

C7800

Počítačová chemie a molekulové modelování I - cvičení

Petr Kulhánek, Jakub Štěpán

kulhanek@chemi.muni.cz

Národní centrum pro výzkum biomolekul, Přírodovědecká fakulta
Masarykova univerzita, Kotlářská 2, CZ-61137 Brno

Obsah

➤ Linux

- terminály, textové editory, midnight commander, moduly, souborový systém klastru WOLF

➤ Programy pro MM I

- vmd, pdb formát, xyz formát

➤ Programy pro MM II

- openbabel

➤ Programy pro MM III

- nemesis, avogadro

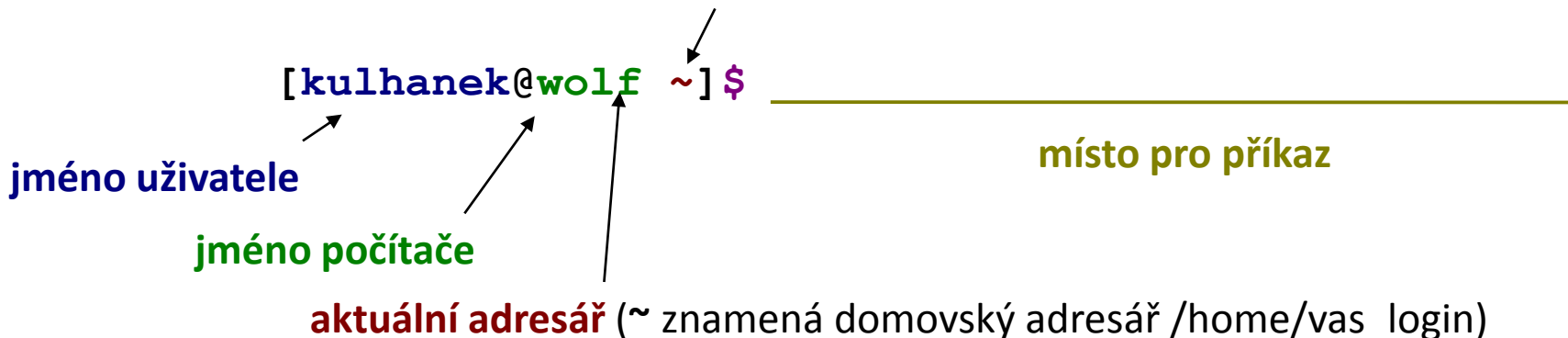
➤ Linux

- vzdálené přihlašování

Linux

Příkazová řádka

Prompt - typ uživatele / výzvy (\$ běžný uživatel, # super uživatel, další možné %, >)



Příkaz se vykoná zmáčknutím klávesy **Enter**.

Historie: pomocí kurzorových šipek nahoru a dolů lze procházet seznamem již zadaných příkazů. Příkaz z historie lze znovu použít nebo upravit a upravený použít. Historie je přístupná i příkazem **history**.

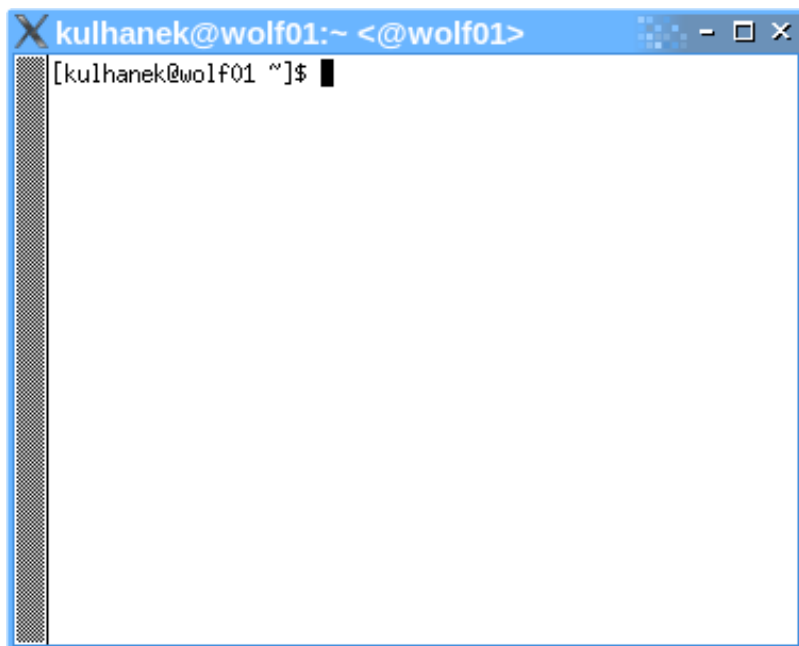
Automatické doplňování: zmáčknutím klávesy Tab (tabulátor) se interpret příkazové řádky snaží dokončit rozepsané slovo. Doplňují se jména příkazů, cesty a jména souborů (pokud jeden stisk nic nevyvolá, existuje více možností doplnění, opakovaný stisk vylistuje možnosti).

Kopírování textu: Ne pomocí Ctrl+C! Pro kopírování textu z terminálu stačí text označit, pro následné vložení stiskněte kolečko myši.

Terminály

Příkazová řádka je přístupná přímo z textových terminálů. V grafickém prostředí X11 je nutné spustit vhodnou aplikaci emulující textový terminál.

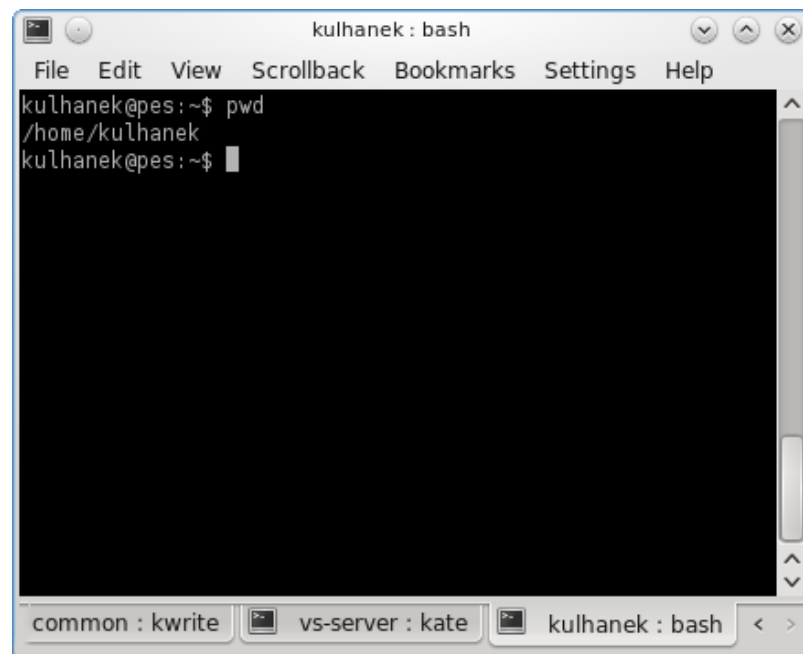
xterm



jednoduché, standard na všech UNIXových systémech

Výchozím adresářem je: **/home/vas_login**

konsole



jednoduché přitom značně konfigurovatelné

Midnight Commander

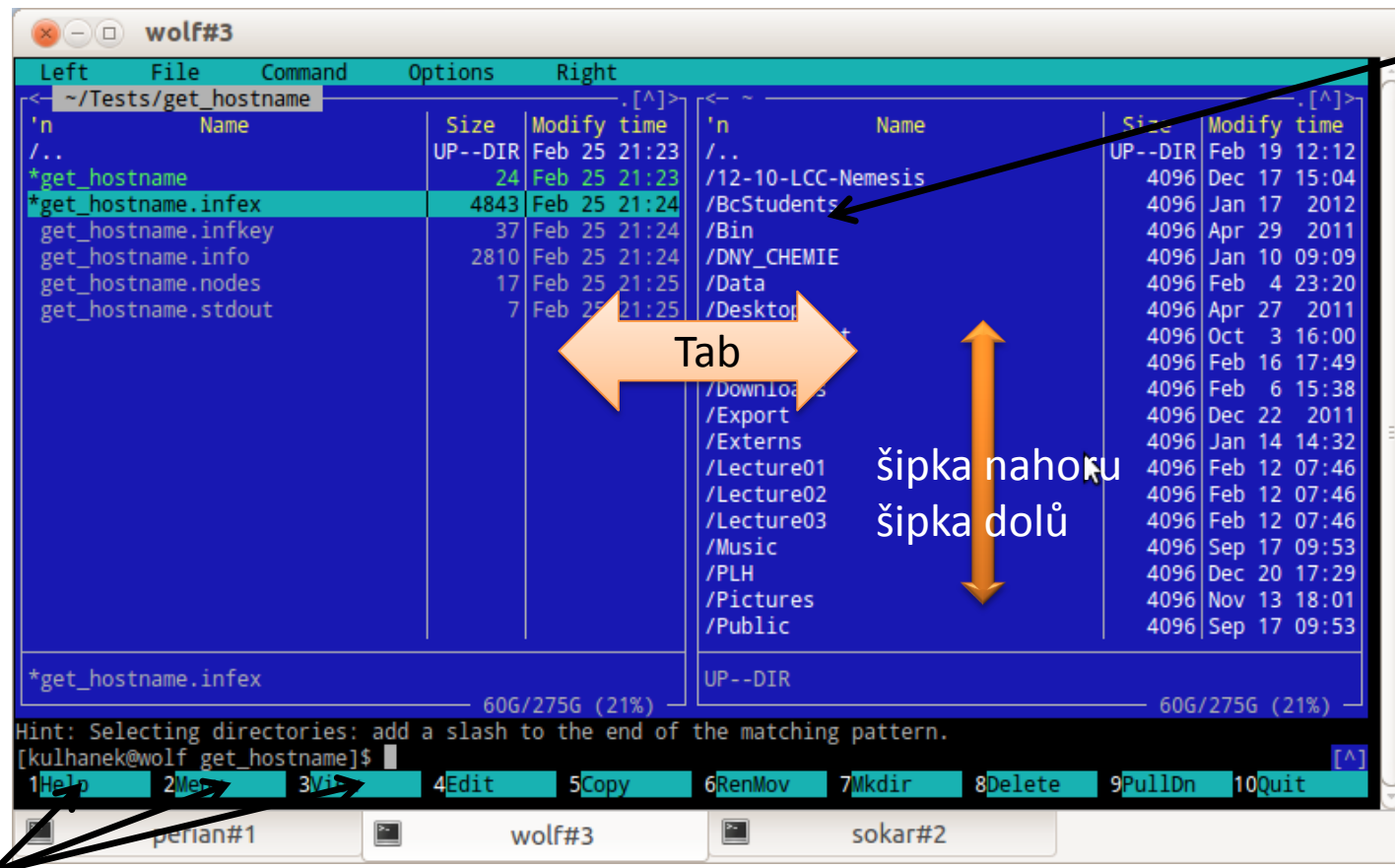
Souborový manažer, který pracuje v terminálu.

<http://www.midnight-commander.org/>

Spuštění:

\$ mc

označení více souborů - **Insert**



Tab

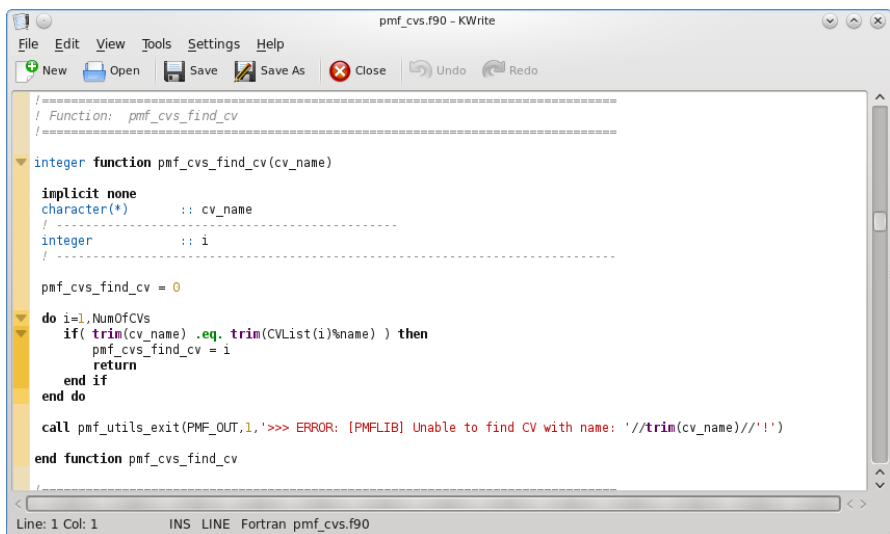
šipka nahoru
šipka dolů

funkční klávesy F1, F2, F3, ...

Skrytí obou panelů: Ctrl+O

Lze použít myš.

Textové editory

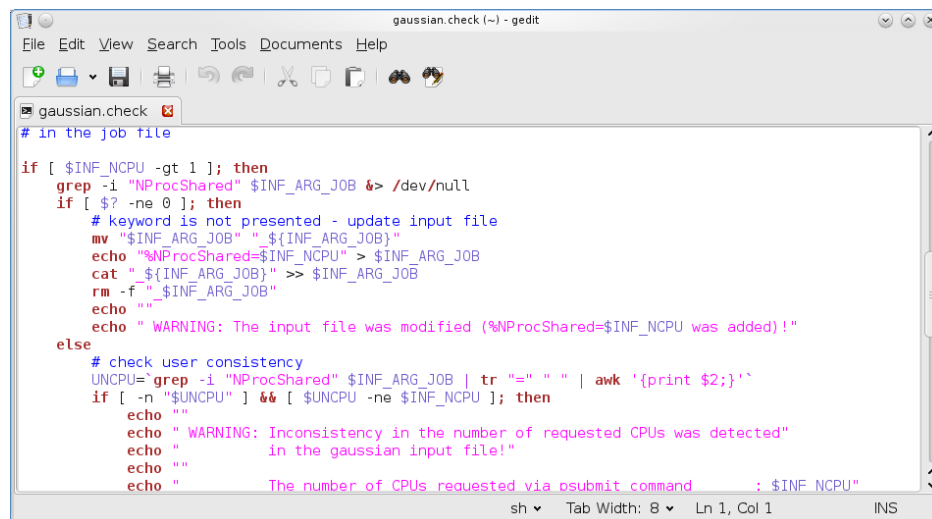


```
pmf_cvs.f90 - KWrite
File Edit View Tools Settings Help
New Open Save Save As Close Undo Redo
! =====
! Function: pmf_cvs_find_cv
! =====
integer function pmf_cvs_find_cv(cv_name)
implicit none
character(*) :: cv_name
! =====
integer :: i
! =====
pmf_cvs_find_cv = 0
do i=1, NumOfCVs
if (trim(cv_name) .eq. trim(CVList(i)%name)) then
pmf_cvs_find_cv = i
return
end if
end do
call pmf_utils_exit(PMF_OUT,1,'>>> ERROR: [PMFLIB] Unable to find CV with name: '//trim(cv_name)//'')
end function pmf_cvs_find_cv
Line: 1 Col: 1 INS LINE Fortran pmf_cvs.f90
```

Spuštění:

\$ kwrite

Rozšířená funkcionality: kate



```
gaussian.check (-) - gedit
File Edit View Search Tools Documents Help
gaussian.check
# in the job file
if [ $INF_NCPU -gt 1 ]; then
grep -i "NProcShared" $INF_ARG_JOB &> /dev/null
if [ $? -ne 0 ]; then
# keyword is not presented - update input file
mv "$INF_ARG_JOB" "${INF_ARG_JOB}"
echo "%NProcShared=$INF_NCPU" > $INF_ARG_JOB
cat "$INF_ARG_JOB" >> $INF_ARG_JOB
rm -f "$INF_ARG_JOB"
echo ""
echo " WARNING: The input file was modified (%NProcShared=$INF_NCPU was added)!"
else
# check user consistency
UNCPU=`grep -i "NProcShared" $INF_ARG_JOB | tr "=" " " | awk '{print $2;}'`
if [ -n "$UNCPU" ] && [ $UNCPU -ne $INF_NCPU ]; then
echo ""
echo " WARNING: Inconsistency in the number of requested CPUs was detected"
echo " in the gaussian input file!"
echo ""
echo " The number of CPUs requested via psubmit command : $INF_NCPU"
fi
sh Tab Width: 8 Ln 1, Col 1 INS
```

Spuštění:

\$ gedit

Vědecko-technické aplikace

Vědeckotechnické aplikace, které jsou instalovány v několika verzích (verze aplikace, typ kompilace, paralelní verze), jsou dostupné ve formě **modulů**. Před použitím aplikace je nutné příslušný modul aktivovat.

Přehled dostupných aplikací:

`$ module`

Přehled dostupných verzí aplikace:

`$ module versions nemesis`

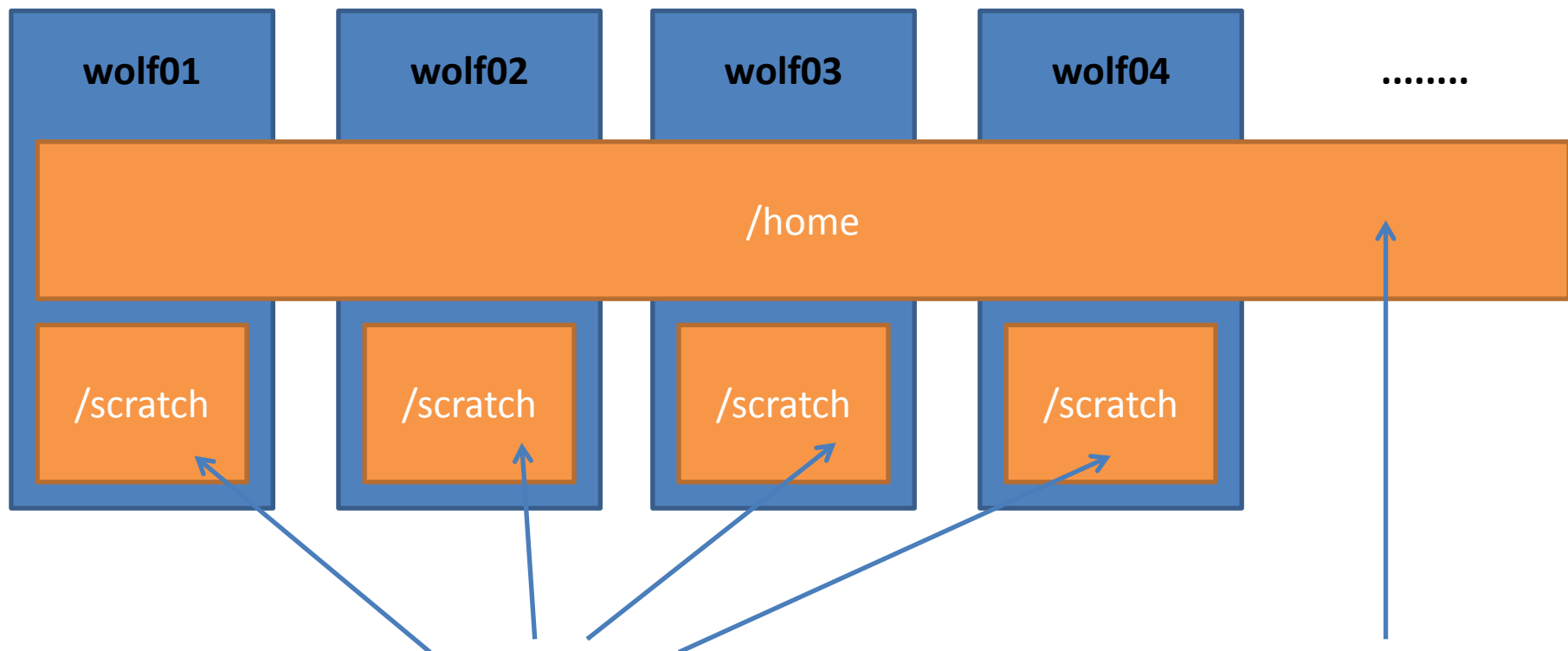
Aktivace aplikace:

`$ module add nemesis`

Spuštění aplikace z modulu nemesis

`$ nemesis`

Souborový systém na klastru WOLF



Rozdílný obsah na každém uzlu.

Data na svaku /scratch se **nezálohuji** a mohou být **kdykoliv smazána** bez předchozího upozornění.

Kapacita **není omezena** kvótou na uživatele.

Sdílený obsah na všech uzlech klastru WOLF.

Data jsou **zálohována**. Kapacita na uživatele omezena na **1,5GB kvótou**.

Programy pro molekulové modelování I

Nemesis

<https://lcc.ncbr.muni.cz/whitezone/development/nemesis/>

Program pro stavbu a vizualizaci molekul. Alfa verze pro Linux. Testovací verze pro MS Windows na vyžádání.

Avogadro

http://avogadro.openmolecules.net/wiki/Main_Page

Program pro stavbu a vizualizaci molekul. Volně dostupný pro MS Windows a Linux.

Přehled funkcionality: <https://www.youtube.com/watch?v=xdmLoBILmq5>

VMD

<http://www.ks.uiuc.edu/Research/vmd/>

Program pro vizualizaci molekul. Po bezplatné registraci dostupný pro MS Windows a Linux.

Spuštění programů

Nemesis

```
$ module add nemesis
```

```
$ nemesis
```

Avogadro

```
$ module add avogadro
```

```
$ avogadro
```

VMD

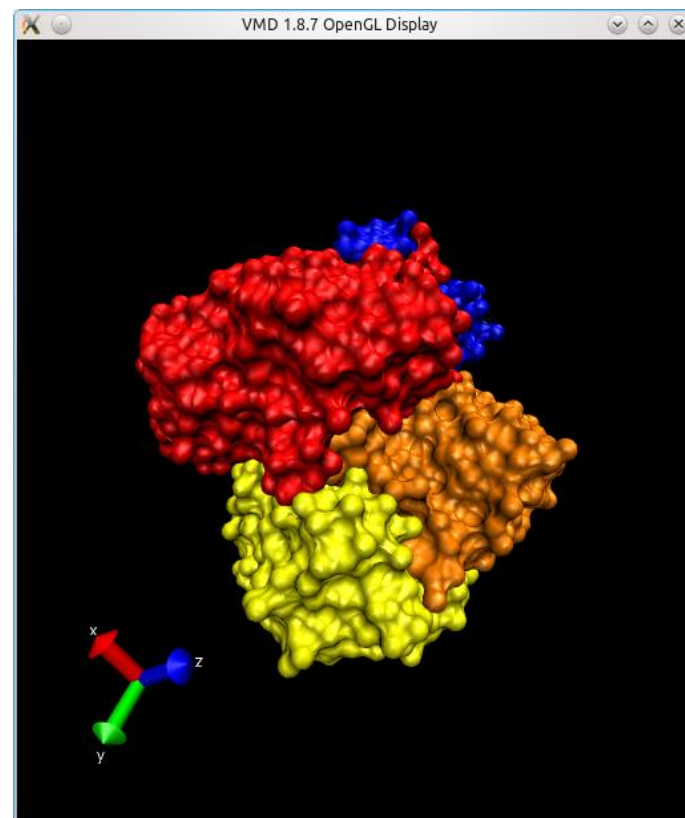
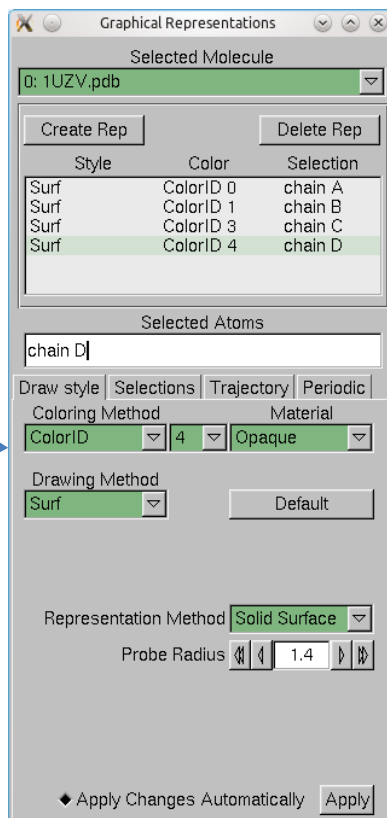
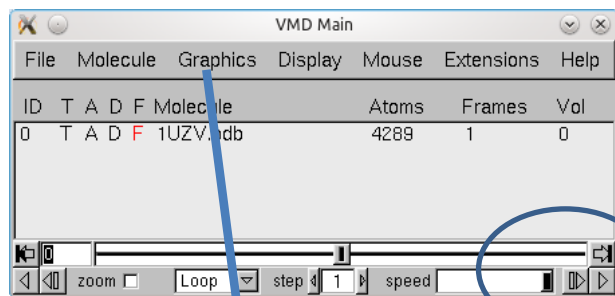
```
$ module add vmd
```

```
$ vmd
```

Program VMD

<http://www.ks.uiuc.edu/Research/vmd/>

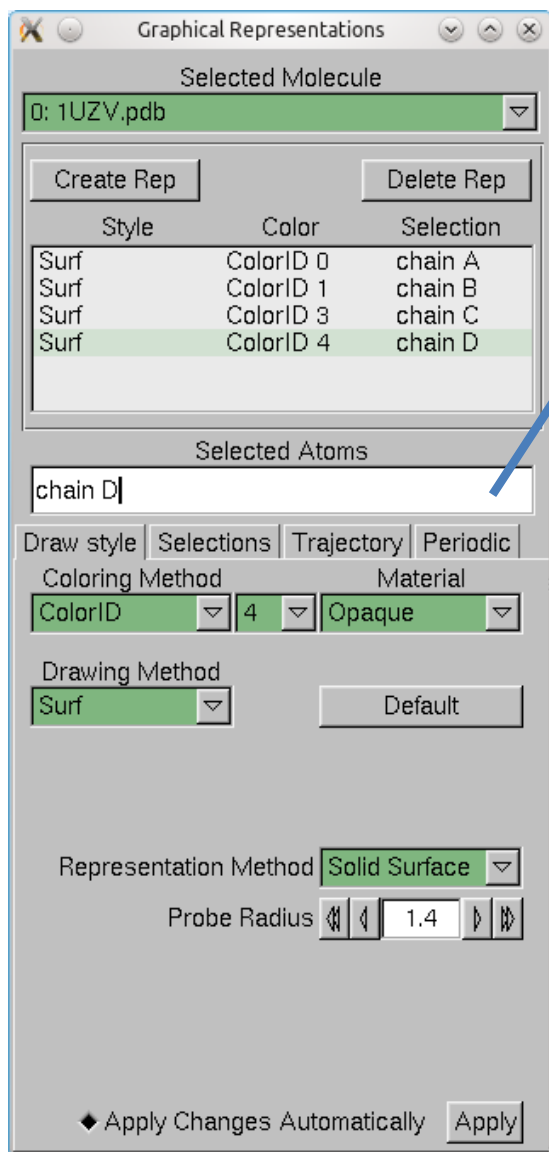
Program slouží k vizualizaci (bio)molekul a k analýze výsledků molekulárně dynamických simulací. Program je volně dostupný (vyžaduje registraci) a je dostupný i pro operační systém MS Windows.



Representation

posouvání mezi snímky

Program VMD



Selekcce (volba) části molekuly:

- protein – zvolí všechny aminokyseliny
- water – zvolí všechny molekuly vody
- chain X – zvolí řetězec X
- resname X – zvolí residuum s názvem X
- resid X – zvolí residuum s číslem X

Příklady:

chain A
chain A B C
resname ASP GLU
resid 1
resid 1 to 100

Bližší informace:

C2150 Zpracování informací a vizualizace v chemii

PDB Databáze

www.pdb.org

The screenshot shows a Mozilla Firefox browser window displaying the PDB entry page for 1UZV. The address bar shows the URL: <http://www.pdb.org/pdb/explore.do?structureId=1UZV>. The page title is "RCSB PDB - 1UZV Structure Summary".

The main content area features the title: **HIGH AFFINITY FUCOSE BINDING OF PSEUDOMONAS AERUGINOSA LECTIN II: 1.0 A CRYSTAL STRUCTURE OF THE COMPLEX**. Below the title is the DOI: [10.2210/pdb1uzv/pdb](https://doi.org/10.2210/pdb1uzv/pdb).

The "Primary Citation" section contains the following text: **High affinity fucose binding of Pseudomonas aeruginosa lectin PA-III: 1.0 A resolution crystal structure of the complex combined with thermodynamics and computational chemistry approaches.** The authors listed are Mitchell, E., Sabin, C.D., Snajdrova, L., Pokorna, M., Gautier, S., Perret, C., Hofr, C., Gilboa-Garber, N., Koca, J., Imberty, M., and Wimmerovaa. The journal information is: **Journal:** (2005) Proteins: Struct., Funct., Bioinf. 58: 735. The PubMed ID is [15573375](https://pubmed.ncbi.nlm.nih.gov/15573375/) and the DOI is [10.1002/prot.20330](https://doi.org/10.1002/prot.20330). There is a link to "Search Related Articles in PubMed".

On the right side, there is a "Biological Assembly" section with a 3D ribbon diagram of the protein structure, colored in green, yellow, and blue. The structure shows a complex of protein chains with a fucose molecule bound to it.

The left sidebar contains navigation links under "Home" and "Deposition". The "Home" section includes links for News & Publications, Usage/Reference Policies, Deposition Policies, Website FAQ, Deposition FAQ, Contact Us, About Us, Careers, External Links, Sitemap, and New Website Features. The "Deposition" section includes links for All Deposit Services, Electron Microscopy, X-ray | NMR, Validation Server, BioSync Beamlines/Facilities, and Related Tools.

Obsahuje struktury biomolekul určené metodami rentgenové strukturní analýzy, nukleární magnetické rezonance, teoretické modely.

Cvičení VMD - I

- Najděte protein s PDB kódem **1UZV** . Kterou metodou byla struktura určena? Jakého rozlišení bylo dosaženo? Uložte strukturu ve formátu **pdb** do vašeho domovského adresáře.
- Zobrazte strukturu **1UZV** v programu **VMD**.
- Zvýrazněte jednotlivé monomerní jednotky komplexu (model: surf, selekce: chain A, chain B, ...)
- Zobrazte sekundární strukturu komplexu (NewCartoon, Color: Secondary Structure). Který strukturní element ve struktuře převládá?
- Zobrazte navázaný ligand (resname FUC). Kolik ligandů je v komplexu obsaženo?
- Zobrazte vápenaté ionty (resname CA). Kolik iontů je v komplexu obsaženo?

Cvičení VMD - II

- Najděte protein s PDB kódem **2KRC**. Kterou metodou byla struktura určena? Uložte strukturu ve formátu **pdb** do vašeho domovského adresáře.
- Zobrazte strukturu **2KRC** v programu **VMD**.
- Zobrazte sekundární strukturu komplexu (NewCartoon, Color: Secondary Structure). Který strukturní element ve struktuře převládá?
- Kolik modelů bylo publikováno? Kvalitativně určete jak se od sebe modely liší.

Cvičení XYZ formát

1. V textovém editoru vytvořte soubor popisující model vody s následujícími parametry. Délka vazeb O-H bude 1 Å. Vazebný úhel H-O-H bude 90°. Uložte jej do domovského adresáře jako **water.xyz**
2. Vytvořený soubor načtěte do programu VMD.
3. Ověřte skutečnou délku vazeb a velikost úhlu H-O-H. (VMD Main >Mouse->Label, správa popisků v VMD Main >Graphics>Labels)
4. Molekulu vody zobrazte v následujících modelech: Lines, CPK, Licorice, VDW.

Programy pro molekulové modelování II

OpenBabel

Open Babel is a chemical toolbox designed to speak the many languages of chemical data. It's an open, collaborative project allowing anyone to search, convert, analyze, or store data from molecular modeling, chemistry, solid-state materials, biochemistry, or related areas.

http://openbabel.org/wiki/Main_Page

Konverze programem openbabel:

```
$ module add openbabel  
$ babel input.xyz output.mol2
```

alternativně


```
$ babel ixyz input.txt -omol2 output.out
```

Seznam podporovaných formátů:

```
$ babel -L formats
```

Nápověda:

```
$ babel -H
```

 **velké H**

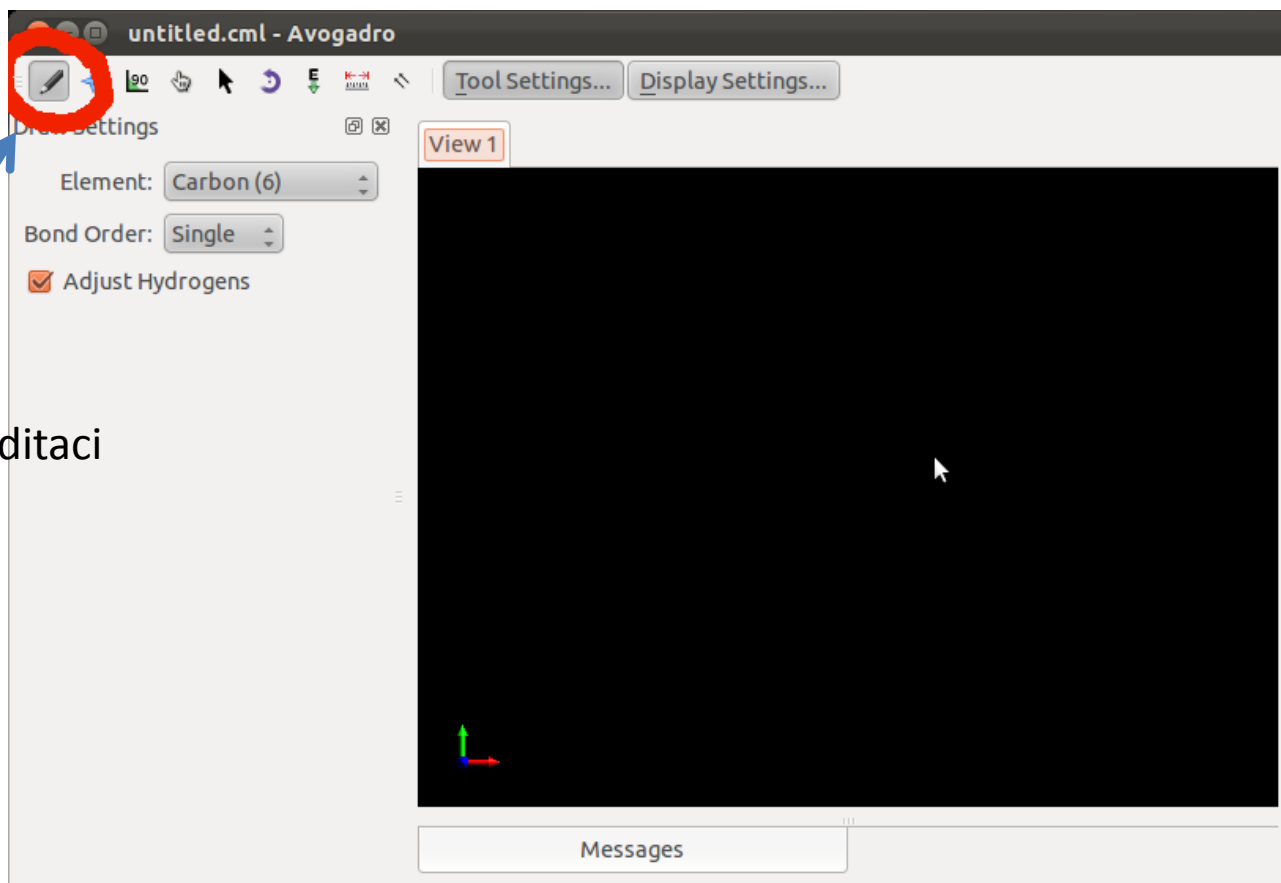
Cvičení

1. Aktivujte modul openbabel.
2. Vypište formáty, které instalovaná verze open babelu podporuje.
3. Zkonvertujte soubor **water.xyz** do formátu Sybyl Mol2 format a uložte jej pod názvem **water.mol2**
4. Otevřete soubor **water.mol2** v textovém editoru a diskutujte význam jeho částí.
5. Zkonvertujte soubor **water.xyz** do formátu InChI a uložte jej po názvem **water.txt**
6. Otevřete soubor **water.txt** v textovém editoru a diskutujte význam jeho částí.

Programy pro molekulové modelování III

Avogadro

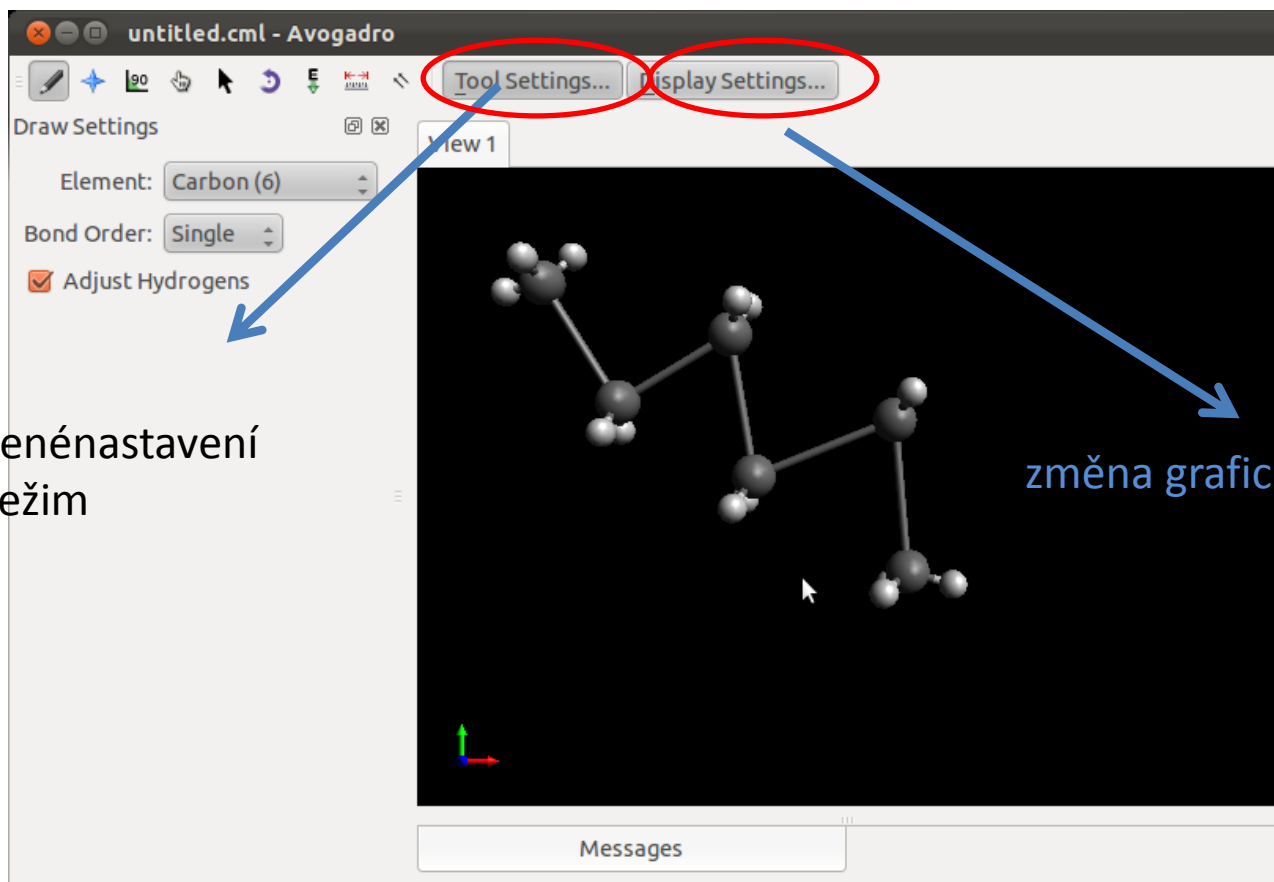
Ke stavbě 3D struktury molekul můžete použít program **Avogadro**. Jedná se o volně šiřitelný program, který lze používat jak pod operačním systémem MS Windows tak i pod Linuxovými klony (např. Ubuntu).



aktivuje editaci

Avogadro

Při stavbě molekuly nejsou délky vazeb, úhly a další parametry molekuly optimální. Je to dáno způsobem, jakým se v programu Avogadro, struktury editují. Draft struktury je proto nutné před dalším použitím upravit pomocí optimalizace geometrie.

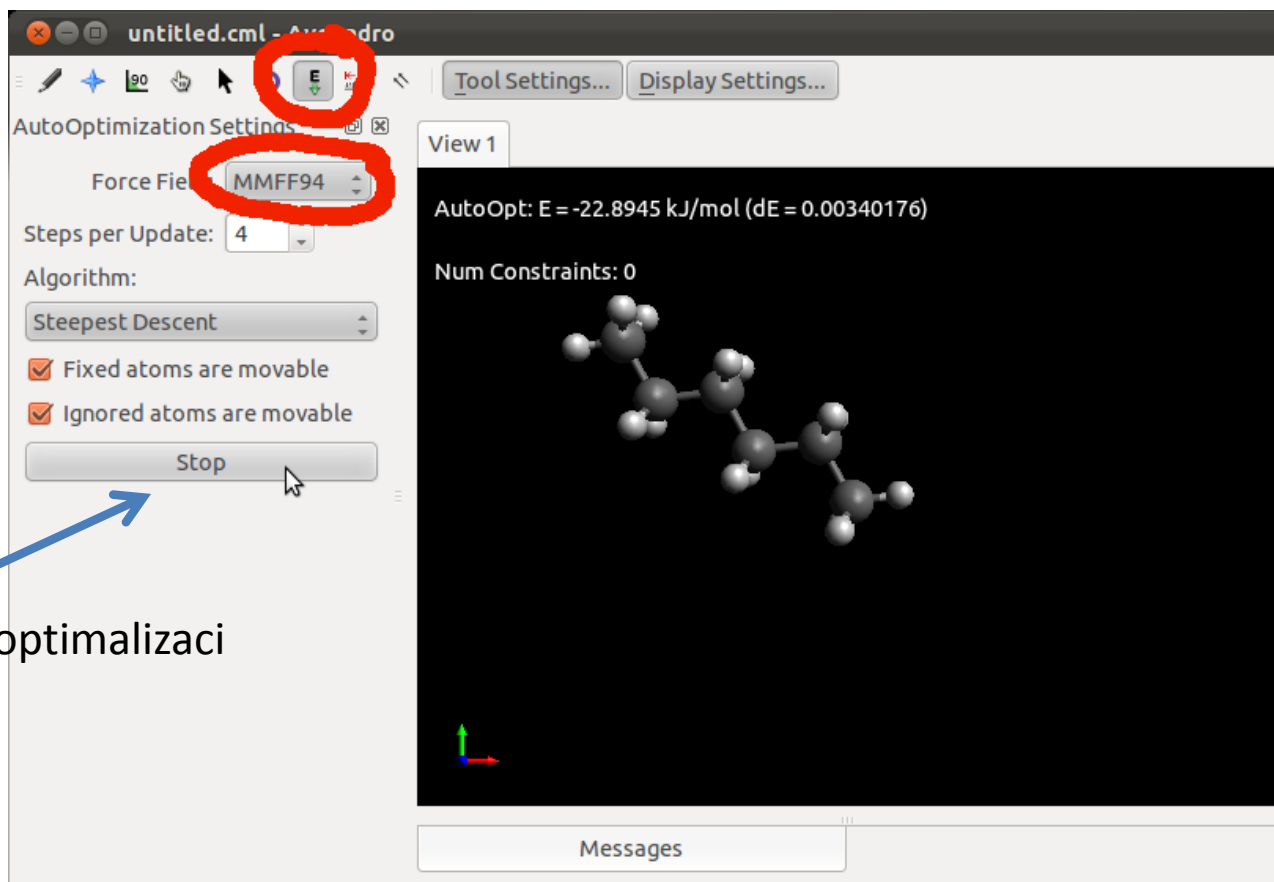


zobrazí rozšířená nastavení
pro zvolený režim

změna grafické vizualizace

Avogadro

Program používá pro optimalizace geometrie metody molekulové mechaniky (MM). Pro její správnou funkci musíte ve struktuře správně uvést řady vazeb. Protože MM je empirickou metodou, musíte zvolit i typ parametrizace. V našem případě budeme používat silové pole MMFF94.



zapíná/vypíná optimalizaci geometrie

Nemesis

The screenshot shows the Nemesis Molecular Modelling Package interface. The main window is titled "Project 1: NEMESIS - Molecular Modelling Package". The interface is divided into several panels:

- Structures panel (left):** Contains a table with columns "Name", "SID", and "Ato". It lists "Structure 1" with SID "1". Below the table are controls for "Number of structures" (set to 1) and "Active Profile" (set to "Profile 1").
- Profile objects panel (bottom left):** Contains a table with columns "Name" and "Type". It lists "Light 1" (Light), "Background 1" (Backgro), "Standard Model 1" (Standar), and "Freezed Atoms 1" (Freezec).
- Build panel (right):** Contains various chemical building blocks (C, O, N, F, Cl, Br, I, S) and buttons for "Delete atom", "Make bond", "Break bond", and "Delete bond". The "Optimize" button is circled in red.
- Geometry panel (bottom right):** Contains buttons for "Position", "Distance", "Angle", and "Torsion".

Annotations with blue arrows point to specific features:

- "vrstvy" points to the Structures panel.
- "stavba/editace molekuly" points to the Build panel.
- "optimalizace geometrie" points to the "Optimize" button in the Build panel.
- "grafické modely" points to the Profile objects panel.
- "měření" points to the Geometry panel.

Nastavení silového pole pro optimalizaci: menu Geometry-> Optimizer Setup

Nemesis

Myš:

Levé tlačítko	selekce
Prostřední tlačítko	rotace
Levé tlačítko	posun
Kolečko	zoom

Modifikátory:

Shift	XZ -> Y pohyby
Ctrl	přepíná mezi sekundárním a primárním manipulátorem

Cvičení I

1. Načtete do programu **Avogadro** molekulu ze souboru **water.xyz**
2. Provedte optimalizaci její geometrie. Jaké je optimální délka vazby a vazebný úhel?
3. Zobrazte molekulu v různých grafických reprezentacích.
4. Načtete do programu **Nemesis** molekulu ze souboru **water.xyz** (Import Structure -> OpenBabel)
5. Zobrazte molekulu v různých grafických reprezentacích.
6. Provedte optimalizaci její geometrie. Jaké je optimální délka vazby a vazebný úhel? Srovnajte s výsledky získanými v programu Avogadro. Vysvětlete případné rozdíly.

Cvičení II

1. V programu **Nemesis** nakreslete strukturní vzorec molekuly benzoové v projektu Sketch Structure
2. Převeďte molekulu do 3D reprezentace. Ohodnoťte kvalitu převodu.
3. V programu **Nemesis** nakreslete strukturní vzorec molekuly cyklohexanu v projektu Sketch Structure.
4. Převeďte molekulu do 3D reprezentace. Ohodnoťte kvalitu převodu.
5. V projektu Sketch Structure programu **Nemesis** vložte molekulu fullerenu C_{60} ve formátu SMILES (View->Insert->SMILES...).
6. Převeďte molekulu do 3D reprezentace. Ohodnoťte kvalitu převodu.
7. Úlohu s molekulou C_{60} zopakujte v programu **Avogadro**. Jakým způsobem lze získat lepší model?

Vzdálené přihlašování

- Příkaz ssh
- Autorizace pomocí ssh klíčů

Vzdálené přihlášení

Existuje několik možností vzdáleného přihlášení (rsh, XDMCP, apod.) avšak nejpoužívanějším a **nejbezpečnějším** je použití příkazu **ssh** (secure shell).

Syntaxe:

```
$ ssh [user@]hostname [command]           [] - možno vynechat
```

jméno uživatele;
pokud není uvedeno, použije se
jméno přihlášeného uživatele

jméno počítače

příkaz, který se má vykonat;
pokud není uveden, zpřístupní se
příkazová řádka

Příklady použití:

```
$ ssh wolf.wolf.inet
```

Ukončení přihlášení:

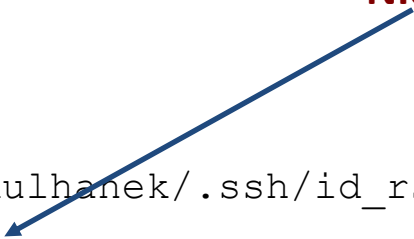
```
$ exit
```

Přihlašování bez hesla v rámci klastru

1. Vytvoření dvojice veřejného a soukromého klíče:

```
[kulhanek@wolf01 ~]$ cd .ssh
[kulhanek@wolf01 .ssh]$ ssh-keygen
Generating public/private rsa key pair.
Enter file in which to save the key (/home/kulhanek/.ssh/id_rsa):
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in /home/kulhanek/.ssh/id_rsa.
Your public key has been saved in /home/kulhanek/.ssh/id_rsa.pub.
The key fingerprint is:
e9:07:0b:fc:17:23:b3:c5:1a:8a:0c:1a:98:8f:fe:28 kulhanek@wolf01.wolf.inet
```

Nic se nezadává!



2. Vložení veřejného klíče do seznamu autorizovaných klíčů:

```
[kulhanek@wolf01 .ssh]$ cat id_rsa.pub >> authorized_keys
```

Výhody:

- nemusí se neustále zadávat heslo
- bezpečnější použití příkazů ssh a scp ve skriptech
- urychlení práce

Nevýhody:

- v případě kompromitace jednoho počítače, jsou kompromitovány všechny počítače se vzájemně autorizovanými veřejnými klíči.

Podrobnější popis: man ssh

Cvičení

1. Přihlaste se příkazem **ssh** na uzel **wolf.wolf.inet**
2. Příkazem **who** zjistěte, kdo je na uzlu přihlášen.
3. Odhlaste se příkazem **exit**.
4. Aktivujte si přihlašování bez hesla v rámci klastru WOLF.
5. Ověřte funkčnost přihlašování bez hesla. Přihlaste se na uzel **wolf.wolf.inet** pomocí příkazu **ssh**.

Pro další cvičení je nutné mít aktivované přihlašování bez hesla v rámci klastru WOLF.