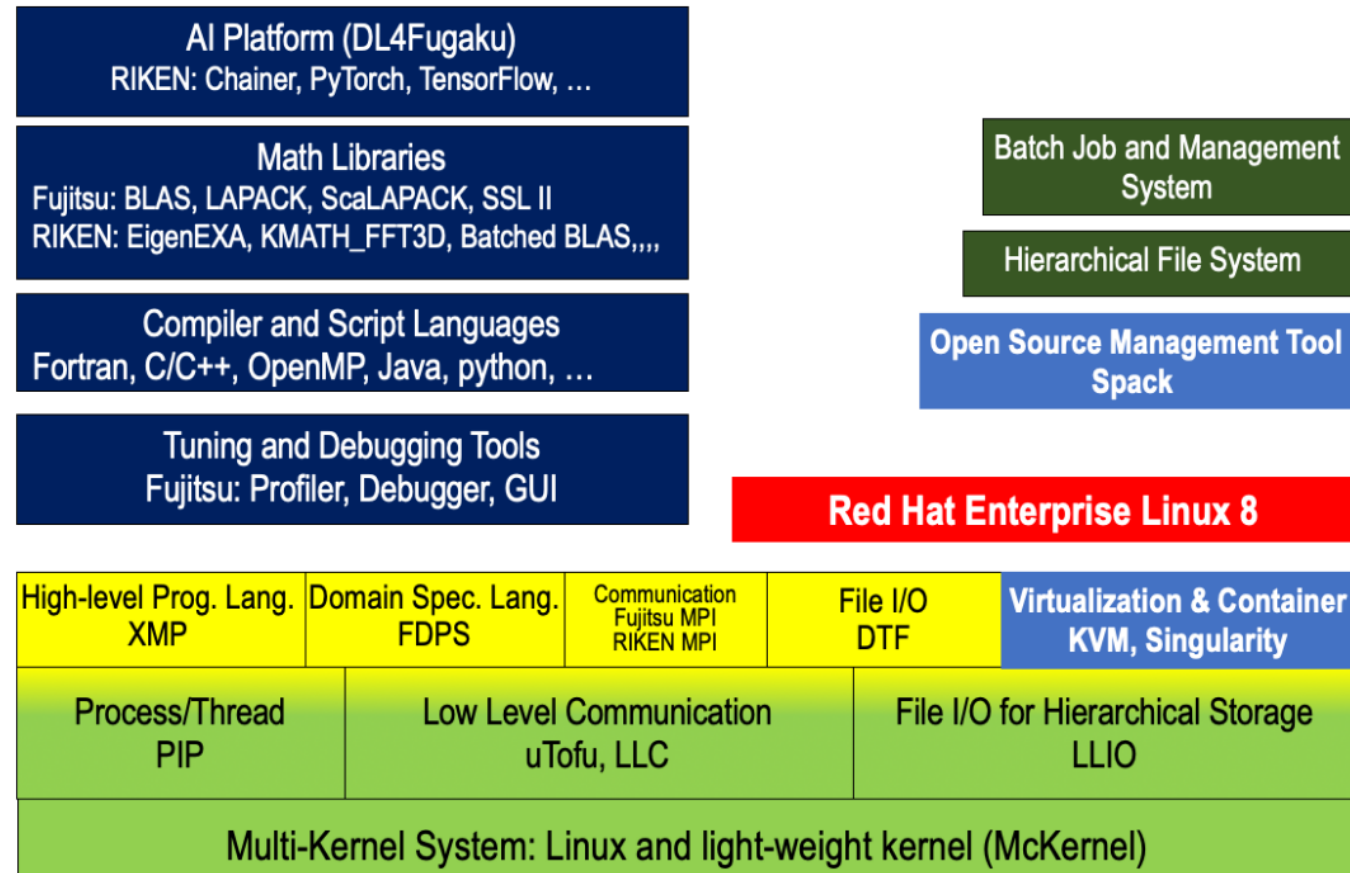


PRE-INSTALLED SOFTWARE

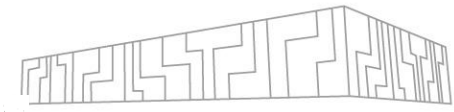


- Environment Module System
 - Modification of the environment paths
 - Software in several versions

Fugaku software stack

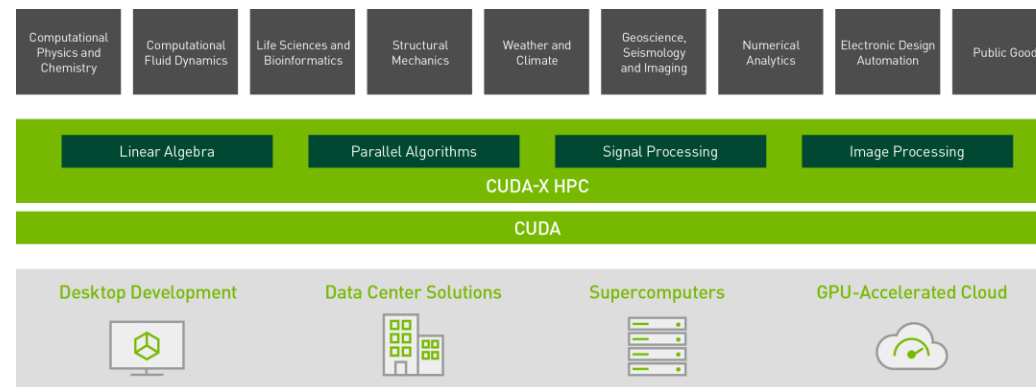
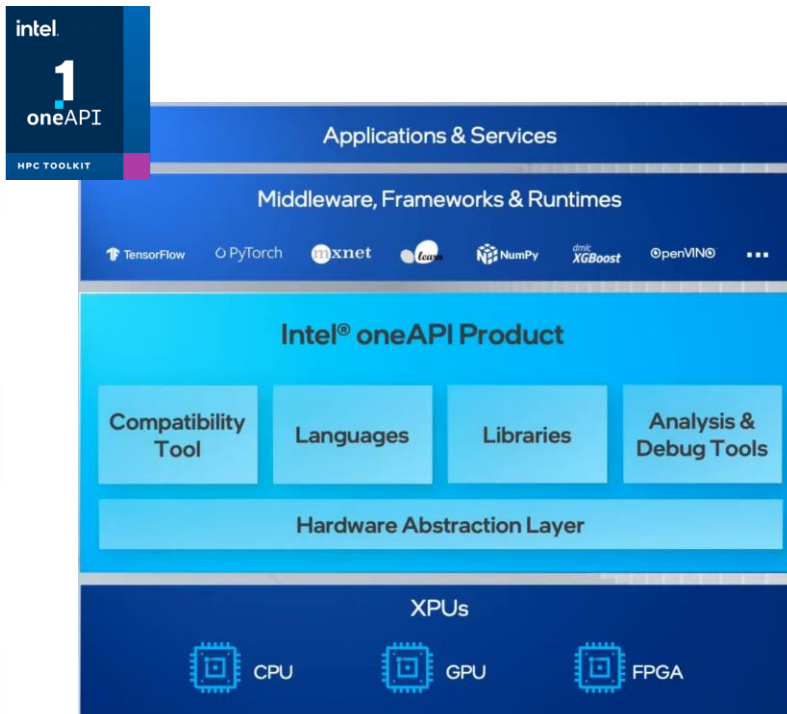


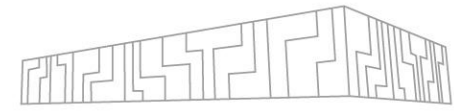
VENDOR'S SOFTWARE STACK



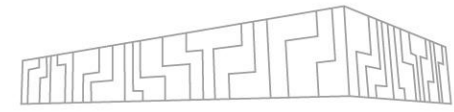
Simplified software development for heterogenous hardware

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





TRENDS



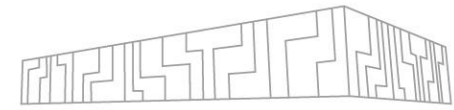
Path to exascale

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 (HPL-MxP)

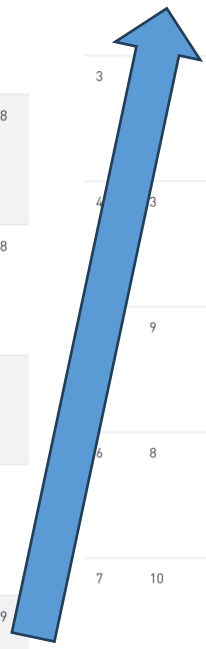


TOP500 LIST HPL + HPCG



Rank	System	Cores	Rmax (PFlop/s)	Rpeak (PFlop/s)	Power (kW)
1	El Capitan - HPE Cray EX255a, AMD 4th Gen EPYC 24C 1.8GHz, AMD Instinct MI300A, Slingshot-11, TOSS, HPE DOE/NNSA/LLNL United States	11,039,616	1,742.00	2,746.38	29,581
2	Frontier - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE Cray OS, HPE DOE/SC/Oak Ridge National Laboratory United States	9,066,176	1,353.00	2,055.72	24,607
3	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
4	JUPITER Booster - BullSequana XH3000, GH Superchip 72C 3GHz, NVIDIA GH200 Superchip, Quad-Rail NVIDIA InfiniBand NDR200, RedHat Enterprise Linux, EVIDEN EuroHPC/FZJ Germany	4,801,344	793.40	930.00	13,088
5	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	
6	HPC6 - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, RHEL 8.9, HPE Eni S.p.A. Italy	3,143,520	477.90	606.97	8,461
7	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
8	Alps - HPE Cray EX254n, NVIDIA Grace 72C 3.1GHz, NVIDIA GH200 Superchip, Slingshot-11, HPE Cray OS, HPE Swiss National Supercomputing Centre (CSCS) Switzerland	2,121,600	434.90	574.84	7,124
9	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
10	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

Rank	TOP500 Rank	System	Cores	Rmax (PFlop/s)	HPCG (TFlop/s)
1	1	El Capitan - HPE Cray EX255a, AMD 4th Gen EPYC 24C 1.8GHz, AMD Instinct MI300A, Slingshot-11, TOSS, HPE DOE/NNSA/LLNL United States	11,039,616	1,742.00	17406.90
2	7	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
3		Frontier - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE Cray OS, HPE DOE/SC/Oak Ridge National Laboratory United States	9,066,176	1,353.00	14054.00
4	3	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
9		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
6	8	Alps - HPE Cray EX254n, NVIDIA Grace 72C 3.1GHz, NVIDIA GH200 Superchip, Slingshot-11, HPE Cray OS, HPE Swiss National Supercomputing Centre (CSCS) Switzerland	2,121,600	434.90	3671.32
7	10	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
8	15	ABCI 3.0 - HPE Cray XD670, Xeon Platinum 8558 48C 2.1GHz, NVIDIA H200 SXM5 141 GB, Infiniband NDR200, Rocky Linux 9, HPE National Institute of Advanced Industrial Science and Technology (AIST) Japan	479,232	145.10	2445.67
9	25	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
10	20	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

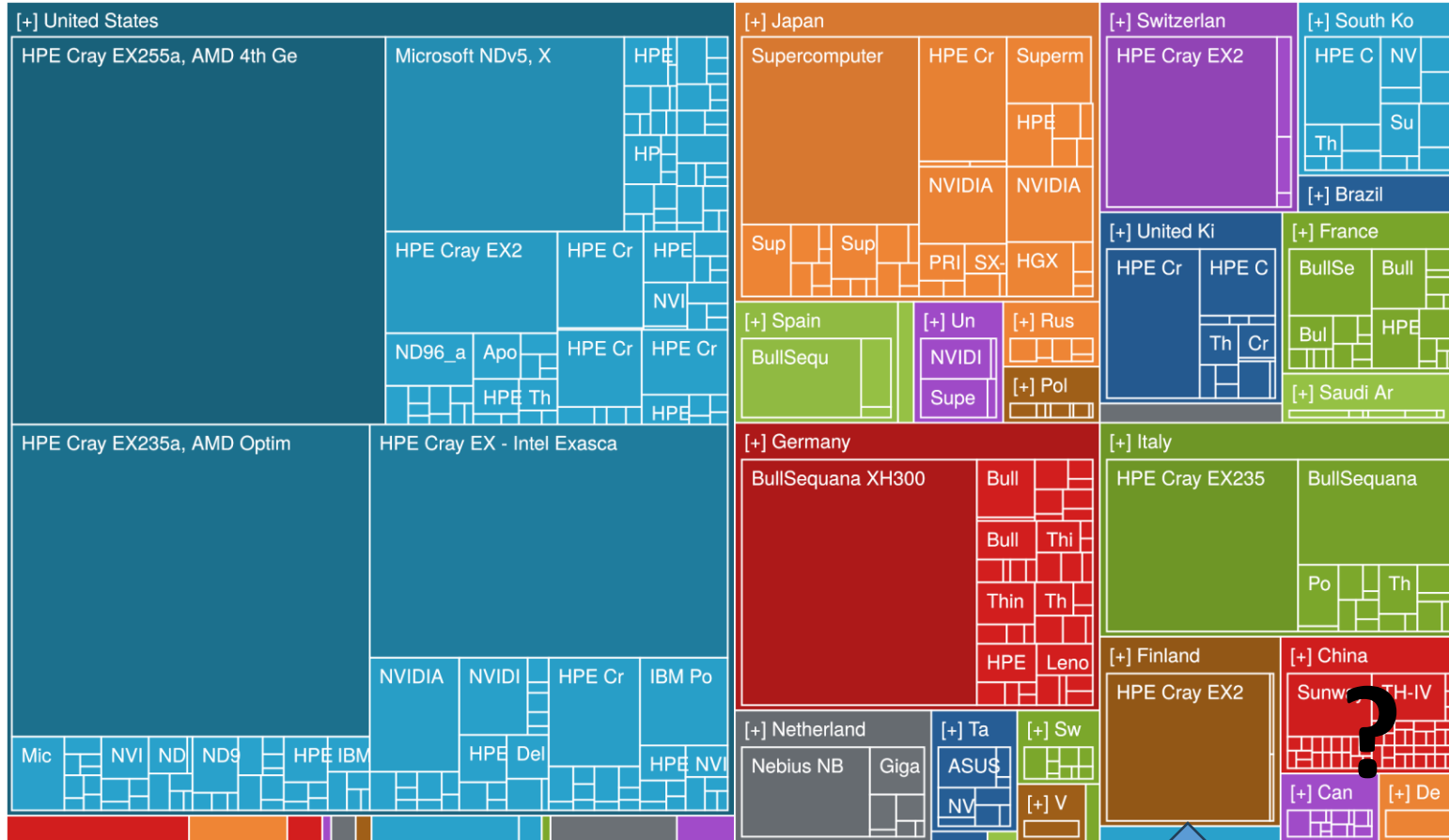
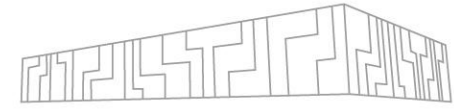


arm



06/2025

TOP500 LIST



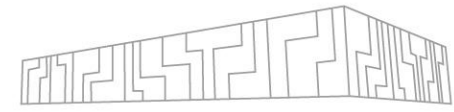
	Countries	Count	System Share (%)	Rmax (GFlops)	Rpeak (GFlops)	Cores
1	United States	175	35	6,698,604,970	10,642,950,168	58,129,200
2	China	47	9.4	281,269,604	471,807,345	18,920,568
3	Germany	41	8.2	1,185,266,660	1,519,298,880	8,943,756
4	Japan	39	7.8	1,229,009,840	1,631,301,398	12,787,024
5	France	25	5	327,543,500	508,630,042	4,676,480
6	Italy	17	3.4	874,883,920	1,128,393,618	7,730,480
7	South Korea	15	3	323,114,900	442,276,740	2,799,412
8	Canada	13	2.6	78,990,060	138,729,945	1,051,472
9	United Kingdom	13	2.6	394,527,564	540,015,334	3,152,784
10	Brazil	9	1.8	67,911,650	122,168,142	874,720
11	Norway	9	1.8	44,075,210	62,065,550	642,528
12	Sweden	9	1.8	69,449,910	98,505,836	491,648
13	Taiwan	8	1.6	116,600,520	170,134,420	753,216
14	Poland	7	1.4	63,412,670	101,087,440	401,120
15	Netherlands	7	1.4	291,308,450	489,421,170	1,403,296
16	Russia	6	1.2	71,457,000	98,725,620	721,488
17	Saudi Arabia	6	1.2	93,713,210	149,854,280	2,435,392
18	India	6	1.2	29,411,690	38,541,096	785,400
19	Singapore	5	1	38,287,830	62,498,340	315,424
20	Australia	4	0.8	55,227,310	73,106,095	496,032
21	United Arab Emirates	4	0.8	122,281,000	201,730,583	816,800
22	Switzerland	4	0.8	471,168,600	628,264,760	2,663,712
23	Czechia	3	0.6	17,991,180	21,784,990	360,192
24	Spain	3	0.6	221,872,600	306,102,991	1,542,016
25	Finland	3	0.6	391,388,310	546,192,585	3,116,992
26	Thailand	2	0.4	21,996,000	35,449,860	134,656
27	Ireland	2	0.4	6,697,330	13,160,450	152,320
28	Slovenia	2	0.4	6,918,000	10,047,000	156,480
29	Turkey	2	0.4	7,366,290	10,100,610	69,760
30	Bulgaria	2	0.4	7,044,870	9,154,160	164,224
31	Austria	2	0.4	34,566,078	57,517,978	199,200
32	Israel	1	0.2	41,500,000	52,679,020	74,880
33	Denmark	1	0.2	66,590,000	100,629,630	223,088
34	Iceland	1	0.2	10,530,000	17,015,380	36,208
35	Morocco	1	0.2	3,158,110	5,014,730	71,232
36	Luxembourg	1	0.2	10,520,000	15,288,000	99,200
37	Belgium	1	0.2	2,776,000	3,966,560	23,200
38	Portugal	1	0.2	3,955,500	5,013,500	78,336
39	Argentina	1	0.2	5,390,150	12,582,910	43,008
40	Hungary	1	0.2	3,105,000	4,508,490	27,776
41	Vietnam	1	0.2	46,650,000	67,443,000	142,240

Czechia/IT4I

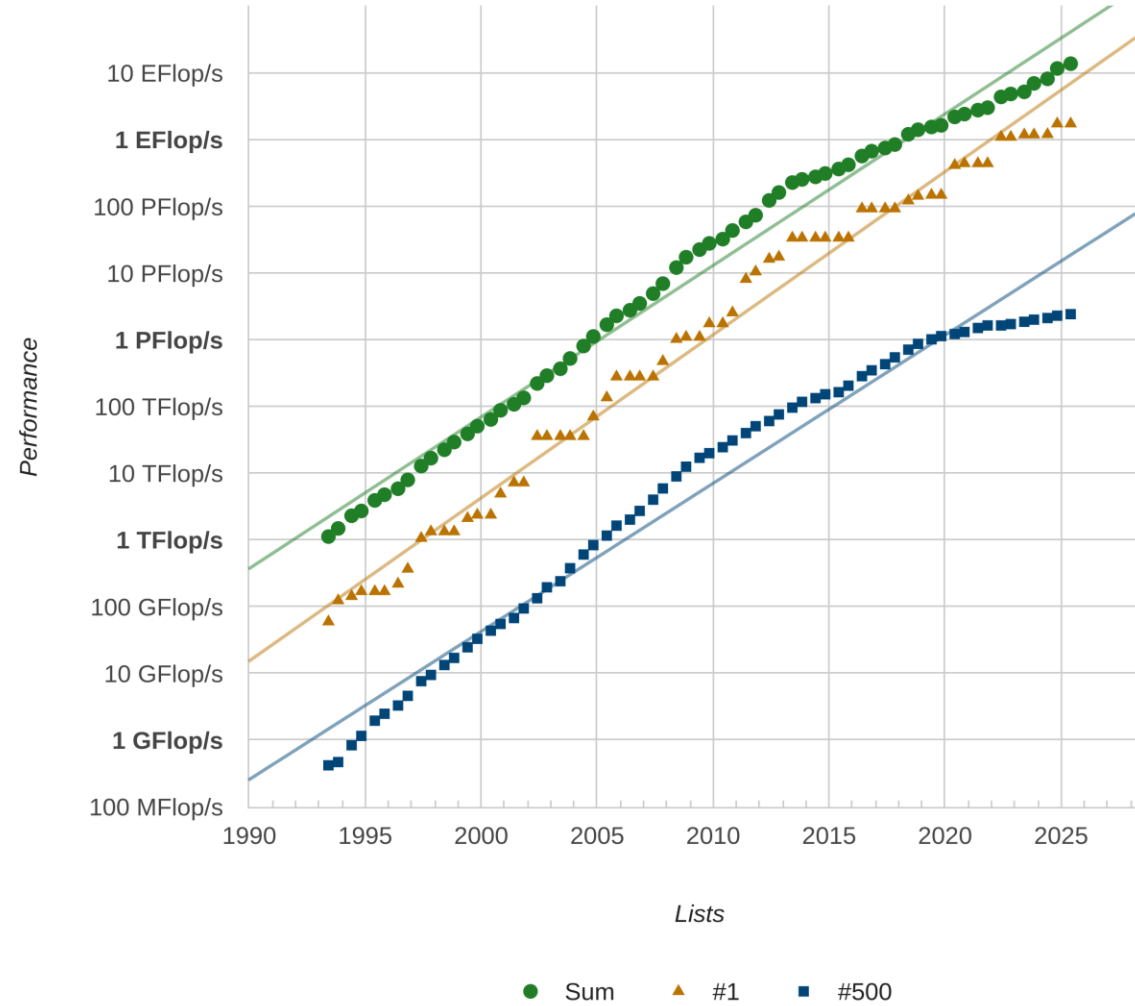
LUMI

6/2025

TOP500 LIST

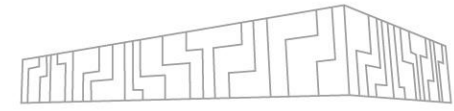


Projected Performance Development

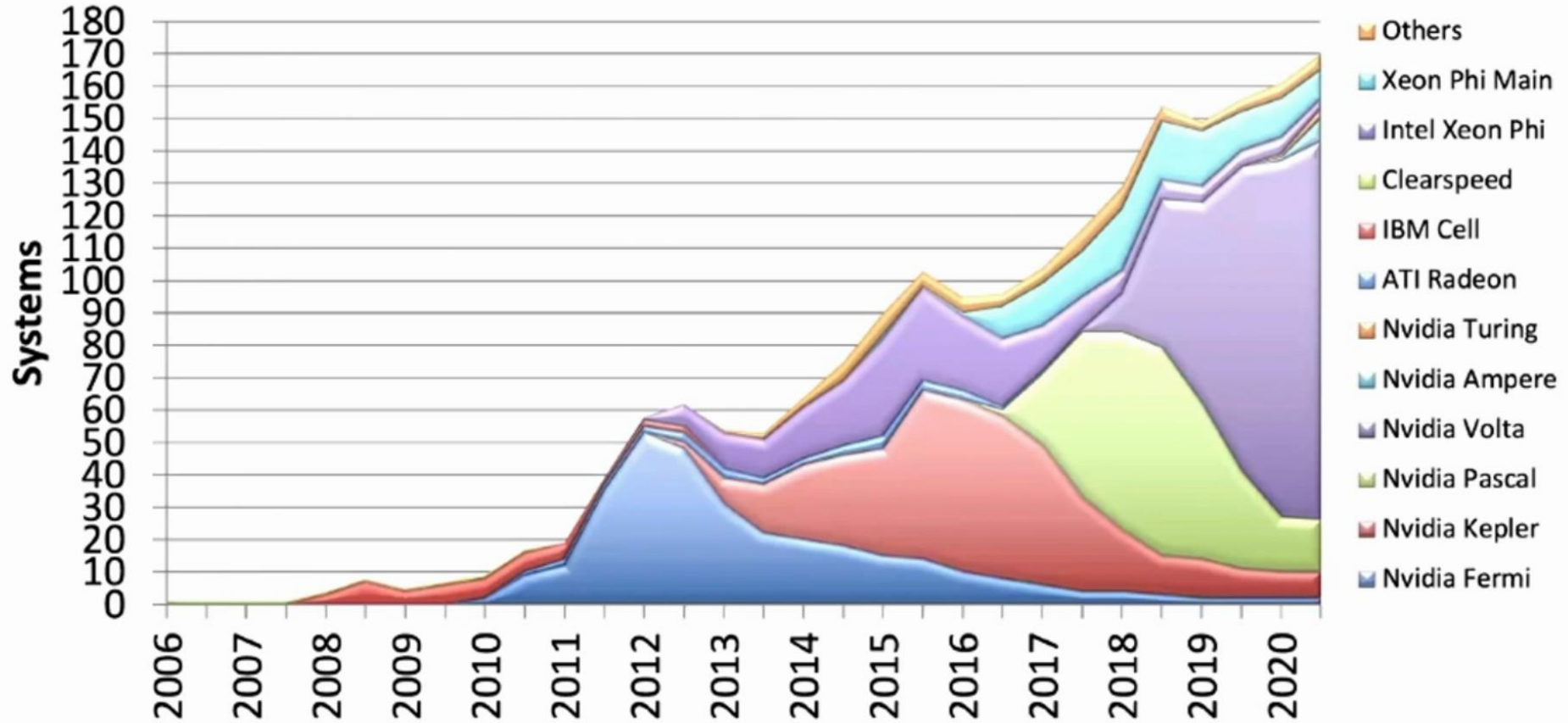


6/2025

TOP500 LIST



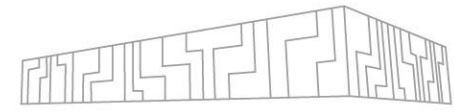
ACCELERATORS



11/2020

6/2024 - 193 out of 500 systems are accelerated

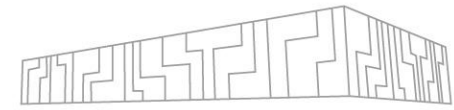
TOP500 LIST HPL / HPL - MXP



Rank	Site	System	Cores	HPL Rmax [TFlop/s]	Rpeak [TFlop/s]	Power [kW]	Efficiency [GFlop/W]	HPL-MxP Rmax [TFlop/s]	HPL-MxP Speedup [-]
1	DOE/NNSA/LLNL United States	El Capitan - HPE Cray EX255a, AMD 4th Gen EPYC 24C 1.8GHz, AMD Instinct MI300A , Slingshot-11, TOSS HPE	11,039,616	1,742.00	2,746.38	29,581	58.89	16,700	9.6
2	DOE/SC/Oak Ridge National Laboratory United States	Frontier - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X , Slingshot-11, HPE Cray OS HPE	9,066,176	1,353.00	2,055.72	24,607	54.98	11,400	8.4
3	DOE/SC/Argonne National Laboratory United States	Aurora - HPE Cray EX - Intel Exascale Compute Blade, Xeon CPU Max 9470 52C 2.4GHz, Intel Data Center GPU Max , Slingshot-11 Intel	9,264,128	1,012.00	1,980.01	38,698	26.15	11,600	11.5
4	EuroHPC/FZJ Germany	JUPITER Booster - BullSequana XH3000, GH Superchip 72C 3GHz, NVIDIA GH200 Superchip , Quad-Rail NVIDIA InfiniBand NDR200, RedHat Enterprise Linux EVIDEN	4,801,344	793.40	930.00	13,088	60.59		
5	Microsoft Azure United States	Eagle - Microsoft NDv5, Xeon Platinum 8480C 48C 2GHz, NVIDIA H100 , NVIDIA Infiniband NDR Microsoft Azure	2,073,600	561.20	846.84				
6	Eni S.p.A. Italy	HPC6 - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X , Slingshot-11, RHEL 8.9 HPE	3,143,520	477.90	606.97	8,461	56.48		
7	RIKEN Center for Computational Science Japan	Supercomputer Fugaku - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D Fujitsu	7,630,848	442.01	537.21	29,899	14.78	2,000	4.5
8	Swiss National Supercomputing Centre (CSCS) Switzerland	Alps - HPE Cray EX254n, NVIDIA Grace 72C 3.1GHz, NVIDIA GH200 Superchip , Slingshot-11, HPE Cray OS HPE	2,121,600	434.90	574.84	7,124	61.05		
9	EuroHPC/CSC Finland	LUMI - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X , Slingshot-11 HPE	2,752,704	379.70	531.51	7,107	53.43	2,350	6.2
10	EuroHPC/CINECA Italy	Leonardo - BullSequana XH2000, Xeon Platinum 8358 32C 2.6GHz, NVIDIA A100 SXM4 64 GB , Quad-rail NVIDIA HDR100 Infiniband EVIDEN	1,824,768	241.20	306.31	7,494	32.19	1,800	7.6

06/2025

TOP500 LIST HPL / HPL - MXP

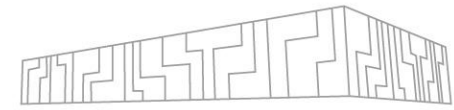









Rank	Site	System	Cores	HPL Rmax [TFlop/s]	Rpeak [TFlop/s]	Power [kW]	Efficiency [GFlop/W]	HPL-MxP Rmax [TFlop/s]	HPL-MxP Speedup [-]
1	DOE/NNSA/LLNL United States	El Capitan - HPE Cray EX255a, AMD 4th Gen EPYC 24C 1.8GHz, AMD Instinct MI300A , Slingshot-11, TOSS HPE	11,039,616	1,742.00	2,746.38	29,581	58.89	16,700	9.6
2	DOE/SC/Oak Ridge National Laboratory United States	Frontier - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X , Slingshot-11, HPE Cray OS HPE	9,066,176	1,353.00	2,055.72	24,607	54.98	11,400	8.4
3	DOE/SC/Argonne National Laboratory United States	Aurora - HPE Cray EX - Intel Exascale Compute Blade, Xeon CPU Max 9470 52C 2.4GHz, Intel Data Center GPU Max , Slingshot-11 Intel	9,264,128	1,012.00	1,980.01	38,698	26.15	11,600	11.5
4	EuroHPC/FZJ Germany	JUPITER Booster - BullSequana XH3000, GH Superchip 72C 3GHz, NVIDIA GH200 Superchip , RedHat Enterprise Linux 8.9 EVIDEN	10,000,000	1,000.00	1,000.00	16,088	60.59		
5	Microsoft Azure United States	Eagle - Microsoft H100 , NV Microsoft							
6	Eni S.p.A. Italy	HPC6 - HPE 64C 2GHz, AMD Instinct MI250X , Slingshot-11, RHEL 8.9 HPE	3,143,520	477.90	606.97	8,461	56.48		
7	RIKEN Center for Computational Science Japan	Supercomputer Fugaku - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D Fujitsu	7,630,848	442.01	537.21	29,899	14.78	2,000	4.5
8	Swiss National Supercomputing Centre (CSCS) Switzerland	Alps - HPE Cray EX254n, NVIDIA Grace 72C 3.1GHz, NVIDIA GH200 Superchip , Slingshot-11, HPE Cray OS HPE	2,121,600	434.90	574.84	7,124	61.05		
9	EuroHPC/CSC Finland	LUMI - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X , Slingshot-11 HPE	2,752,704	379.70	531.51	7,107	53.43	2,350	6.2
10	EuroHPC/CINECA Italy	Leonardo - BullSequana XH2000, Xeon Platinum 8358 32C 2.6GHz, NVIDIA A100 SXM4 64 GB , Quad-rail NVIDIA HDR100 Infiniband EVIDEN	1,824,768	241.20	306.31	7,494	32.19	1,800	7.6

Exascale-ready machines need to achieve efficiency over 50 GFlop/W.
(to be hosted in typical 20MW datacenter)

06/2025

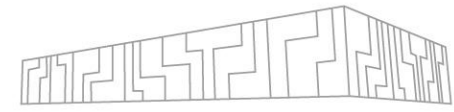
GREEN500



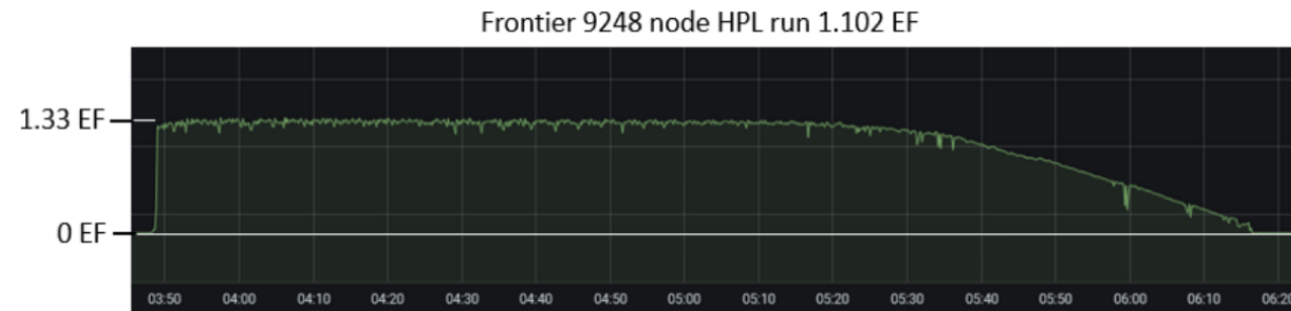
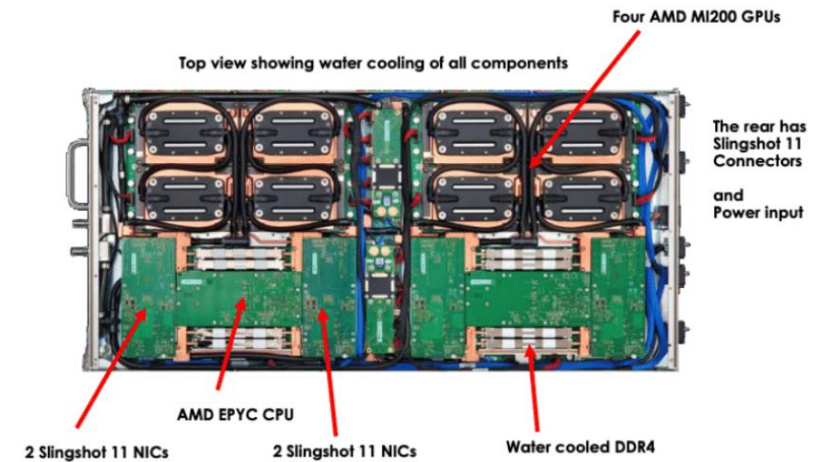
Rank	TOP500 Rank	System	Cores	Rmax [PFlop/s]	Power [kW]	Energy Efficiency [GFlops/W]
	1	JEDI - BullSequana XH3000, Grace Hopper Superchip 72C 3GHz, NVIDIA GH200 Superchip , Quad-Rail NVIDIA InfiniBand NDR200 , ParTec/EVIDEN, EuroHPC/FZJ Germany	19,584	4.50	67	72.733
	2	ROMEO-2025 - BullSequana XH3000, Grace Hopper Superchip 72C 3GHz, NVIDIA GH200 Superchip , Quad-Rail NVIDIA InfiniBand NDR200, EVIDEN, ROMEO HPC Center France	47,328	9.86	160	70.912
	3	Adastra 2 - HPE Cray EX255a, AMD 4th Gen EPYC 24C 1.8GHz, AMD Instinct MI300A , Slingshot-11, RHEL , HPE, GENCI-CINES France	16,128	2.53	37	69.098
	4	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
	5	Otus (GPU only) - ThinkSystem SD665-N V3, AMD EPYC 9655 96C 2.6GHz, NVIDIA H100 SXM5 80GB , Infiniband NDR, Rocky Linux 9.4 , Lenovo , Universitaet Paderborn - PC2 Germany	19,440	4.66		68.177
	6	Capella - Lenovo ThinkSystem SD665-N V3, AMD EPYC 9334 32C 2.7GHz, Nvidia H100 SXM5 94GB , Infiniband NDR200, AlmaLinux 9.4 , MEGWARE , TU Dresden, ZIH Germany	85,248	24.06	445	68.053
	7	SSC-24 Energy Module - HPE Cray XD670, Xeon Gold 6430 32C 2.1GHz, NVIDIA H100 SXM5 80GB , Infiniband NDR400, RHEL 9.2 , HPE , Samsung Electronics South Korea	11,200	3.82	69	67.251
	8	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
	9	AMD Ouranos - BullSequana XH3000, AMD 4th Gen EPYC 24C 1.8GHz, AMD Instinct MI300A , Infiniband NDR200, RedHat Enterprise Linux , EVIDEN , Atos France	16,632	2.99	48	66.464
	10	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

06/2025

FRONTIER

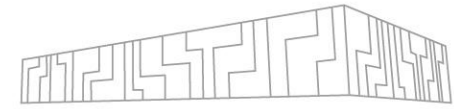


- 74 HPE Cray EX cabinets, 9 408 nodes
- 1 AMD Milan “Trento” 7A53 Epyc CPU + 4 AMD Instinct MI250X GPUs
- 512GiB DDR4 + 512GiB HMB2e (128GiB per GPU) coherent memory across node
- HPE Slingshot-11 interconnect (200 Gbit/s)
- Similar architecture used for LUMI-G
- 1.102 exaflops of Linpack, 21.1 MW



Wall Clock time (Eastern) May 27, 2022 (2.45 hour run time)

USA ROADMAP



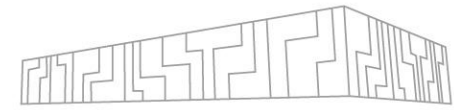
Pre-Exascale Systems

Future Exascale Systems



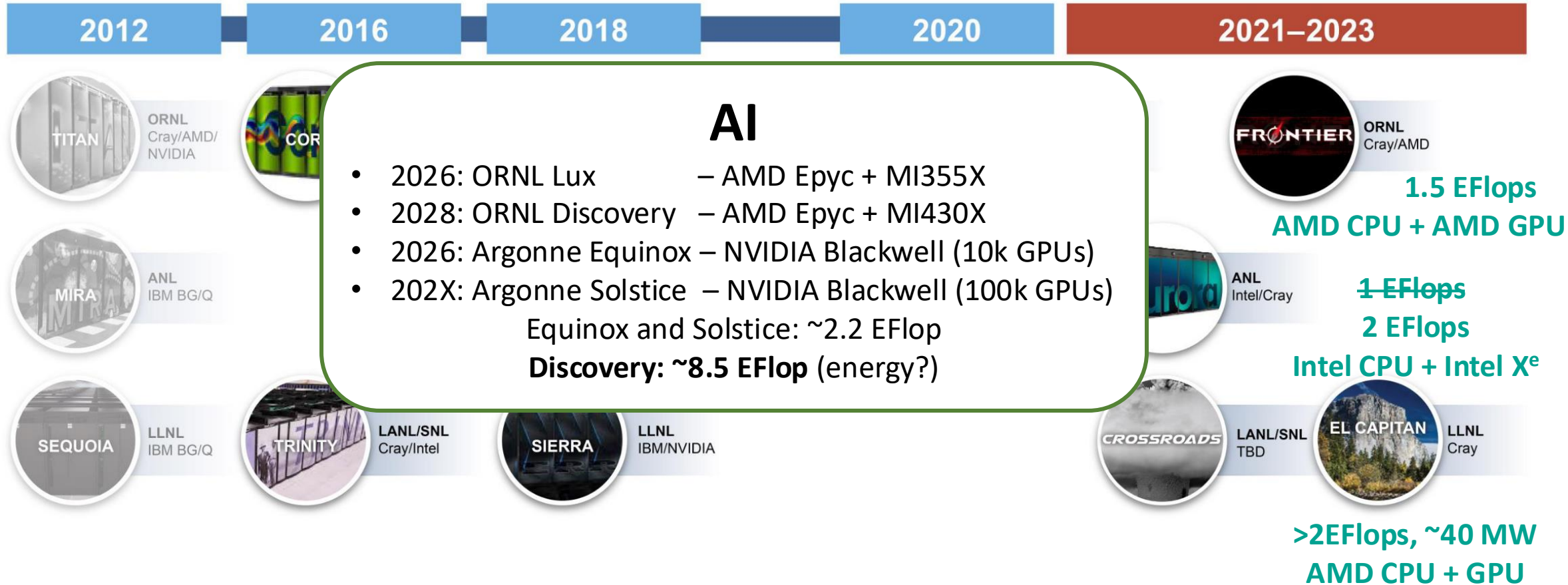
High variability of CPU and GPU vendors

USA ROADMAP



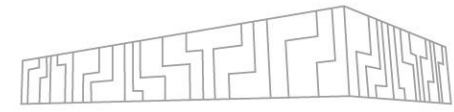
Pre-Exascale Systems

Future Exascale Systems



High variability of CPU and GPU vendors

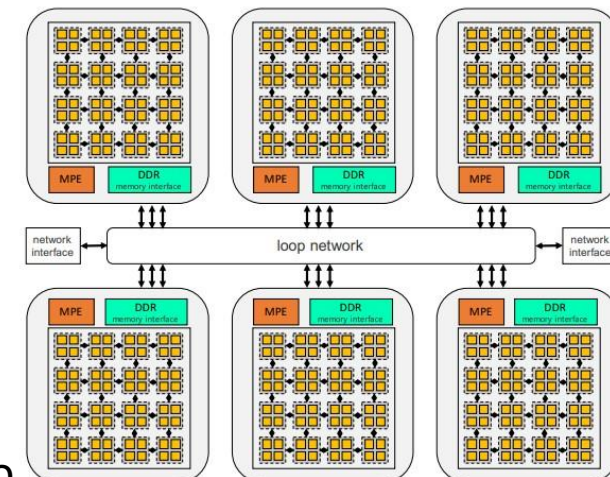
SUPERCOMPUTER #1 ?!



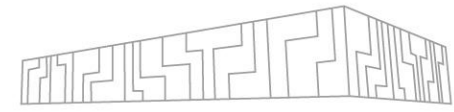
- El Capitan (USA) 06/2025 - 1.742 exaflops of Linpack, 29.6 MW

Meanwhile in China:

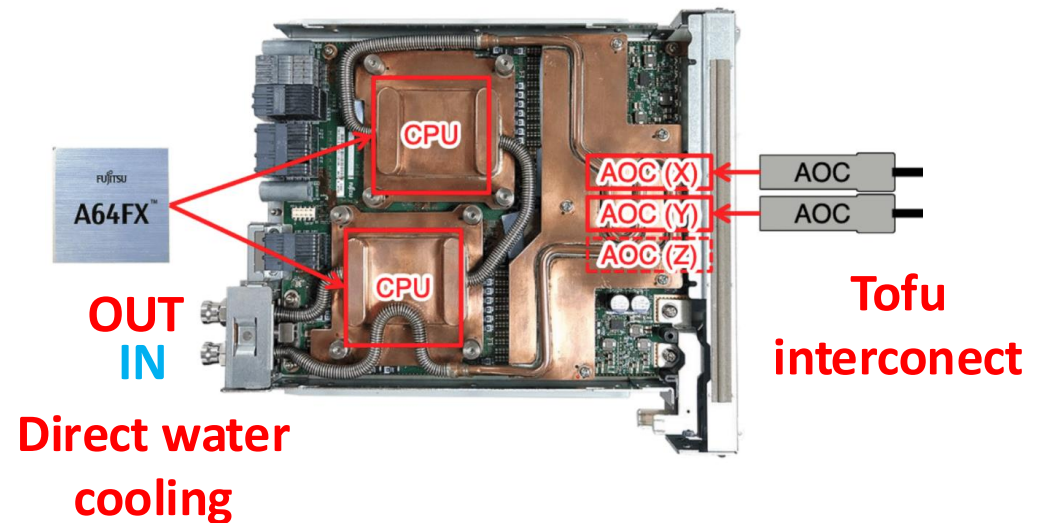
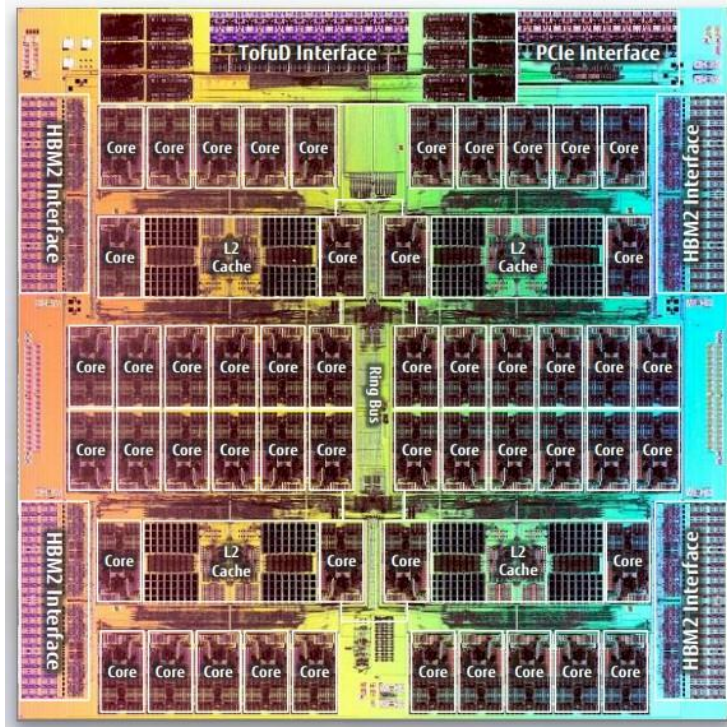
- Sunway Oceanlite (03/2021) - 1.05 exaflops of Linpack, ~35MW
 - ShenWei post-Alpha CPU ISA, 512-bit IS
 - 96 cabinets, 98 304x SW39010 390-core CPU, 14nm
 - Not in the top500.org list
- Tianhe-3 (10/2021) - 1.3 exaflops Linpack
 - 2x Phytium 2000+ FTP ARM CPU (16nm) + Matrix 2000+ MTP accelerator
 - Not in the top500.org list
- Shenzhen Phase 2 - scheduled for 2025
 - 2 exaflops
 - Sugon's Hygon CPU - delayed



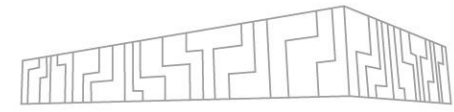
FUGAKU SUPERCOMPUTER



- 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

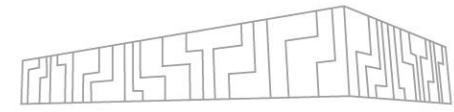
- **Bulgaria** (6PF, AMD+Nvidia), **Czech Republic** (15PF, AMD+Nvidia), **Luxembourg** (18PF, AMD+Nvidia), **Portugal** (10PF, A64FX, AMD+Nvidia), **Slovenia** (6.8PF, AMD+Nvidia)

High-range Pre-Exascale Supercomputers

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

Next generations of systems planned

EUROPEAN PRE-EXASCALE SYSTEMS



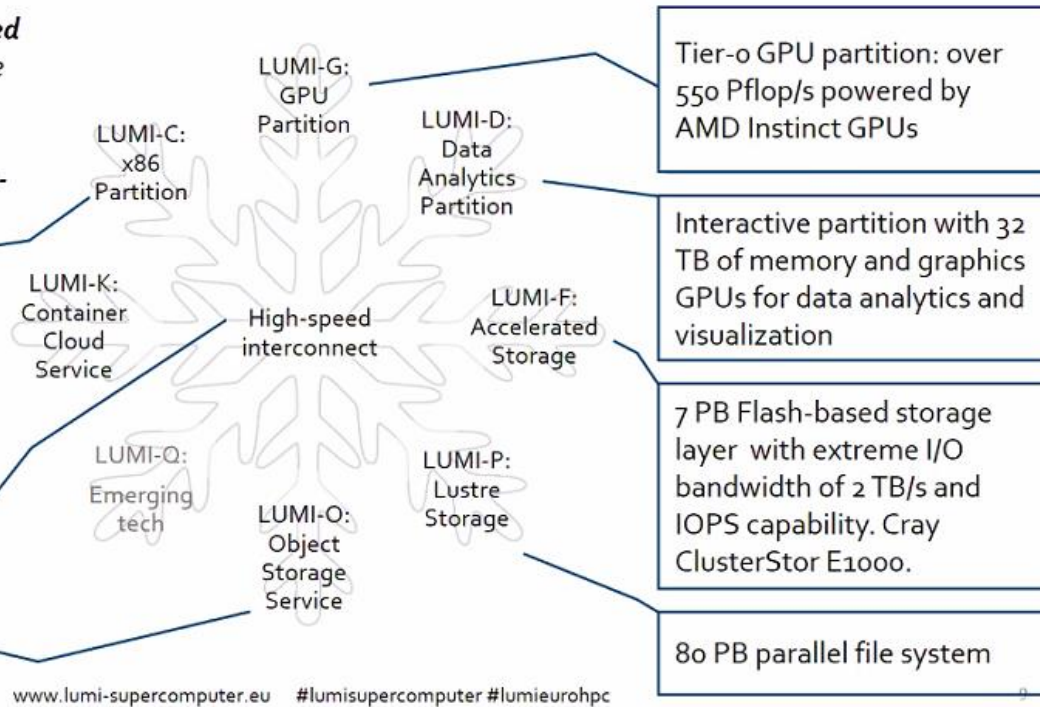
LUMI

LUMI is a Tier-0 GPU-accelerated supercomputer that enables the convergence of high-performance computing, artificial intelligence, and high-performance data analytics.

- Supplementary CPU partition
- ~200,000 AMD EPYC CPU cores

Possibility for combining different resources within a single run. HPE Slingshot technology.

30 PB encrypted object storage (Ceph) for storing, sharing and staging data



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



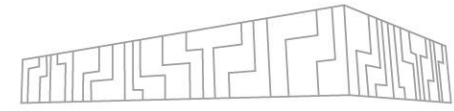
- H2 2022
- 240M €, 248 PFlops
- 3456 accelerated nodes
2x Intel Xeon Ice Lake CPUs
+ 4 Nvidia A100 GPUs
- 1536 non-accelerated nodes
2x Intel Xeon Sapphire Rapids

MareNostrum V

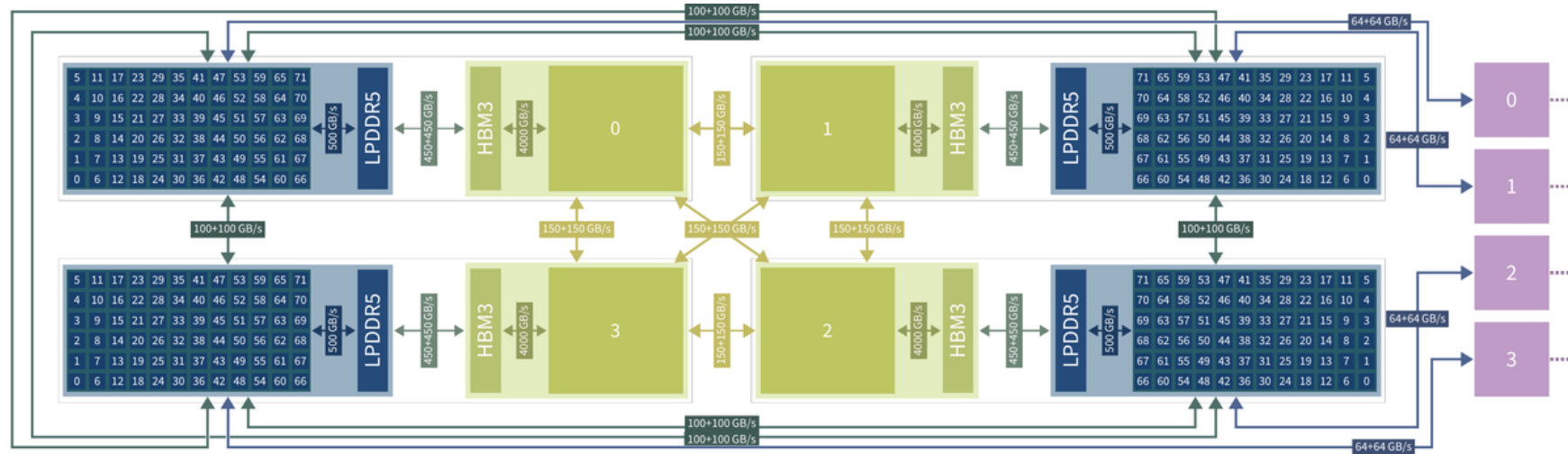
- H2 2023
- 223M €, 200 PFlops
- 2x Intel Sapphire Rapids +
4x Nvidia H100



JUPITER SUPERCOMPUTER

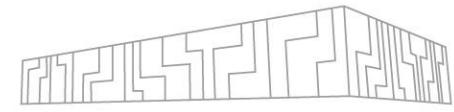


- ~6000 nodes of Nvidia Grace Hopper, 1 ExaFLOP/s



- >1300 nodes of SiPearl Rhea, 5 PetaFLOP/s
 - 2x Rhea1 (80 Arm Neoverse Zeus cores) + 2x 64 GB HBM2e + 512 GB DDR5
 - Similar approach to A64FX (ARM cores with wide SVE vector units)

EUROPEAN PROCESSOR INITIATIVE (EPI)

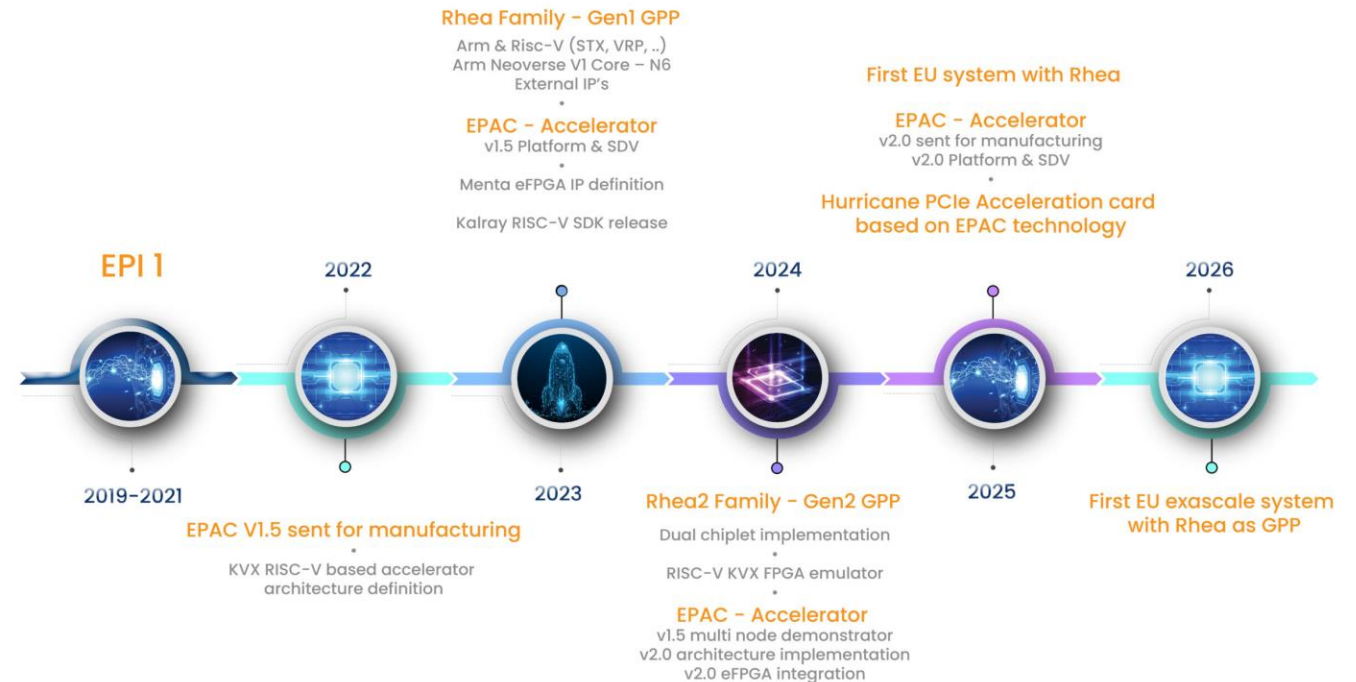


Europe invests into development of a new processor

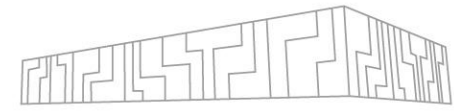
- Security
- Competitiveness

Design a roadmap of future European low power processors

- common platform
- general purpose processor
- accelerator
- automotive



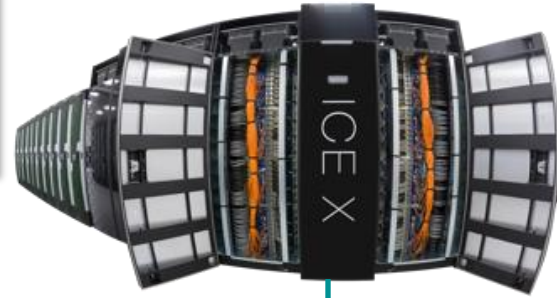
HISTORY OF THE IT4INNOVATIONS



Anselm



Salomon



NVIDIA DGX-2



ARTIFICIAL INTELLIGENCE



Barbora



5/2011

7/2014

7/2015

3/2019

10/2019

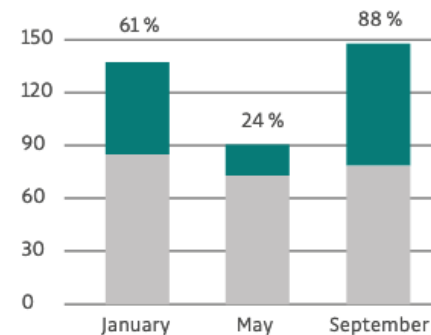
7/2021

6/2013



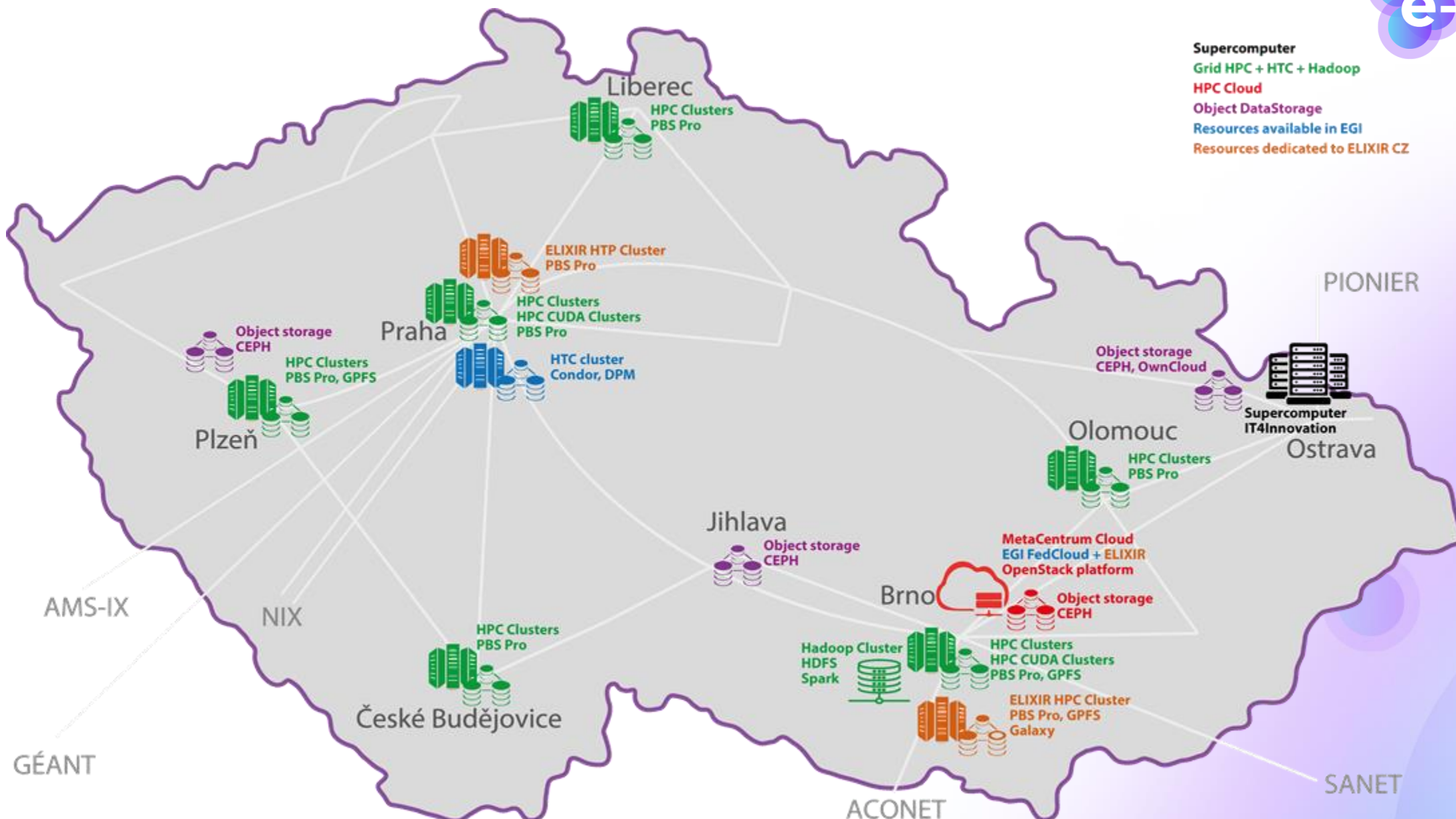
Open Access Grant Competitions in 2020

- Granted allocation
- Difference between demand and granted allocation

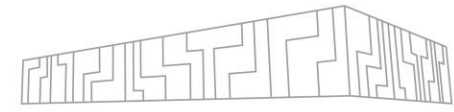


KAROLINA





IT4I – A MODERN DATA CENTER



500 sq.m.



OxyReduct fire prevention

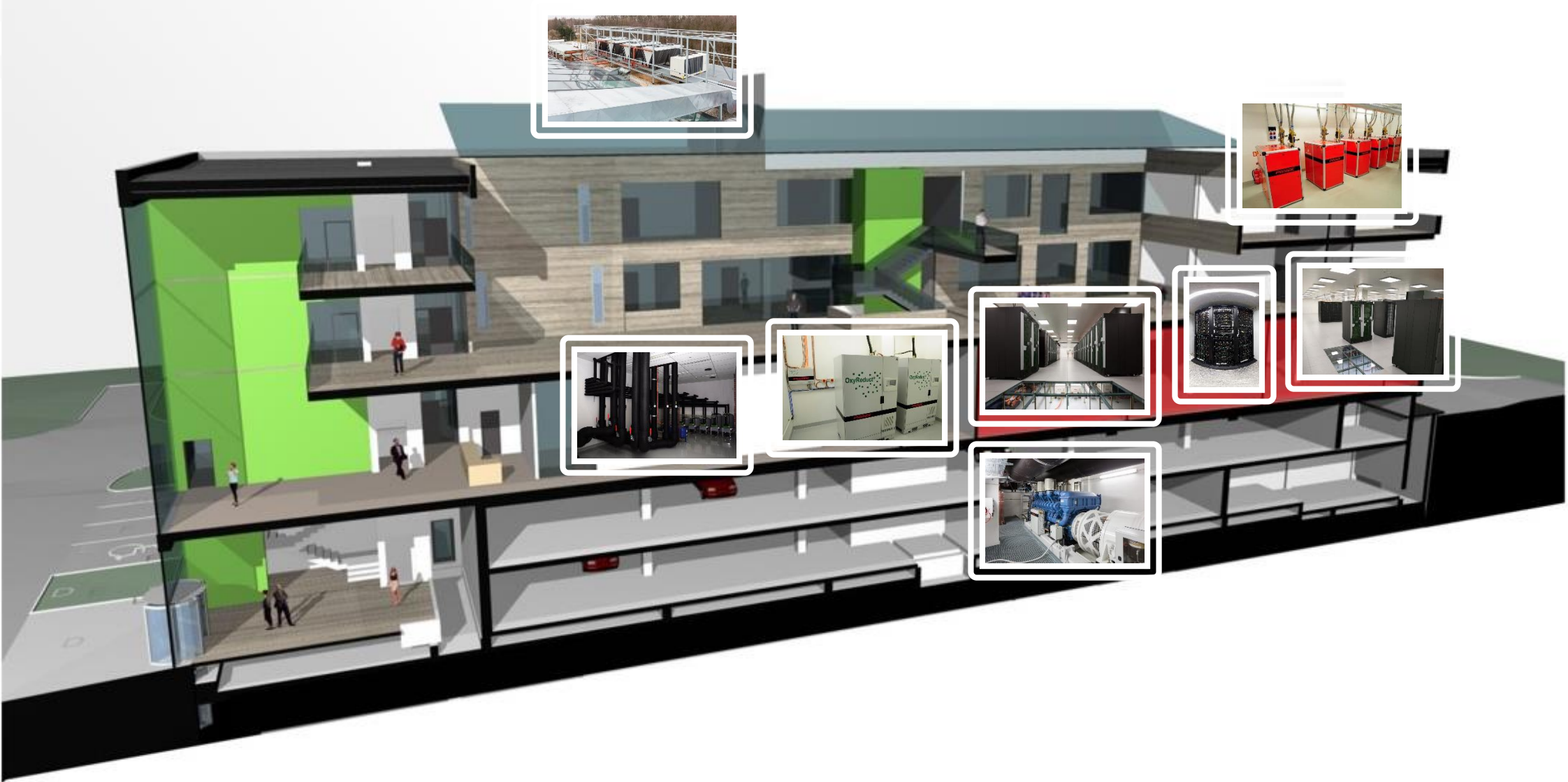
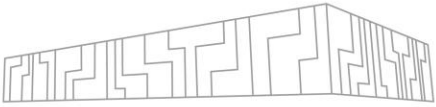
Dynamic rotating UPS 2x2,5MVA



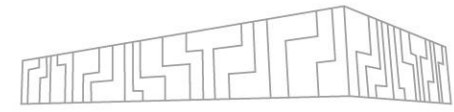
Cold and Hot water cooling



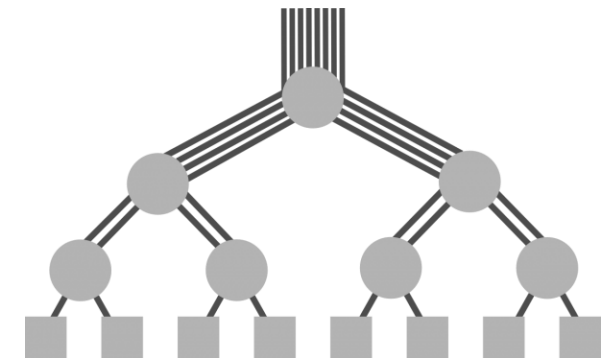
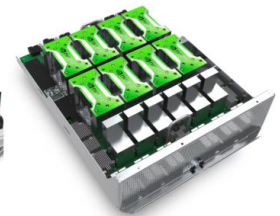
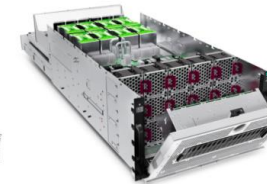
SUPPLEMENTARY INFRASTRUCTURE



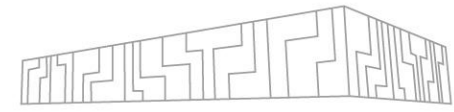
KAROLINA SUPERCOMPUTER



- **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



KAROLINA SUPERCOMPUTER

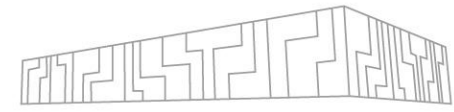


- 720x compute nodes, universal partition
 - **3833** TFLOPS Peak performance
- 72x compute nodes, accelerated partition
 - **8645** TFLOPS Peak performance

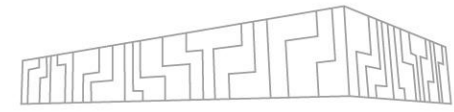


BARBORA SUPERCOMPUTER

- 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

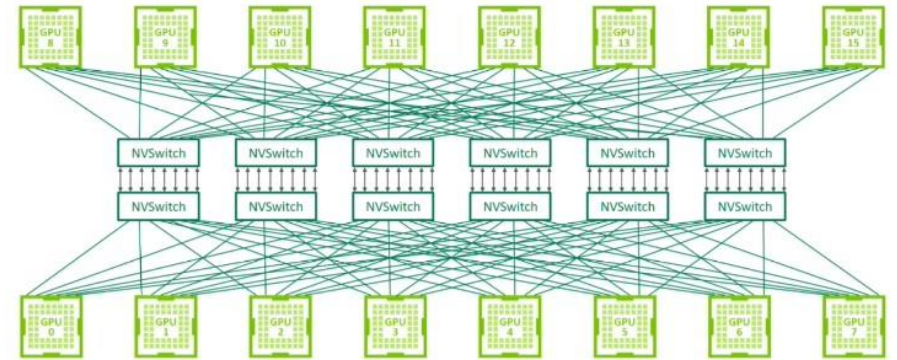


NVIDIA DGX PLATFORM



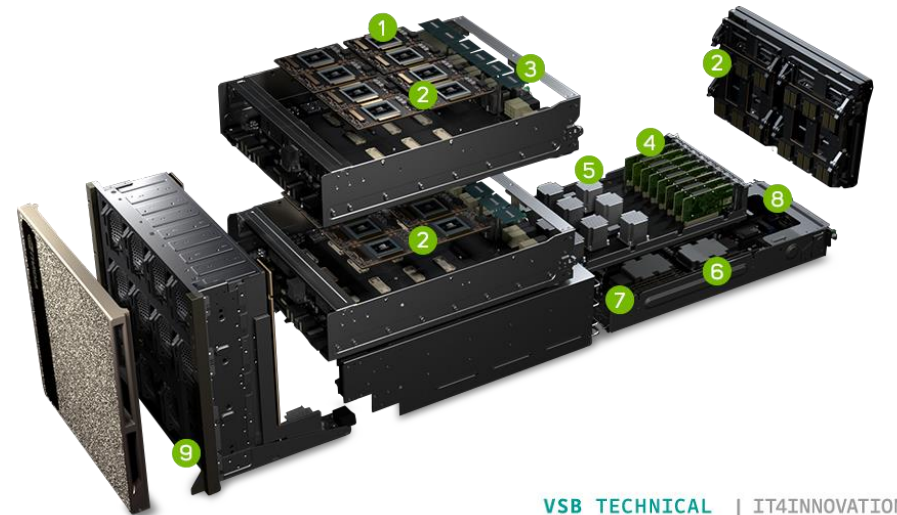
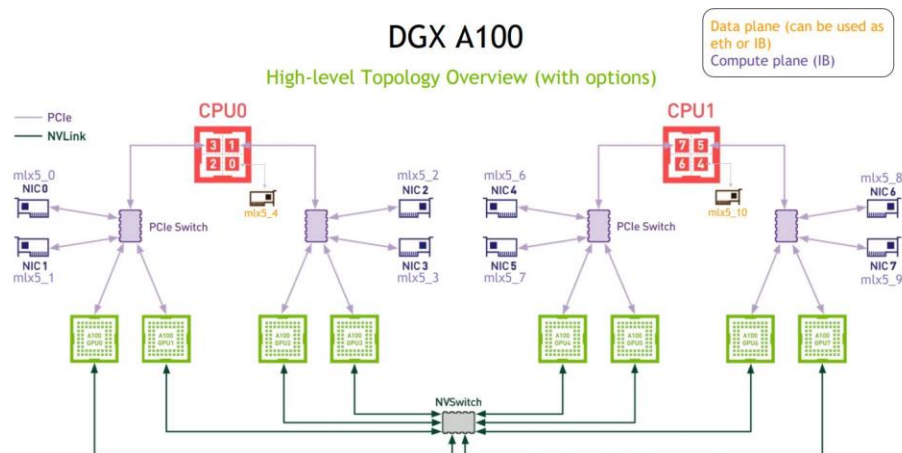
DGX-2

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

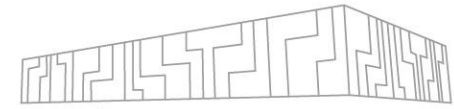


DGX-A100

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



IT4I IN THE TOP500.ORG

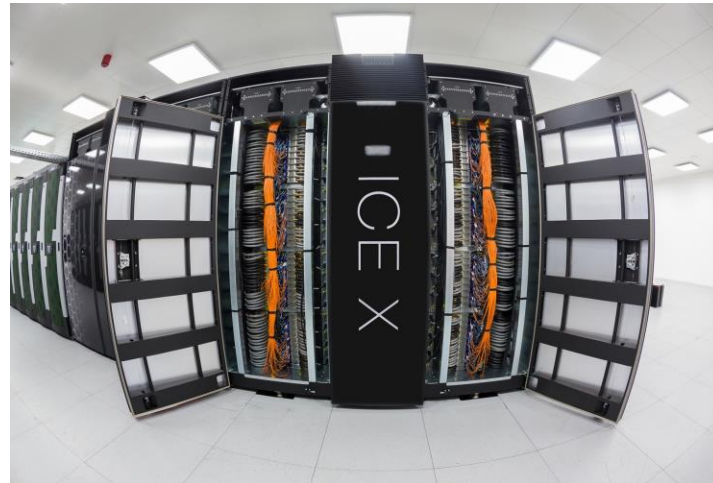


Salomon ranking

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

40 **Salomon** - SGI ICE X, Xeon E5-2680v3 12C 2.5GHz, Infiniband FDR, Intel Xeon Phi 7120P, HPE
IT4Innovations National Supercomputing Center, VSB-
Technical University of Ostrava
Czechia

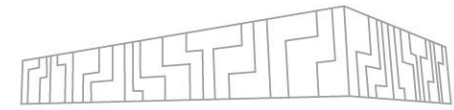
76,896 CPU cores
1,457.73 Rmax [TFlop/s]
2,011.64 Rpeak [TFlop/s]
4,806 power [kW]



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 CPU cores
6.75 Rmax [PFlop/s]
9.08 Rpeak [PFlop/s]
311 power [kW]

IT4I IN THE TOP500.ORG

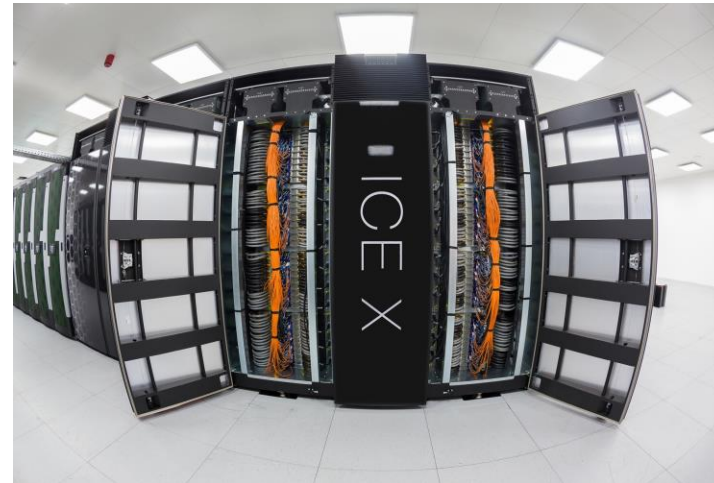


Karolina GPU ranking

List	Rank
06/2025	195
11/2024	165
06/2024	135
11/2023	112
06/2023	95
11/2022	85
06/2022	79
11/2021	71

40 **Salomon** - SGI ICE X, Xeon E5-2680v3 12C 2.5GHz, Infiniband FDR, Intel Xeon Phi 7120P, HPE
IT4Innovations National Supercomputing Center, VSB-
Technical University of Ostrava
Czechia

76,896 CPU cores
1,457.73 Rmax [TFlop/s]
2,011.64 Rpeak [TFlop/s]
4,806 power [kW]



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 CPU cores
6.75 Rmax [PFlop/s]
9.08 Rpeak [PFlop/s]
311 power [kW]



Ondřej Vysocký, Filip Vaverka
Ondrej.vysocky@vsb.cz, filip.vaverka@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

