High Performance Computing

Term 4 2018/2019

Lecturer: Associate Professor Sergey Rykovanov

Teaching Assistant: Daniil Stefonishin

Course logistics

- Tuesdays/Thursdays 12:30-15:30 lectures/hands-on sessions
- Fridays 12:30-15:30 seminars/computer labs

Activity Type	Activity Weight, %
Attendance	10
Essay	10
Homework Assignments	10
Computer Labs	40
Final Project	30

- Recommended textbooks:
  - Sterling, Anderson, Brodowicz. High Performance Computing. Modern Systems and Practices. Morgan Kaufmann Publishers, 2018
  - В.П. Гергель. Высокопроизводительные вычисления для многоядерных многопроцессорных систем. Издательство Нижегородского госуниверситета, 2010
  - В.Л. Баденко. Высокопроизводительные вычисления. Учебное пособие. Издательство Политехнического университета, Санкт-Петербург. 2010
  - High Performance Computing for Dummies.
- Video resources:
  - Coursera: Введение в параллельное программирование с использованием OpenMP и MPI
  - Udacity: Introduction to High Performance Computing

Course logistics

Prerequisites:

- Knowledge of Unix-like systems (working in terminal)
- C/C++ programming language and preferably Python
- Basic undergrad mathematics (calculus, linear algebra, ODEs and PDEs)
- Laptop
  - we will work in terminal
  - gcc, openmp, openmpi

Please fill out the questionnaire before the end of today‘s lecture.

You will need to get account on Skoltech‘s Pardus supercomputer: through **IT helpdesk**

Please fill out the User Agreement

Instructors

Associate Professor
Sergey Rykovanov

2000-2006 Lomonosov Moscow State University, Physics Faculty
2006-2009 PhD from Ludwig-Maximilians University Munich, Germany (Theoretical and computational plasma physics)
2009-2018 Postdoc at Ludwig-Maximilians University Munich, Lawrence Berkeley National Laboratory USA, Group Leader at Helmholtz Institute Jena, Germany
July 2018-… Skoltech

s.rykovanov@skoltech.ru

Research Intern
Daniil Stefonishin

2009-2014 Lomonosov Moscow State University, Faculty of Computational Mathematics and Cybernetics
2014-2019 PhD student, MSU and Institute of Numerical Mathematics, Russian Academy of Sciences

d.stefonishin@skoltech.ru

Goals of the course

Not to be afraid of supercomputers.

Be able to write parallel programs on modern multi-core CPUs and GPUs.

Learn how to use large computing infrastructure.

Have fun learning.

Four paradigms of modern science

1. Experiment

2. Theory

$$i\hbar\frac{\partial}{\partial t}\Psi = \hat{H}\Psi$$

Limitations:

- experiments/theory too complicated
- experiments too expensive (car/airplane construction)
- experiments too slow (evolution of galaxies)
- experiments too dangerous (explosives, chemicals, climate)
- experiments not possible at the moment

Four paradigms of modern science

1. Experiment

2. Theory

$$i\hbar\frac{\partial}{\partial t}\Psi=\hat{H}\Psi$$

3. Modeling

Mathematical models

Numerical methods

Computers and supercomputers

Four paradigms of modern science

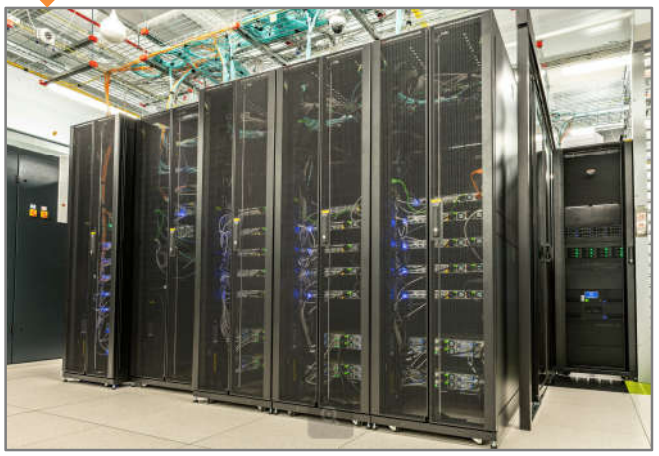
1. Experiment

2. Theory

$$i\hbar\frac{\partial}{\partial t}\Psi = \hat{H}\Psi$$

3. Modeling

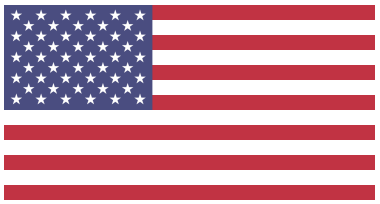
4. Big Data and AI



High Performance Computing is a strategic necessity

- Main players like USA, China, European Union are investing **billions of dollars** into HPC
- HPC drives scientific discovery (plasma physics, computational chemistry, novel materials, artificial intelligence).
- HPC has entered almost all areas of human activity, for example:
  - new perfumes development
  - automobile tires development
  - entertainment industry, movies (Disney had collaboration with plasma physicists on particle simulations using GPUs), cybersport
  - PayPal saved **~1 billion dollars** detecting online fraud with HPC
- HPC is one of the main drivers of the modern world innovation.

HPC share compared to the country GDP



24% GDP

38% HPC power

15% GDP

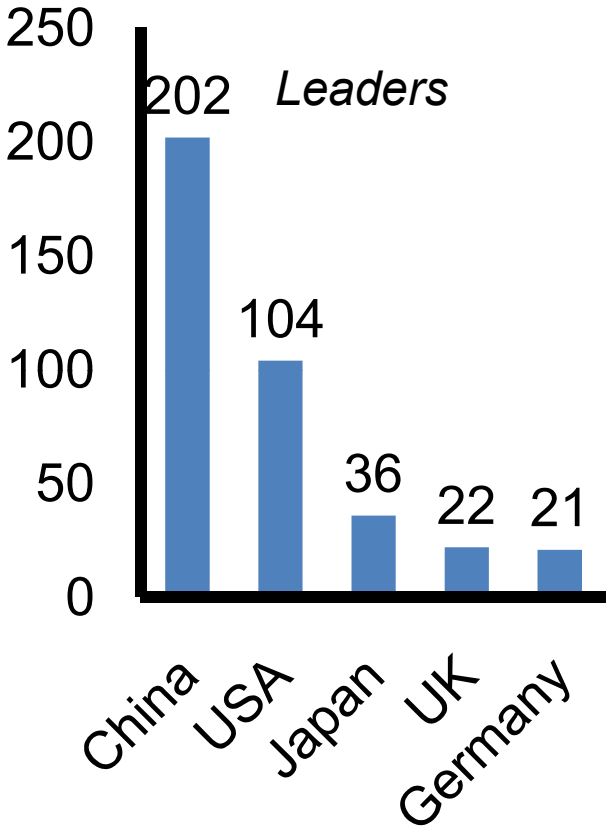
29% HPC power

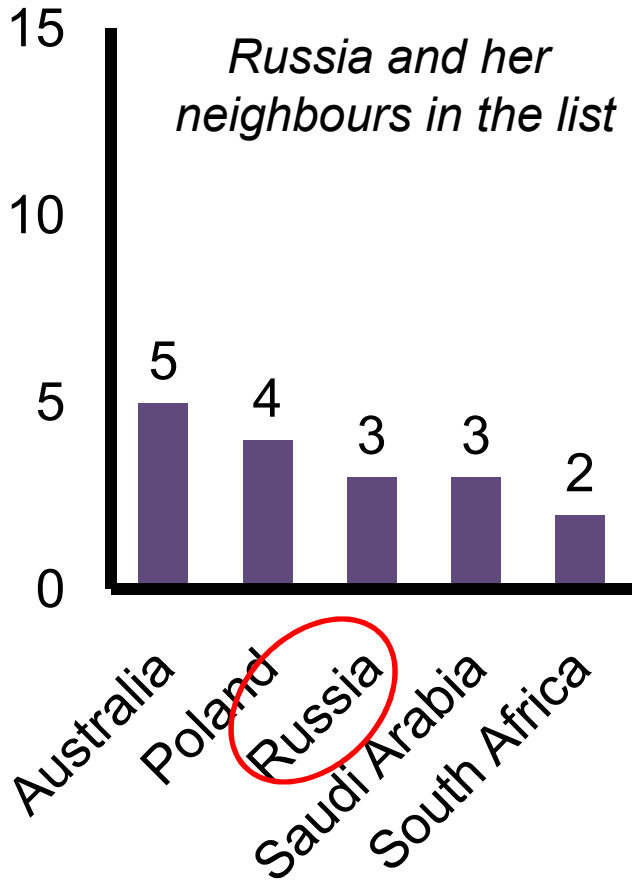
2% GDP

0.3% HPC power
(19th place, Poland – 18th)

Russia's position on the global HPC arena

Number of supercomputers in the world's top 500 list





Russia's position:

2009 year - 8
2010 year - **7**
2011 year - 9
2012 year - 9
2013 year - 13
2014 year - 9
2015 year - 9
2016 year - 10
2017 year -18
2018 year - **17**

* https://www.top500.org/ list of the most powerful supercomputers in the world

Units for measuring supercomputer performance

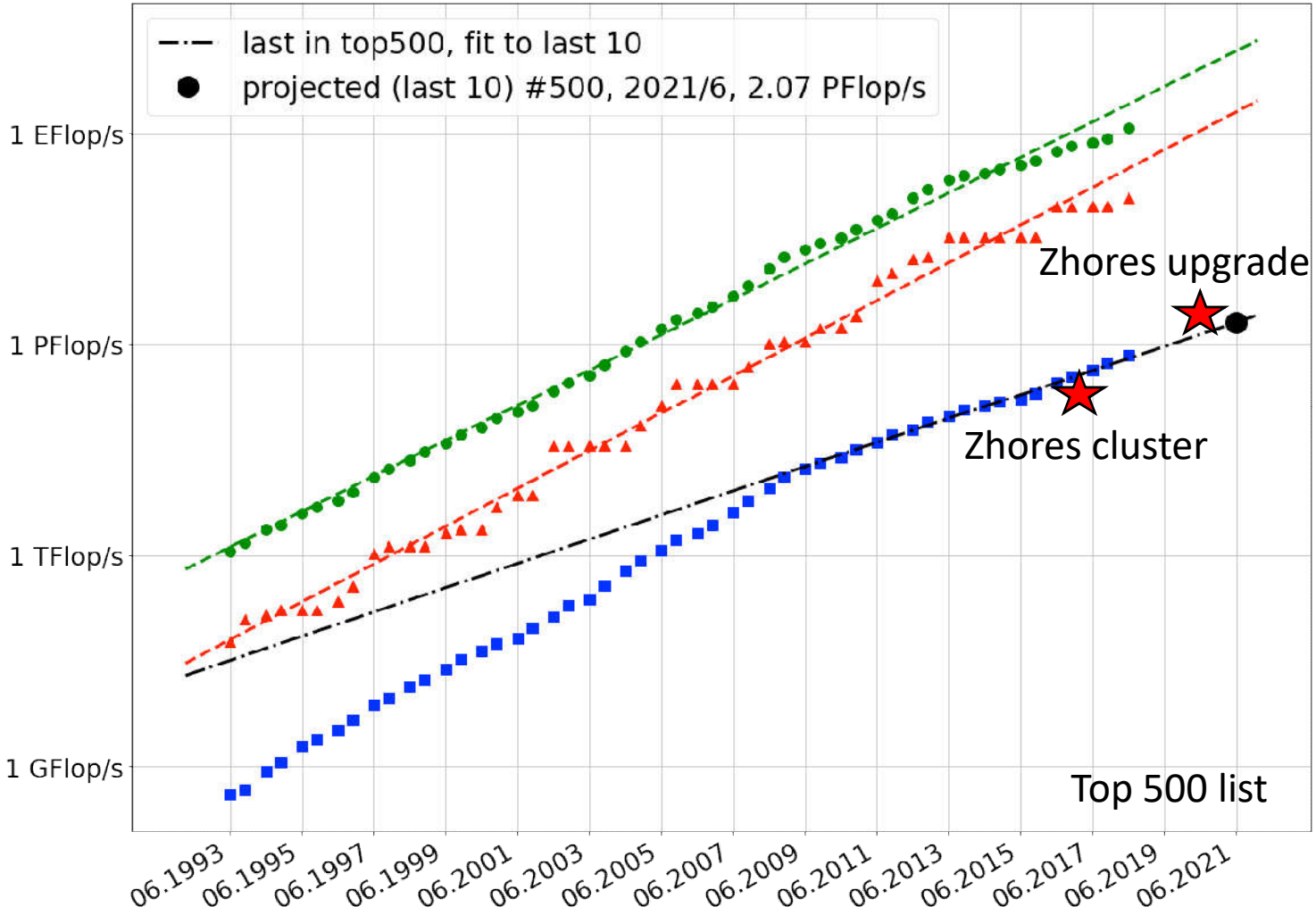
- High Performance Computing (HPC) units are:
  - Flop: floating point operation, usually double precision unless noted
  - Flop/s: floating point operations per second
  - Bytes: size of data (a double precision floating point number is 8 bytes)
- Typical sizes are millions, billions, trillions…

Kilo	Kflop/s = 10^3 flop/sec	Kbyte = $10^3 \sim 2^{10}$ = 1024 bytes (KiB)
Mega	Mflop/s = 10^6 flop/sec	Mbyte = $10^6 \sim 2^{20}$ bytes (MiB)
Giga	Gflop/s = 10^9 flop/sec	Gbyte = $10^9 \sim 2^{30}$ bytes (GiB)
Tera	Tflop/s = 10^{12} flop/sec	Tbyte = $10^{12} \sim 2^{40}$ bytes (TiB)
Peta	Pflop/s = 10^{15} flop/sec	Pbyte = $10^{15} \sim 2^{50}$ bytes (PiB)
Exa	Eflop/s = 10^{18} flop/sec	Ebyte = $10^{18} \sim 2^{60}$ bytes (EiB)
Zetta	Zflop/s = 10^{21} flop/sec	Zbyte = $10^{21} \sim 2^{70}$ bytes (ZiB)
Yotta	Yflop/s = 10^{24} flop/sec	Ybyte = $10^{24} \sim 2^{80}$ bytes (YiB)

- Current fastest (public) machines are petaflop systems
  - Up-to-date list at www.top500.org

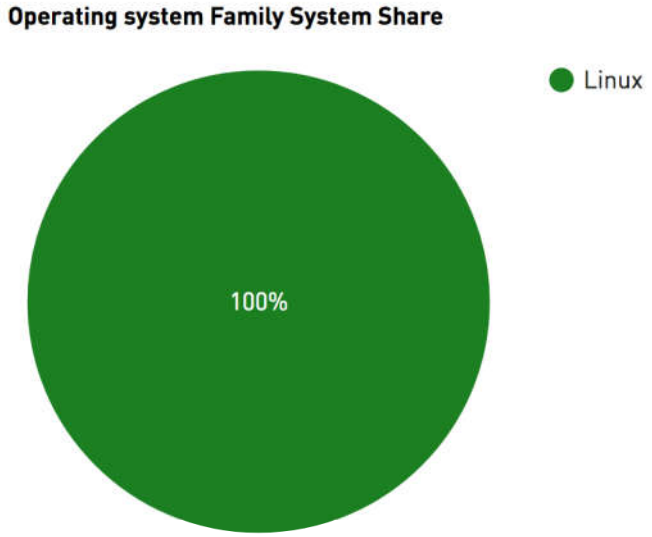
~ 10 Mflop/Watt

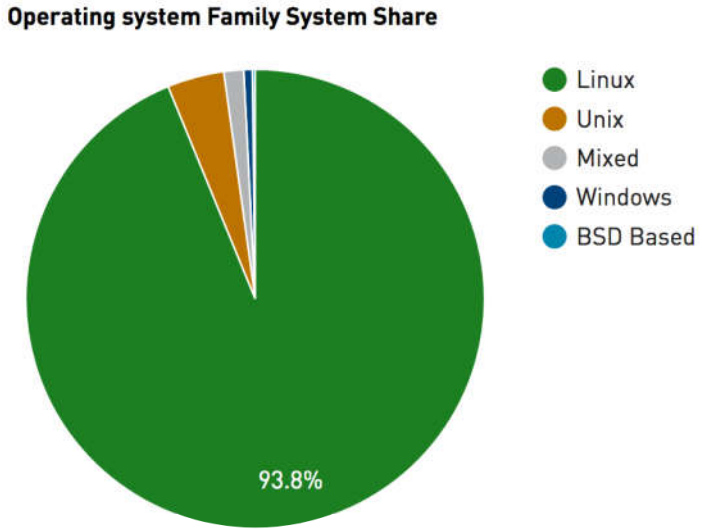
Top 500 – list of most powerful supercomputers

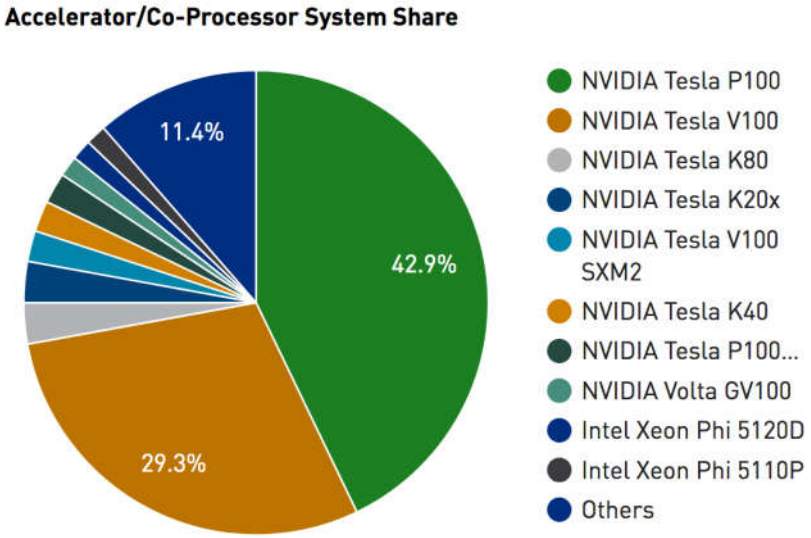


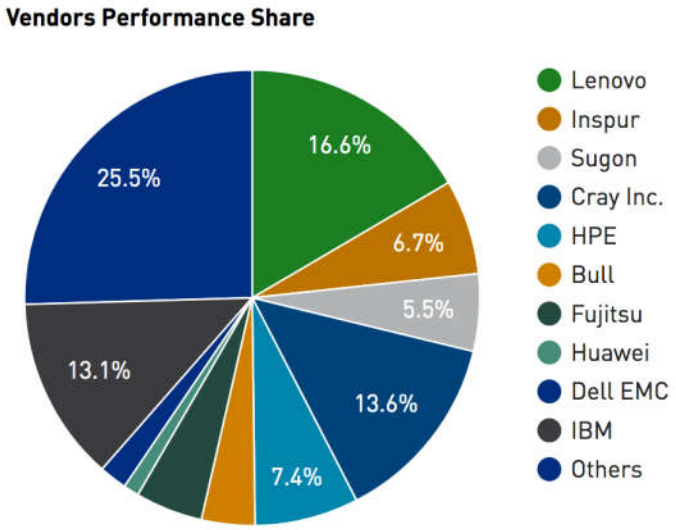
Currently only 3 Russian supercomputers are in the list

Top 500 – list of most powerful supercomputers









Summit (#1 machine) System Overview

System Performance

- Peak performance of 200 petaflops for modeling & simulation
- Peak of 3.3 ExaOps for data analytics and artificial intelligence

Each node has

- 2 IBM POWER9 processors
- 6 NVIDIA Tesla V100 GPUs
- 608 GB of fast memory
- 1.6 TB of NVMe memory

The system includes

- 4608 nodes
- Dual-rail Mellanox EDR InfiniBand network
- 250 PB IBM Spectrum Scale file system transferring data at 2.5 TB/s

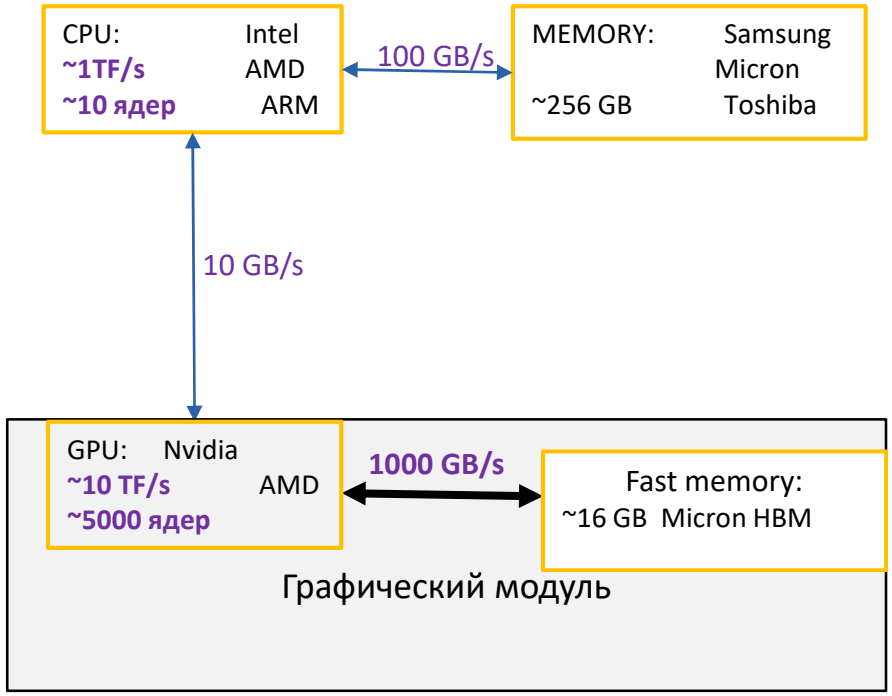
Russia in Top 500

3 entries found.

Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
79	**Lomonosov 2** - T-Platform A-Class Cluster, Xeon E5-2697v3 14C 2.6GHz,Intel Xeon Gold 6126, Infiniband FDR, Nvidia K40m/P-100 , T-Platforms Moscow State University - Research Computing Center Russia	64,384	2,478.0	4,946.8	
283	Cray XC40, Xeon E5-2697v4 18C 2.3GHz, Aries interconnect , Cray Inc./T-Platforms Main Computing Center of Roshydromet Russia	35,136	1,200.3	1,293.0	
487	**Lomonosov** - T-Platforms T-Blade2/1.1, Xeon X5570/X5670/E5630 2.93/2.53 GHz, Nvidia 2070 GPU, PowerXCell 8i Infiniband QDR , T-Platforms Moscow State University - Research Computing Center Russia	78,660	901.9	1,700.2	2,800

Functioning of a single compute node

- *Von Neumann Principle*
  - CPU loads data from memory, operates on it and puts the result back to memory: classical PC
  - Bottleneck is memory transfer CPU-memory which limits the computation speed



- *Specialized accelerator*
  - Applications need to be rewritten
  - Example: GPU computing (CUDA, OpenCL);
  - “Fast” memory inside the GPU;
  - Memory volume is bounded

    (V100: 16 ГБ)

  Memory hierarchy

Measuring the cluster performance

- Benchmarck for performance measurement (LINPACK): $Ax = b$

  where *A* is an *N x N matrix, N is large (10^6)*

  Performance: benchmark – R_{max}; theoretical or peak – R_{peak}

Системы на ключевых местах в списке ТОП-500

№	Cluster	Country	Rmax [Pflop/s]	Rpeak [Pflop/s]	Power [MW]
1	Summit - IBM	USA	122,3	187,6	8,8
2	Sunway TaihuLight	China	93	125,4	15,3
…	…	…			
500	CSCS Cray	Switzerland	0,7	0,84	0,3
…	…	…			
…	**Zhores - Dell**	**Russia, Skoltech**	**0,5**	**0,8**	**0,1**

Factors influencing the performance

- **Technology of semiconductors**
  - CMOS technology;
  - Decreasing the size of the semiconductor structures leads to better performance:

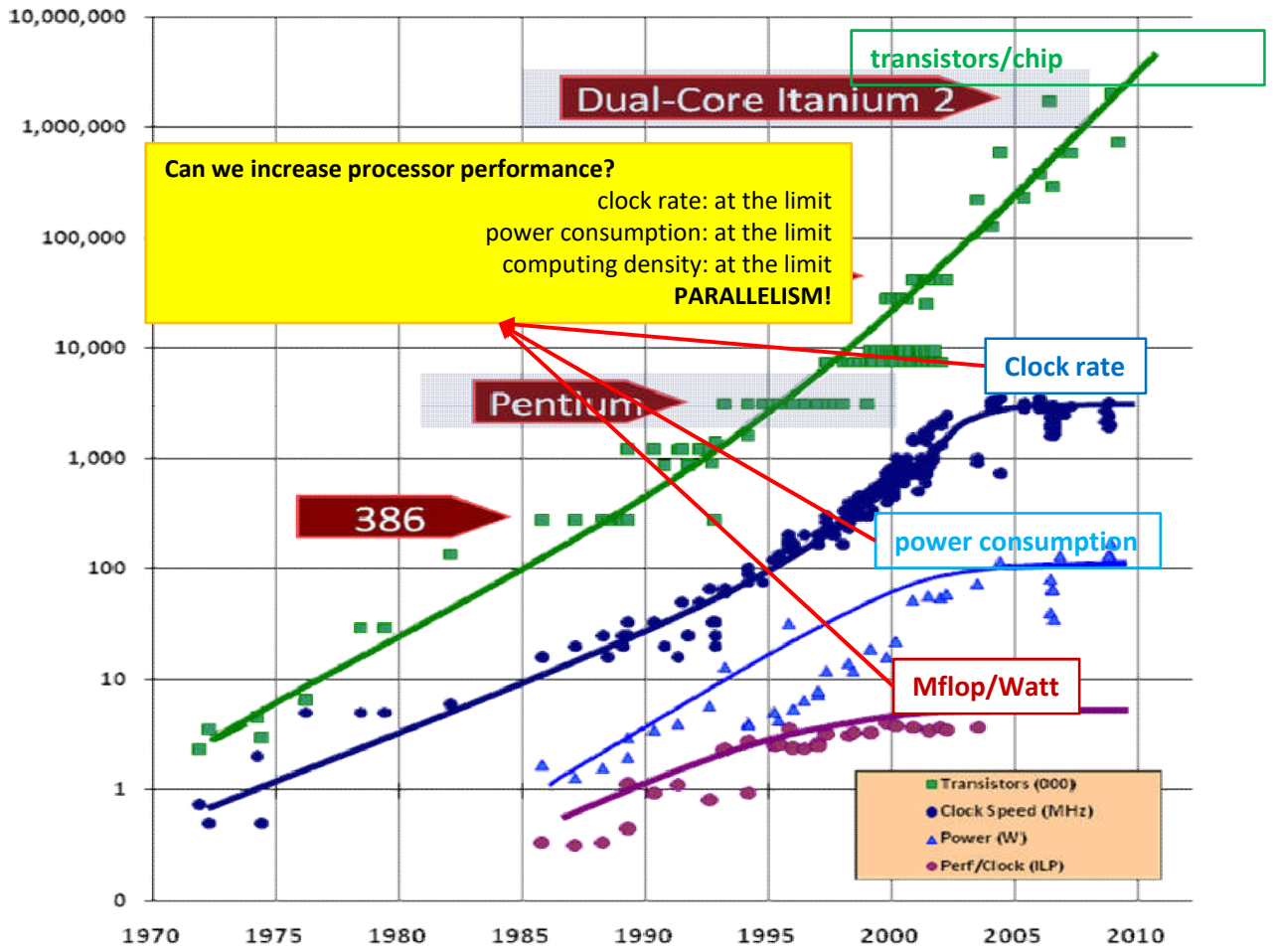
    size, energy consumption, clock rate, price

- **Compute node architecture**
  - Connection of CPUs and GPUs to memory, memory bandwidth, interconnect bandwidth

- **Infrastructure architecture**
  - Energy efficiency;
  - Cooling efficiency (influences the clock rate)
  - Computational density (less path for the signal to take)

Processor technology overview

Scale	Year Intro
130 nm	2001
90 nm	2004
65 nm	2006
45 nm	2008
32 nm	2010
22 nm	2012
14 nm	2014
10 nm	2017 (-19 Intel)
7 nm	2019 AMD (Roma)
5 nm	~*2020 (???)*



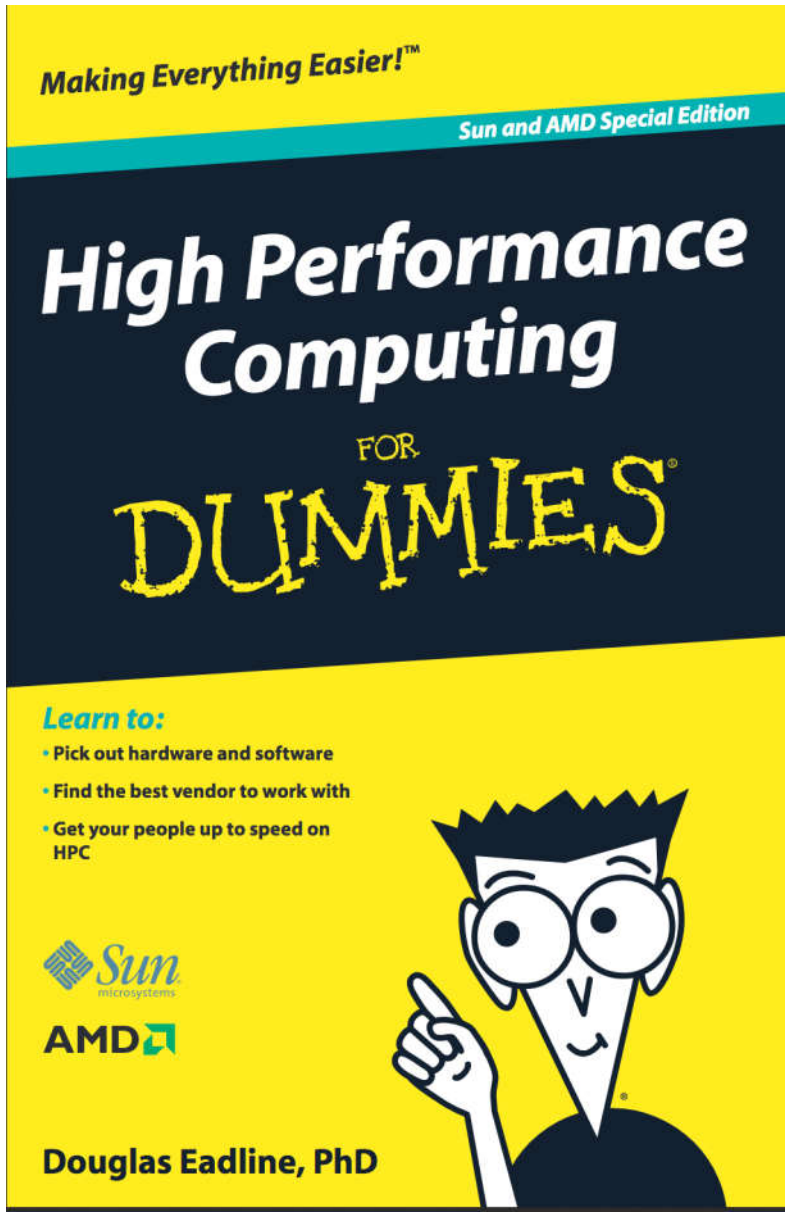
Why the ~~Fastest~~ all (since 2005) Computers are Parallel Computers

Including laptops and cell phones

Think parallel!

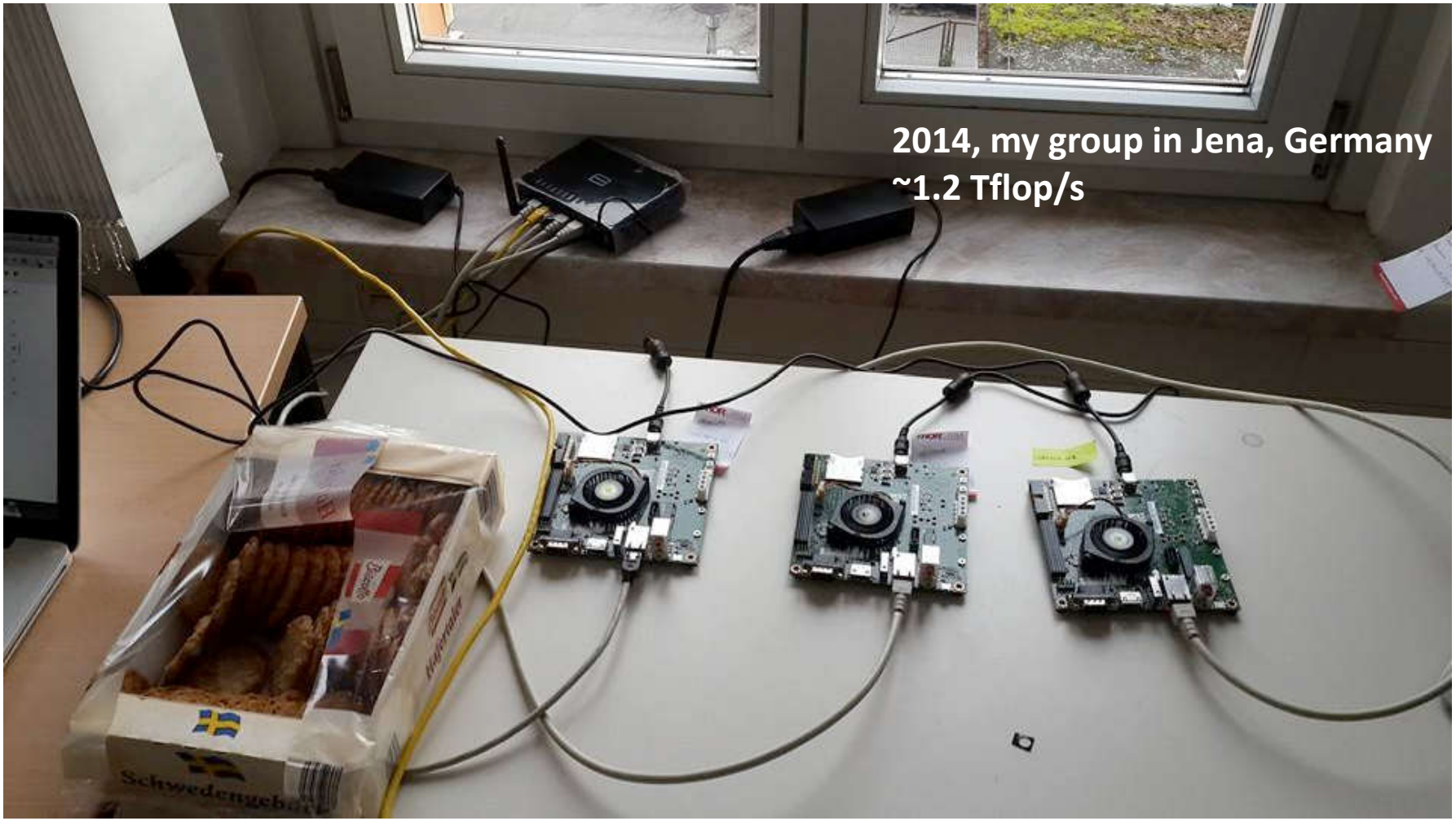
What is a supercomputer and HPC?

What is a supercomputer and HPC?



Highly recommended book
(google for pdf)

What is a supercomputer and HPC?



What is a supercomputer and HPC?

Visionary quotes about computers and HPC

Thomas Watson (chairman of IBM), 1943:
"I think there is a world market for maybe five computers."

Ken Olson (chairman of DEC), 1977:
"There is no reason for any individual to have a computer in his home."

Bill Gates, 1981:
"640K ought to be enough for anybody"

Popular mechanics 1949:
«Where a calculator on the ENIAC is equipped with 18,000 vacuum tubes and weighs 30 tons, computers in the future may have only 1,000 vacuum tubes and weigh only 1.5 tons.»

History of HPC

Abacus, 5 BC

17th century, Pascaline (1642)

Leibniz's Stepped reckoner (1671)

History of HPC

19th century

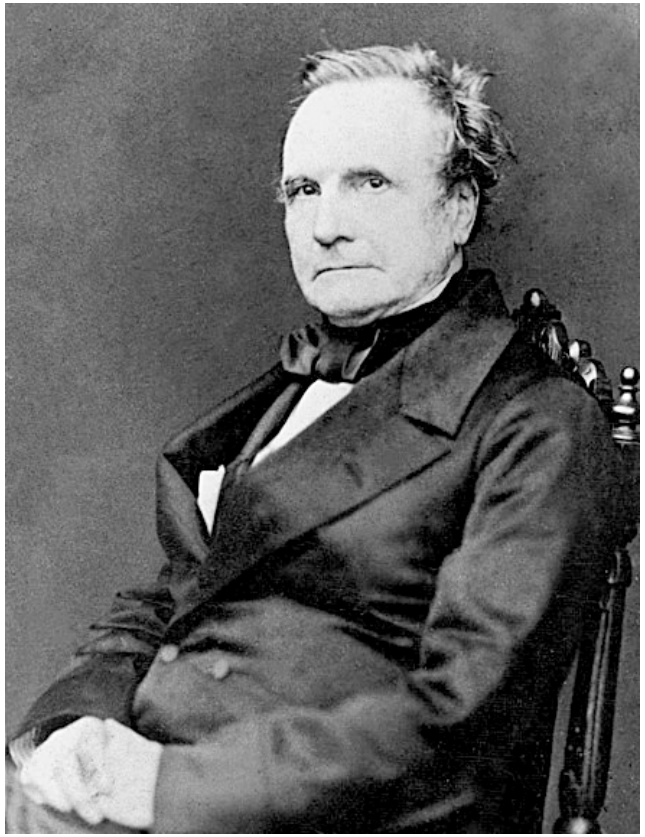
Arithmometer, Charles de Calmar (1820)

Punched card for weaving (1801)

Punched card + calculator, Herman Holerith (1890)

History of HPC

19th century



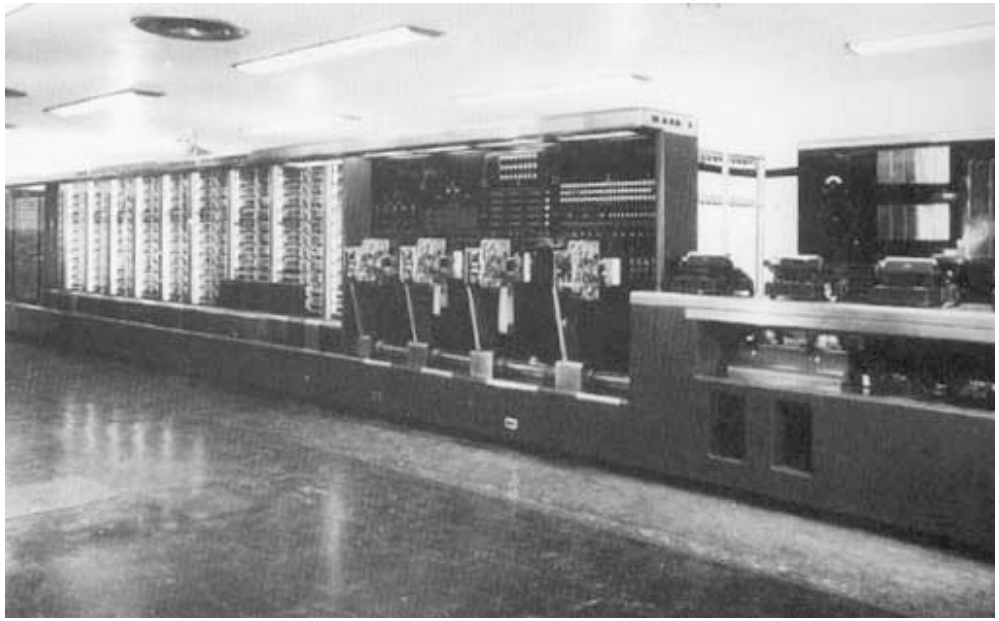
Charles Babbage, concept of fully automated calculation by mechanical means

Ada Lovelace, first programmer, first algorithm

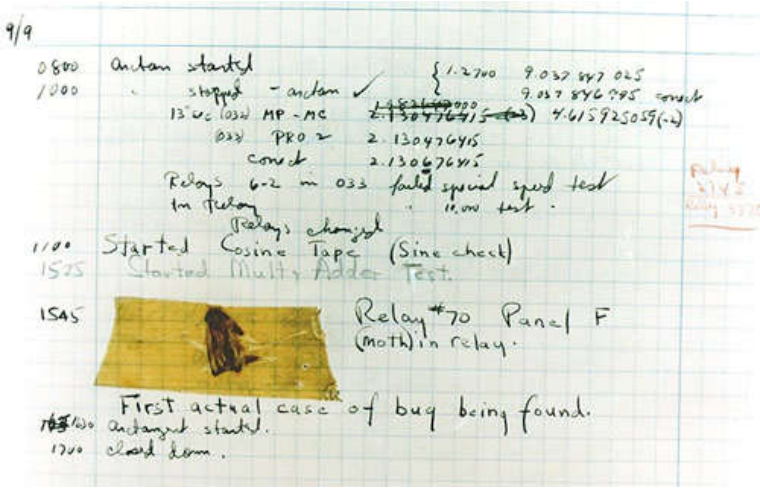
Punched card for weaving (1801)

History of HPC

Konrad Zuse 1938, Z1 – first programmable mechanical computer

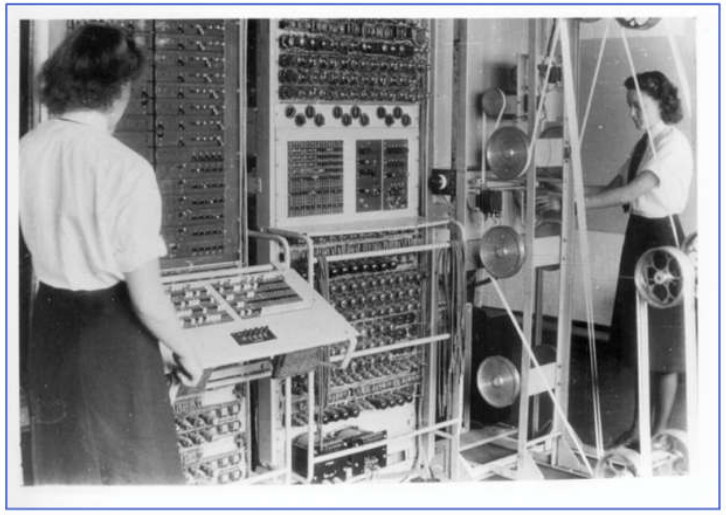


Harvard Mark I

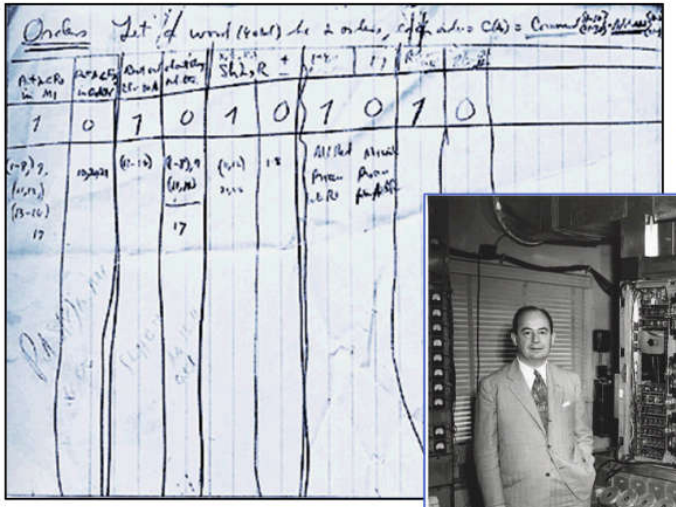


History of HPC

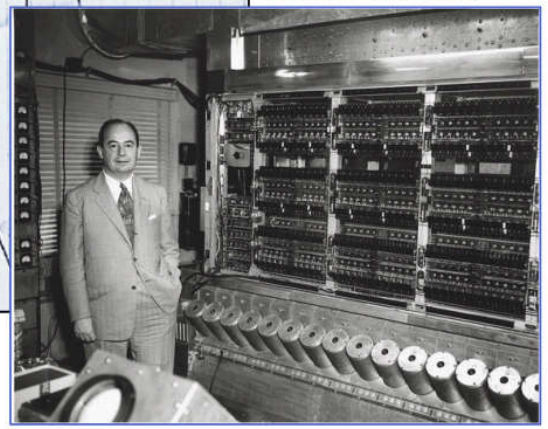
- **Electronic computing developed to meet military needs in WWII**
  - Colossus, Bletchley Park, 1943 … code breaking, dedicated function
  - ENIAC, U. Pennsylvania, 1945 … ballistics tables, plugboard programming
  - Von Neumann
    - Initiated 1946 at Institute for Advanced Studies (IAS)
    - Vacuum tubes, oscilloscopes, assembly language ... many operational challenges
    - But momentous – flexible stored program, reliability architecture, hydrogen bomb
    - IAS, Princeton, 1951
    - MANIAC, LANL, 1952
    - ORACLE, ORNL, 1953



Women's Royal Naval Service operating Colossus during World War II



First line of code written for the von Neumann Digital Computing Project



von Neumann with the IAS machine

History of HPC

Cray-1, 1976, specialized supercomputer
133 MFlops

History of HPC

Scalable „Beowulf“ cluster made of similar compute nodes interconnected between each other.

1994, NASA (Thomas Sterling)

Flagship supercomputer “Zhores” for AI, Big data and HPC

Hybrid energy-efficient architecture

- 74 compute nodes
- 26 nodes with powerful graphic cards Nvidia Tesla V100 (NVLink + RDMA)
- tensor cores for deep learning;
- 90 kWatt power consumption;
- 1PFlops peak performance;
- 0.5 Pbytes storage system
- #6 in Russia
- was installed by our own small team

«Zhores» is a unique for Russia supercomputer capable of solving a wide range of interdisciplinnary problems in machine learning, data science and mathematical modeling in such areas as: biomedicine, image classification, Digital Pharma, Photonics, predictive maintenance, new X- and gamma-ray sources

ZHORES
SUPERCOMPUTER
HIGH PERFORMANCE COMPUTING AND BIG DATA
SKOLTECH CDISE

Other supercomputers

General purpose cluster “Arkuda”

“Arkuda” is a general purpose supercomputer administered by CDISE HPC&Big Data team:

- 54 regular compute blades
- 3 big memory nodes
- 12 nodes with powerful GPUs (NVidia K80 and NVidia M40)
- Performance ~150 TFLOP/s
- Storage system 0.9 PBytes

General purpose cluster “Pardus”

“Pardus” is a general purpose supercomputer administered by CDISE HPC&Big Data team:

- 27 compute nodes in total
- 1 node with powerful GPU (NVidia K80)
- Performance ~25 TFLOP/s
- Storage system 70 Tbytes

We will add a node with 10 GTX 1080 Ti this week.

You should get an account!