C2115 Practical Introduction to Supercomputing

2. lekce

Petr Kulhánek, Jakub Štěpán

kulhanek@chemi.muni.cz

National Centre for Biomolecular Research, Faculty of Science
Masaryk University, Kotlářská 2, CZ-61137 Brno

INVESTMENTS IN EDUCATION DEVELOPMENT

CZ.1.07/2.2.00/15.0233

Contents

- **History, usage and future of computing**
- **Real problems examples**
  Laboratory of Computational Chemistry projects
- **CZ Supercomputer centres overview**
  MetaCentrum, CERIT-SC, IT4 Innovation
- **Foreigin supercomputer centres**
  centres accessible from applications from CZ, Top500
- **Exercise**

History

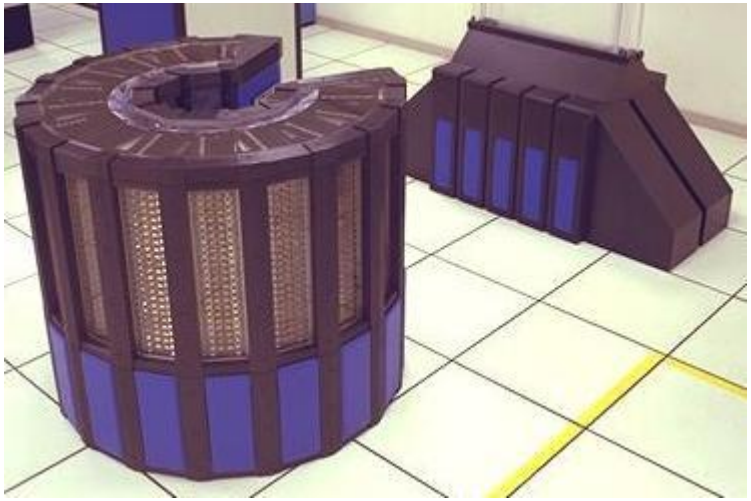
http://en.wikipedia.org/wiki/History_of_computing_hardware

http://www.root.cz

Pavel Tišnovský, Unixové vykopávky

1800 punched cards beginings
1946 ENIAC
1947 transistor discovery
1971 Intel 4004 (4 bit)
1974 Intel 8080 (8 bit)
1976 Intel 8086 (16 bit)
1985 Intel 80386 (32 bit)
2001 IA-64 (64 bit)
2003 AMD64/EM64T (64 bit)

2010 Intel Core i7 980X: @3,33 GHz (6C/12T, Turbo@3,46 GHz): **109 GFLOPS**



1985 Cray 2 **1,9 GFLOPS**

proprietary vector CPU

source: wikipedia.org, intel.com

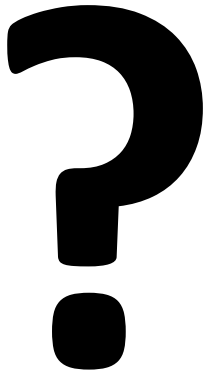
Computing devices usage

Computing devices (computers) influenced all human activities and became undispensable part of our life. **Rapid development** in last 20 years caused that. Computing is used in entertainment, processing and concumtion of information.

Computing devices (at most supercomputers) are used to **solve demanding numerical** problems as:

- Wheather forecast simulations, climete and geologycal simulations (floods, tsunami, earthquake)
- New material and drug design
- Economy modelling
- **Scientific-technical calculations** (**chemistry**, fyzics, mathematics)
- Military purposes (nuclear veapons simulations)
- .....

Future



Maasive usage GPGPU ...

Human brain simulator:

http://www.humanbrainproject.eu/

... quantum computers, ...

Laboratory of Computational Chemistry

- **Project overview**

Group of Computational Chemistry

prof. RNDr. Jaroslav Koča, DrSc.

1 professor
5 senior researchers
2 post-doc students
11 doktoral students
7 bachelor and magister students

Mgr. Zdeněk Kříž, Ph.D.

E-mail: zdenek.kriz@ceitec.muni.cz
Expertise: QM, MD, Docking

RNDr. Petr Kulhánek, Ph.D.

E-mail: petr.kulhanek@ceitec.muni.cz
Expertise: QM, QM/MM, MD, Free Energy

Mgr. Martin Prokop, Ph.D.

E-mail: martin.prokop@ceitec.muni.cz
Expertise: Software dev, Dockig

RNDr. Radka Svobodová, Ph.D.

E-mail: radka.svobodova@ceitec.muni.cz
Expertise: Chemo and Bioinformatics

RNDr. Robert Vácha, Ph.D.

E-mail: robert.vacha@ceitec.muni.cz
Expertise: MD, MC, Coarse Grain, Free Energy

Mgr. Stanislav Kozmon, Ph.D.

E-mail: stano@chemi.muni.cz
Expertise: QM, QM/MM

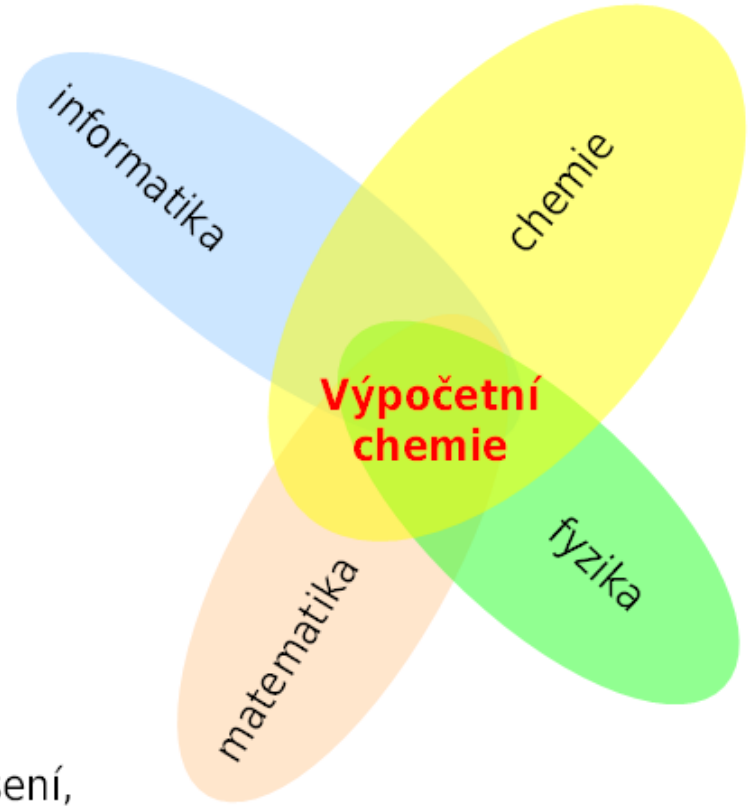
Computational chemistry

Computational Chemistry

is a branch of chemistry that uses **computer** simulation to assist in solving **chemical problems**. It uses methods of **theoretical chemistry**, incorporated into efficient **computer programs**, to calculate the structures and properties of molecules and solids. While computational results normally complement the information obtained by chemical experiments, it can in some cases predict hitherto unobserved chemical phenomena. It is widely used in the **design of new drugs and materials**.

www.wikipedia.org

Interdisciplinary field



algoritmy, CPU/GPU, cluster/grid, symbolické výpočty

(bio)chemické problémy, experimenty, ověřování

analytické řešení, numerická řešení, aproximace

teorie, aproximace

What we do and offer …

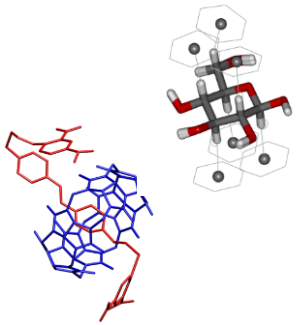
- **expertize in field of computational chemistry**

  Molecular dynamics simulations, docking, chemo a bioinformatics, bonding energy, quantum mechanics, coarse grained models …

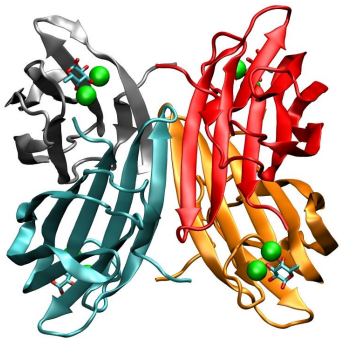
- **Development of computational tools, approaches and software**

  in silico drug design and protein mutagenesis, free energy calulations, fast charge calculations, virtual screening, tunnel and cavity prediction in biomolecules …

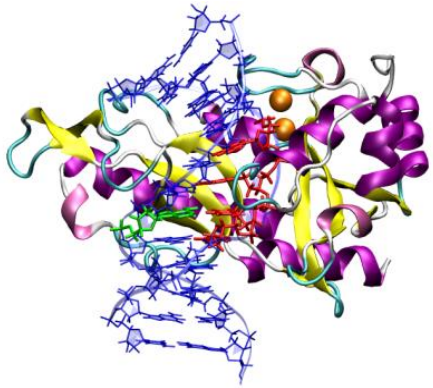
- **studium systémů různých velikostí**



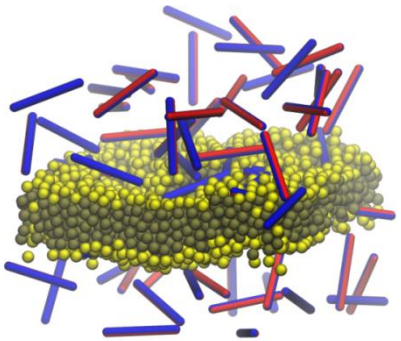
Small complexes



lectins

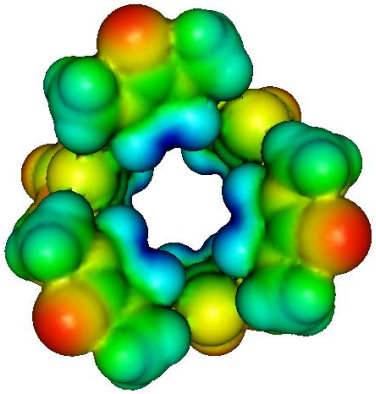


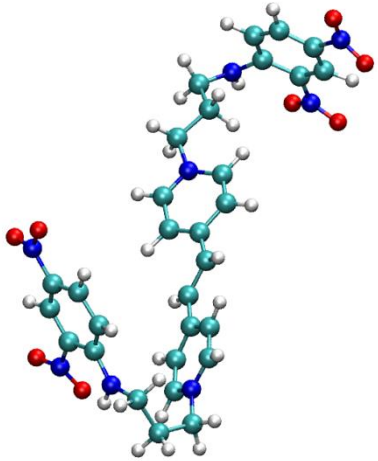
enzymes

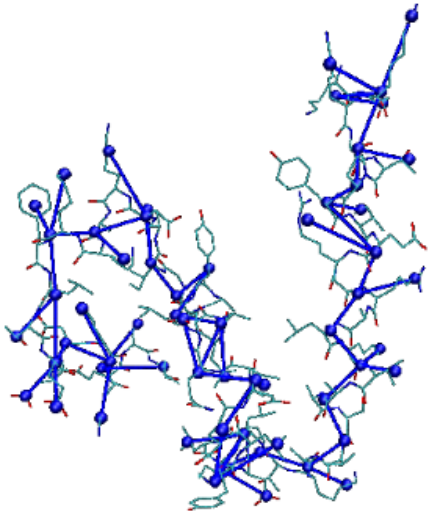


large biomolecular complexes

Theory levels



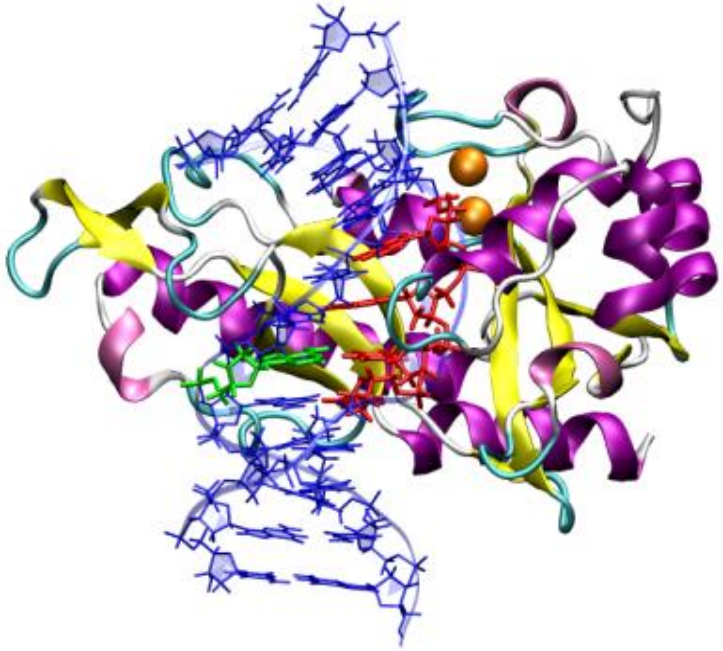




Quantum mechanics	Molecular mechanics	*Coarse-grained* mechanics
atomic resolution		bead resolution
reactivity	Conformational changes	Domains movement, folding
Up to 1'000 atoms *	Up to 1'000'000 atoms *	Up to 1'000'000 beads *
Up to 100 ps *	Up to 1 µs *	Up to ms *

Projects

(bio)molecular systems studies



Quantum chemical calculations

Time independent Schrödinger equation

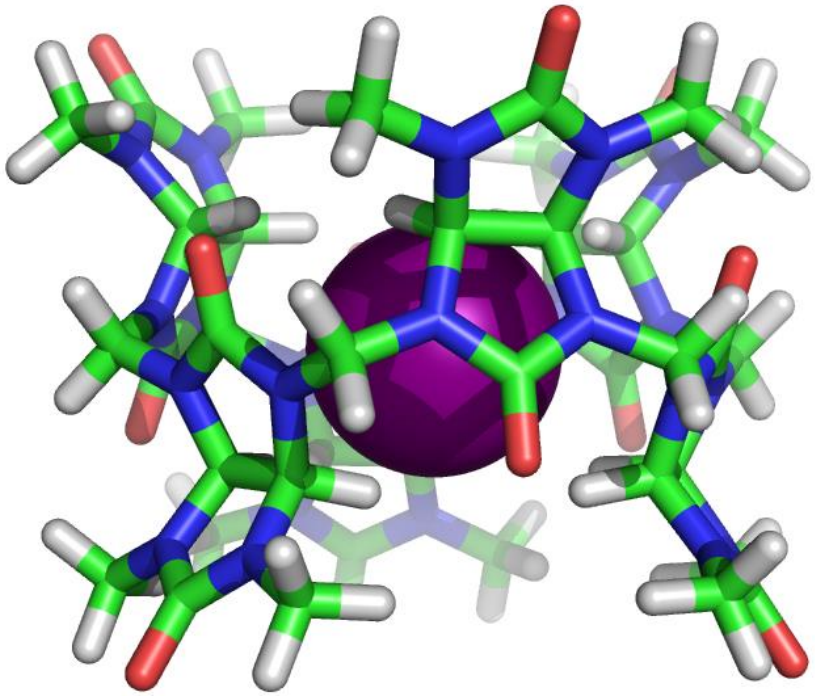
$$\hat{H}\,\psi_k(\mathbf{r}) = E_k\,\psi_k(\mathbf{r})$$

		Metods		
Formal scaling	**HF**	**CI metods**	**MP metods**	**CC metods**
N^4 -> N^2 -> N^1	HF,DFT			
N^5			MP2	CC2 (iterative)
N^6		CISD	MP3, MP4(SDQ)	CCSD (iterative)
N^7			MP4	CCSD(T), CC3 (iterative)
N^8		CISDT	MP5	CCSDT
N^9			MP6	
N^{10}		CISDTQ	MP7	CCSDTQ (iterative)

Scaling, time demands: http://en.wikipedia.org/wiki/Time_complexity
HF - Hartree–Fock metod, DFT – density functional theory,
CI – configuration interaction methods, MP - Møller–Plesset perturbation teory,
CC – coupled cluster metod, N – base function number

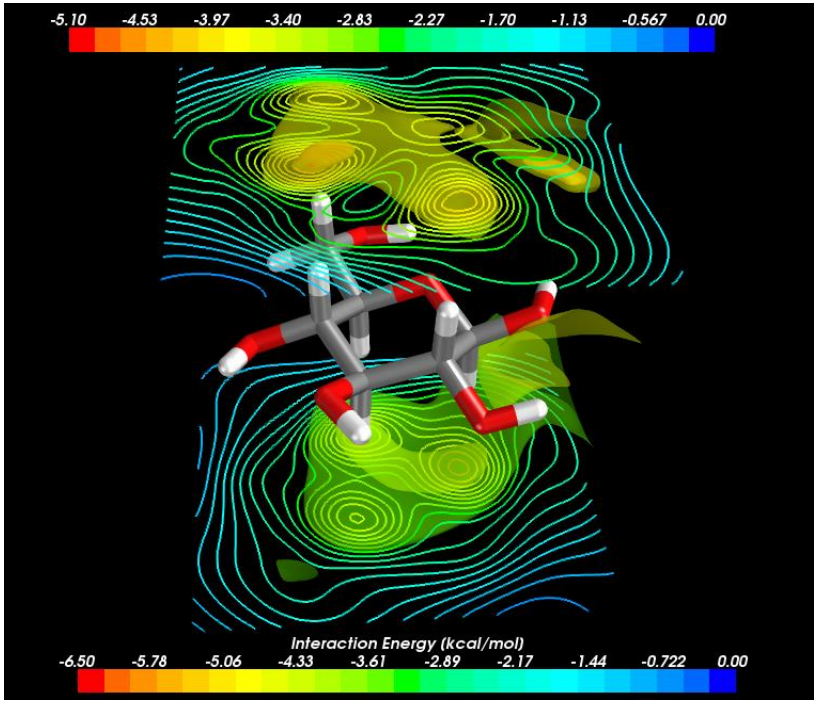
Jensen, F. Introduction to computational chemistry; 2nd ed.; John Wiley & Sons: Chichester, England; Hoboken, NJ, 2007.

Quantum chemical calculations



supramoleculr complexes

Sacharides binding abilities



Molecular mechanics

Schrödinger equation => quantum mechanics view

$$H_a \Psi(r) = E(R) \Psi(r)$$

Approximation uses classical physics not considering explicit electron movement (electrons are included in empirical parameters)

$$E(R) = E_{bonds} + E_{angles} + E_{torsions} + E_{el} + E_{vdw}$$

Classical physics => mechanical view

Bonding term

Non-bonding terms

Formal scaling: N^2 -> $N \log_2 N$

N – atom number

Molecular dynamics

$$-\frac{\partial E(\mathbf{R})}{\partial \mathbf{R}} = \mathbf{F} \qquad \boxed{\mathbf{F}_i = m_i \mathbf{a}_i} \qquad \mathbf{a}_i = \frac{d^2 \mathbf{r}_i}{dt^2}$$

II. Newton law of motion (law of force)

$$-\frac{\partial^2 E(\mathbf{R})}{\partial \mathbf{r}_i^2} = m_i \frac{d^2 \mathbf{r}_i}{dt^2}$$

Equation system of second order needs numerical solution

↓

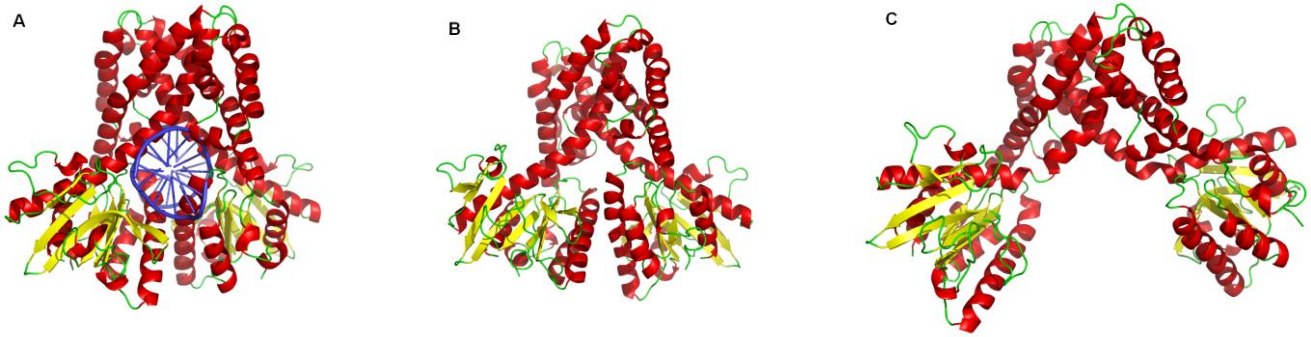
Discrete molecular movement in short time steps

Determined by fastest movement (bond vibration) → **1 fs** **Typical integration step**

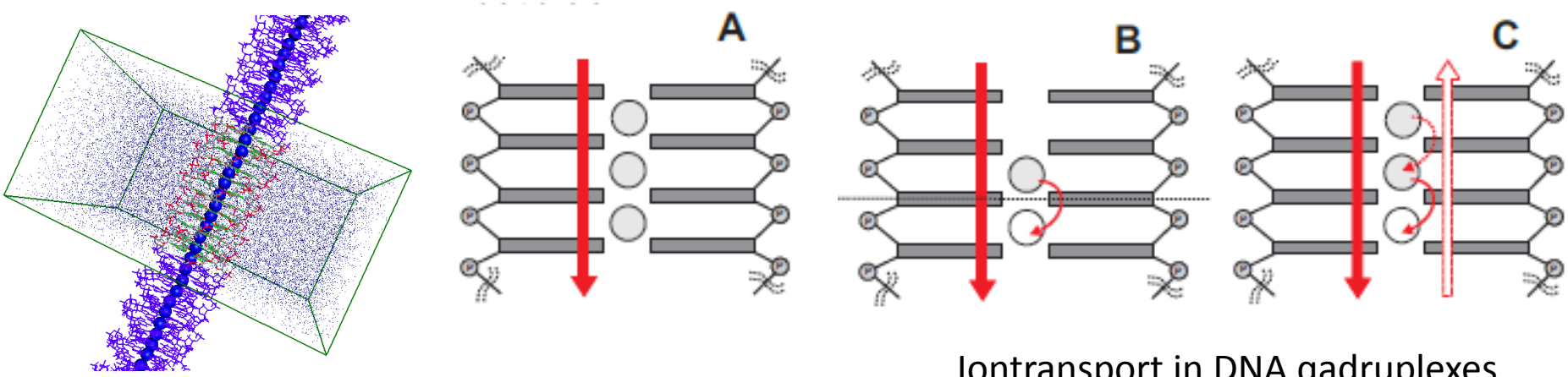
Integration imperfections are corrected by usage of **thermostats** and **barostats**, that ensure given simulation conditions.

Conformational changes, transport

- **Conformational changes and transport are important for biomolecular system function**

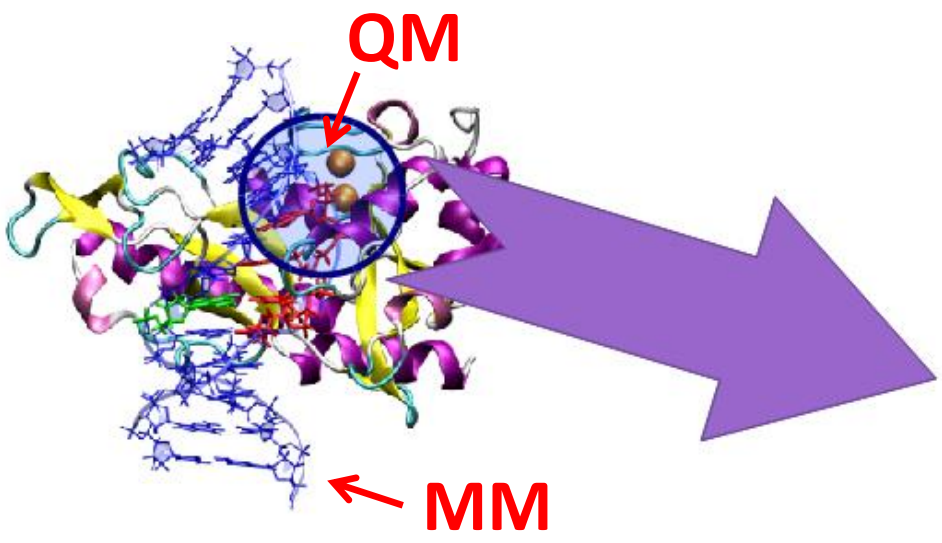


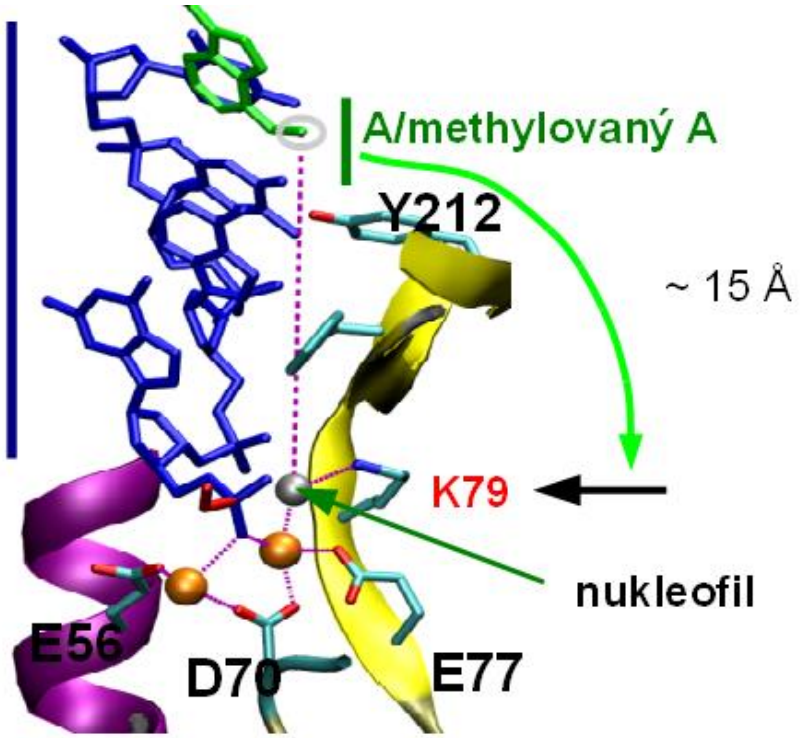
endonuclease BsoBI opening



Iontransport in DNA qadruplexes

Enzyme reaction mechanisms QM/MM



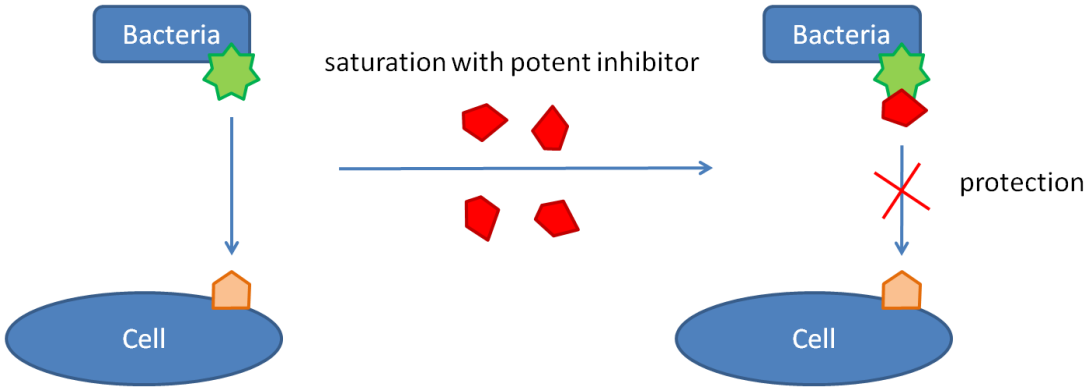


Je součastí opravných mechanismů poškozené DNA v bakteriích.

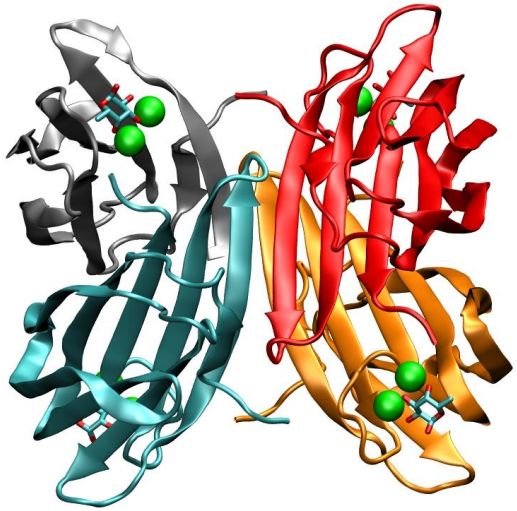
Katalyzovaná reakce – hydrolýza fosfodiesterové vazby

Virtual screening

Motivation

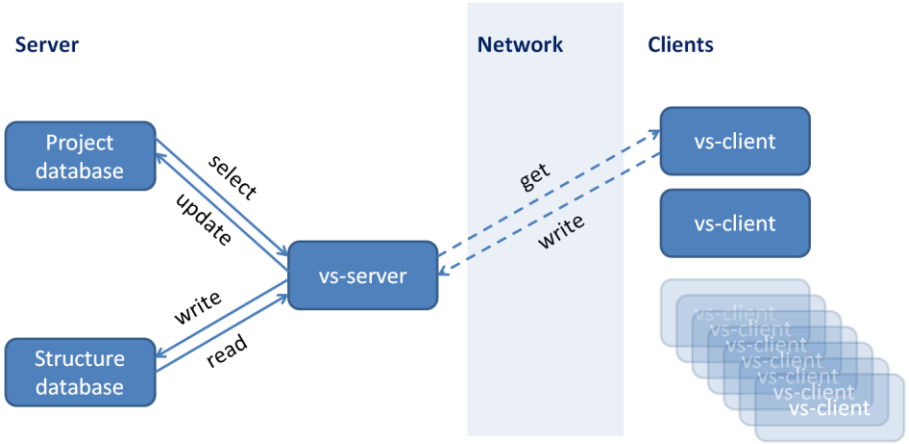


Cíl



Lectin PA-IIL

Software tools



Performance

- Autodock Vina
- Heterogennous computational sources
- One docking cca 1-10 min/ligand
- ca **900 parallel tasks**
- Search speed cca **250 000 ligands per day**

Contacts

Laboratory of computational chemistry

National Centre for Biomolecular Research, UKB, Pavilon A4

http://lcc.ncbr. muni.cz

Seminars LCC group are on Thursday at 10 A.M. in room 2.11/A4.

Computing centres in CZ

- MetaCentrum
- CERIT-SC
- MetaCentrum VO
- IT4 Innovation

MetaCentrum

http://www.metacentrum.cz/

Long-term goal of the **MetaCentrum** project is operation and coordination of distributed computing and data storage infrastructure accompanied by an appropriate support environment and continual expansion of available computational capacities.

The main aim of the project is constitution of a virtual computer that allows effective utilization of installed facilities in the frame of supercomputing project and solving tasks whose memory and/or CPU requirements exceeds possibility of individual single supercomputing centers.

MetaCentrum is **CESNET** activity.

CERIT-SC

http://www.cerit-sc.cz/

Center CERIT-SC (CERIT Scientific Cloud) is national center providing flexible storage and computional facilities and conected services, including support of their experimental support. Moreover center does research in field of flexible e-infrastructure and cooperates on research with users.

Center CERIT-SC developed from Supercomputing center Brno (SCB), that is part of Institute Ústav výpočetní techniky (ÚVT) of Masaryk University (MU).

Center CERIT-SC plans to provide more than 3500 cores and 3,5 PB storage by 2013. These resources will be continuously put into operation and are accessible through National Grid Infrastructure, cloud and other interfaces.

MetaCentrum VO

http://metavo.metacentrum.cz/

Associates resources provided by MetaCentre, project CERIT-SC and other partners.

- National Grid Infrastructure
- OS Debian
- ca **2500 CPU** cores
- **CEITEC/NCBR own resources cca 850** CPU cores
- 3 x 100 TB storage capacity
- cca **3 TB per user**

- Account may be provided to any CZ university student.
- Account is not limited to particular project and is granted for one year.
- Extension is done based on year report.

MetaCentrum VO

http://metavo.metacentrum.cz/

Associates resources provided by MetaCentre, project CERIT-SC and other partners.

- National Grid Infrastructure
- OS Debian
- ca **2500 CPU** cores
- **CEITEC/NCBR own resources cca 850** CPU cores
- 3 x 100 TB storage capacity
- cca **3 TB per user**

Prakticky: infrastruktura, klastry, datové úložiště, zadávání úloh

- Account may be provided to any CZ university student.
- Account is not limited to particular project and is granted for one year.
- Extension is done based on year report.

IT4 Innovation

http://www.it4i.cz/

IT4Innovations project goals are to build national centre of excelent research in information technologies. Part of the project is purchase of supercomputer, that is to be started by year 2014.

Centre basis will be in **computing**, that is devided into three interconnected research fields:

- **IT4People (Information Technology for People)** – research directed to quality of life improvement using modern information technologies.
- **SC4Industry (Supercomputing for Industry)** – supercomputing in solving industry problems, modelling in field of nature science and nanotechnologies (shape optimisations, material design, biomechanical simulations, ...).
- **Theory4IT (Theory for Information Technology)** – field directed to basic research, development of new calculation methods (data mining, anthill theory).

Current call: to 4. 3 2013

Project is supported by five subjects: Vysoká škola báňská-Technická univerzita Ostrava, Ostravská univerzita v Ostravě, Slezská univerzita v Opavě, Vysoké učení technické v Brně a Ústav geoniky AV ČR.

Foreign computing centres

- PRACE
- TOP500

PRACE

PRACE: Partnership for Advanced Computing in Europe **http://www.prace-ri.eu/**

Project Types:

- **Multi-year Access** is available to major European projects or infrastructures that can benefit from PRACE resources and for which Project Access is not appropriate.
- **Project Access** is intended for individual researchers and research groups including multi-national research groups and has a one year duration. Calls for Proposals for Project Access are issued twice yearly (February and September).
- **Preparatory Access** is intended for resource use required to prepare proposals for Project Access. Applications for Preparatory Access are accepted at any time.

Next call: 13. FEB – 26. MAR 2013

PRACE - členové

Austria: JKU - Johannes Kepler University of Linz
Belgium: DGO6-SPW - Direction générale opérationnelle de l'Économie, de l'Emploi et de la Recherche – Service Public de Wallonie
Bulgaria: NCSA - Executive agency "Electronic communication networks and information systems"
Cyprus: CaSToRC – Computation-based Science and Technology Research Center, The Cyprus Institute
Czech Republic: VŠB - Technical University of Ostrava
Denmark: DeIC - Danish e-Infrastructure Cooperation
Finland: CSC - IT Center for Science Ltd.
France: GENCI - Grand Equipement National de Calcul Intensif
Germany: GCS - GAUSS Centre for Supercomputing e.V
Greece: GRNET - Greek Research and Technology Network S.A.
Hungary: NIIFI - National Information Infrastructure Development Institute
Ireland: ICHEC - Irish Centre for High-End Computing
Israel: IUCC - Inter-University Computation Center
Italy: CINECA - Consorzio Interuniversitario
Norway: SIGMA – UNINETT Sigma AS – The Norwegian Metacenter for Computational Science
The Netherlands: SURFSARA: SARA Computing and Networking Services
Poland: PSNC – Instytut Chemii Bioorganicznej Pan – Institute of Bioorganic Chemistry – Poznan Supercomputing and Networking Center
Portugal: Universidade de Coimbra
Serbia: IPB - Institute of Physics Belgrade
Slovenia: ULFME - University of Ljubljana, Faculty of Mechanical Engineering
Spain: BSC – Barcelona Supercomputing Center – Centro Nacional de Supercomputación
Sweden: Vetenskapsrådet – Swedish Research Council
Switzerland: ETH – Eidgenössische Technische Hochschule Zürich – Swiss Federal Institute of Technology, Zürich
Turkey: UYBHM – Ulusal Yuksek Basarimli Hesaplama Merkezi, Istanbul Technical University – National Center for High Performance Computing
UK: EPSRC – The Engineering and Physical Sciences Research Council

PRACE Summer of HPC

http://summerofhpc.prace-ri.eu/

Summer of HPC is a PRACE programme that offers summer placements at HPC centres across Europe. Up to twenty top applicants from across Europe will be selected to participate. Participants will spend two months working on projects related to PRACE technical or industrial work to produce a visualisation or video. The programme will run from July 1st, to August 30th 2013 and will include a kick-off training week.

Flights, accommodation & a stipend will be provided to all successful applicants; all you need to bring is your interest in computing and some enthusiasm! Prizes will be awarded for the best participants.

Call: 25. JAN – 17. FEB 2013

PRACE Summer of HPC

http://summerofhpc.prace-ri.eu/

Summer of HPC is a PRACE programme that offers summer placements at HPC centres across Europe. Up to twenty top applicants from across Europe will be selected to participate. Participants will spend two months working on projects related to PRACE technical or industrial work to produce a visualisation or video. The programme will run from July 1st, to August 30th 2013 and will include a kick-off training week.

Flights, accommodation & a stipend will be provided to all successful applicants; all you need to bring is your interest in computing and some enthusiasm! Prizes will be awarded for the best participants.

Home work:
1. Exercise of Code Test in Fortran

Call: 25. JAN – 17. FEB 2013

TOP500

http://www.top500.org/

TOP500 is project, that lists 500 most powerfull computers on the Earth.

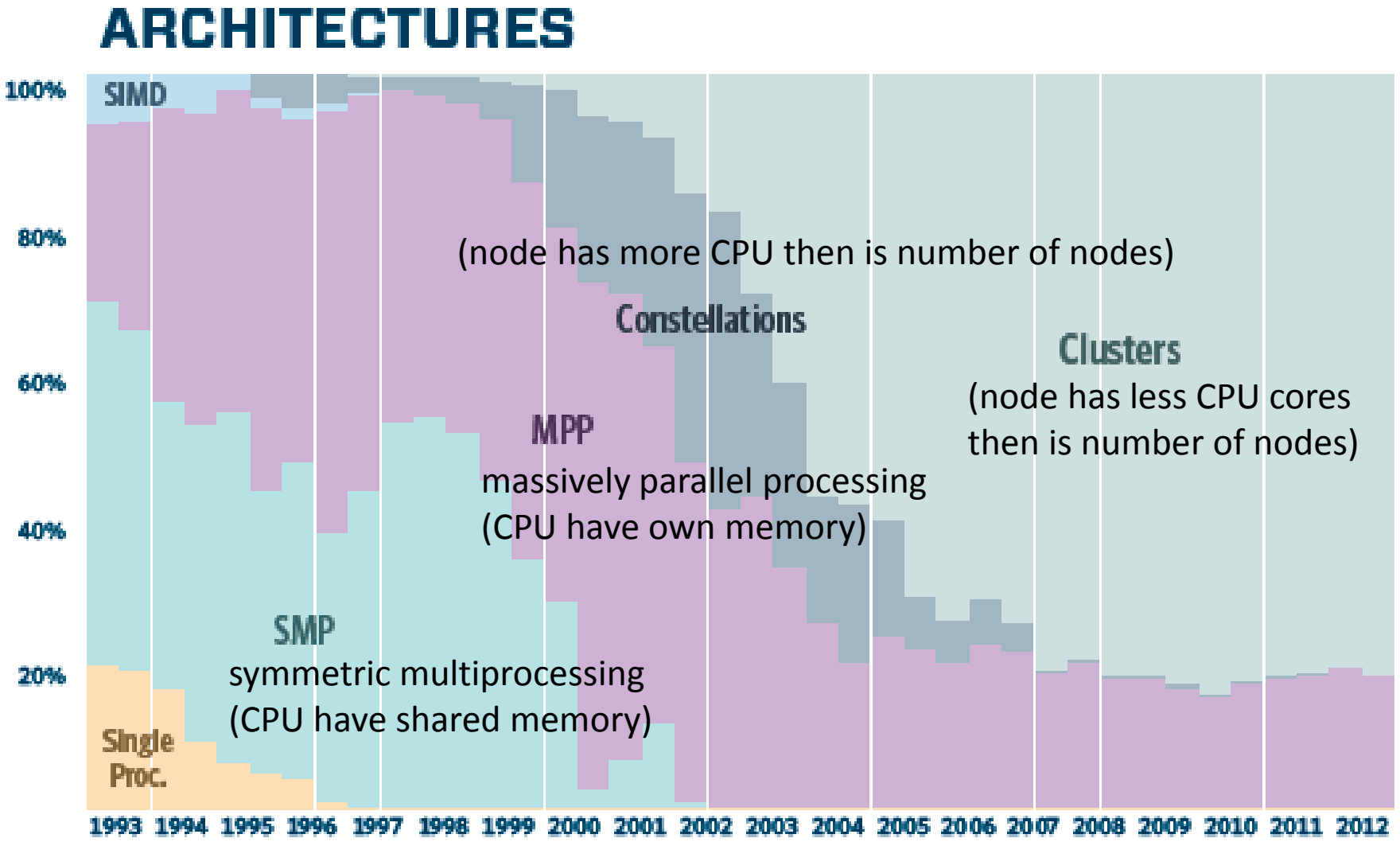
TOP500 benchmark

Our simple **TOP500** approach does not define "supercomputer" as such, but we use a benchmark to rank systems and to decide on whether or not they qualify for the TOP500 list. The benchmark we decided on was **Linpack**, which means that systems are ranked only by their ability to solve a set of linear equations, **A x = b**, using a **dense random matrix A**.

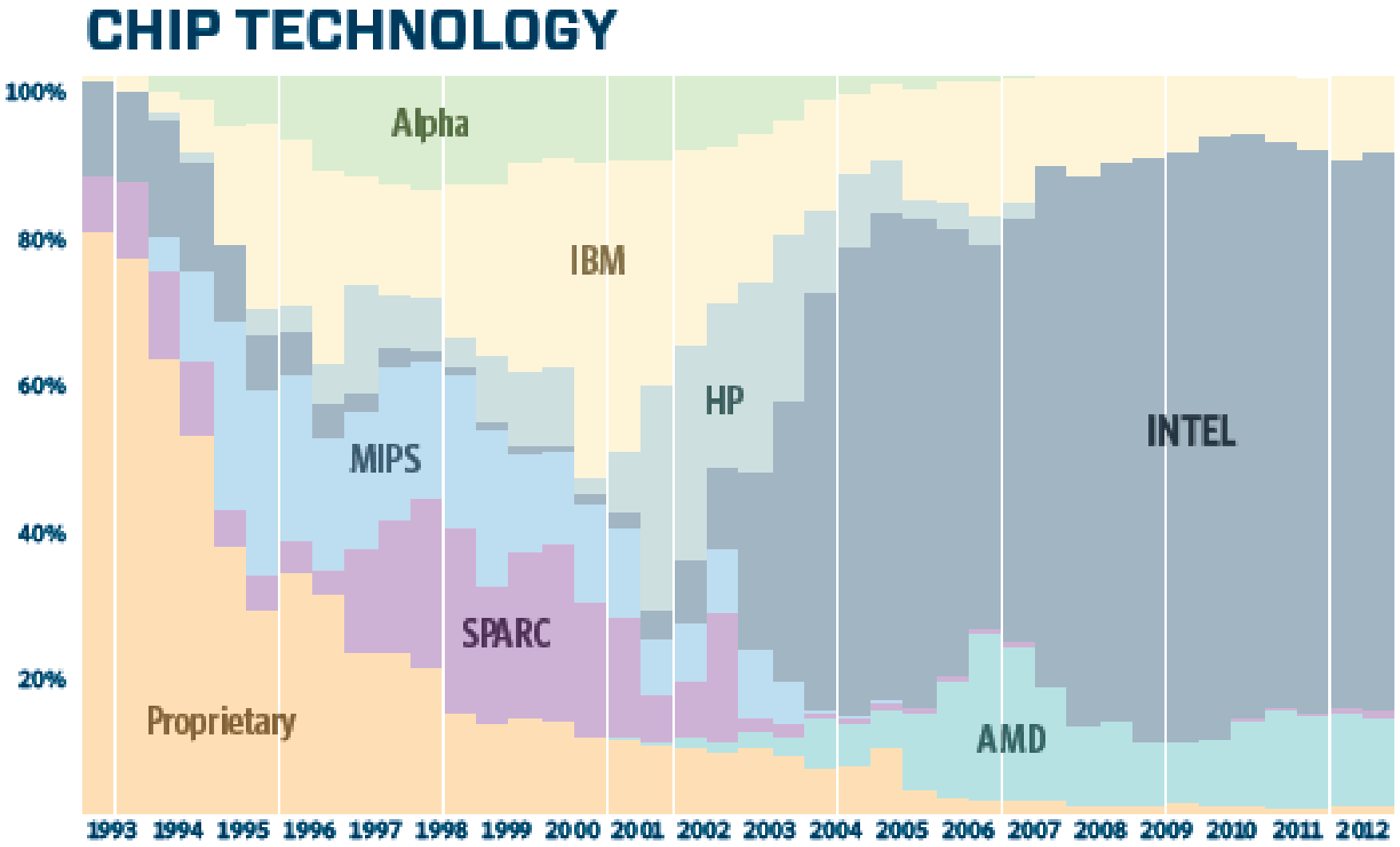
November 2012

	NAME	SPECS	SITE	COUNTRY	CORES	RMAX PFLOP/S	POWER MW
1	TITAN	Cray XK7, Operon 6274 16C 2.2 GHz + Nvidia Kepler GPU, Custom Interconnect	DOE/OS/ORNL	USA	560,640	**17.6**	8.3
2	SEQUOIA	IBM BlueGene/Q, Power BQC 16C 1.60 GHz, Custom Interconnect	DOE/NNSA/LLNL	USA	1,572,864	**16.3**	7.9
3	K COMPUTER	Fujitsu SPARC64 VIIIfx 2.0GHz, Custom Interconnect	RIKEN AICS	Japan	705,024	**10.5**	12.7
4	MIRA	IBM BlueGene/Q, Power BQC 16C 1.60 GHz, Custom Interconnect	DOE/OS/ANL	USA	786,432	**8.16**	3.95
5	JUQUEEN	IBM BlueGene/Q, Power BQC 16C 1.60 GHz, Custom Interconnect	Forschungszentrum Jülich	Germany	393,216	**4.14**	1.97

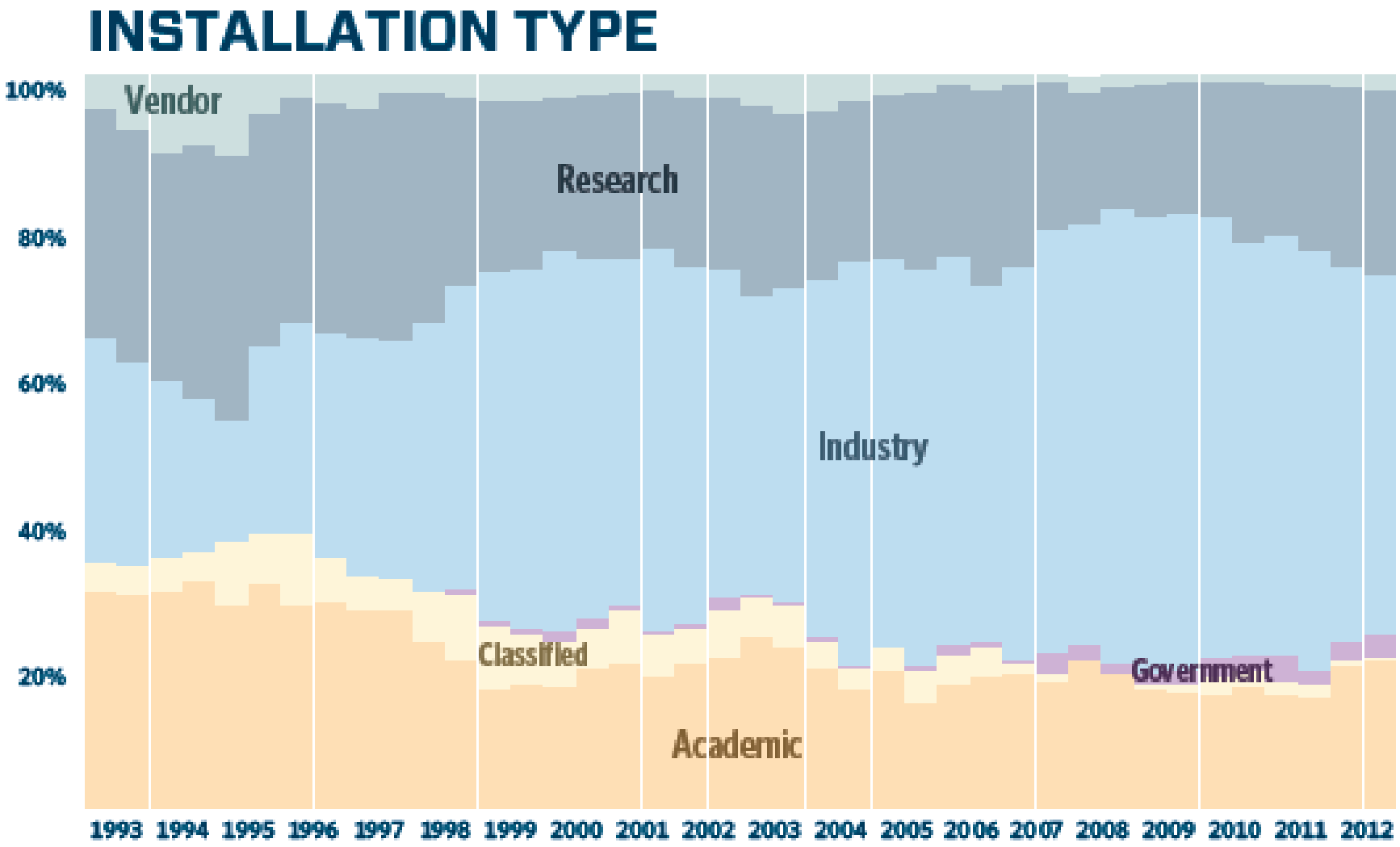
TOP500 – Topology



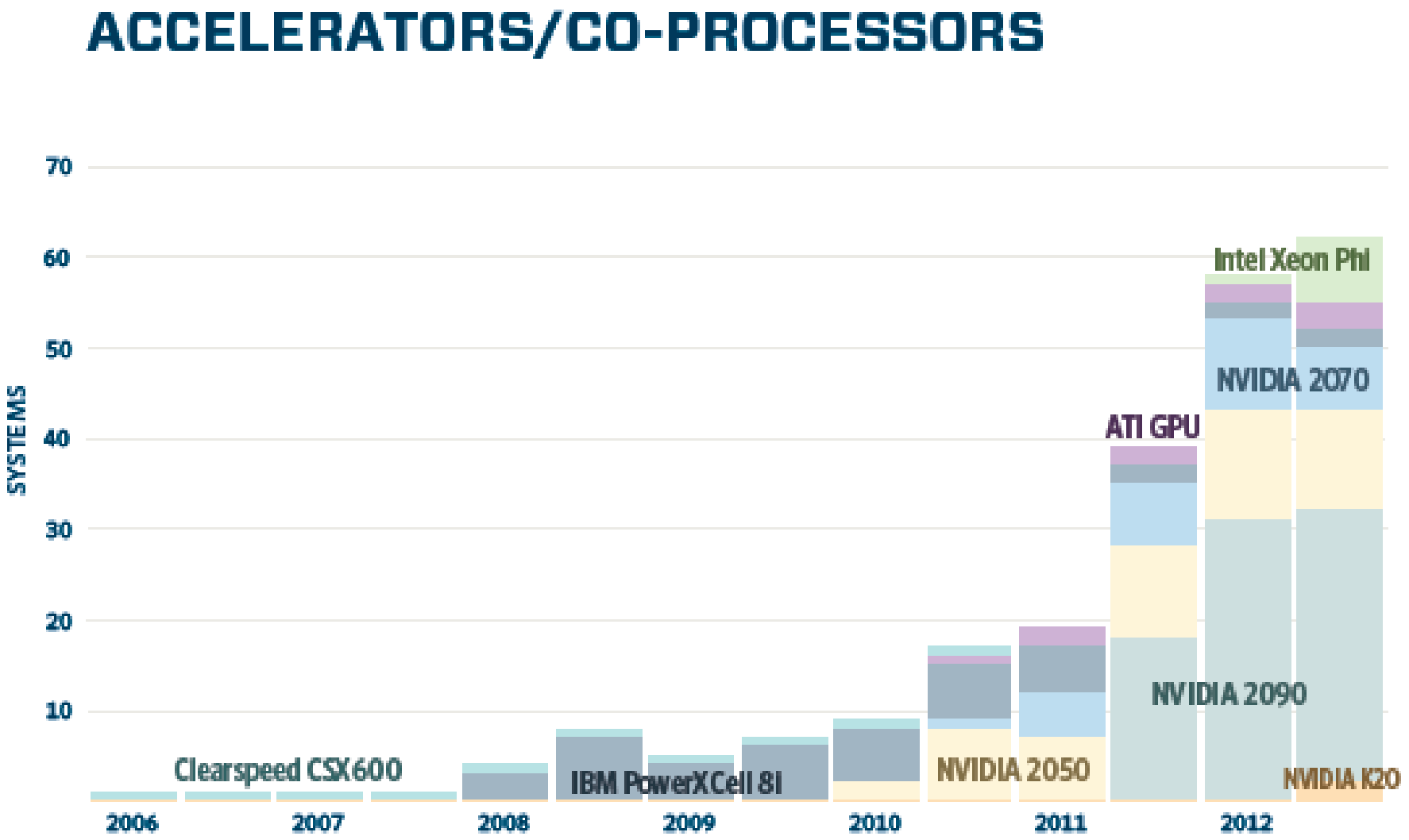
TOP500 – CPU architecture



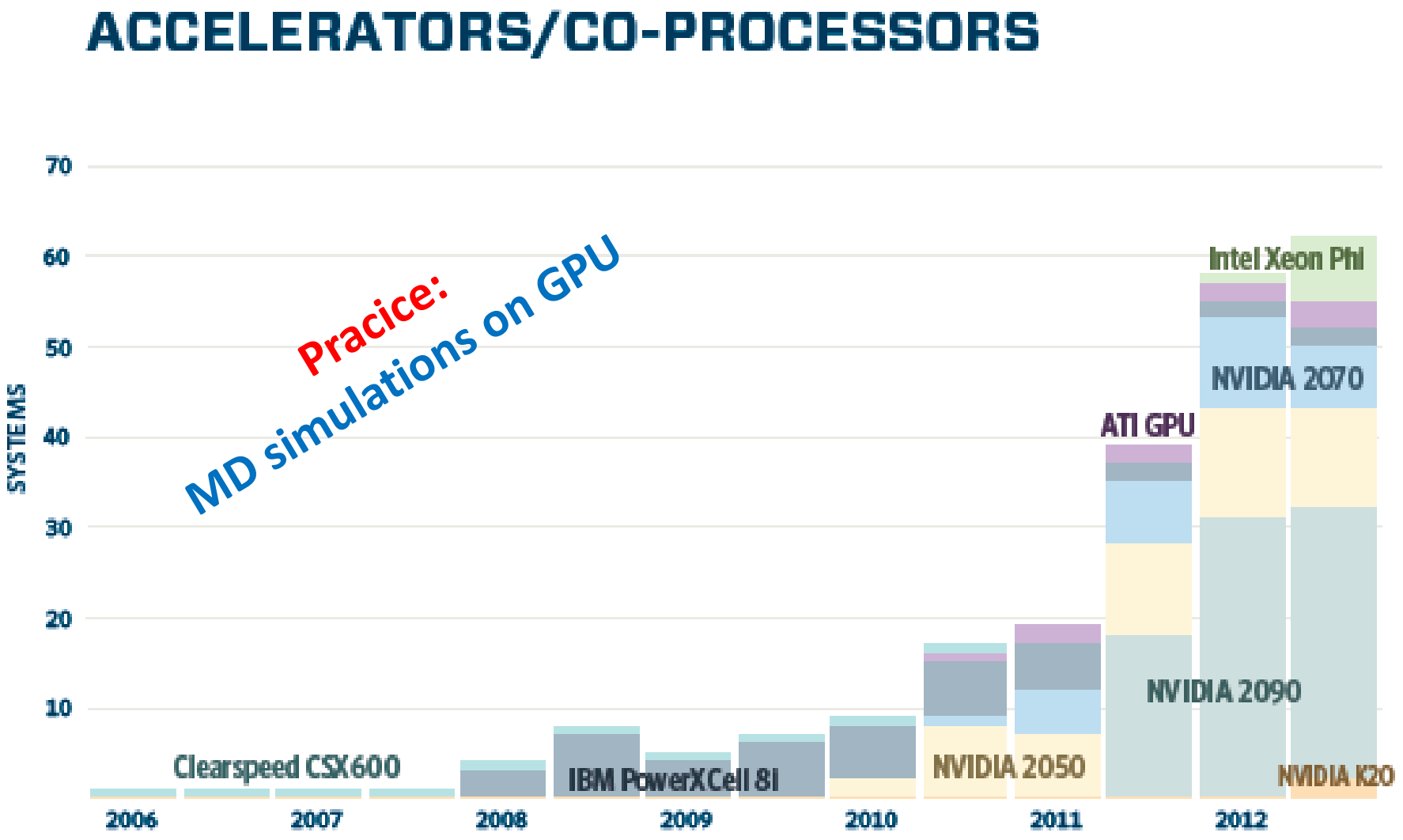
TOP500 – Usage types



TOP500 – Accelerators/Coprocessors



TOP500 – Accelerators/Coprocessors



K – computer, 3rd place

http://www.youtube.com/watch?v=UJPsIu9OaTc

Show video

Exercise LI.1

1. How many times longer will be energy calculation of benzene molecule by quantum chemistry method CCSD(T) between base aug-cc-pVDZ and aug-cc-pVTZ? Base function number is 192 in aug-cc-pVDZ and 414 in aug-cc-VTZ.
2. Kow many days will take simulation of celulose fragment with length 1 μs in explicit solvent box with total atom number 408609, on graphics card GTX680 in NPT conditions? Use data provided on: Http://ambermd.org/gpus/benchmarks.htm#Benchmarks
3. In 2009 average energy consumption on one CZ inhabitant was 1398 kWh (Energetic regulative office statistics). How many people would take same year energy as consumes most powerfull computer Titan with performance of cca 17 PFLOPS?