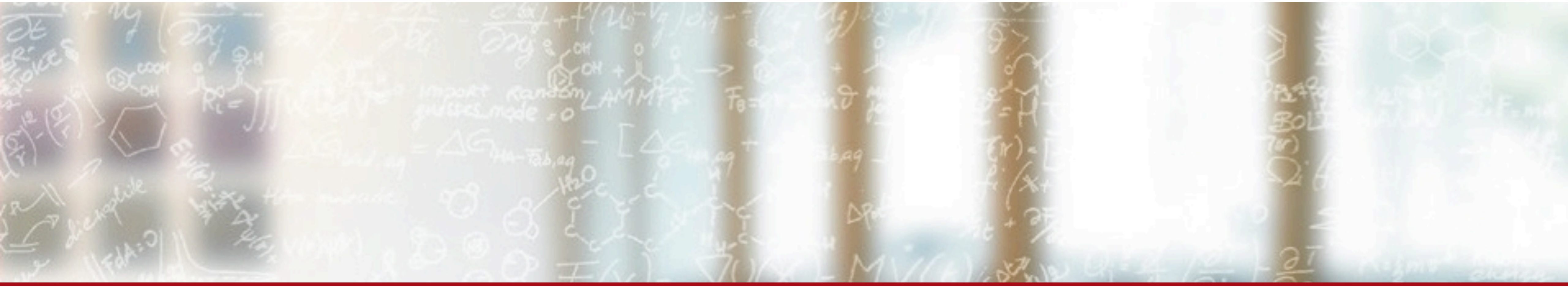
CSCS
Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

ETH zürich



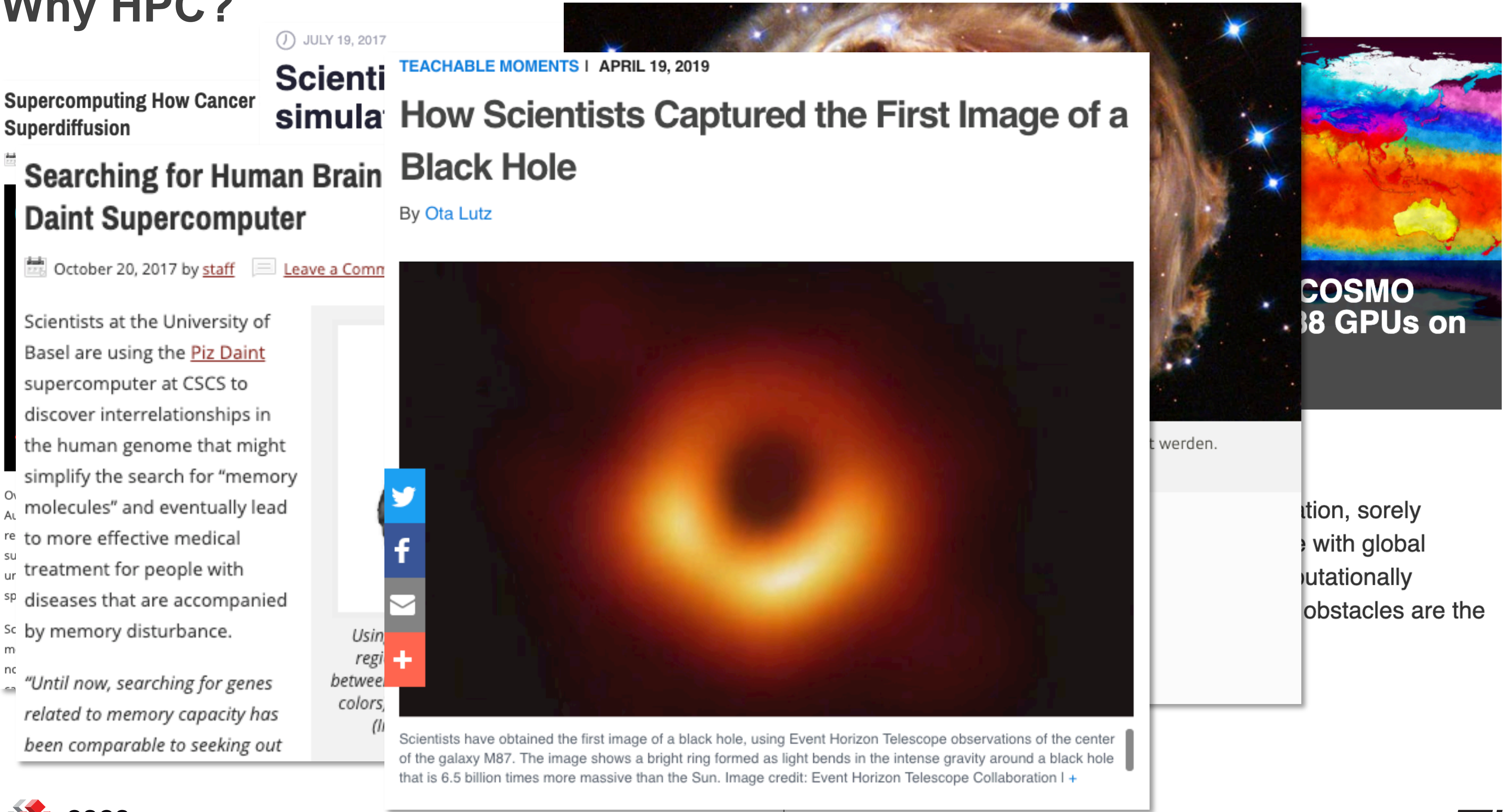
Introduction to High Performance Computing

CSCS-USI Summer School 2019

Vasileios Karakasis, CSCS

July 15, 2019

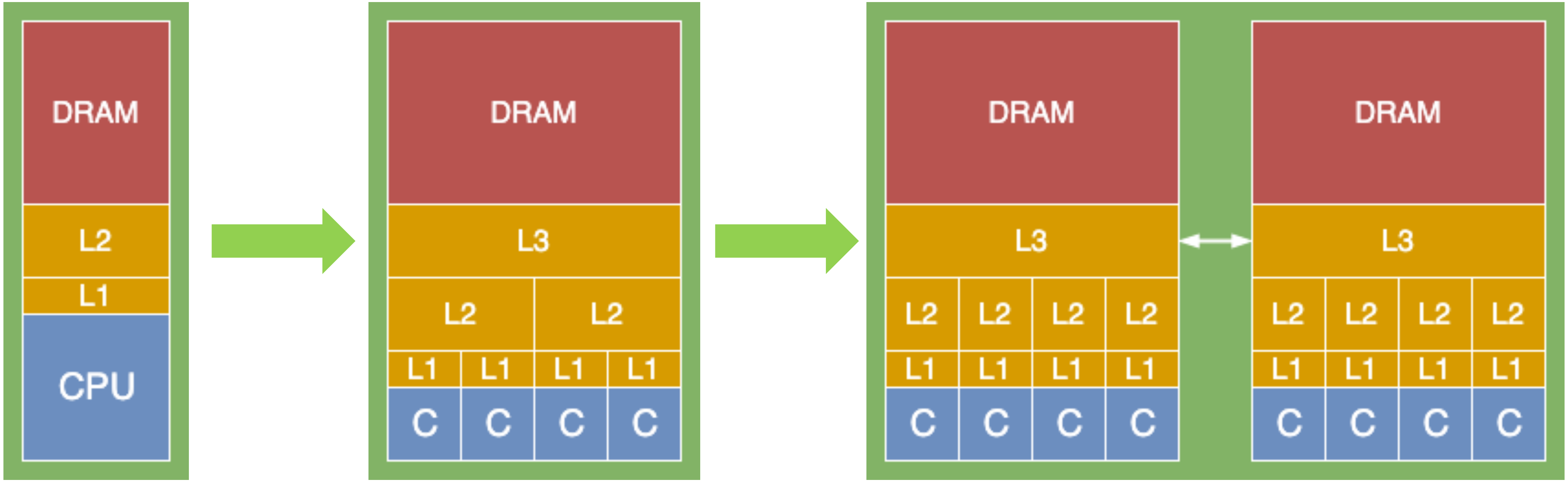
Why HPC?


JULY 19, 2017
Scienti
simula
Supercomputing How Cancer Superdiffusion
Searching for Human Brain
Daint Supercomputer
October 20, 2017 by staff
Leave a Comm
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
TEACHABLE MOMENTS | APRIL 19, 2019
How Scientists Captured the First Image of a Black Hole
By Ota Lutz
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 | +
t werden.
COSMO
8 GPUs on
tion, sorely
with global
utationally
obstacles are the

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



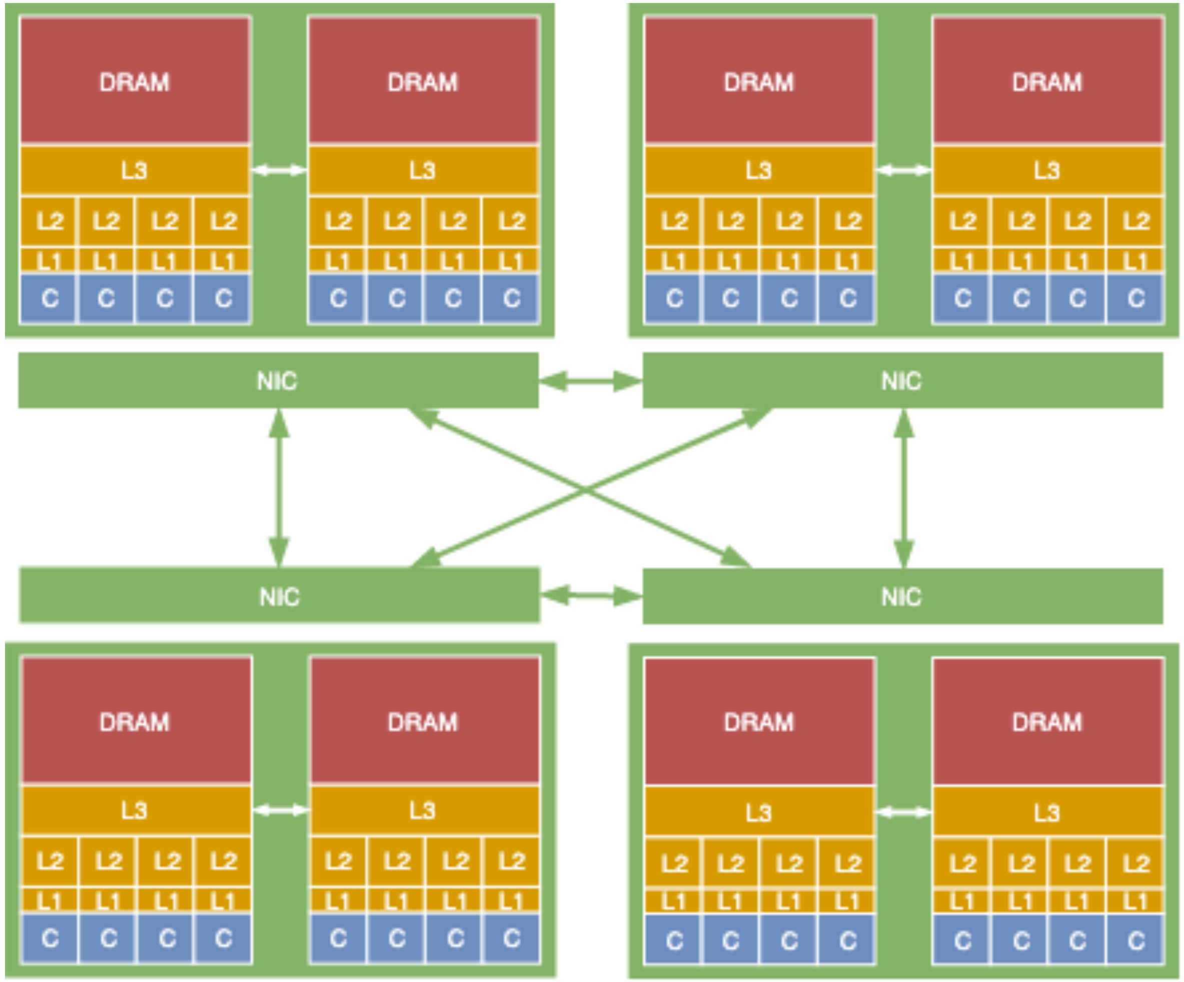
Single-core

Multicore SMP

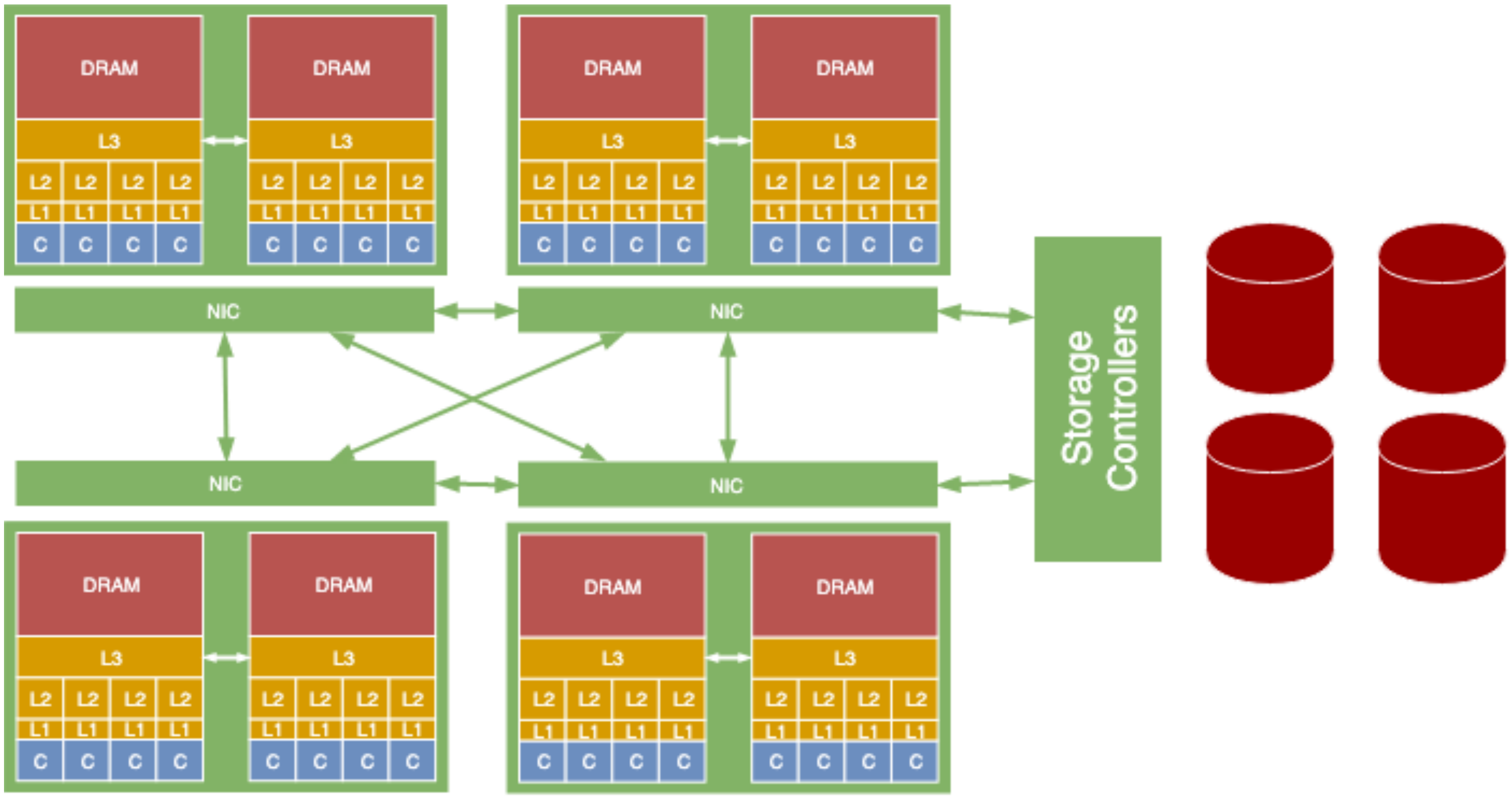
Multicore NUMA

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

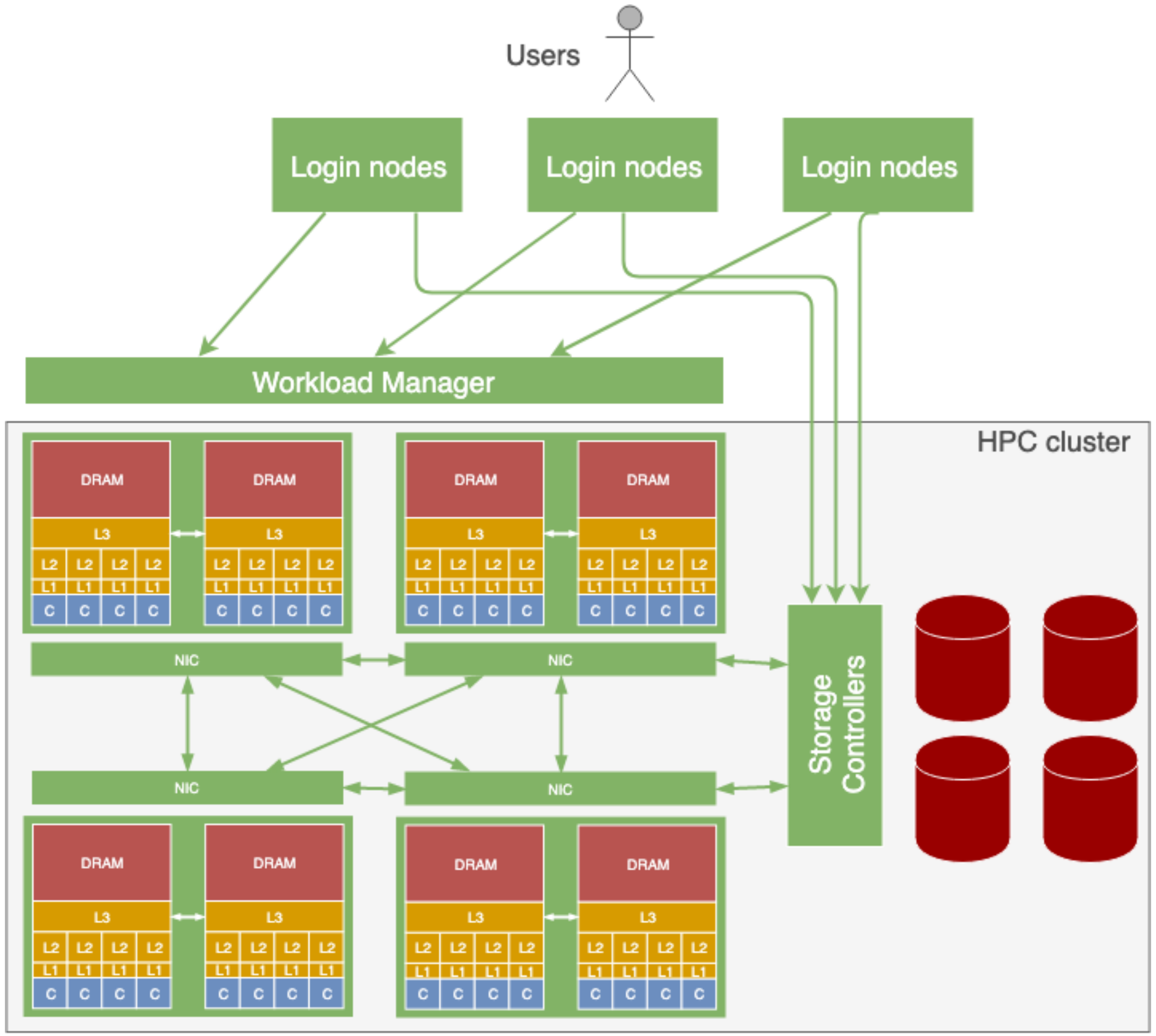
Building blocks for HPC systems: the network



Building blocks for HPC systems: the storage



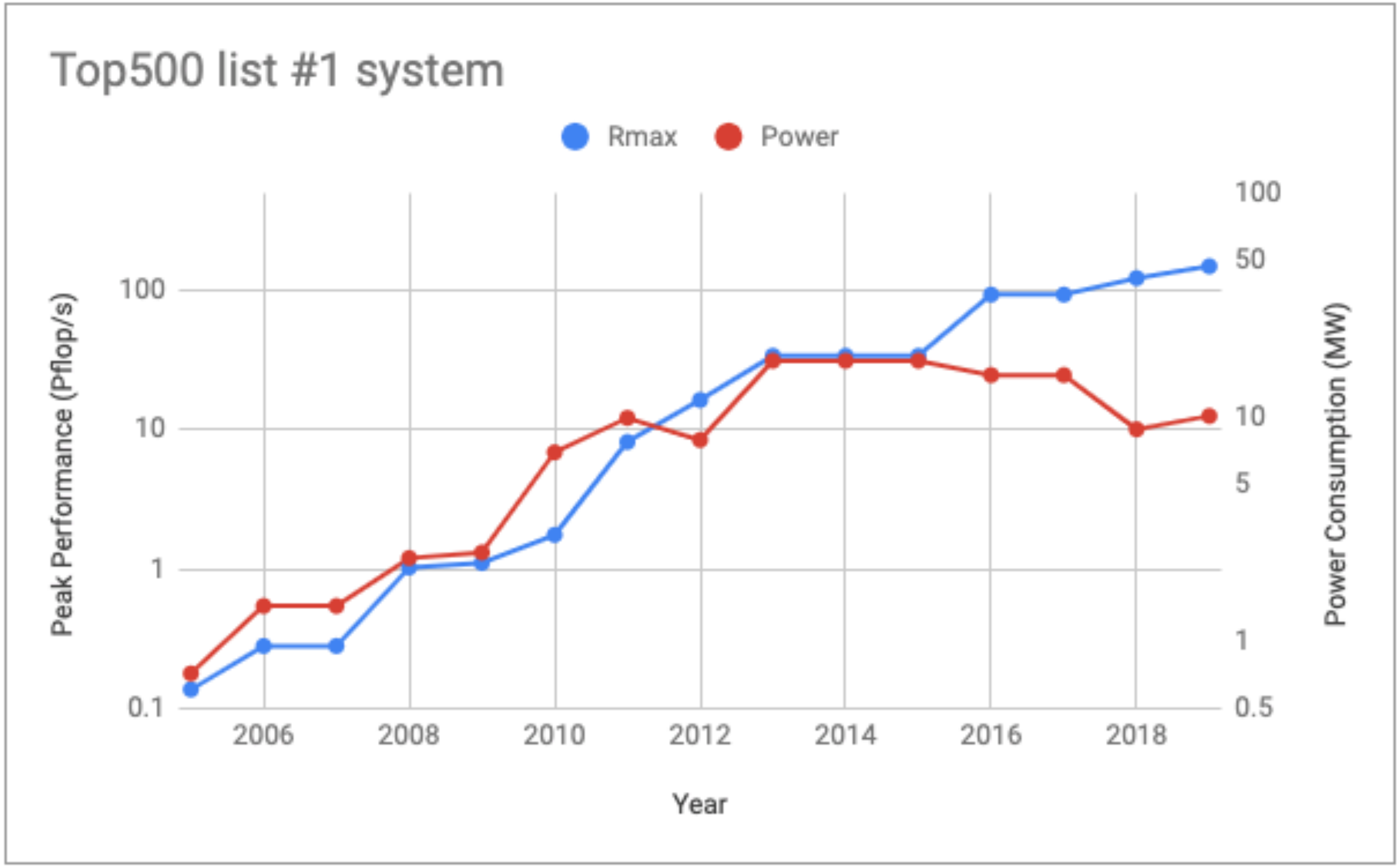
Building blocks for HPC systems: login nodes & workload manager



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



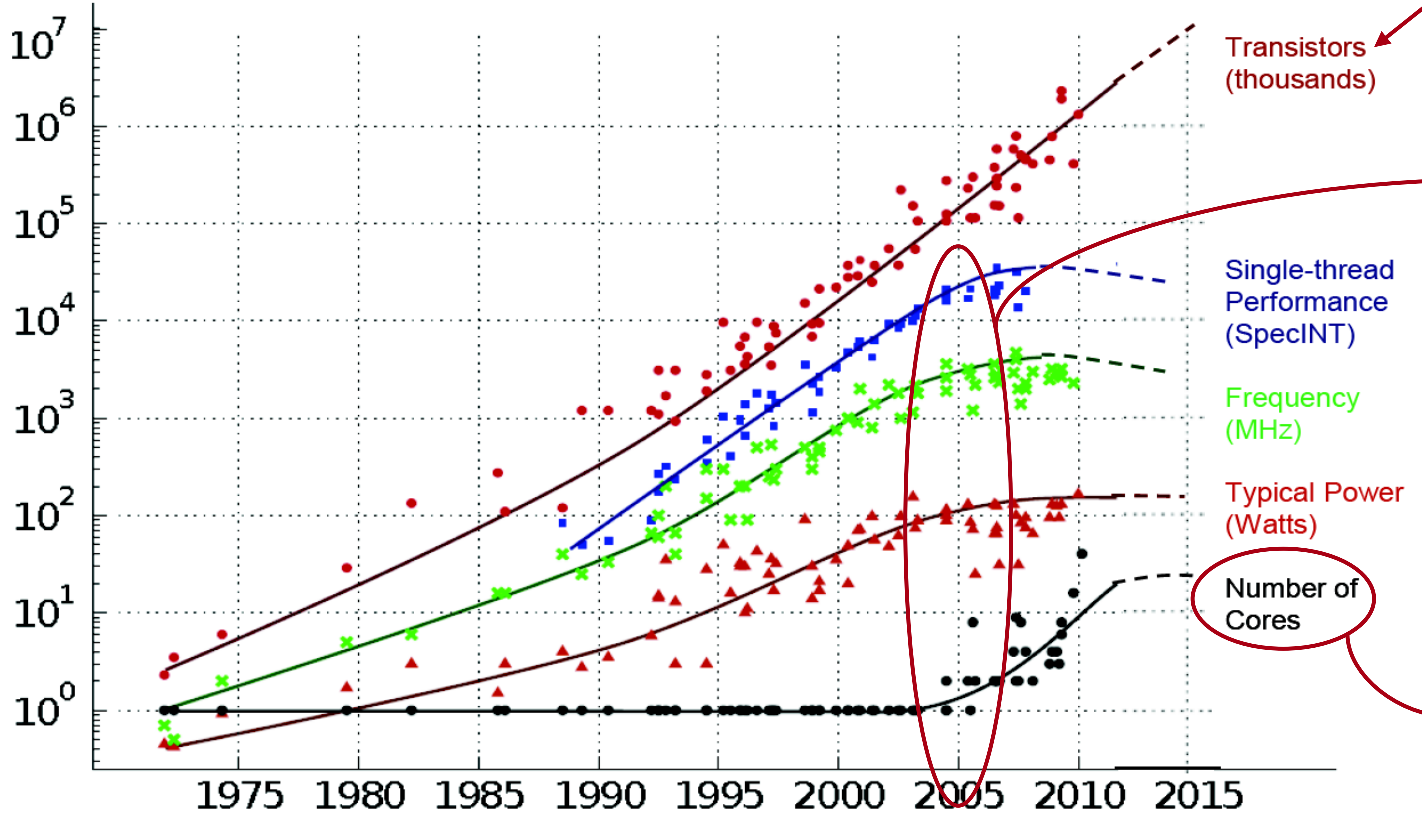
Power consumption has been following closely the exponential growth in performance, but at a lower rate.

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



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!

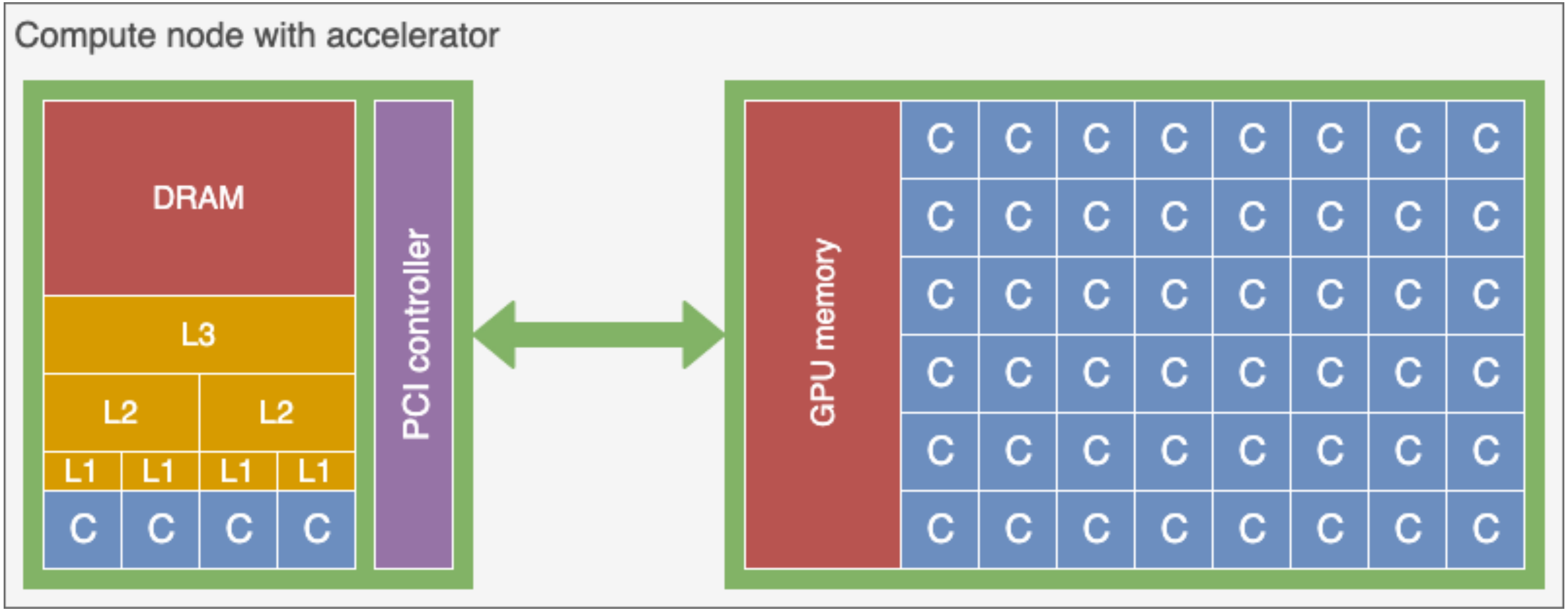
Performance can only grow through node-level parallelism!

Beyond multicores

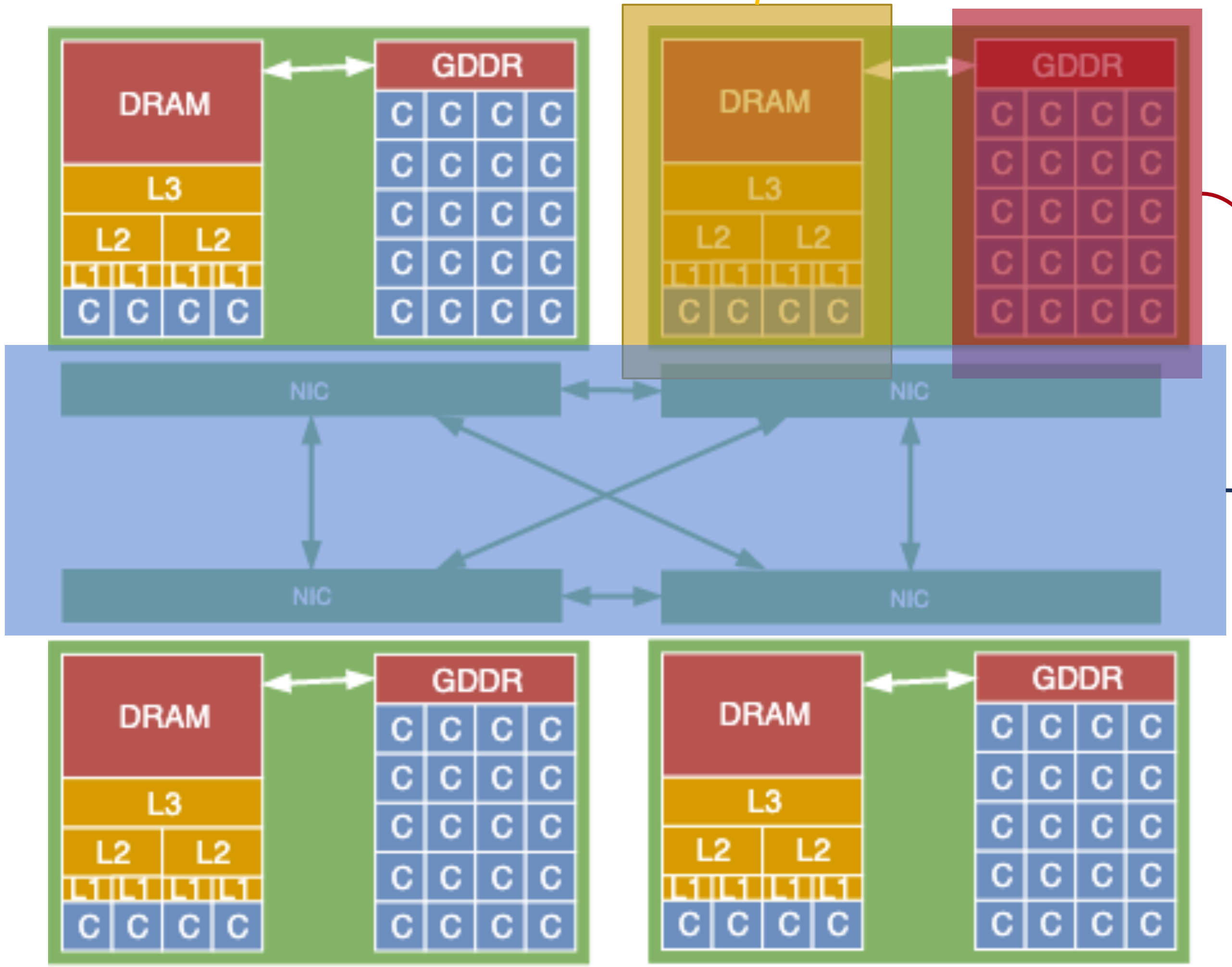
- Multicores have limitations
  - Fat cores (branch prediction, out-of-order execution, large caches)
    - Optimized for latency and multiprocessing
  - Still high frequencies
  - Still high power consumption
  - But programming is easy; matches better our brain's serial way of thinking

- 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: #6 in the world, #1 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

Rank	Top500 rank	System	Rmax (Tflop/s)	Power (kW)	Efficiency (Gflops/W)
1	472	**Shoubu system B** - ZettaScaler-2.2, Xeon D-1571 16C 1.3GHz, Infiniband EDR, PEZY-SC2	1,063.3	60	17.604
2	470	**DGX SaturnV Volta** - NVIDIA DGX-1 Volta36, Xeon E5-2698v4 20C 2.2GHz, Infiniband EDR, NVIDIA Tesla V100	1,070.0	97	15.113
3	1	**Summit** - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband	148,600.0	10,096	14.719
4	8	**AI Bridging Cloud Infrastructure (ABCI)** - PRIMERGY CX2570 M4, Xeon Gold 6148 20C 2.4GHz, NVIDIA Tesla V100 SXM2, Infiniband EDR	19,880.0	1,649	14.423
5	394	**MareNostrum P9 CTE** - IBM Power System AC922, IBM POWER9 22C 3.1GHz, Dual-rail Mellanox EDR Infiniband, NVIDIA Tesla V100	1,145.0	81	14.131

Green500 list, June 2019

Summary

- HPC has an important societal impact
- Very high complexity in 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