

Introduction to High Performance Computing

CSCS-USI Summer School 2020
Vasileios Karakasis, CSCS
July 13, 2020

<https://github.com/eth-cscs/SummerSchool2020>

Why HPC?

Supercomputing How Cancer Superdiffusion

Searching for Human Brain Daint Supercomputer

October 20, 2017 by [staff](#) [Leave a Comment](#)

Scientists at the University of Basel are using the [Piz Daint](#) supercomputer at CSCS to discover interrelationships in the human genome that might simplify the search for "memory molecules" and eventually lead to more effective medical treatment for people with diseases that are accompanied by memory disturbance.

"Until now, searching for genes related to memory capacity has been comparable to seeking out

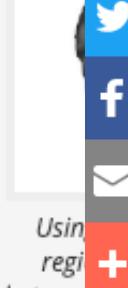
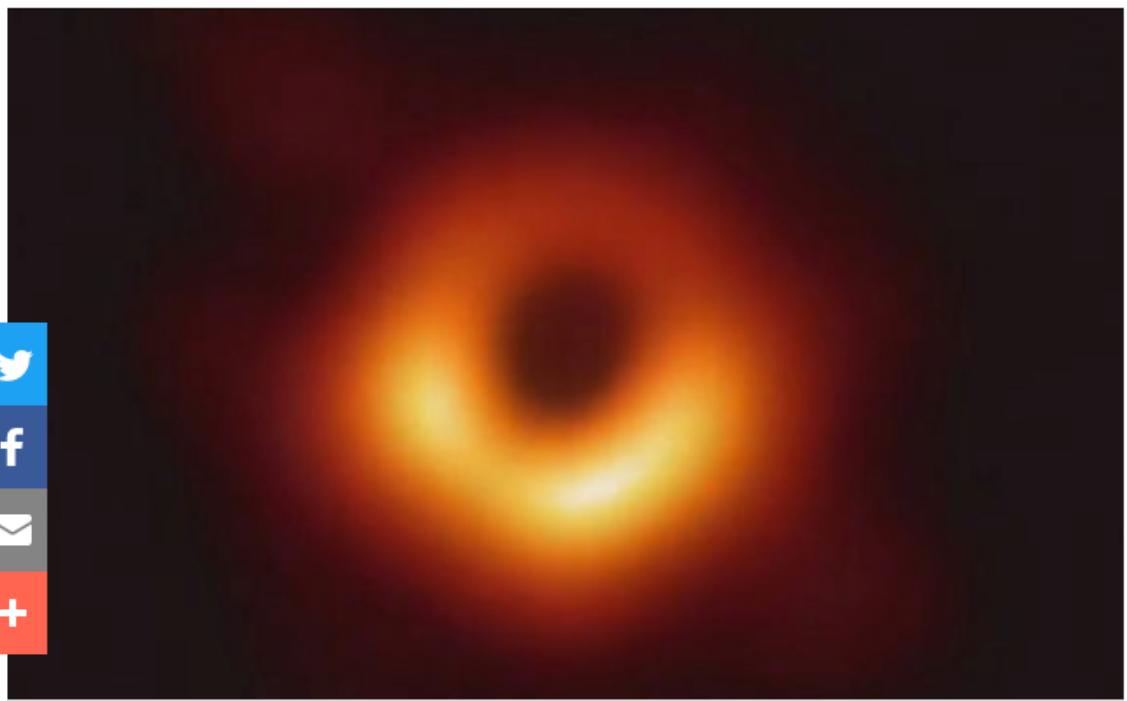
JULY 19, 2017

TEACHABLE MOMENTS

APRIL 19, 2019

Scientific Simulation How Scientists Captured the First Image of a Black Hole

By [Ota Lutz](#)



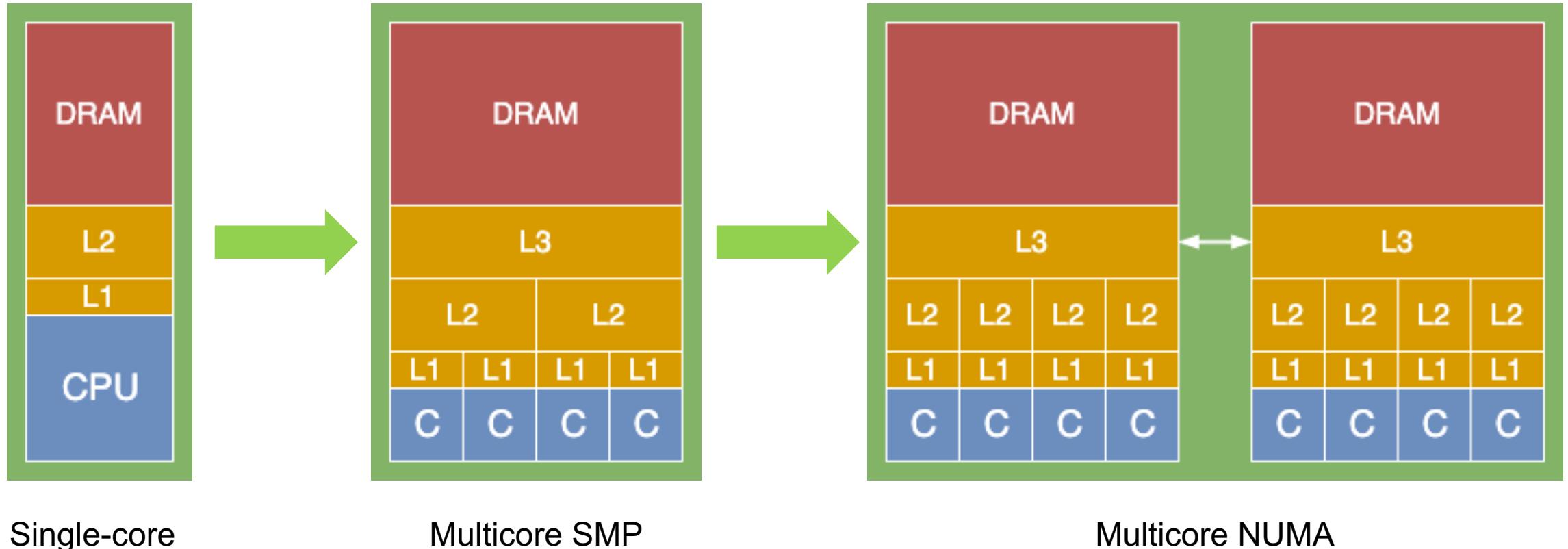
Using
regional
between
colors,
(II)

Scientists have obtained the first image of a black hole, using Event Horizon Telescope observations of the center of the galaxy M87. The image shows a bright ring formed as light bends in the intense gravity around a black hole that is 6.5 billion times more massive than the Sun. Image credit: Event Horizon Telescope Collaboration +

Why HPC?

- Complex workloads
 - Computationally intensive algorithms
 - Latency-sensitive, high communication needs
 - Heavy post-processing of data
 - Machine learning and AI
 - Demanding visualization processes
- Huge amounts of data
 - Efficient stage-in and stage-out of data
 - Checkpointing
 - Parallel reading and writing to filesystem at high speeds
- Sophisticated solutions are required; No. 1 requirement is **high performance**
 - Processors and memory subsystem
 - Interconnection networks and communication protocols
 - Storage and filesystems
 - Libraries, Software, Applications

Building blocks for HPC systems: the CPU



Single-core

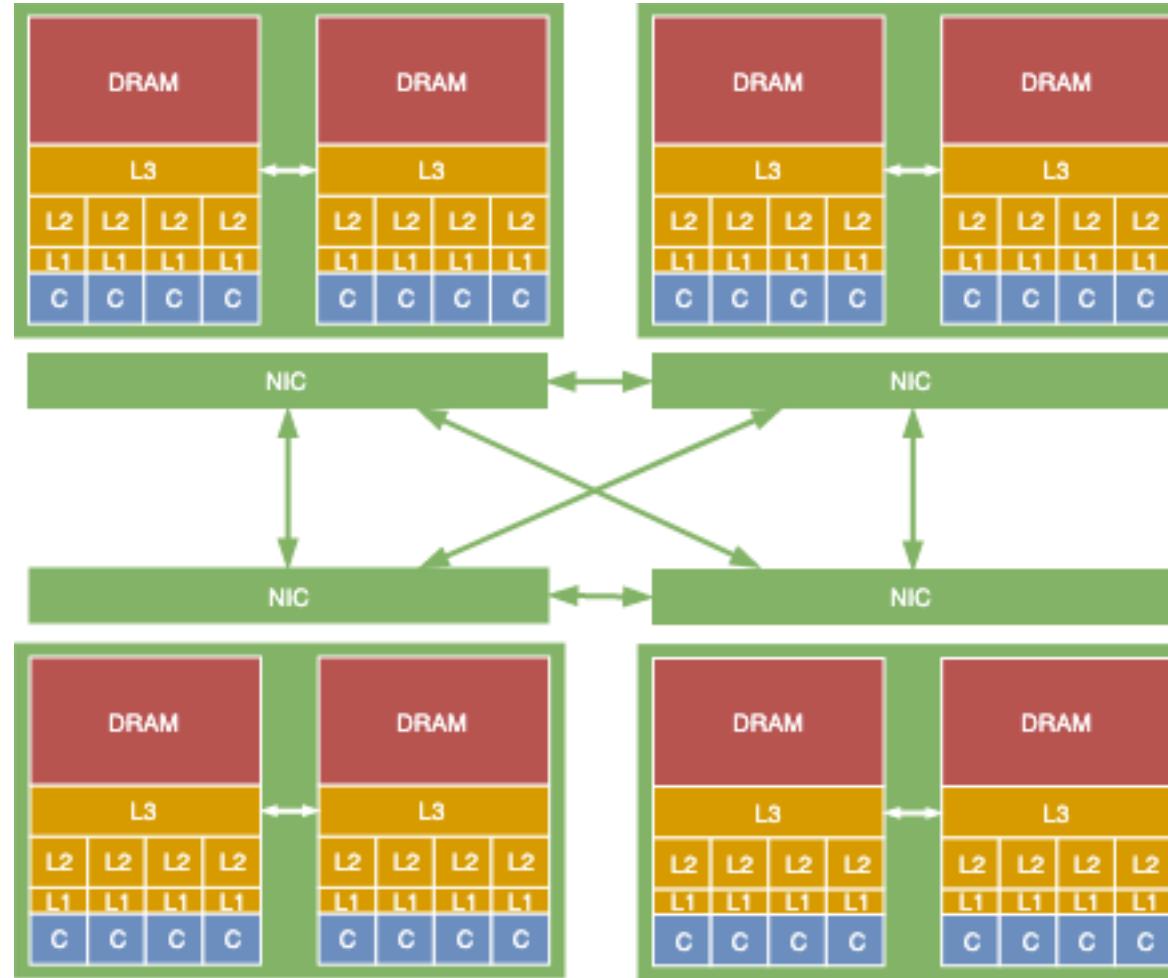
Multicore SMP

Multicore NUMA

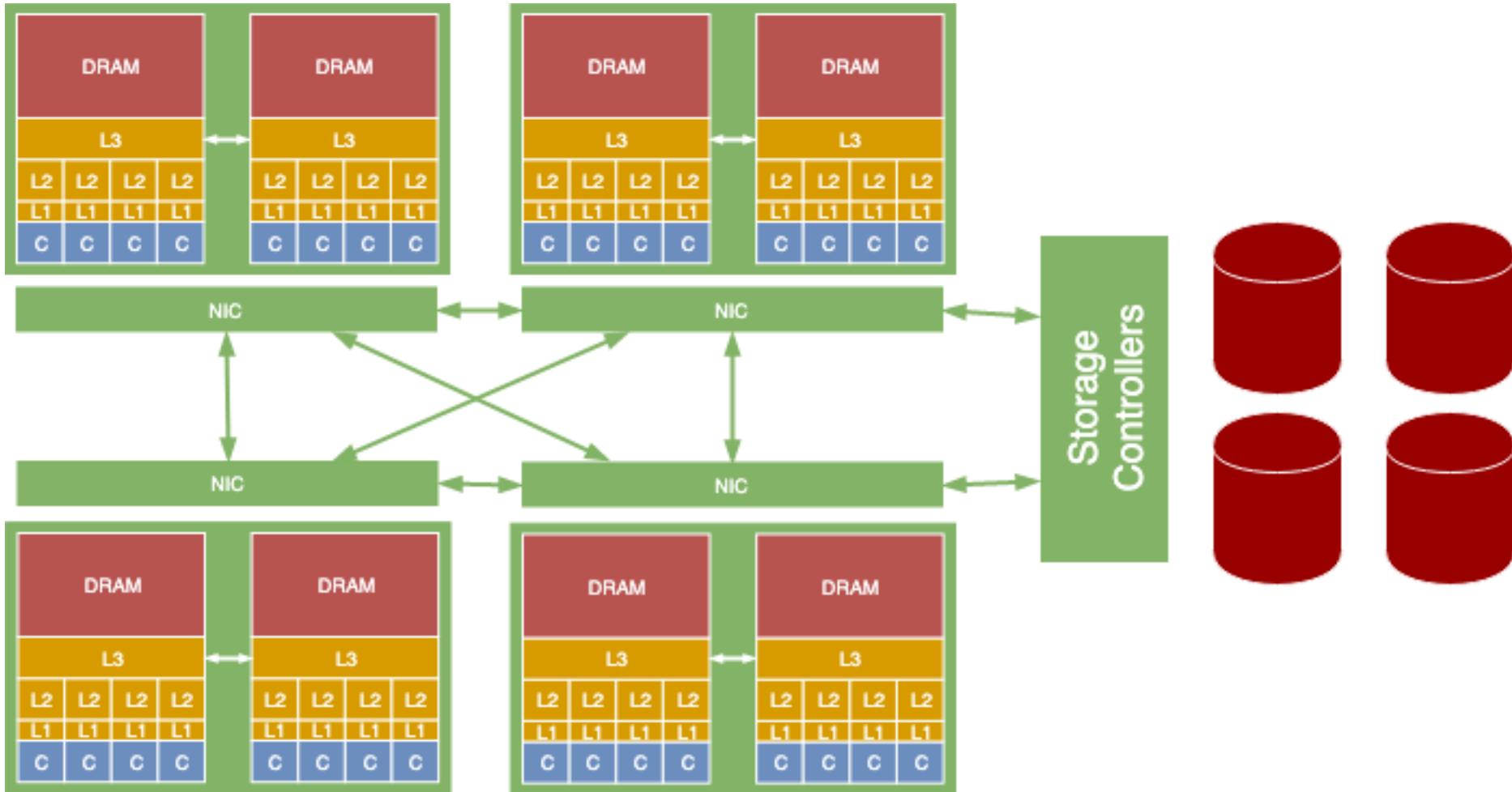
SMP: Symmetric Multi-Processor

NUMA: Non-Uniform Memory Access

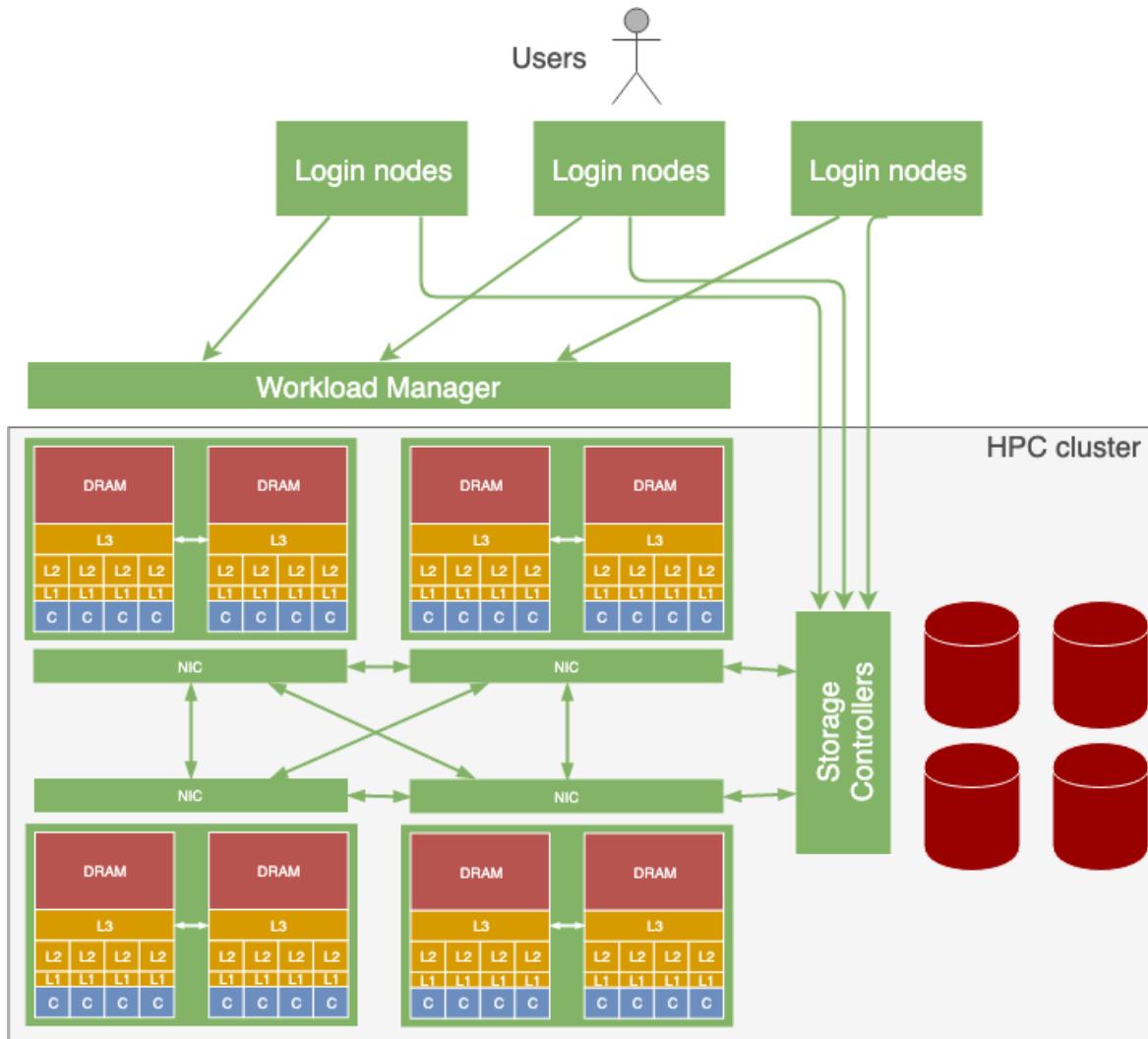
Building blocks for HPC systems: the network



Building blocks for HPC systems: the storage



Building blocks for HPC systems: login nodes & workload manager

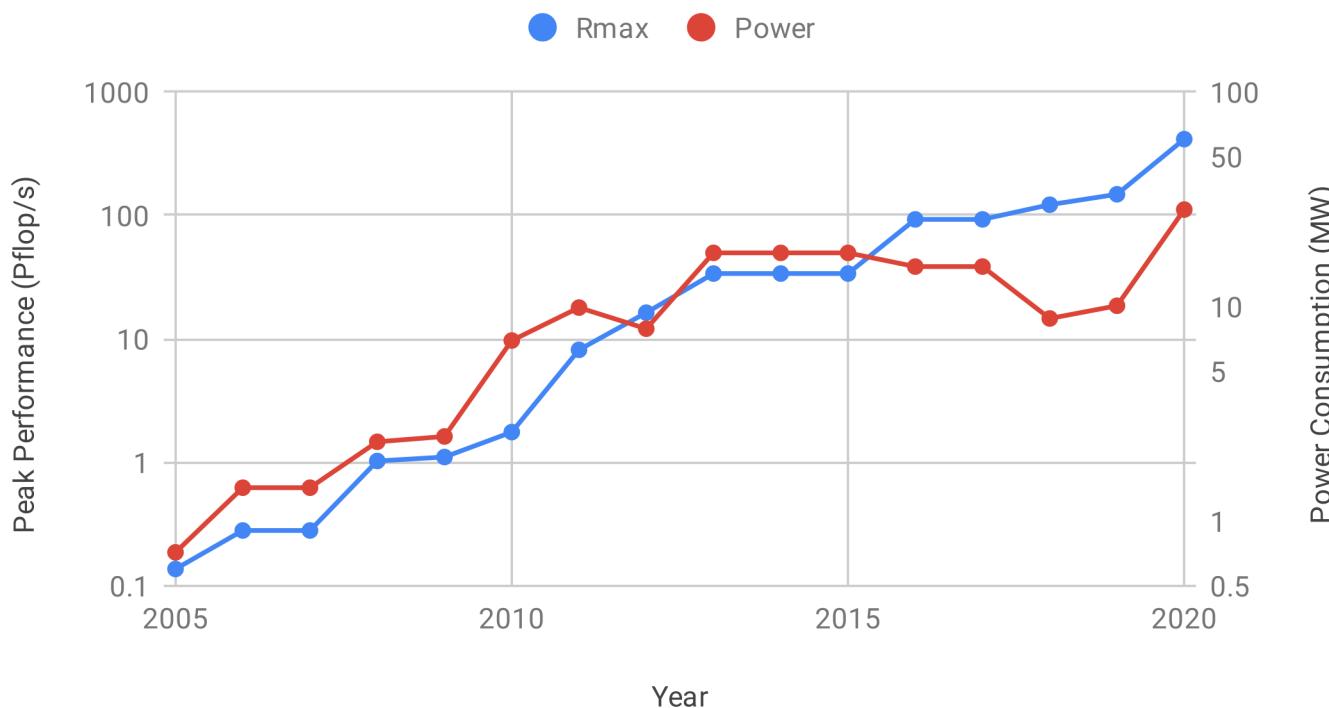


Issues and limitations

- HPC systems are expensive!
 - Power costs
 - Cooling and infrastructural costs
 - Technology costs
 - High-end processors
 - Fast, low-latency networks
 - Fast storage
 - ...

Performance and power consumption evolution

Top500 list #1 system



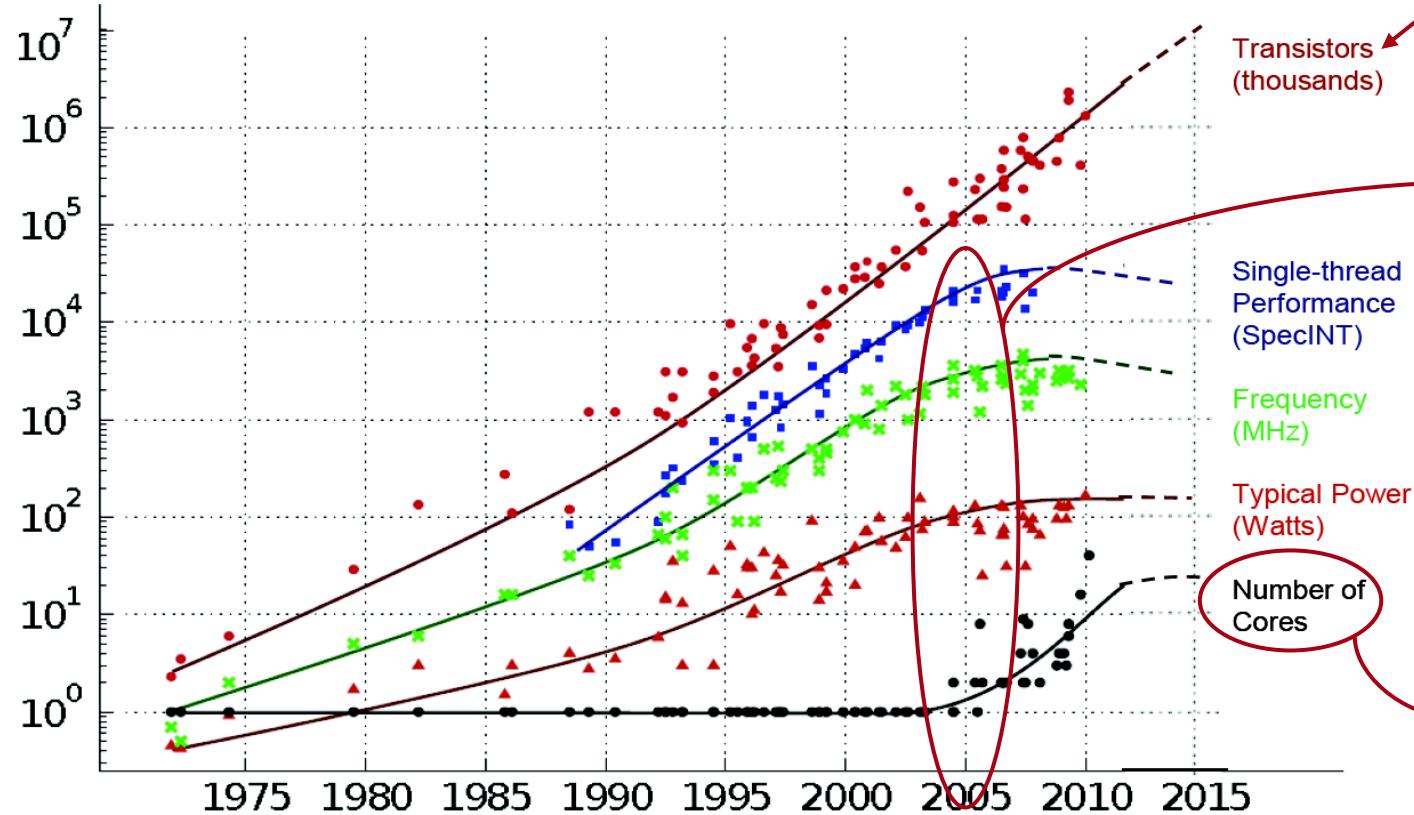
Power consumption has been following closely the exponential growth in performance, but at a lower rate.

Exascale is now “ante portas” !

- EU's EuroHPC
- DoE's ECP
- Japan's Post-K
- China's ??

How did we reach here?

35 YEARS OF MICROPROCESSOR TREND DATA



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!

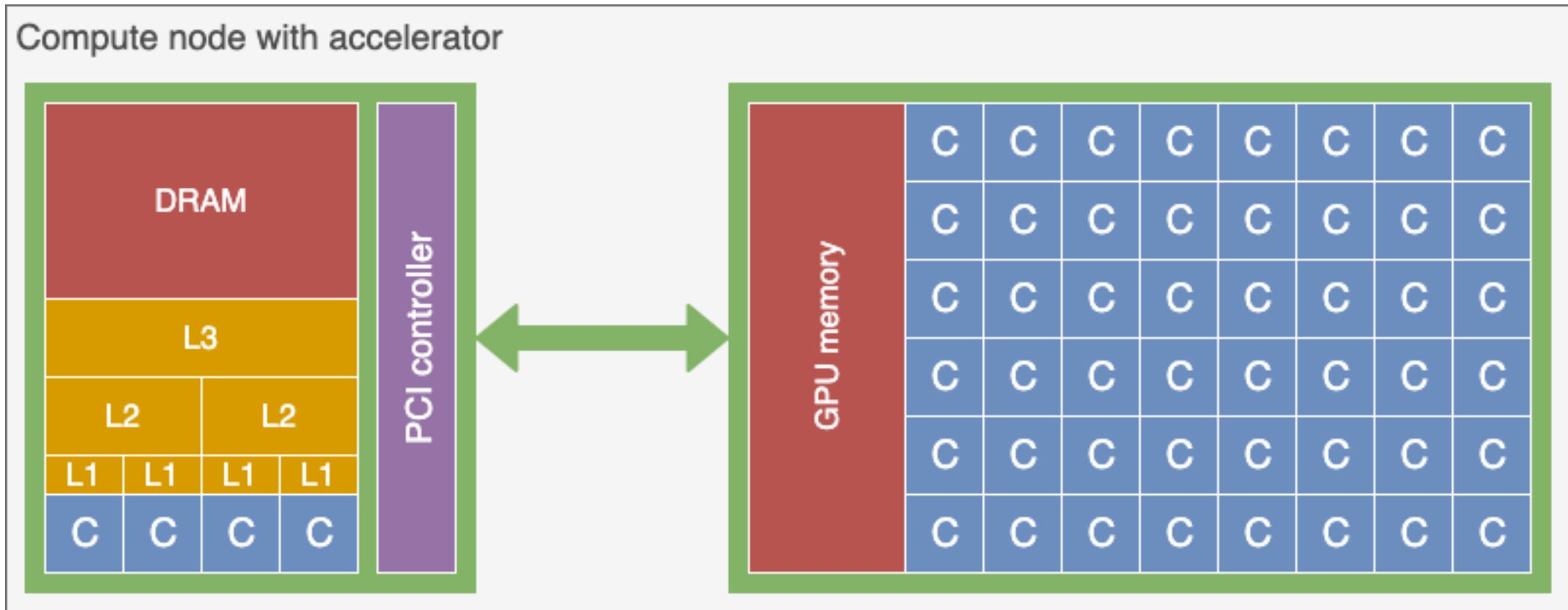
Performance can only grow through node-level parallelism!

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

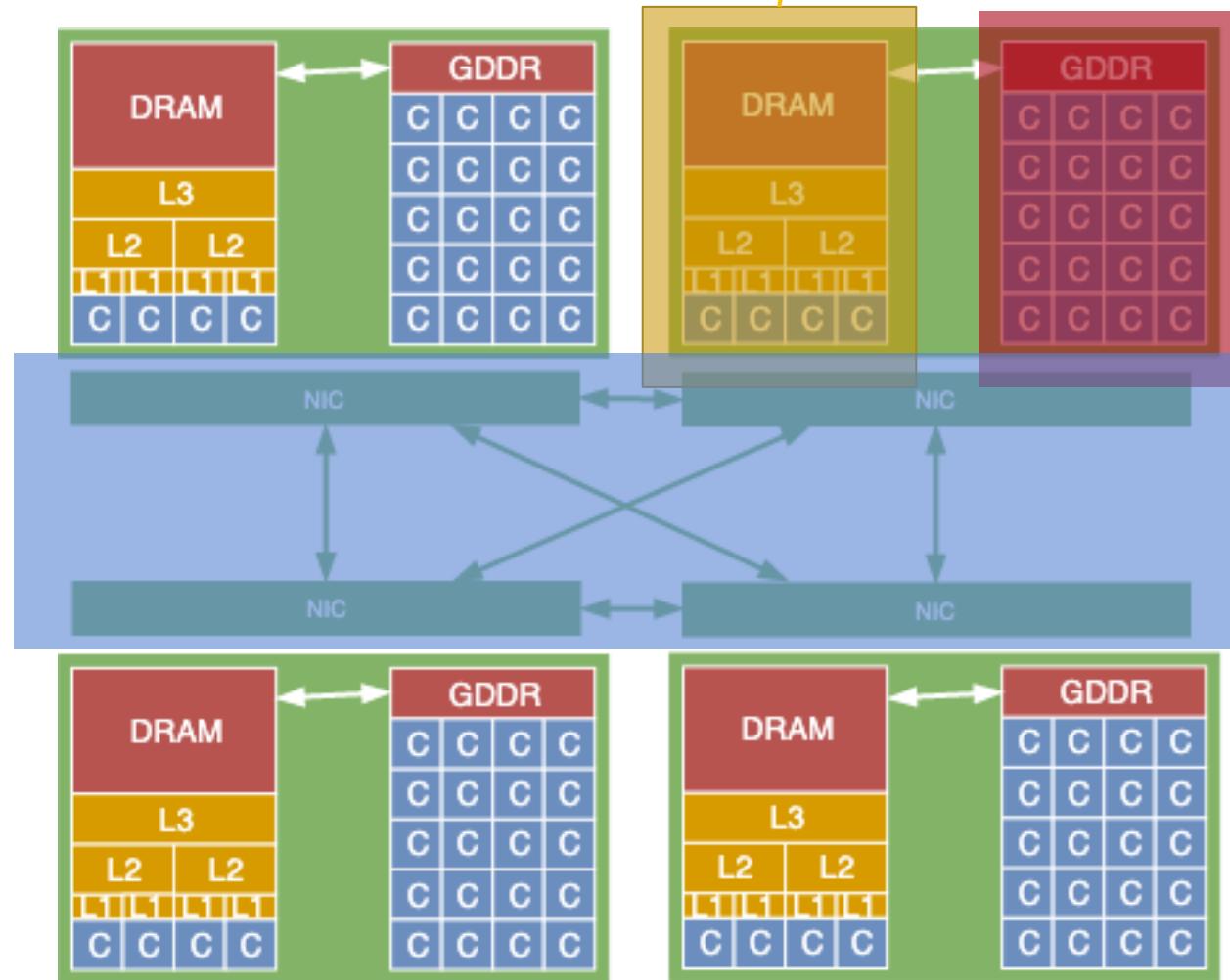
Beyond multicores

- 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

Accelerators in a HPC system



How do we program for HPC?



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

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

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

Piz Daint

- Cray XC40/XC50 system
 - Top500: #10 in the world, #3 in Europe
- 5320 XC50 nodes
 - 1x 12-core Haswell (64 GB DRAM) + 1x Nvidia Tesla P100 (Pascal) GPU (16 GB HBM2)
- 1813 XC40 nodes
 - 2x 18-core Broadwell (64/128 GB DRAM)
- Dragonfly network + Aries routing
- Filesystems
 - 6.2 PB Lustre filesystem for scratch data
 - GPFS for users home and long-term data



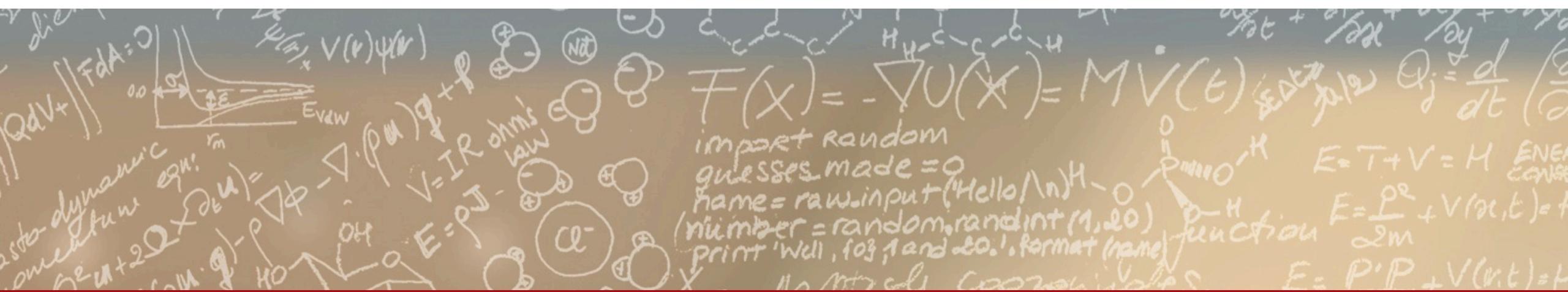
Energy efficiency of accelerators and specialized processors

Rank	Top500 rank	System	Rmax (Tflop/s)	Power (kW)	Efficiency (Gflops/W)
1	393	MN-3 - MN-Core Server, Xeon 8260M 24C 2.4GHz, MN-Core, RoCEv2/MN-Core DirectConnect, Preferred Networks	1,621.1	77	21.108
2	7	Selene - DGX A100 SuperPOD, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Mellanox HDR Infiniband, Nvidia	27,580	1,344	20.518
3	468	NA-1 - ZettaScaler-2.2, Xeon D-1571 16C 1.3GHz, Infiniband EDR, PEZY-SC2 700Mhz, PEZY Computing / Exascaler Inc.	1,303.2	80	18.433
4	204	A64FX prototype - Fujitsu A64FX, Fujitsu A64FX 48C 2GHz, Tofu interconnect D, Fujitsu	1,999.5	118	16.876
5	26	AiMOS - IBM Power System AC922, IBM POWER9 20C 3.45GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM	8,339.0	512	16.285

Green500 list, June 2020

Summary

- HPC has an important societal impact
- Very high complexity at all levels of integration; from the infrastructure up to the software stack
- Learning how to efficiently use and program such a system can open new horizons to research



Q & A