

EDA284: **Parallel Computer Architecture**

Miquel Pericàs
Chalmers University of Technology

miquelp@chalmers.se

Office EDIT 4106

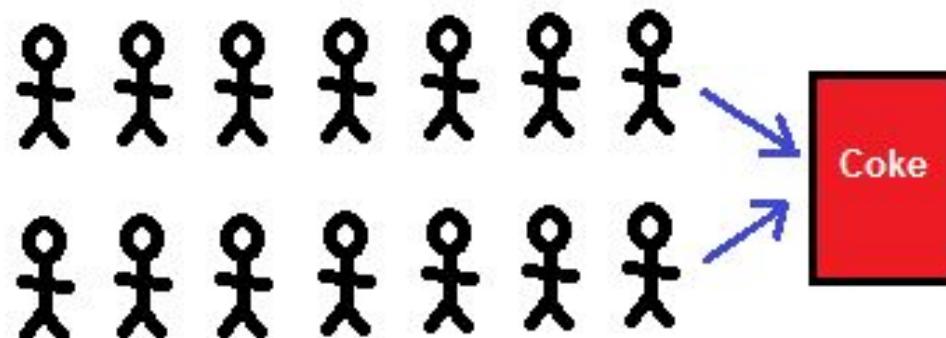
Course staff

- **Examiner:** Miquel Pericàs,
miquelp@chalmers.se
- **Teaching Assistant:**
 - Jing Chen chjing@chalmers.se
 - Mustafa Abduljabbar musabdu@chalmers.se
- **Guest Lecturers:**
 - Yiannis Sourdis, *On-chip Networks*
 - Bhavysha Goel, Mahmoud Eljammaly, *European Processor Initiative*

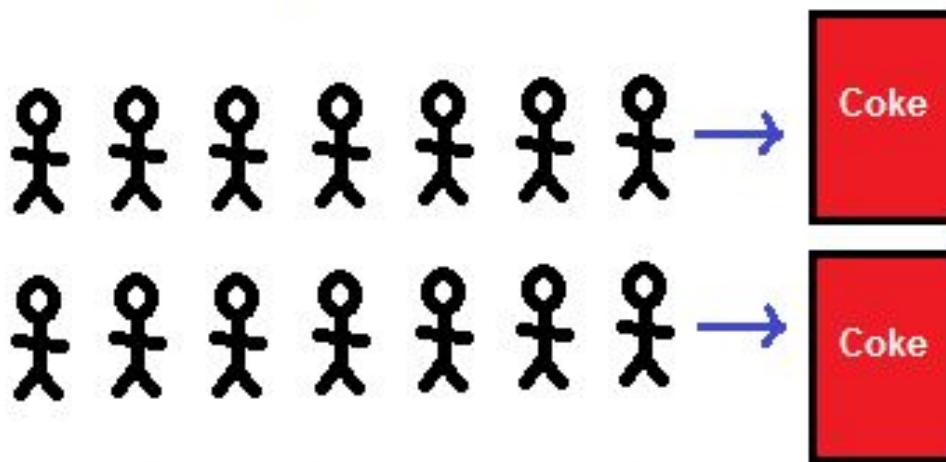
Today's plan

- **Introduction to EDA284 (part 1)**
- **Basic concepts in Parallel Computer Architecture (part 2)**

Concurrency vs Parallelism



Concurrent: 2 queues, 1 vending machine



Parallel: 2 queues, 2 vending machines

Concurrency vs Parallelism

- **Definition from Oracle Multithreaded Programming Guide:**
 - **Concurrency**: “A condition that exists when at least two threads are making progress. A more generalized form of parallelism that can include time-slicing as a form of virtual parallelism.”
 - **Parallelism**: “A condition that arises when at least two threads are executing simultaneously.”
- In other words:
 - **Concurrency** exists when the execution periods of two or more tasks are overlapping. There may, however, never be more than one task executing in the same instant. E.g.: multitasking on a single-core machine.
 - **Parallelism** is when at least two tasks execute simultaneously.

Heterogeneous Computing

- What is heterogeneous computing?
- Examples of heterogeneous computers?

Why parallelism? Why heterogeneity?

- To reach a certain level of **performance** given a set of constraints, such as:
 - chip area
 - **power/energy**
 - budget
 - time to market
 - complexity
 - reliability
 - others?

Range of parallel/heterogeneous computers: from embedded to HPC



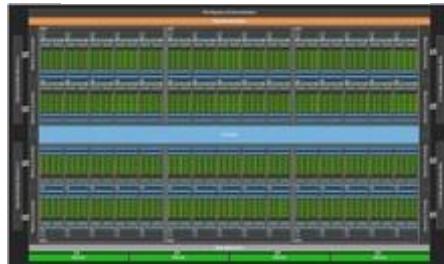
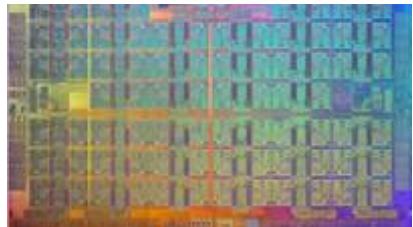
Supercomputers



Data Centers



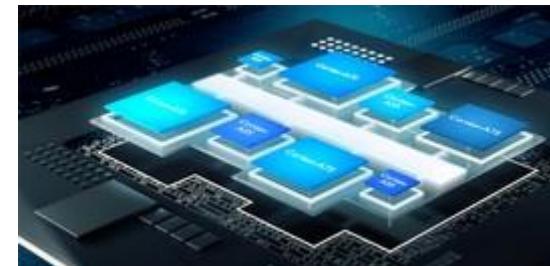
Multiprocessor Blades



Manycore



Multicore



Heterogeneous CMP

and more...

Supercomputers: an example

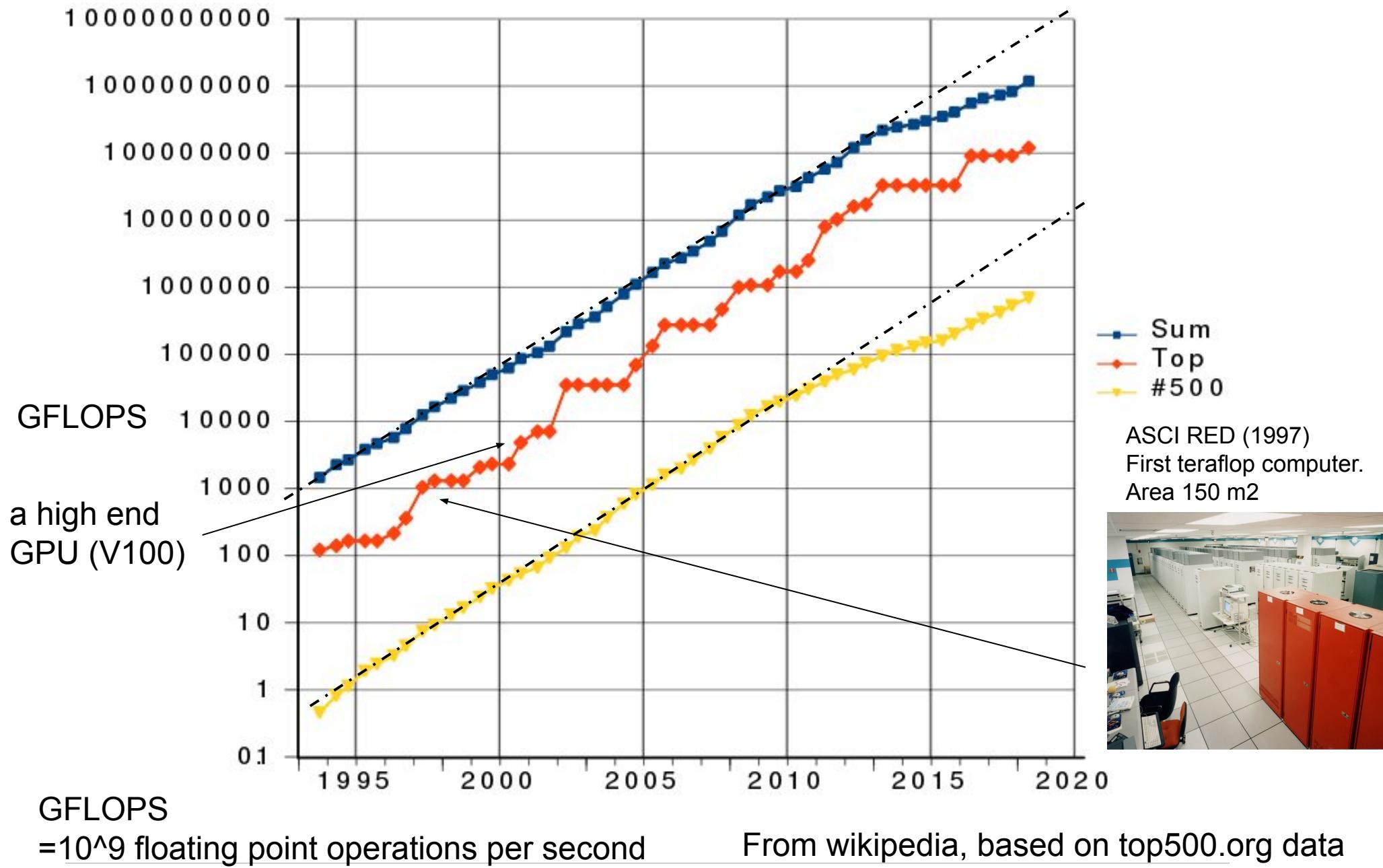
Top500 list published every 6 months since 1986

World's largest supercomputers are ranked according to one particular benchmark: **Linpack**

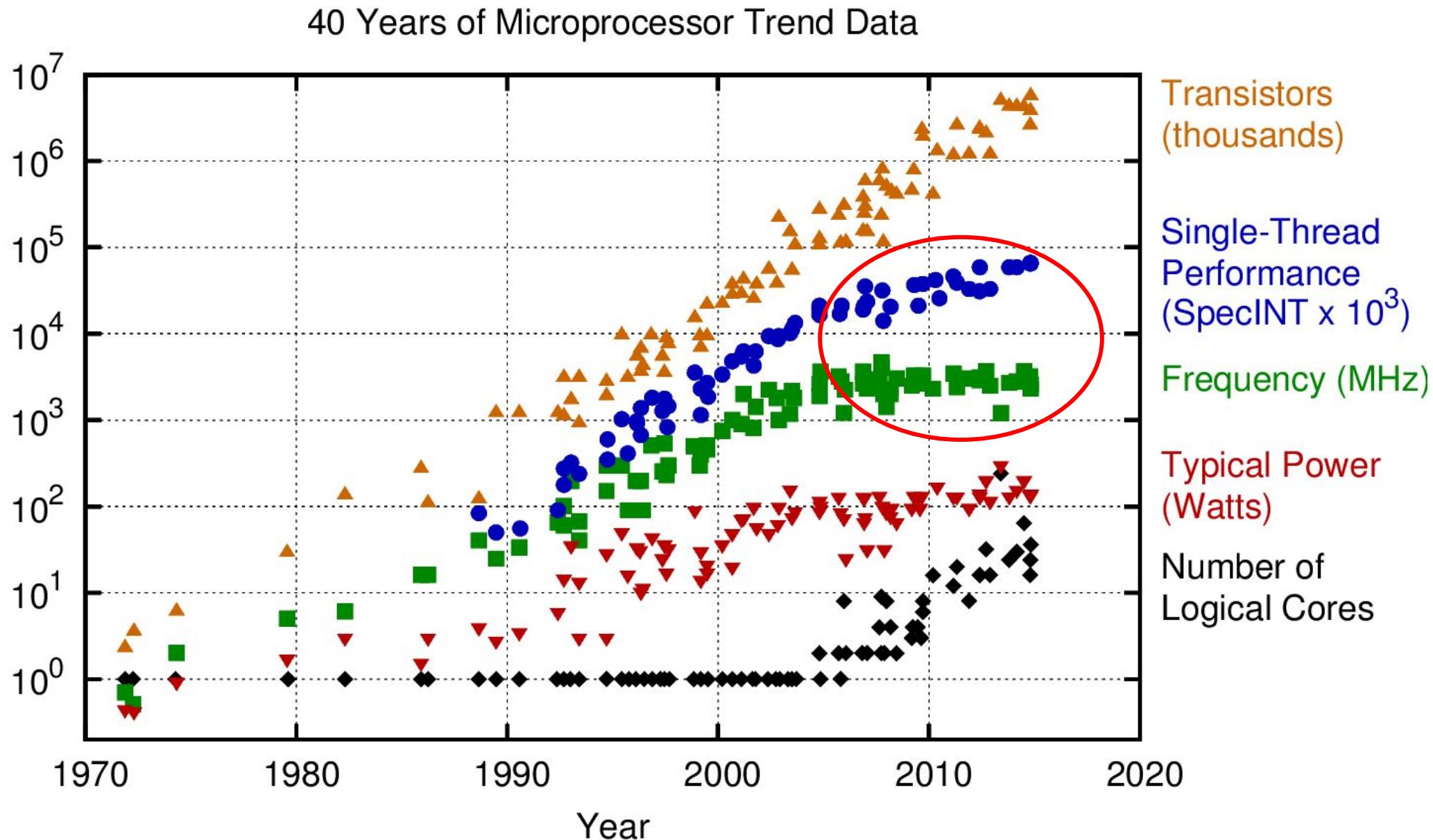
The table on the right is the *November 2019* edition

Rank	Site	System	Cores	Rmax [TFlop/s]	Peak [TFlop/s]	Power [kW]
1	DOE/SC/Oak Ridge National Laboratory United States	Summit - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband IBM	2,414,592	148,600.0	200,794.9	10,096
2	DOE/NNSA/LLNL United States	Sierra - IBM Power System AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband IBM / NVIDIA / Mellanox	1,572,480	94,640.0	125,712.0	7,438
3	National Supercomputing Center in Wuxi China	Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway NRCPC	10,649,600	93,014.6	125,435.9	15,371
4	National Super Computer Center in Guangzhou China	Tianhe-2A - TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000 NUDT	4,981,760	61,444.5	100,678.7	18,482
5	Texas Advanced Computing Center/Univ. of Texas United States	Frontera - Dell C6420, Xeon Platinum 8280 28C 2.7GHz, Mellanox InfiniBand HDR Dell EMC	448,448	23,516.4	38,745.9	
6	Swiss National Supercomputing Centre (CSCS) Switzerland	Piz Daint - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect, NVIDIA Tesla P100 Cray/HPE	387,872	21,230.0	27,154.3	2,384
7	DOE/NNSA/LANL/SNL United States	Trinity - Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect Cray/HPE	979,072	20,158.7	41,461.2	7,578

Performance evolution of Top500 supercomputers

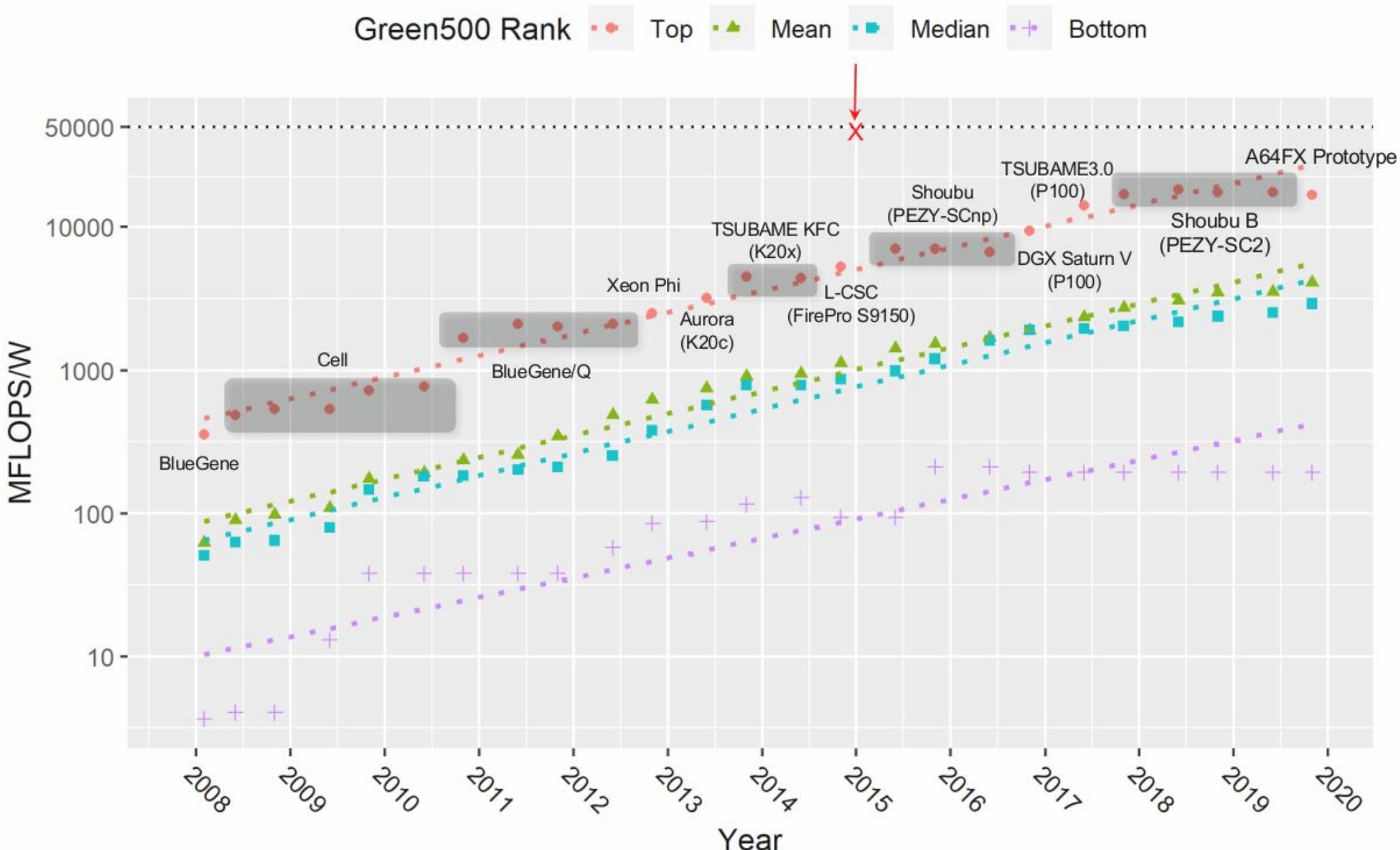


Performance is saturating. Why?



Original data up to the year 2010 collected and plotted by M. Horowitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammond, and C. Batten
New plot and data collected for 2010-2015 by K. Rupp

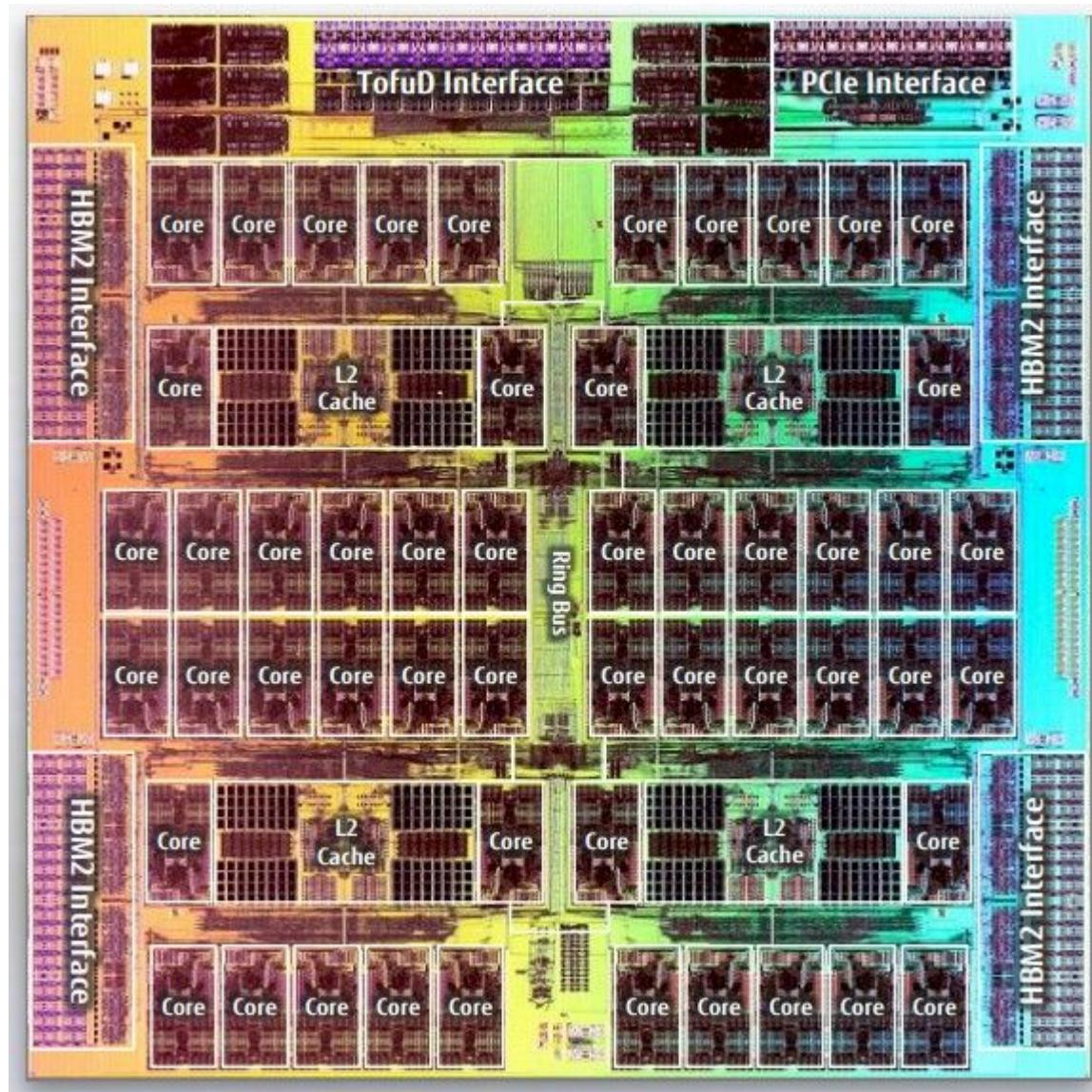
Trends: How Energy Efficient Are We?



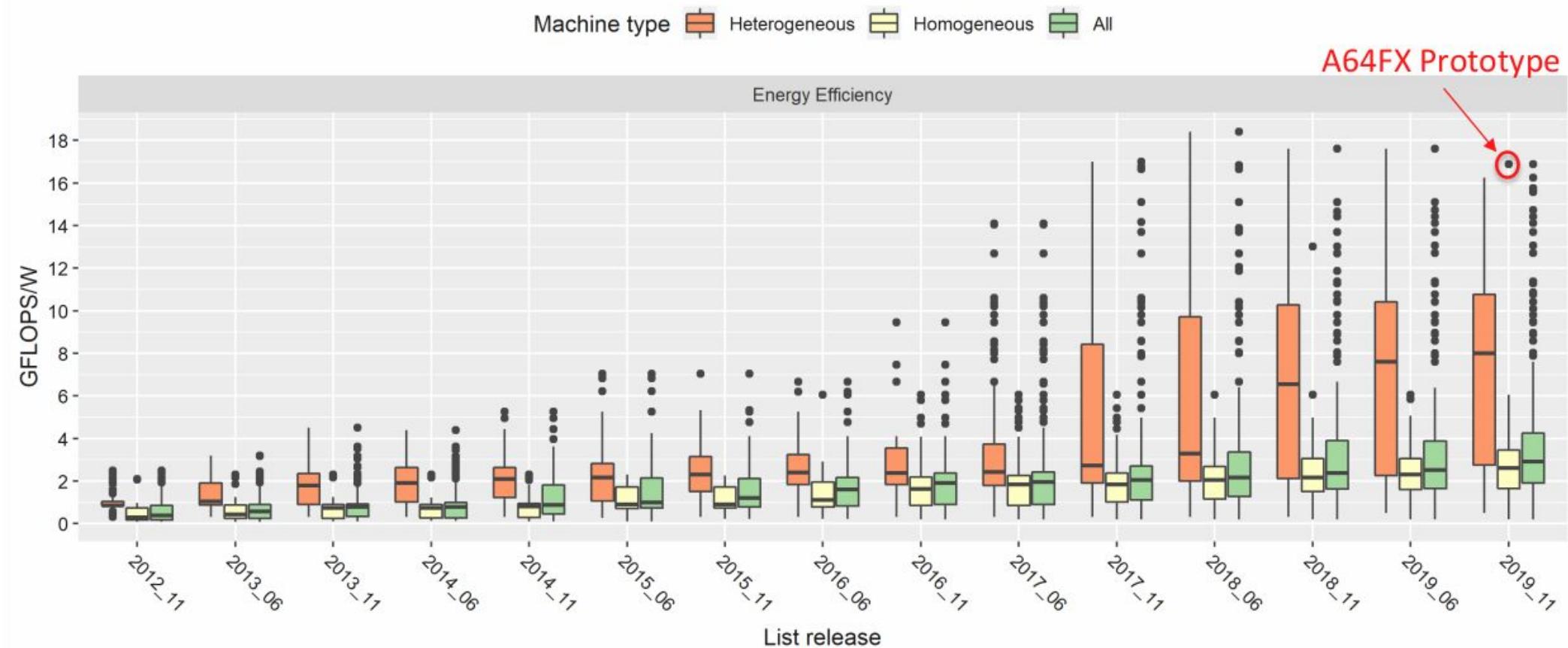
2019 data

TOP500				Cores	Power		
Rank	Rank	System	Rmax [TFlop/s]		Power (kW)	Efficiency (GFlops/watts)	
1	159	A64FX prototype - Fujitsu A64FX, Fujitsu A64FX 48C 2GHz, Tofu interconnect D , Fujitsu Fujitsu Numazu Plant Japan	36,864	1,999.5	118	16.876	
2	420	NA-1 - ZettaScaler-2.2, Xeon D-1571 16C 1.3GHz, Infiniband EDR, PEZY-SC2 700Mhz , PEZY Computing / Exascaler Inc. PEZY Computing K.K. Japan	1,271,040	1,303.2	80	16.256	
3	24	AiMOS - IBM Power System AC922, IBM POWER9 20C 3.45GHz, Dual-rail Mellanox EDR Infiniband, NVIDIA Volta GV100 , IBM Rensselaer Polytechnic Institute Center for Computational Innovations (CCI) United States	130,000	8,045.0	510	15.771	
4	373	Satori - IBM Power System AC922, IBM POWER9 20C 2.4GHz, Infiniband EDR, NVIDIA Tesla V100 SXM2 , IBM MIT/MGHPCC Holyoke, MA United States	23,040	1,464.0	94	15.574	
5	1	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	148,600.0	10,096	14.719	
6	8	AI Bridging Cloud Infrastructure (ABCi) - PRIMERGY CX2570 M4, Xeon Gold 6148 20C 2.4GHz, NVIDIA Tesla V100 SXM2, Infiniband EDR , Fujitsu National Institute of Advanced Industrial Science and Technology (AIST) Japan	391,680	19,880.0	1,649	14.423	

Fujitsu A64FX prototype



Trends in Efficiency (2012 – Present): Homogeneous vs. Heterogeneous Systems



Race to Exascale

European Program to Develop Supercomputing Chips Begins to Take Shape

Michael Feldman | July 5, 2018 22:27 CEST

@ E-mail

 Tweet

 Like

 G +1

 in Share

The European Processor Initiative (EPI), an ambitious program to develop a pair of chips for domestic supercomputers, is poised to change the way Europe does HPC. And although the work is still very much in its early stages, it looks like the Europeans have selected their preferred processor architectures: Arm and RISC-V.



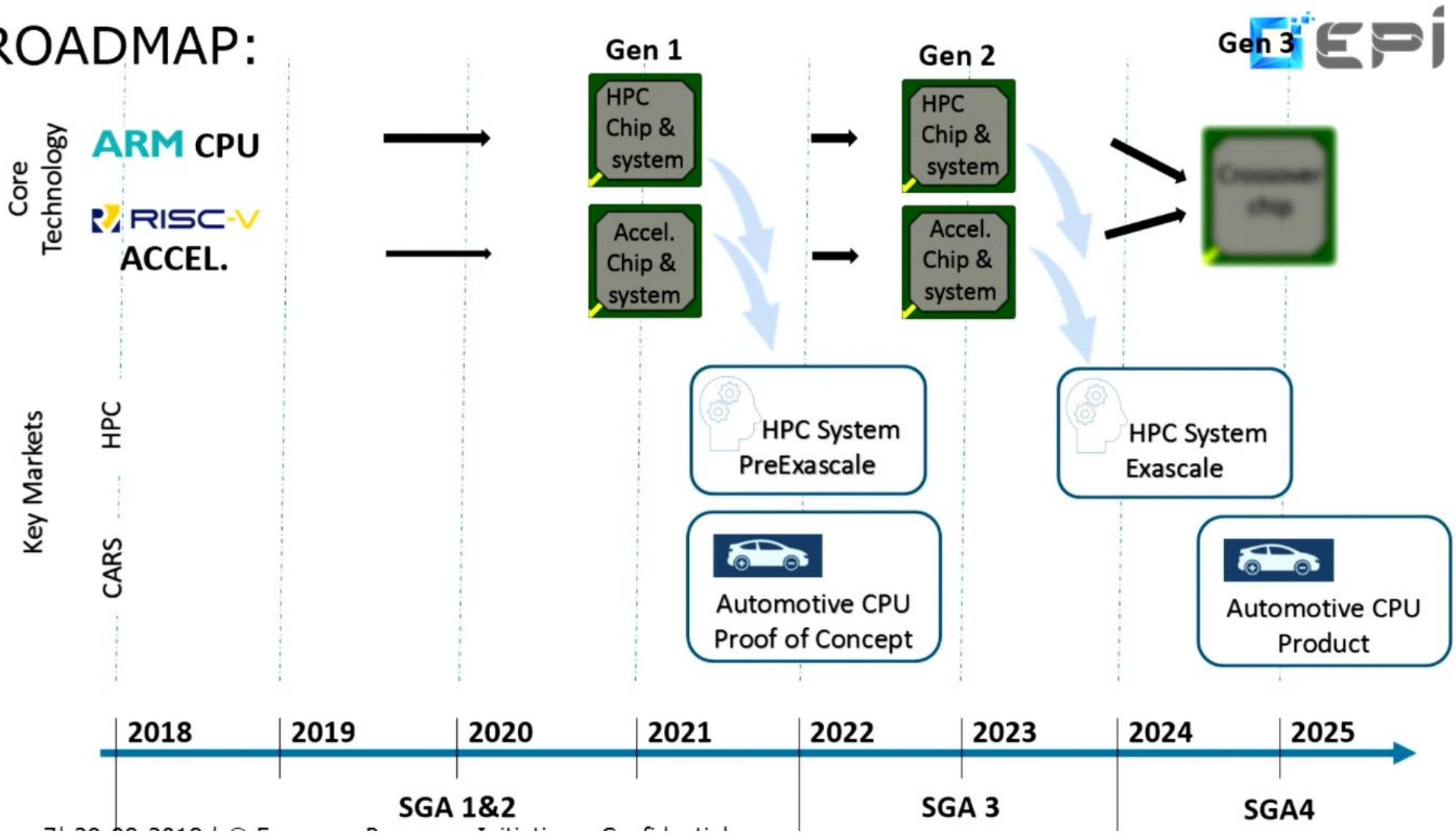
Launched in March 2018 by the European Commission, the EPI's overall aim is to develop domestically produced low-power microprocessors for the European market. Even though the work is focused on delivering chips for HPC, and in particular for exascale supercomputers, the technology will also be applied to the broader datacenter market, as well as the automotive industry. The rationale for this more expansive

strategy is to provide higher volume markets that can economically sustain the considerable effort involved in chip R&D and support..

The first generation of these HPC processors are expected to be delivered toward the end of the decade, in time to form the basis for pre-exascale supercomputers scheduled to be deployed across the EU in the 2020 to 2021 timeframe. The second-generation chips will power Europe's first exascale systems in 2023 and 2024. The system work is being led by EuroHPC, a group formed to bring Europe on par with the US, China, and Japan in high performance computing technology. Part of the mission involves developing home-grown componentry so that EU members have more control over what goes into their supercomputers.

The most central element of these systems is the processor, which puts EPI in the critical path for the EuroHPC work. In a statement delivered at the launch of EPI in March, Vice-President Andrus Ansip, who heads the Digital Single Market, and Mariya Gabriel, the Commissioner for Digital Economy and Society, summed up the strategy as follows:

ROADMAP:



**We will have an invited lecture
on the EPI project**

High Performance Computing needs Parallel Programming!

- Parallel computer programs are more difficult to write than sequential programs:
 1. need to manage larger amount of *state*, often asynchronous → makes debugging difficult
 2. *parallelism* introduces new types of bugs, for example, race conditions (unordered accesses to shared variables)
 3. need to worry about *communication* and *synchronization*:
 1. where is the data?
 2. who/when to provide it?
 4. Maintaining parallel code requires higher effort
- Covered in DAT400...

Goals of this course

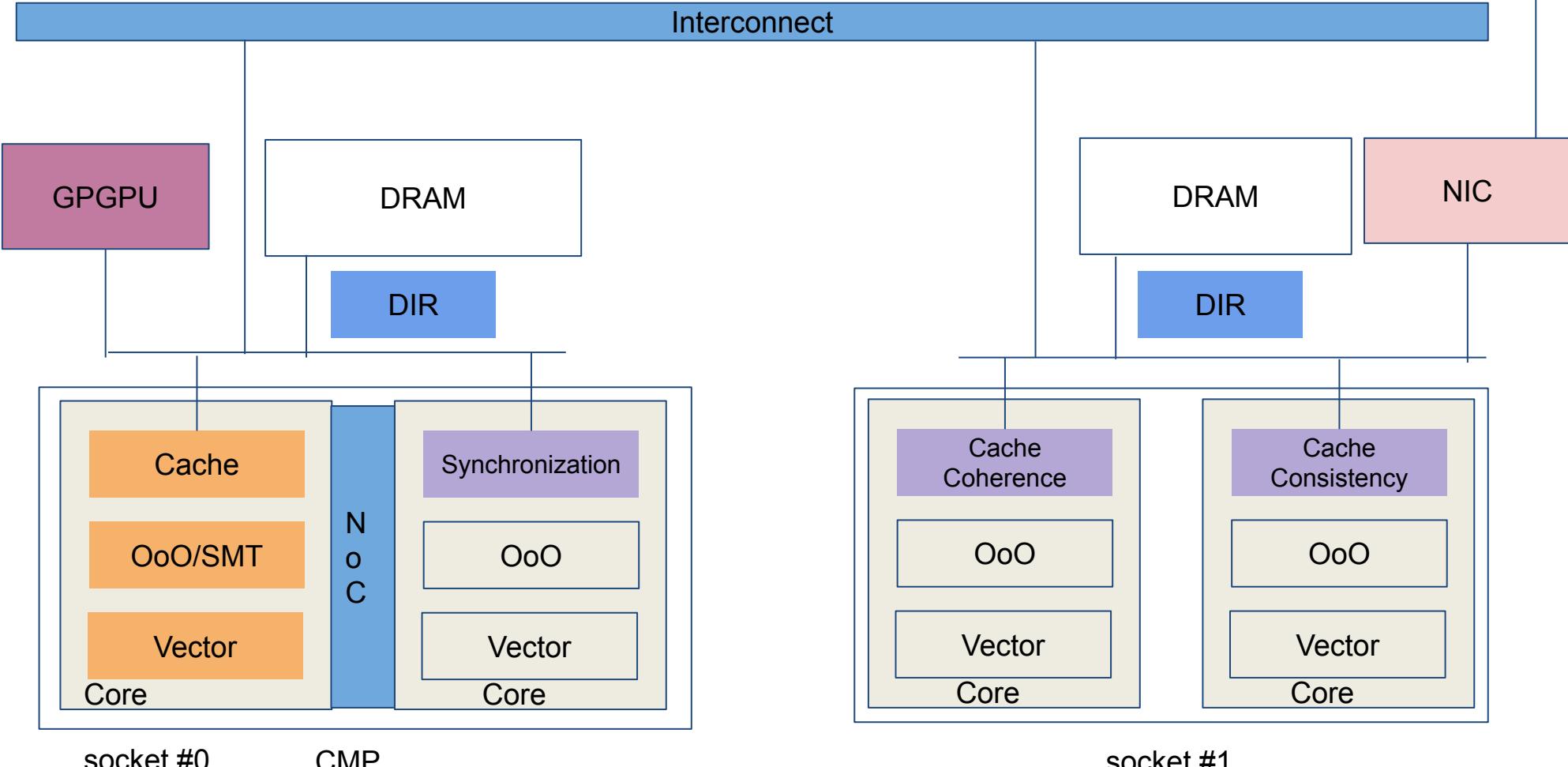
- Understand the trade-offs across the HW/SW interface to meet *functional, performance and cost requirements* of parallel computers (*lectures*)
- Learn how to model parallel computers and use simulators to co-design hardware and software (*labs*)
- Comprehend modern literature and design HW solutions for computationally intensive applications (*project*)

Preliminary schedule

- up-to-date in Canvas
- 14 lectures + 4 exercise + 2 guest lectures
 - Part 1: Metrics
 - Part 2: SIMD + OoO
 - Part 3: SMPs + Multicore
 - Part 4: Message Passing
 - Part 5: GPUs
 - Part 6: Synchronization
- 3 Lab sessions

Session Date	Type	Contents
24/1 8h-11h	Lecture 1	Intro to course + Technology
28/1 8h-11h	Lecture 2	Performance Metrics + Vector
29/1 8h-10h	Lecture 3	Out-of-Order Execution + Multilevel Cache Hierarchy
31/1 x2	No lecture	
4/2 8h-11h	Lecture 4	Parallel programming recap. Multiprocessor Systems (1)
5/2 8h-10h	Lecture 5	Multiprocessor Systems (2)
7/2 8h-10h	Lecture 6	Roofline Model
7/2 10h-12h	Practice #1	SMPs
11/2 8h-10h	Lecture 7	Core Multithreading
11/2 10h-12h	Lab preparation	Intro to GEM5
12/2 8h-10h	Lecture 8	Chip Multiprocessors
14/2 8h-12h	Lab #1	Lab 1 Gem5 + Roofline
18/2 10h-12h	Lecture 10	Guest Lecture: Network-on-Chip (Yiannis Sourdis)
19/2 8h-10h	Practice #2	cc-NUMA
21/2 8h-10h	Lecture 11	Message Passing Hardware
21/2 10h-12h	Practice #3	CMP, NoC
25/2 8h-10h	Lecture 12	GPGPU architectures
25/2 10h-12h	Lab preparation	GEM5 Vector

Course structure



1. Metrics

2. SIMD + OoO

3. Shared Mem

4. Message Passing

5. GPGPU

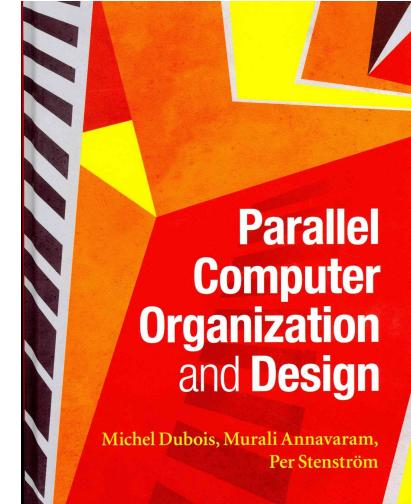
6. Synchronization

Evaluation

- **Three parts**
 - **Project** (1,5c Grades: F, 3, 4, 5)
 - **Written Examination** (4,5c Grades: F, 3, 4, 5)
 - **Lab** (1,5c Grades: Pass/Fail)
- **Project** → Parallel computer Design
 - Described Tuesday Next week
- **Written Exam** → 21/3 (am), covering the course contents and labs.
 - *Not covering the projects*
- **Labs** → Pass/Fail (need to pass all lab sessions)

Course Materials

- Course Book
 - *Parallel Computer Organization and Design*,
 - Michel Dubois, Murali Annavaram, Per Stenström
- Course Slides
 - Evolution of book slides + new materials. Will be published in Canvas ahead of lecture
- Practice Sessions
 - Exercises from book and previous exams
 - Solutions will be posted on Canvas after class



Course representatives

- School Selection

- MPCSN Maria Aguilar Romero <mafernandaguilar@outlook.com>
- MPHPC Alice Gunnarsson <galice@student.chalmers.se>
- MPHPC William Hjelm <hjelmw@student.chalmers.se>
- MPHPC Marcus Karegren <karegren@student.chalmers.se>
- MPHPC Fredrik Lindberg <fredlin@student.chalmers.se>

Ready?