Design of Parallel and High-Performance Computing

Fall 2017

Lecture: Introduction

Instructor: Torsten Hoefler & Markus Püschel

TA: Salvatore Di Girolamo

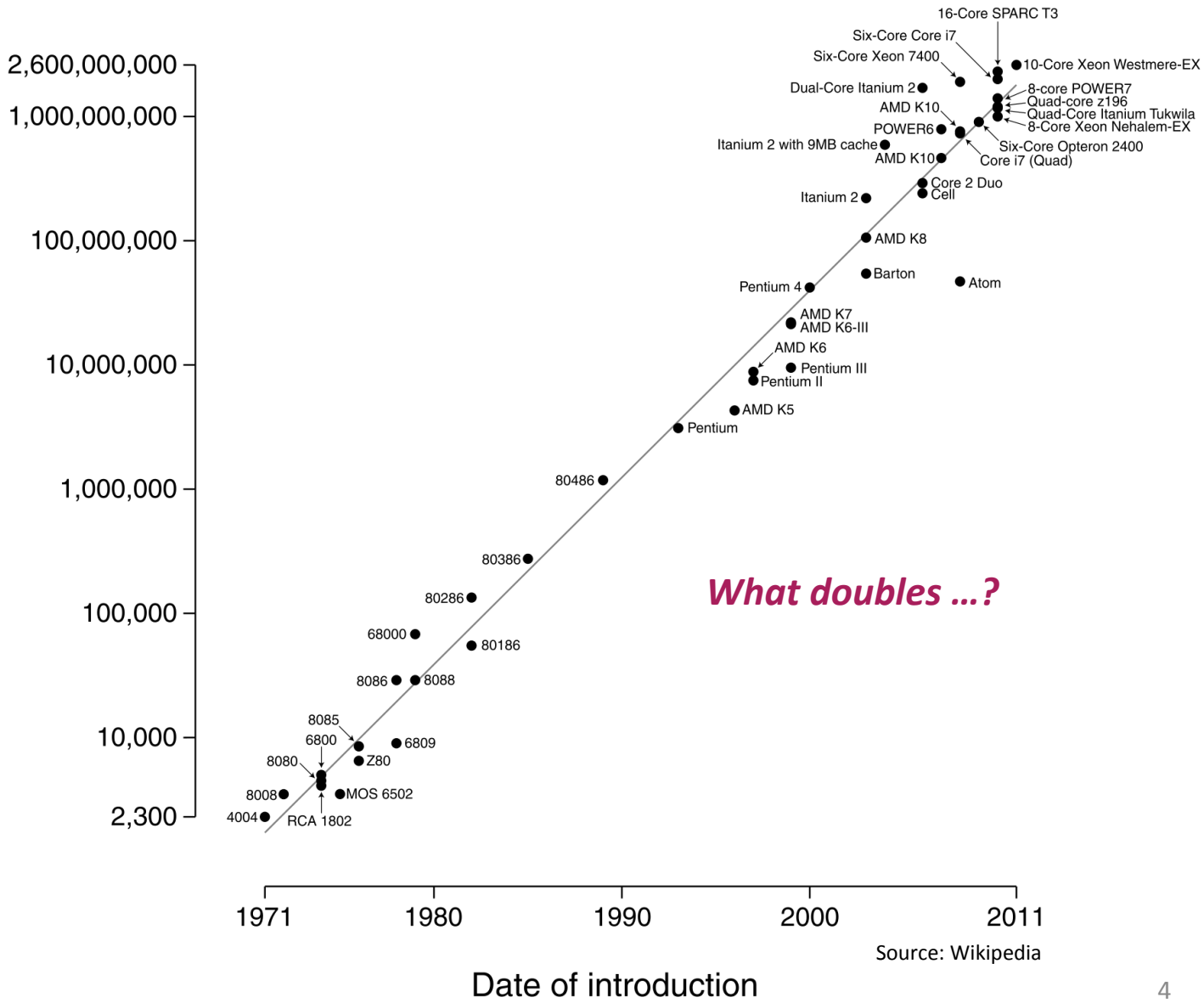
ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Goals of this lecture

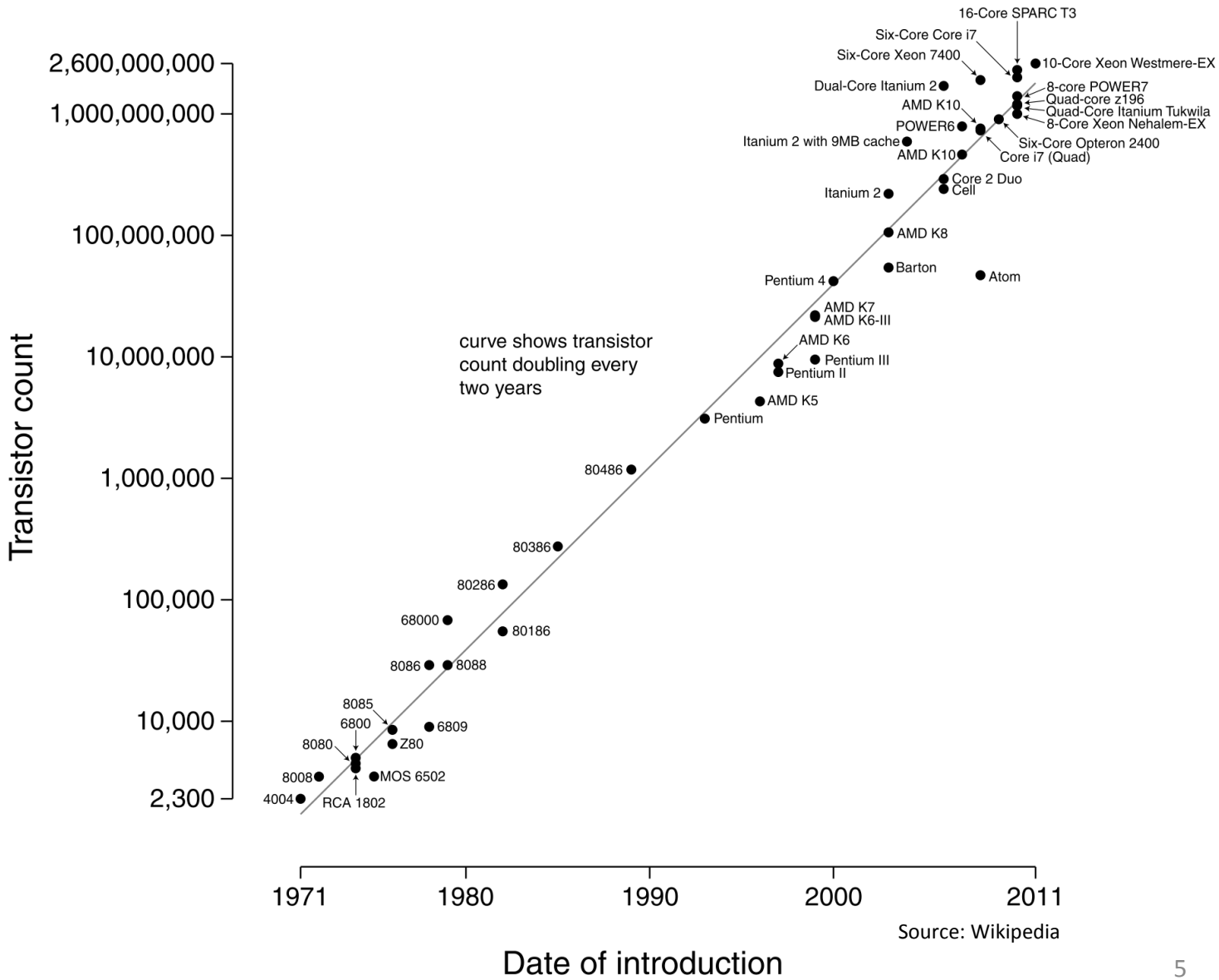
- **Motivate you!**
- **Trends**
- **High performance computing**
- **Programming models**
- **Course overview**

Trends



Source: Wikipedia

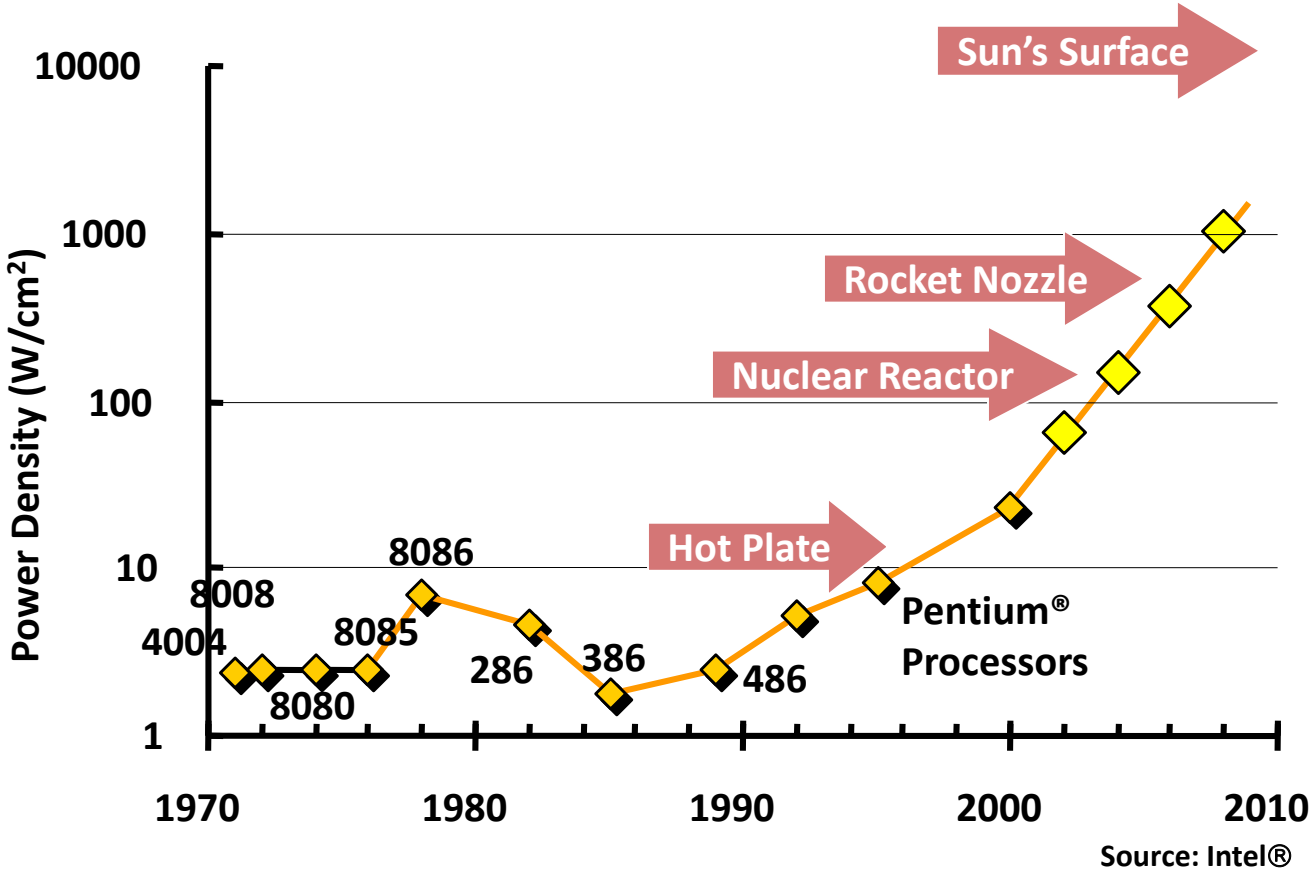
Microprocessor Transistor Counts 1971-2011 & Moore's Law



Source: Wikipedia

How to increase the compute power?

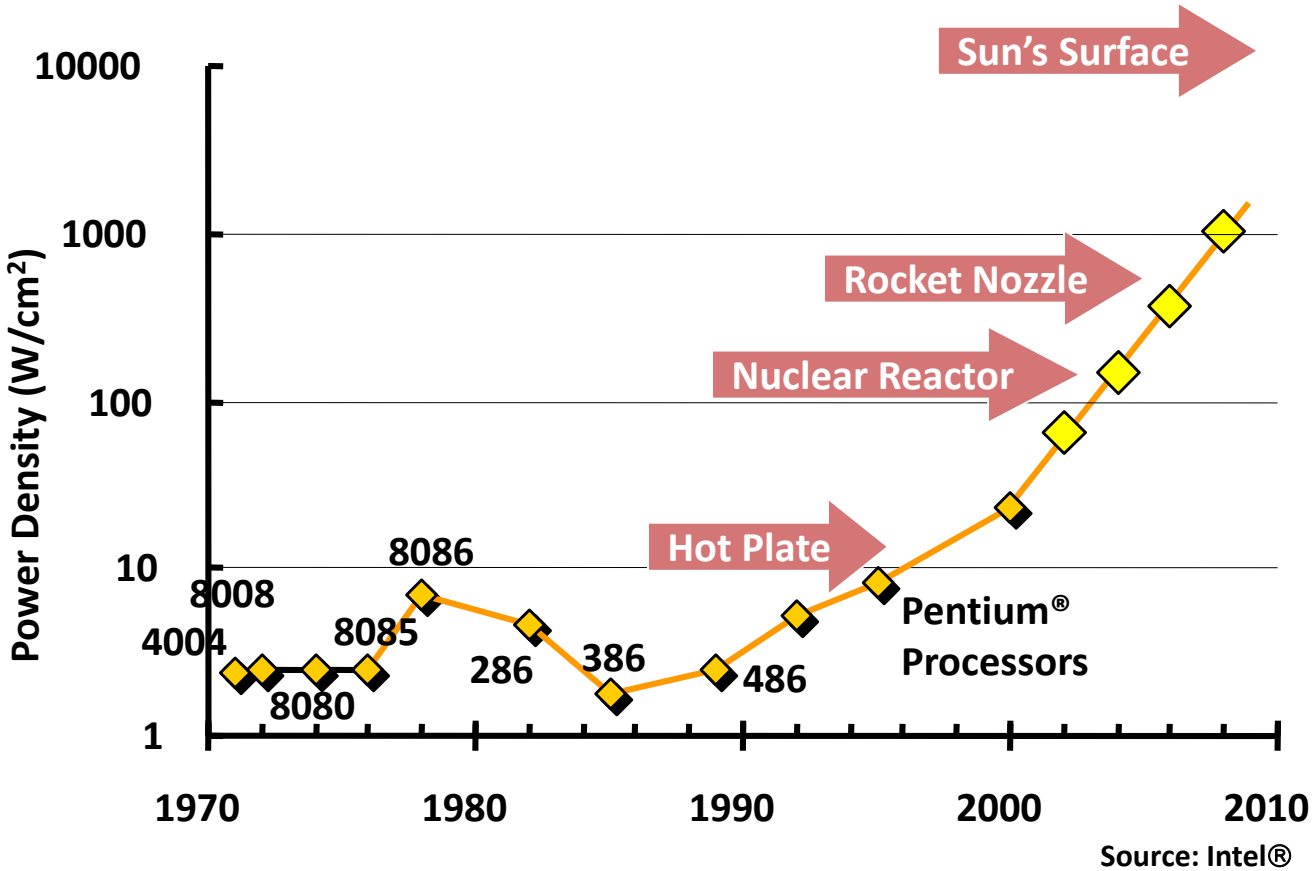
Clock Speed!



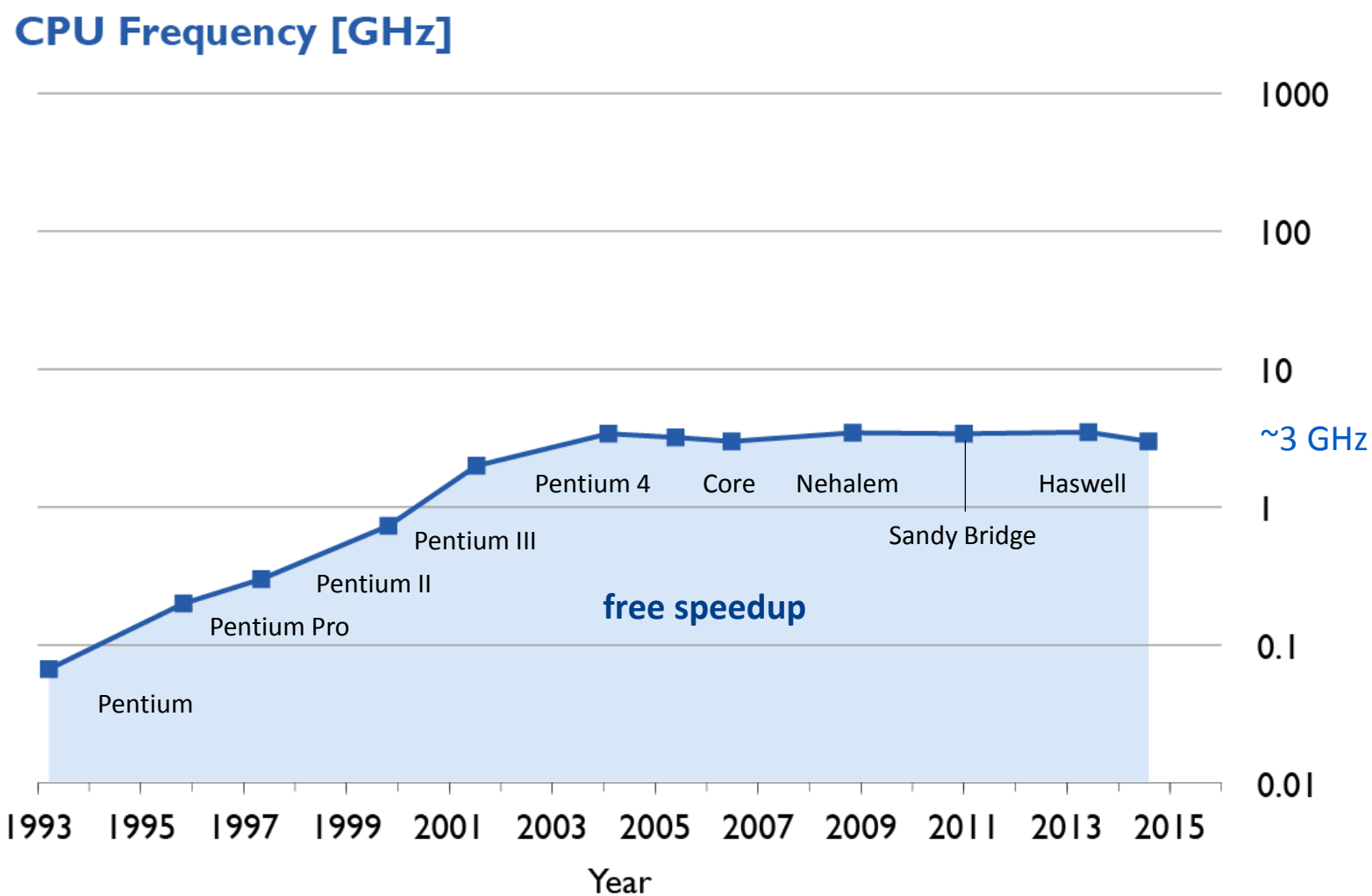
How to increase the compute power?

Not an option anymore!

~~Clock Speed!~~

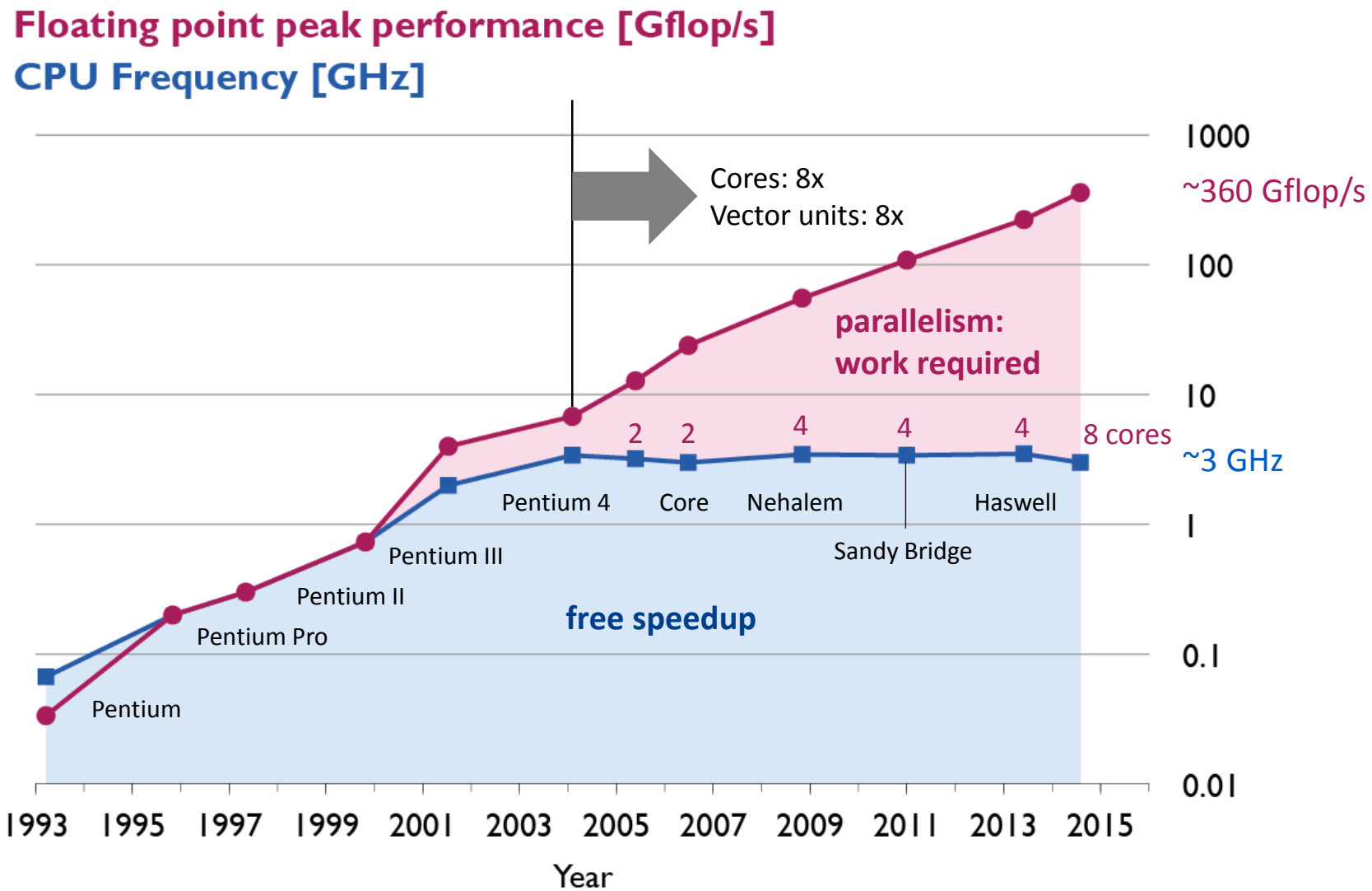


Evolutions of Processors (Intel)



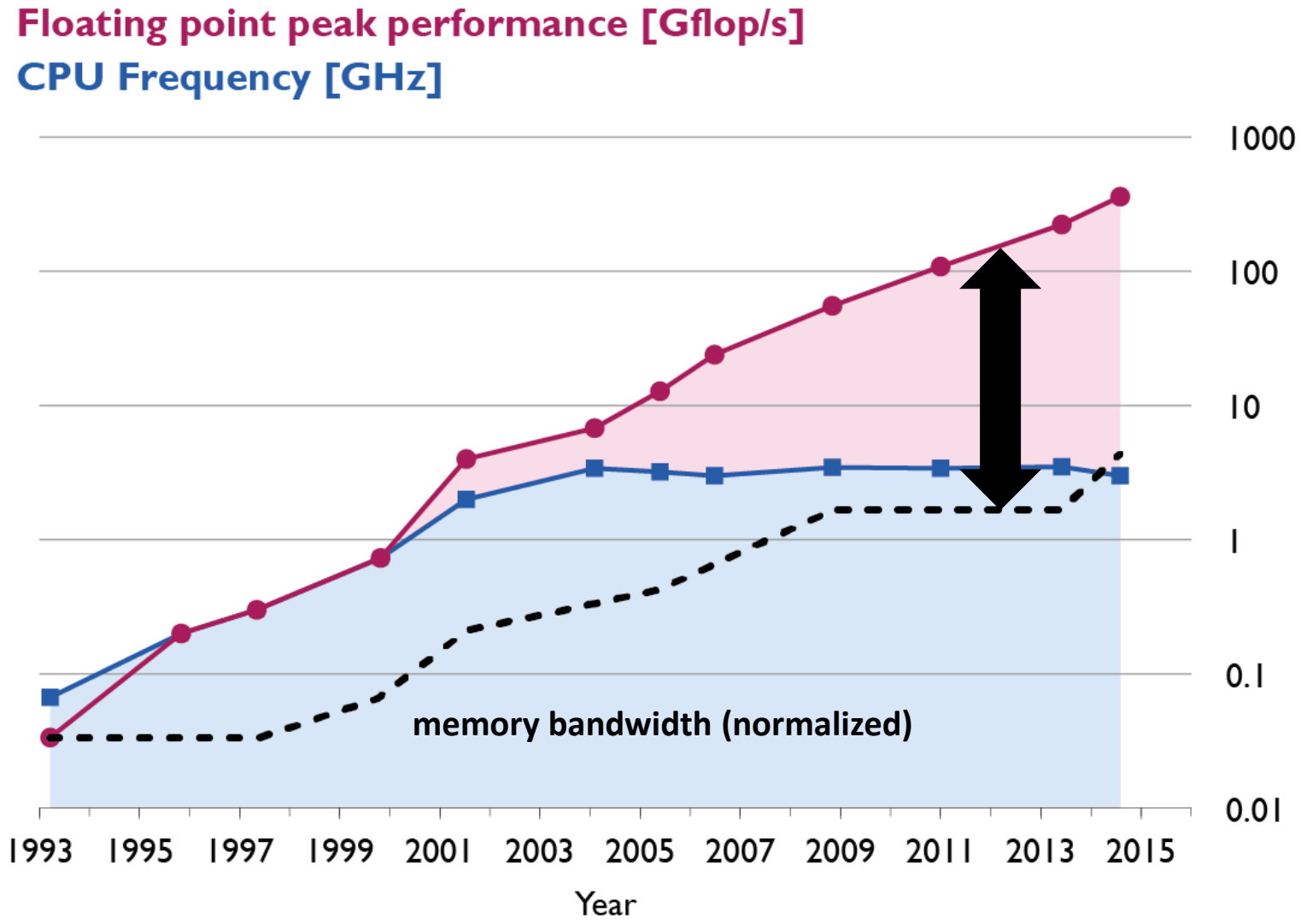
Source: Wikipedia/Intel/PCGuide

Evolutions of Processors (Intel)



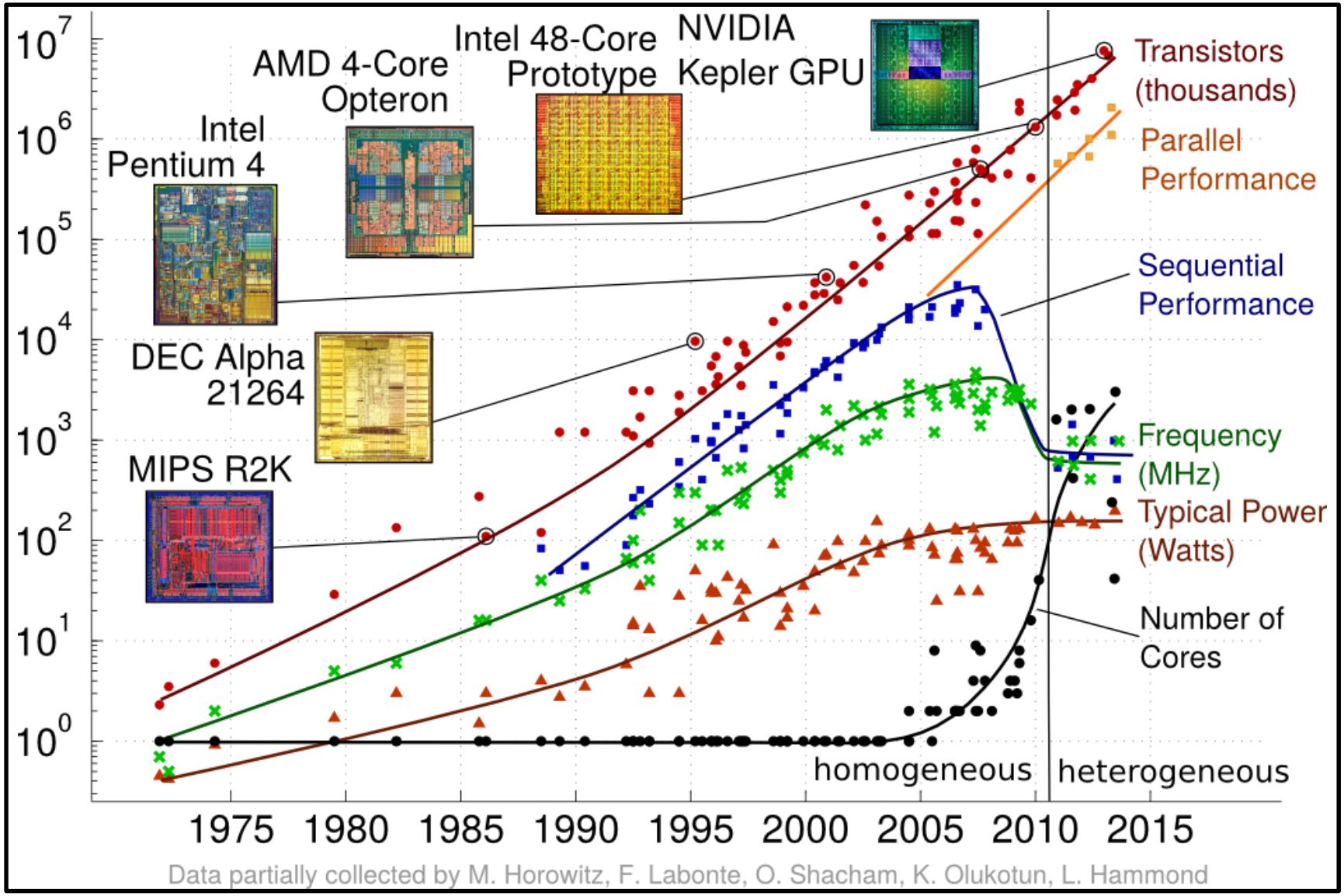
Source: Wikipedia/Intel/PCGuide

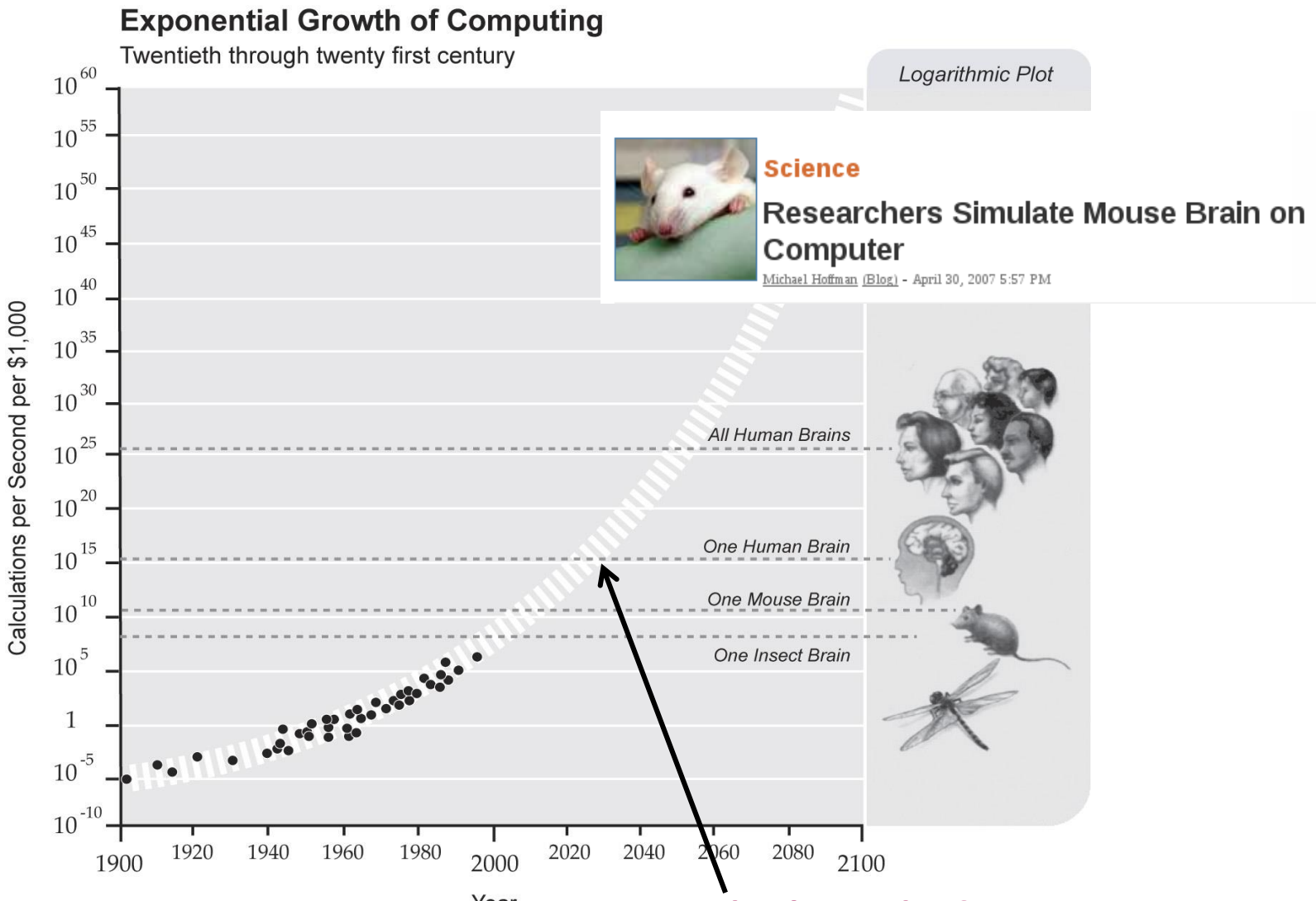
Evolutions of Processors (Intel)



Source: Wikipedia/Intel/PCGuide

A more complete view





Source: www.singularity.com

Can we do this today?

High-Performance Computing

High-Performance Computing (HPC)

- **a.k.a. “Supercomputing”**
- **Question: define “Supercomputer”!**

High-Performance Computing (HPC)

- **a.k.a. "Supercomputing"**
- **Question: define "Supercomputer"!**
  - "A supercomputer is a computer at the frontline of contemporary processing capacity--particularly speed of calculation." (Wikipedia)
  - Usually quite expensive ($s and kWh) and big (space)
- **HPC is a quickly growing niche market**
  - Not all "supercomputers", wide base
  - Important enough for vendors to specialize
  - Very important in research settings (up to 40% of university spending)

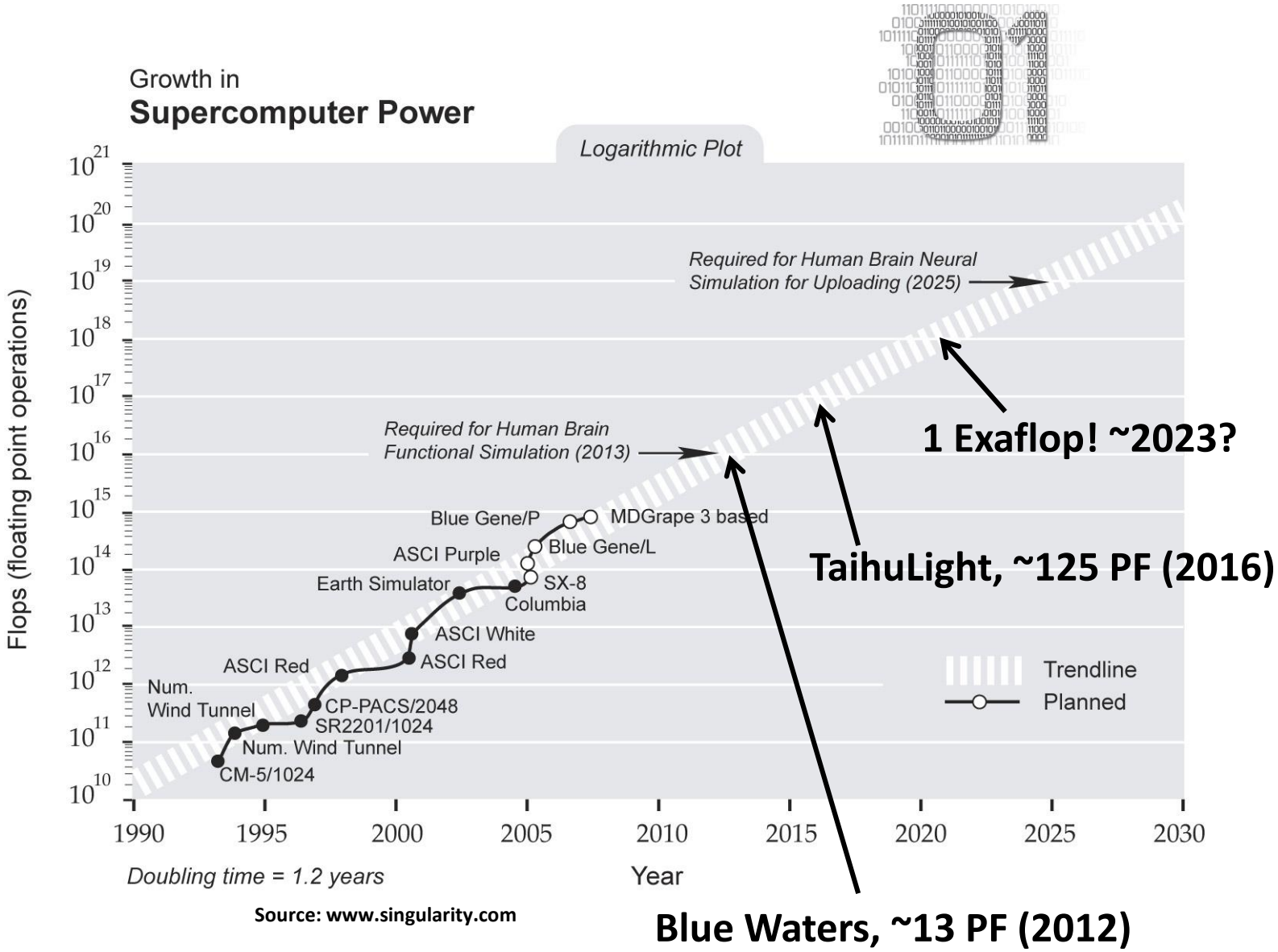
    "Goodyear Puts the Rubber to the Road with High Performance Computing"

    "High Performance Computing Helps Create New Treatment For Stroke Victims"

    "Procter & Gamble: Supercomputers and the Secret Life of Coffee"

    "Motorola: Driving the Cellular Revolution With the Help of High Performance Computing"

    "Microsoft: Delivering High Performance Computing to the Masses"



Source: www.singularity.com

Blue Waters in 2012

Simulating 1 second of human brain activity takes 82,944 processors

By Ryan Whitwam on August 5, 2013 at 1:34 pm | 21 Comments

Share This Article

436 Like | 123 Tweet | 228 | 108 +1 | 24 Share

The brain is a deviously complex biological computing device that even the fastest supercomputers in the world fail to emulate. Well, that's not entirely true anymore. Researchers at the Okinawa Institute of Technology Graduate University in Japan and Forschungszentrum Jülich in Germany have managed to simulate a single second of human brain activity in a very, very powerful computer.

Source: extremetech.com

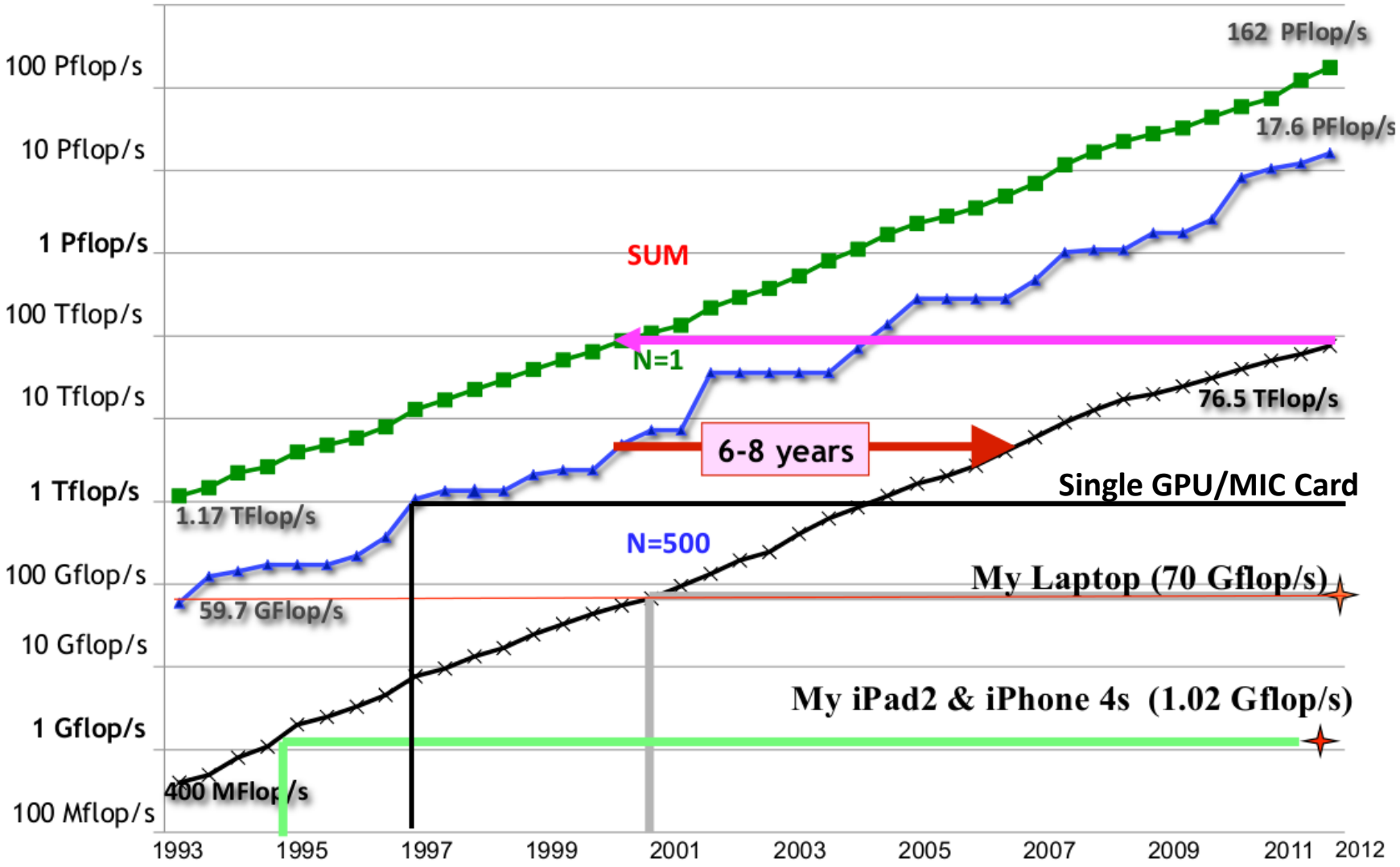
The Top500 List

- **A benchmark, solve Ax=b**
  - As fast as possible! → as big as possible ☺
  - Reflects **some** applications, not all, not even many
  - Very good historic data!
- **Speed comparison for computing centers, states, countries, nations, continents ☹**
  - Politicized (sometimes good, sometimes bad)
  - Yet, fun to watch

The Top500 List (June 2017)

Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	**Sunway TaihuLight** - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway , NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371
2	**Tianhe-2 (MilkyWay-2)** - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P , NUDT National Super Computer Center in Guangzhou China	3,120,000	33,862.7	54,902.4	17,808
3	**Piz Daint** - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100 , Cray Inc. Swiss National Supercomputing Centre (CSCS) Switzerland	361,760	19,590.0	25,326.3	2,272
4	**Titan** - Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x , Cray Inc. DOE/SC/Oak Ridge National Laboratory United States	560,640	17,590.0	27,112.5	8,209
5	**Sequoia** - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom , IBM DOE/NNSA/LLNL United States	1,572,864	17,173.2	20,132.7	7,890
6	**Cori** - Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect , Cray Inc. DOE/SC/LBNL/NERSC United States	622,336	14,014.7	27,880.7	3,939
7	**Oakforest-PACS** - PRIMERGY CX1640 M1, Intel Xeon Phi 7250 68C 1.4GHz, Intel Omni-Path , Fujitsu Joint Center for Advanced High Performance Computing Japan	556,104	13,554.6	24,913.5	2,719
8	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect , Fujitsu RIKEN Advanced Institute for Computational Science (AICS) Japan	705,024	10,510.0	11,280.4	12,660

Top500: Trends



Source: Jack Dongarra

Green Top500 List (June 2017)

Rank	TOP500 Rank	System	Cores	Rmax (TFlop/s)	Power (kW)	Power Efficiency (GFlops/watts)
1	61	**TSUBAME3.0** - SGI ICE XA, IP139-SXM2, Xeon E5-2680v4 14C 2.4GHz, Intel Omni-Path, NVIDIA Tesla P100 SXM2 , HPE GSIC Center, Tokyo Institute of Technology Japan	36,288	1,998.0	142	14.110
2	465	**kukai** - ZettaScaler-1.6 GPGPU system, Xeon E5-2650Lv4 14C 1.7GHz, Infiniband FDR, NVIDIA Tesla P100 , ExaScalar Yahoo Japan Corporation Japan	10,080	460.7	33	14.046
3	148	**AIST AI Cloud** - NEC 4U-8GPU Server, Xeon E5-2630Lv4 10C 1.8GHz, Infiniband EDR, NVIDIA Tesla P100 SXM2 , NEC National Institute of Advanced Industrial Science and Technology Japan	23,400	961.0	76	12.681
4	305	**RAIDEN GPU subsystem** - NVIDIA DGX-1, Xeon E5-2698v4 20C 2.2GHz, Infiniband EDR, NVIDIA Tesla P100 , Fujitsu Center for Advanced Intelligence Project, RIKEN Japan	11,712	635.1	60	10.603
5	100	**Wilkes-2** - Dell C4130, Xeon E5-2650v4 12C 2.2GHz, Infiniband EDR, NVIDIA Tesla P100 , Dell University of Cambridge United Kingdom	21,240	1,193.0	114	10.428
6	3	**Piz Daint** - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100 , Cray Inc. Swiss National Supercomputing Centre (CSCS) Switzerland	361,760	19,590.0	2,272	10.398

Piz Daint @ CSCS

More pictures at: http://spcl.inf.ethz.ch/Teaching/2015-dphpc/

HPC Applications: Scientific Computing

- **Most natural sciences are simulation driven or are moving towards simulation**
  - Theoretical physics (solving the Schrödinger equation, QCD)
  - Biology (Gene sequencing)
  - Chemistry (Material science)
  - Astronomy (Colliding black holes)
  - Medicine (Protein folding for drug discovery)
  - Meteorology (Storm/Tornado prediction)
  - Geology (Oil reservoir management, oil exploration)
  - and many more … (even Pringles uses HPC)

HPC Applications: Commercial Computing

- **Databases, data mining, search**
  - Amazon, Facebook, Google
- **Transaction processing**
  - Visa, Mastercard
- **Decision support**
  - Stock markets, Wall Street, Military applications
- **Parallelism in high-end systems and back-ends**
  - Often throughput-oriented
  - Used equipment varies from COTS (Google) to high-end redundant mainframes (banks)

HPC Applications: Industrial Computing

- **Aeronautics (airflow, engine, structural mechanics, electromagnetism)**
- **Automotive (crash, combustion, airflow)**
- **Computer-aided design (CAD)**
- **Pharmaceuticals (molecular modeling, protein folding, drug design)**
- **Petroleum (Reservoir analysis)**
- **Visualization (all of the above, movies, 3d)**

What can faster computers do for us?

- **Solving bigger problems than we could solve before!**
  - E.g., Gene sequencing and search, simulation of whole cells, mathematics of the brain, …
  - The size of the problem grows with the machine power

    → *Weak Scaling*

- **Solve today's problems faster!**
  - E.g., large (combinatorial) searches, mechanical simulations (aircrafts, cars, weapons, …)
  - The machine power grows with constant problem size

    → *Strong Scaling*

Towards the age of massive parallelism

- **Everything goes parallel**
  - Desktop computers get more cores

    2,4,8, soon dozens, hundreds?
  - Supercomputers get more PEs (cores, nodes)

    > 10 million today

    > 50 million on the horizon

    ➢*1 billion in a couple of years (after 2020)*
- **Parallel Computing is inevitable!**

Parallel vs. Concurrent computing

Concurrent activities ***may*** be executed in parallel

Example:

A1 starts at T1, ends at T2; A2 starts at T3, ends at T4

Intervals (T1,T2) and (T3,T4) may overlap!

Parallel activities:

A1 is executed ***while*** A2 is running

Usually requires separate resources!

Programming Models

Flynn's Taxonomy

SISD **Standard Serial Computer** **(nearly extinct)**	**SIMD** **Vector Machines or Extensions** **(very common)**
MISD **Redundant Execution** **(fault tolerance)**	**MIMD** **Multicore** **(ubiquituous)**

Parallel Resources and Programming

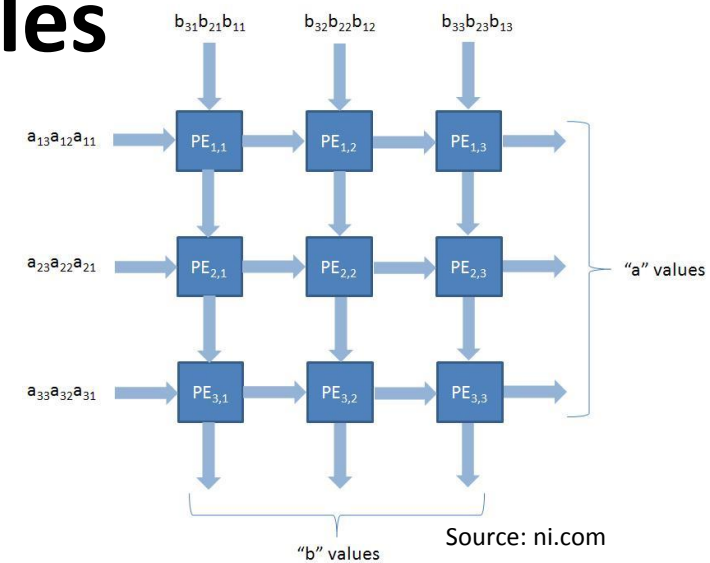
Parallel Resource	Programming

▪ Instruction-level parallelism
  ▪ Pipelining
  ▪ VLIW
  ▪ Superscalar
	▪ Compiler
	  ▪ (inline assembly)
	  ▪ Hardware scheduling

▪ SIMD operations ▪ Vector operations	▪ Compiler (inline assembly) ▪ Intrinsics ▪ Libraries

▪ Instruction sequences
  ▪ Multiprocessors
  ▪ Multicores
  ▪ Multithreading
	▪ Compilers (very limited)
	▪ Expert programmers
	  ▪ Parallel languages
	  ▪ Parallel libraries
	  ▪ Hints

Historic Architecture Examples

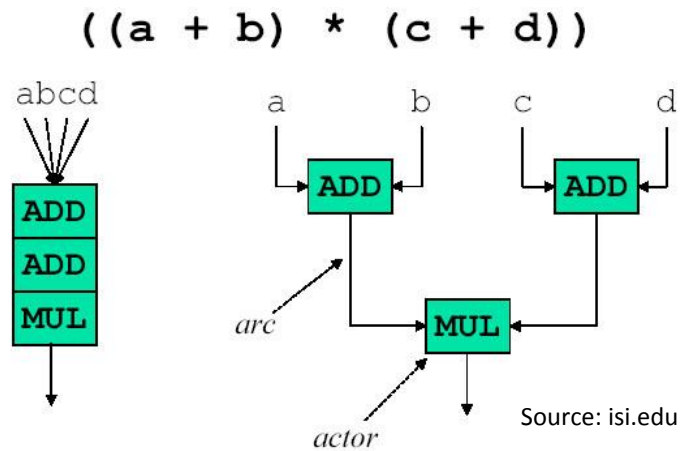


Source: ni.com

- **Systolic Array**
  - Data-stream driven (data counters)
  - Multiple streams for parallelism
  - Specialized for applications (reconfigurable)

- **Dataflow Architectures**
  - No program counter, execute instructions when all input arguments are available
  - Fine-grained, high overheads

    *Example: compute f = (a+b) * (c+d)*

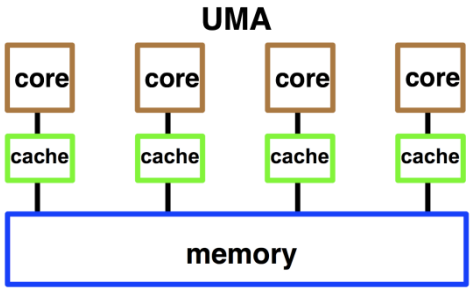


Source: isi.edu

- Both come-back in FPGA computing
  - Interesting research opportunities!

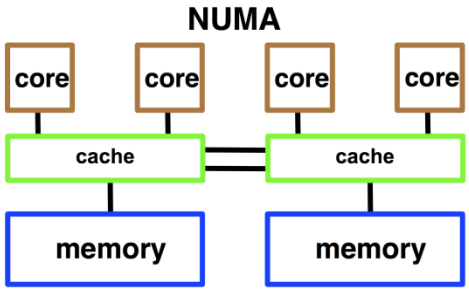
Parallel Architectures 101

Uniform memory access



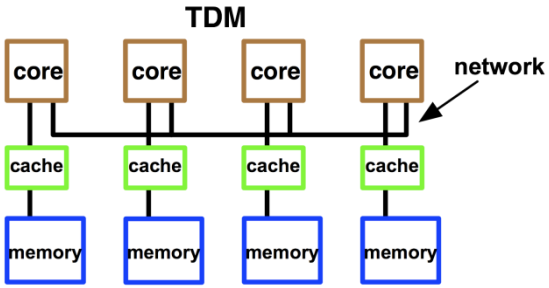
Today's laptops

Non-uniform memory access



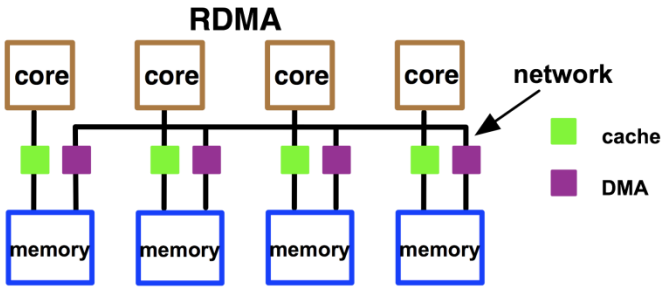
Today's servers

Time-division multiplexing



Yesterday's clusters

Remote direct-memory access

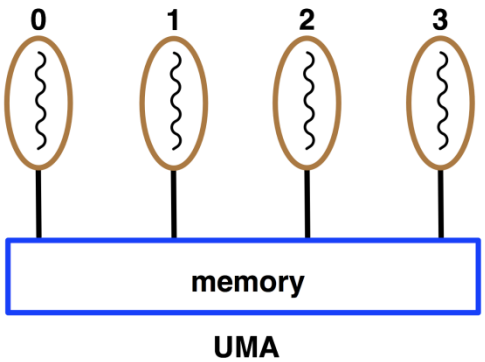


Today's clusters

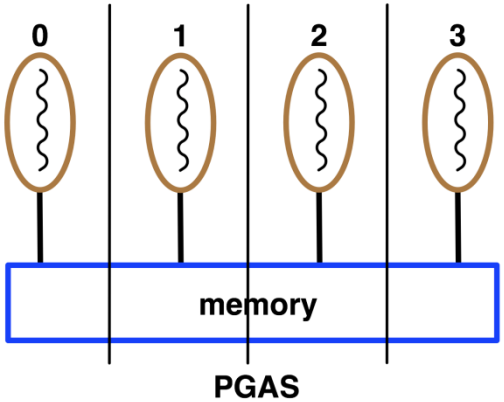
- **… and mixtures of those**

Programming Models

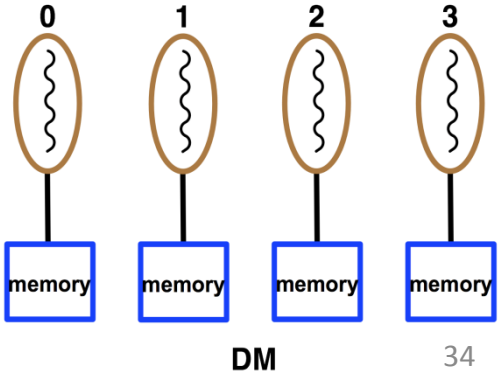
- **Shared Memory Programming (SM)**
  - Shared address space
  - Implicit communication
  - Hardware for cache-coherent remote memory access
  - Cache-coherent Non Uniform Memory Access (cc NUMA)
  - Pthreads, OpenMP



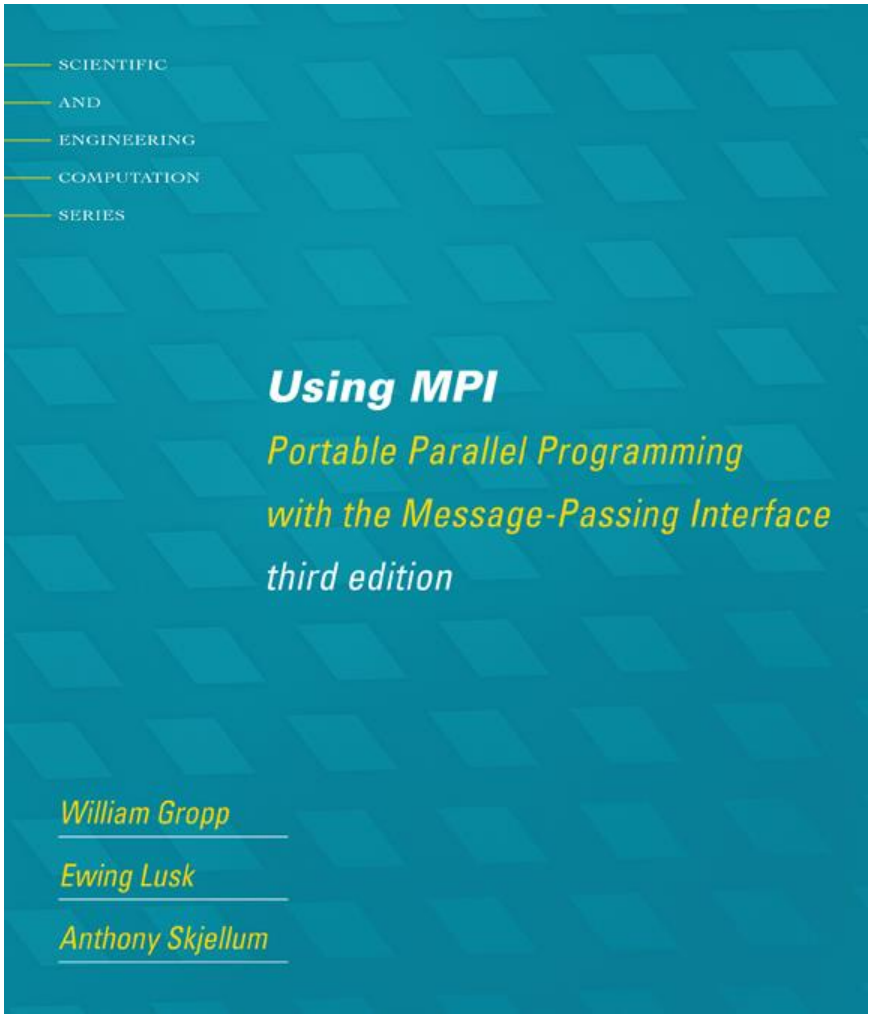
- **(Partitioned) Global Address Space (PGAS)**
  - Remote Memory Access
  - Remote vs. local memory (cc. ncc-NUMA)



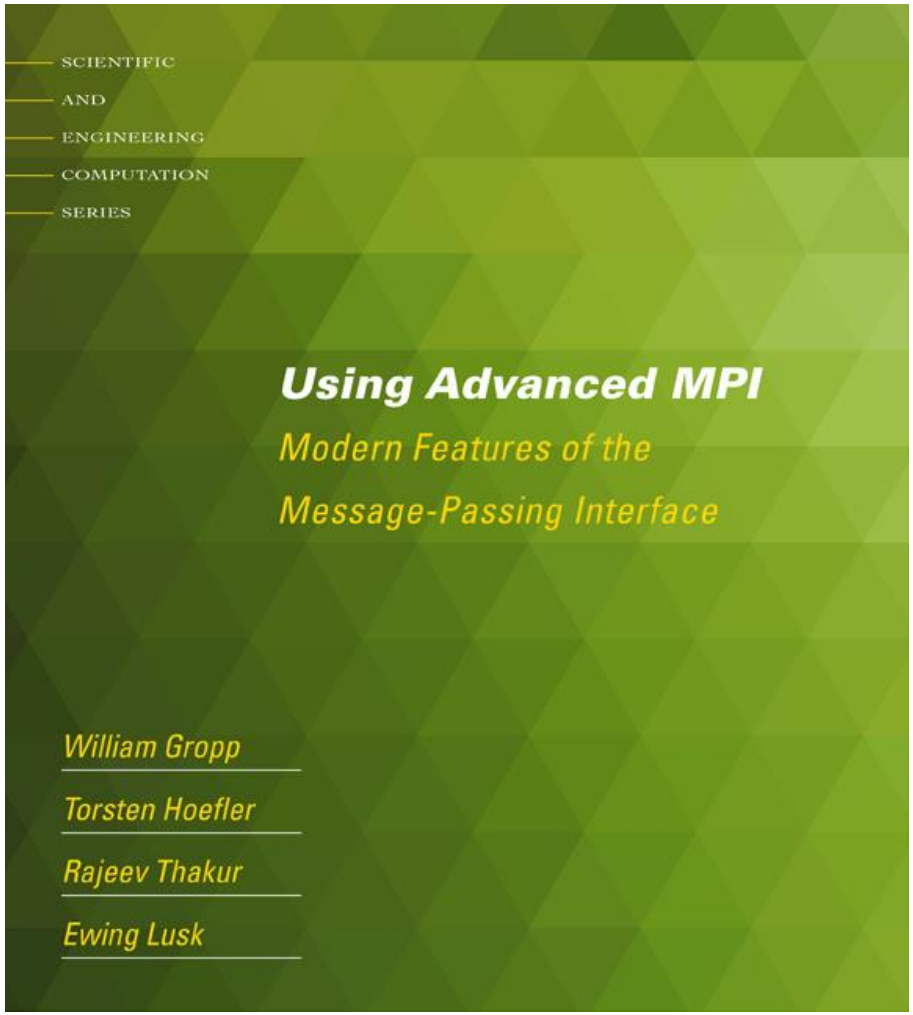
- **Distributed Memory Programming (DM)**
  - Explicit communication (typically messages)
  - Message Passing



MPI: de-facto large-scale prog. standard

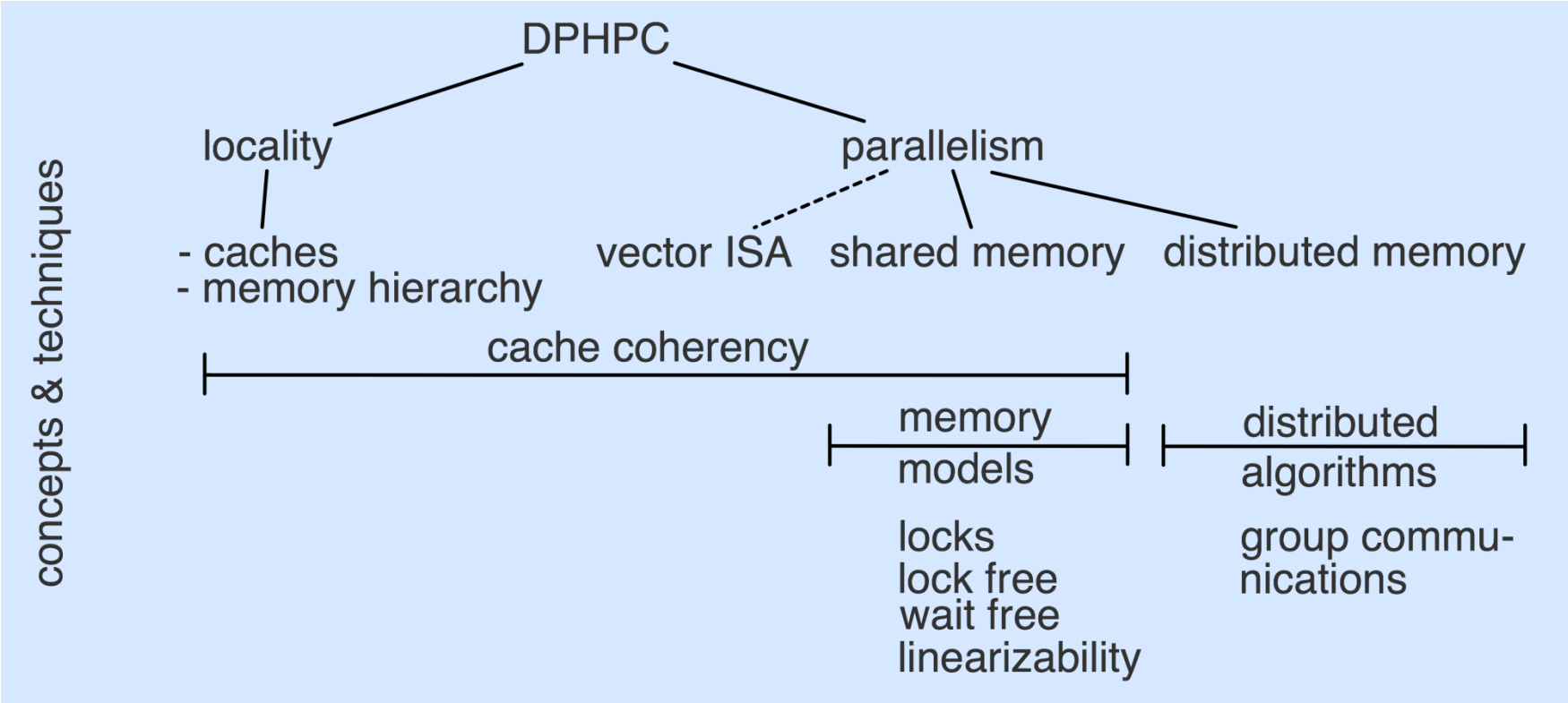


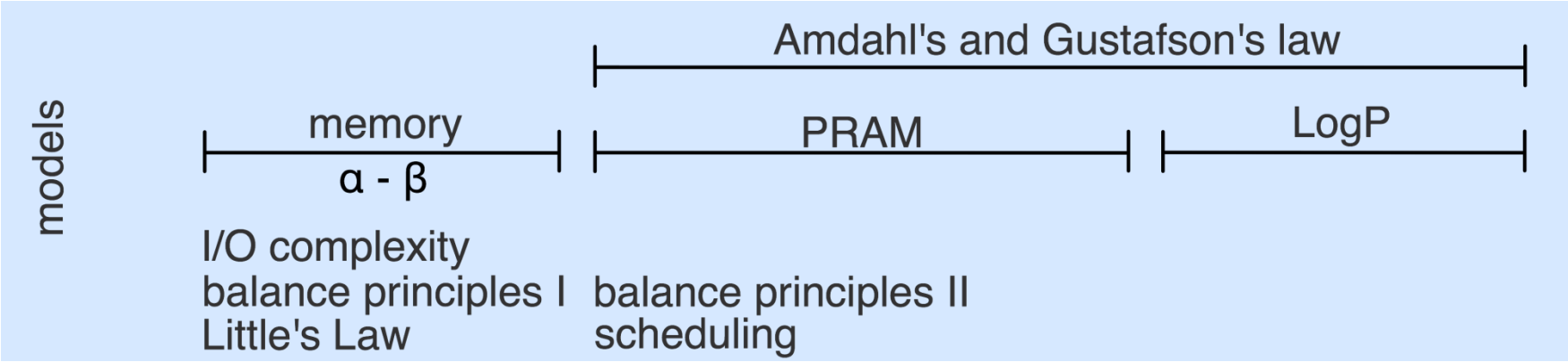
Basic MPI



Advanced MPI, including MPI-3

DPHPC Overview





Schedule of Last Year

k	Monday	Thursday
0	09/19: no lecture	09/22: MPI Tutorial (white bg)
1	09/26: Organization - Introduction (1pp) (6pp)	09/29: Projects - Advanced MPI Tutorial
2	10/03: Cache Coherence & Memory Models (1pp) (6pp)	10/06: Cache Organization - Introduction to OpenMP
3	10/10: Memory Models (1pp) (6pp)	10/13: Sequential Consistency + OpenMP Synchronization
4	10/17: Linearizability (1pp) (6pp)	10/20: Linearizability
5	10/24: Languages and Locks (1pp) (6pp)	10/27: Locks
6	10/31: Amdahl's Law (1pp) (6pp) - Notes	11/03: Amdahl's Law
7	11/07: Project presentations	11/10: No recitation session
8	11/14: Roofline Model (1pp) (6pp) - Notes	11/17: Roofline Model
9	11/21: Balance Principles (1pp) (6pp) - Notes on Balance Principles / Scheduling (1pp) (6pp) - Nodes on Scheduling	11/24: Balance Priciples & Scheduling
10	11/28: Locks and Lock-Free (1pp) (6pp)	12/01: SPIN Tutorial
11	12/05: Lock-Free and distributed memory (1pp) (6pp)	12/08: Benchmarking - (paper)
12	12/12: Guest lecture - Dr. Tobias Grosser	12/15: Network Models
13	12/19: Final Presentations	

Related classes in the SE/PL focus

- **263-2910-00L Program Analysis**
  http://www.srl.inf.ethz.ch/pa.php
  Spring 2017
  Lecturer: Martin Vechev

- **263-2300-00L How to Write Fast Numerical Code**
  http://www.inf.ethz.ch/personal/markusp/teaching/263-2300-ETH-spring16/course.html
  Spring 2017
  Lecturer: Markus Püschel

- **This list is not exhaustive!**