

C2110 UNIX and programming

Lesson 1 / Module 2

PS / 2020 Distance form of teaching: Rev2

Petr Kulhanek

kulhanek@chemi.muni.cz

National Center for Biomolecular Research, Faculty of Science
Masaryk University, Kamenice 5, CZ-62500 Brno

Focus of the Subject

- Covered topics
- Motivation

Covered Topics

- **Basic work with Linux (Unix-type operating system)**
- **Working with the command line**
- **Scripting in the language**
 - bash
 - gnuplot
 - awk
- **Virtualization**

What is the use of this knowledge?

- Growing amounts of data require efficient, ideally automated processing, including visualization of the key results.
- The design of new drugs, materials and devices requires computationally intensive simulations. The simulations are performed in supercomputer centers that use Unix-type operating systems.

Supercomputing

MetaCentrum and CERIT-SC (<http://metavo.metacentrum.cz>)

Situation in 2016

- National grid infrastructure, OS Debian, OS CentOS
- approx **11000 CPU** cores, **1100 TiB** disk array, **17 PiB** hierarchical repositories

An account can be obtained by a student of any university in the Czech Republic.

IT4Innovations (<http://it4i.cz>)

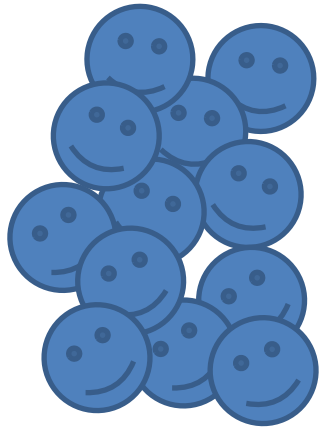
- National Supercomputer Center, OS CentOS
- salomon (approx. 24192 CPU cores, 129TB RAM, Intel Xeon Phi BALL)
 - 40. the most powerful supercomputer in the world at the time of commissioning (TOP500, **139th position 06/2018**)
- anselm (approx. 3000 CPU cores, 15TB RAM)

Machine time is requested in the form of a grant competitions.

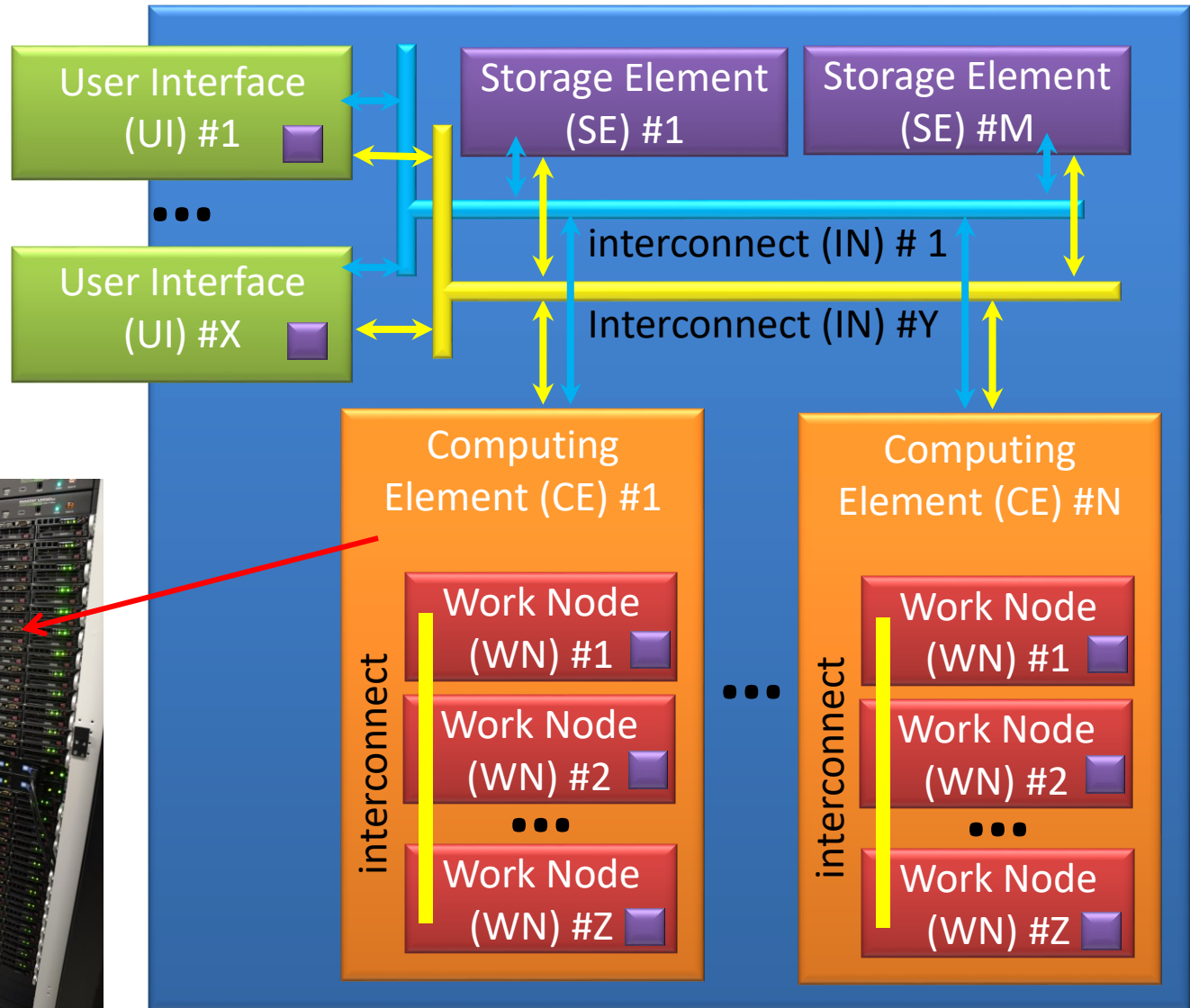
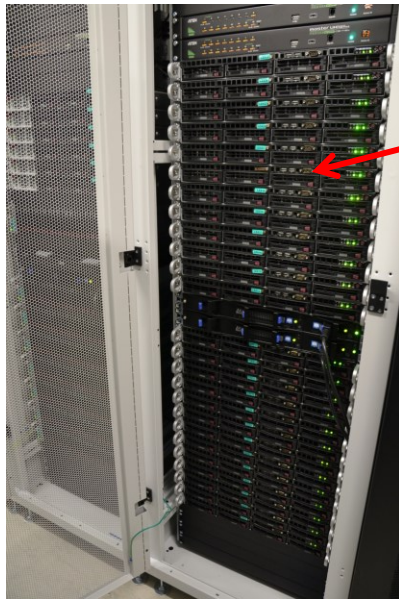
Local computing clusters

- LCC - Computational Chemistry Laboratory

Supercomputers



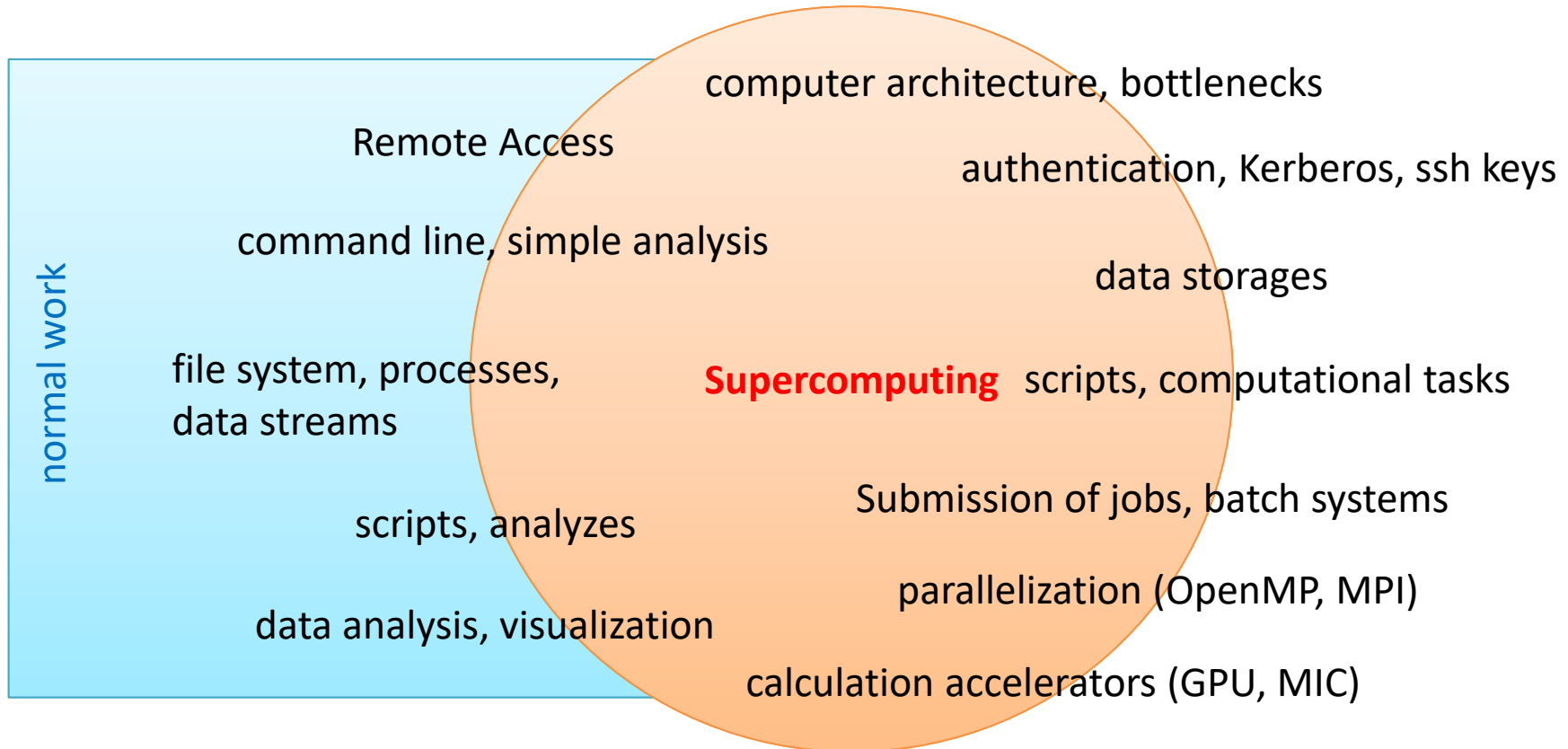
users



C2110 vs C2115

C2110 UNIX and programming

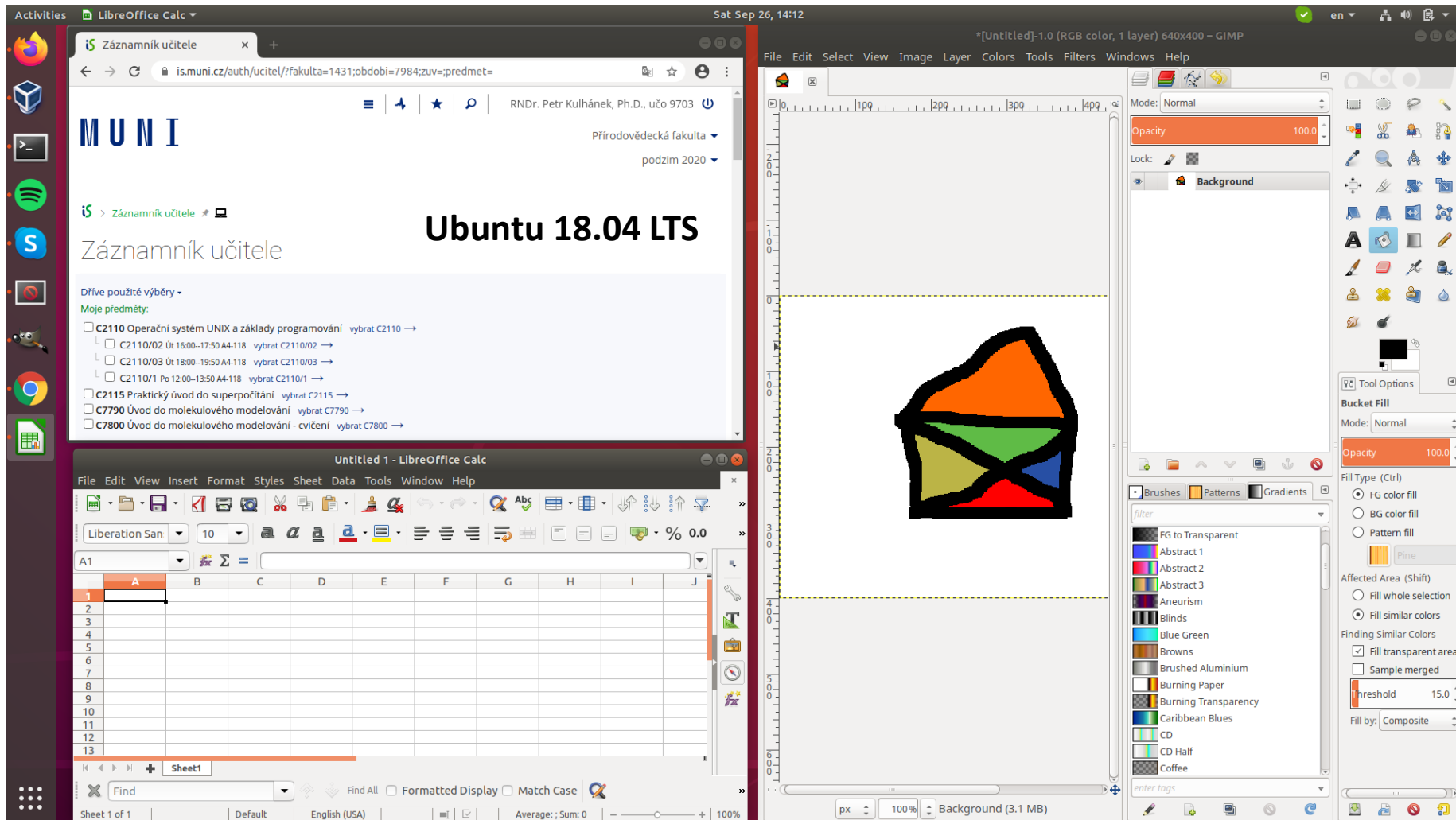
C2115 Practical Introduction to supercomputing



C2115 is a follow-up subject to C2110

- autumn semester - block teaching (February 2021)
- I grant exception to students who have C2110 registered at the same time

Linux - Desktop System



Advantages: free, flexible and extensible, scripting option

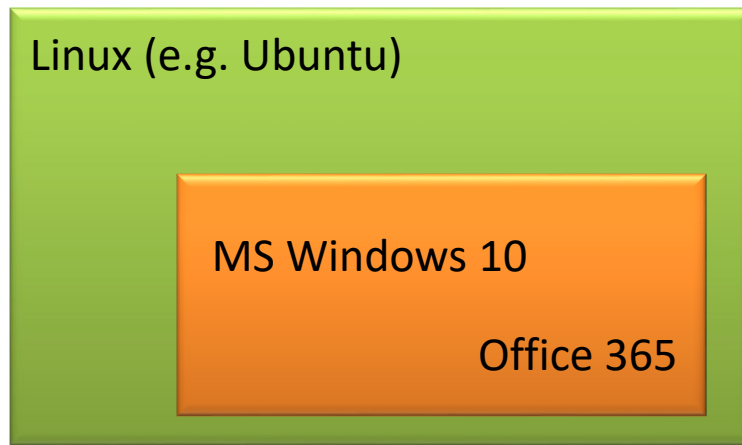
Which desktop to use?

History:

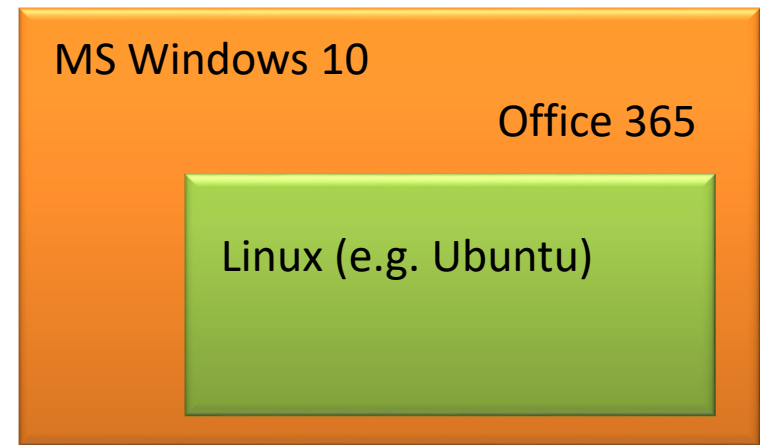
- dual boot

Present

- different types of virtualization



VirtualBox



VirtualBox

Cygwin

Windows Support For Linux (WSL)

Note:

MU students have access to Office 365 (possibility of installation on home computers).

<https://it.muni.cz/sluzby/microsoft-office-365>

Virtualization

The image shows a virtualization environment. The host OS is Ubuntu 18.04 LTS, running a web browser (Google Chrome) displaying the 'Záznámník učitele' (Teacher's Record) page. The guest OS is MS Windows 10, running a presentation slide titled 'C2110 vs C2115'. The slide compares two courses: 'C2110 Operační systém UNIX a základy programování' and 'C2115 Praktický úvod do superpočítání'. The slide content is as follows:

C2110 Operační systém UNIX a základy programování	C2115 Praktický úvod do superpočítání
suborový systém, procesy, datové proudy	architektura počítačů, úzká hrdla
skripty, analýzy	autentizace, Kerberos, ssh klíče
analýza dat, vizualizace	datové úložiské
C2115 je navazující předmět na C2110 - podzimní semestr - bloková výuka (leden)	Superpočítání - skripty, výpočetní úlohy
	zadávání úloh, dávkové systémy
	paralelizace (OpenMP, MPI)
	akceleratory výpočtu (GPU, MIC)
	C2115 je navazující předmět na C2110 - podzimní semestr - bloková výuka (leden)

Two blue arrows point from the text below to the host and guest OS windows. The host OS window is titled 'Host: Ubuntu 18.04 LTS' and the guest OS window is titled 'Guest: MS Windows 10 (virtual machine)'. The guest OS window also shows a 'Click to add notes' button.

Host: Ubuntu 18.04 LTS

Guest: MS Windows 10 (virtual machine)

Processing a Text File

```
.....  
.....  
NSTEP =      6000      TIME (PS) =      206.000      TEMP (K) =      291.69      PRESS =      0.0  
Etot   =      160.8627      EKtot   =      18.5486      EPtot   =      142.3142  
BOND   =      7.2673      ANGLE  =      17.6964      DIHED   =      13.5633  
1-4 NB =      4.8403      1-4 EEL =      199.3739      VDWAALS =      1.5430  
EELEC  =     -101.9700      EHBOND =      0.0000      RESTRAINT =      0.0000  
.....  
.....
```

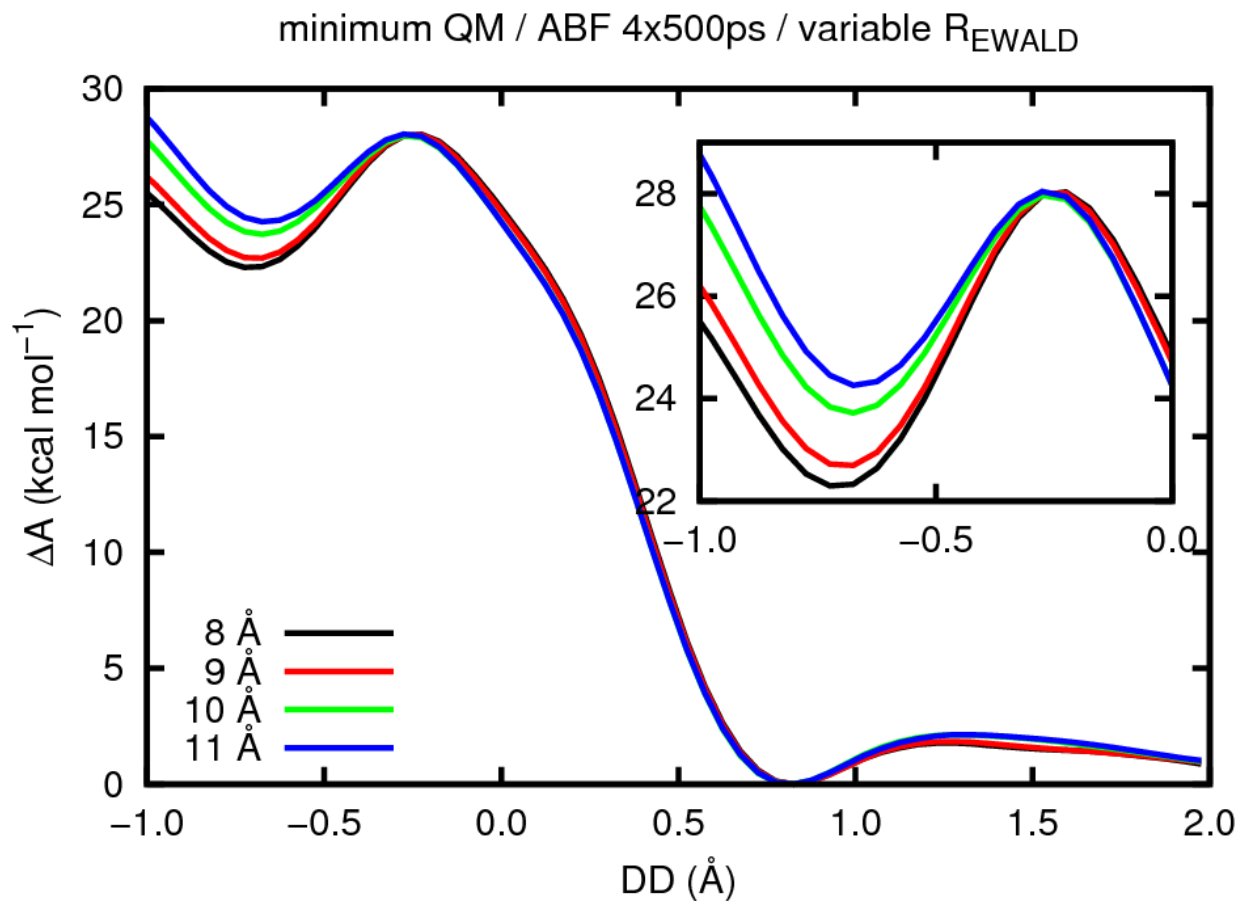
Easy data extraction using AWK language.
(output of a computer programs)

```
.....  
.....  
206.000 291.69  
.....  
.....
```

Data: /home/kulhanek/Documents/C2110/Lesson01/dat/rst.out

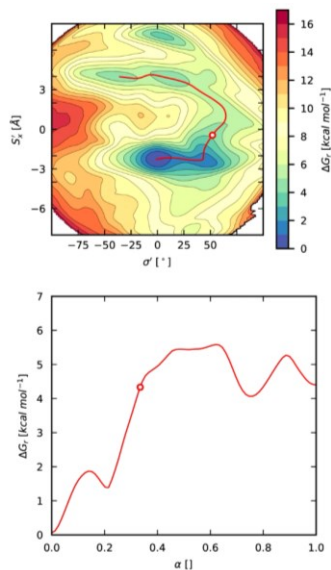
```
$ grep TIME rst.out | awk '{ print $6, $9 }'  
$ awk '/TIME/{print $6, $9}' rst.out
```

Visualization of the Results

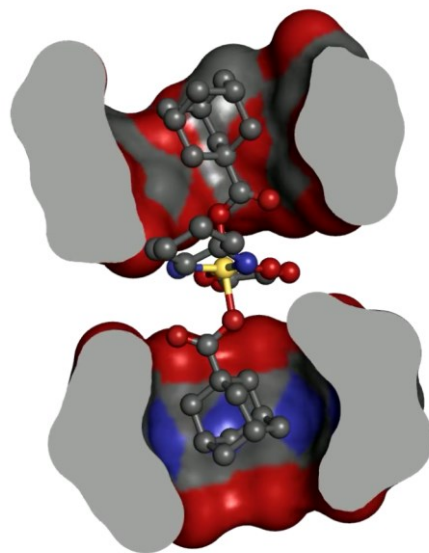
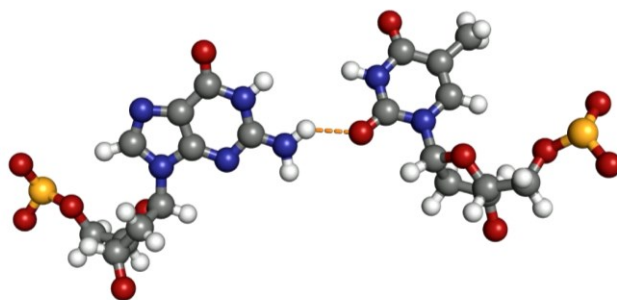


Displaying extracted data in the form of graphs (gnuplot).

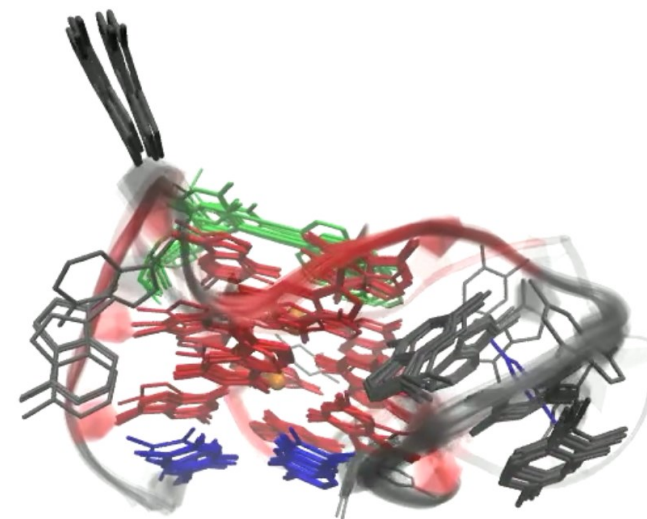
Automation



aG/aT mismatch: Tomas Bouchal



Show video



Quadruplex: Ivo Durník

Supramolecular complex: Ivo Durník

Data: /home/kulhanek/Documents/C2110/Lesson01/video

Exercise 1

1. List the names of the front nodes (front-ends) that can be used to access the MetaCentrum supercomputer center.
2. What supercomputers are available in IT4I?
3. On what position is the Salomon supercomputer in the TOP500 list?
4. What operating systems are used on the world's top 10 fastest supercomputers?