

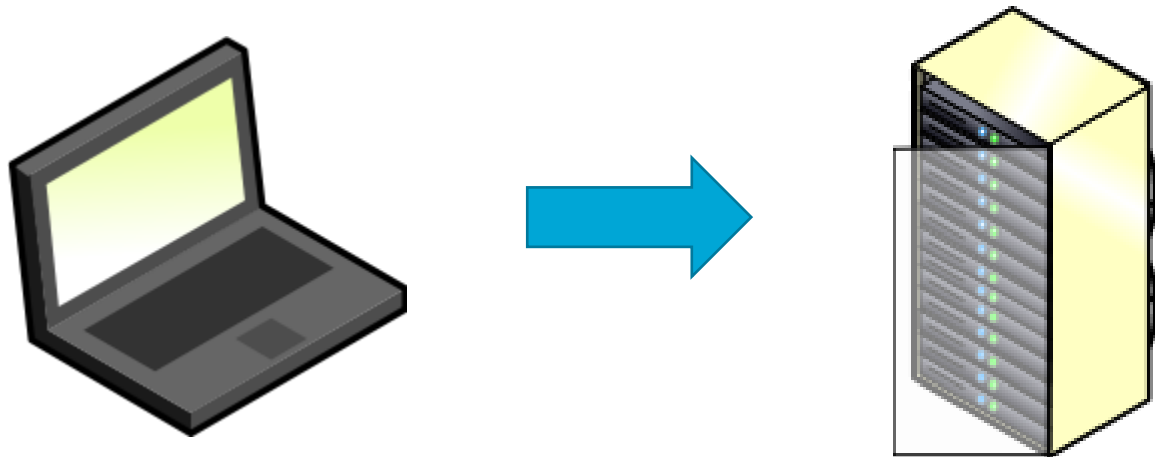
Modelling, Uncertainty and Data for Engineers (MUDE)

Introduction to Linux and HPC

Denis Voskov, Ilshat Saifullin

Why we need calculations on clusters?

- Frequently, research problems that use computing can outgrow the capabilities of the desktop or laptop computer where they started
 - The numerical validation of the model requires resolution study which usually increases the computation load from seconds or minutes to hours or days
 - Even when your forward numerical model can run in seconds, optimization or uncertainty quantification will require to run thousands or even millions simulations
 - Running things in parallel requires optimal communication between different processing units
 - Intensive calculations requires robust hardware solution with optimized cooling and power option



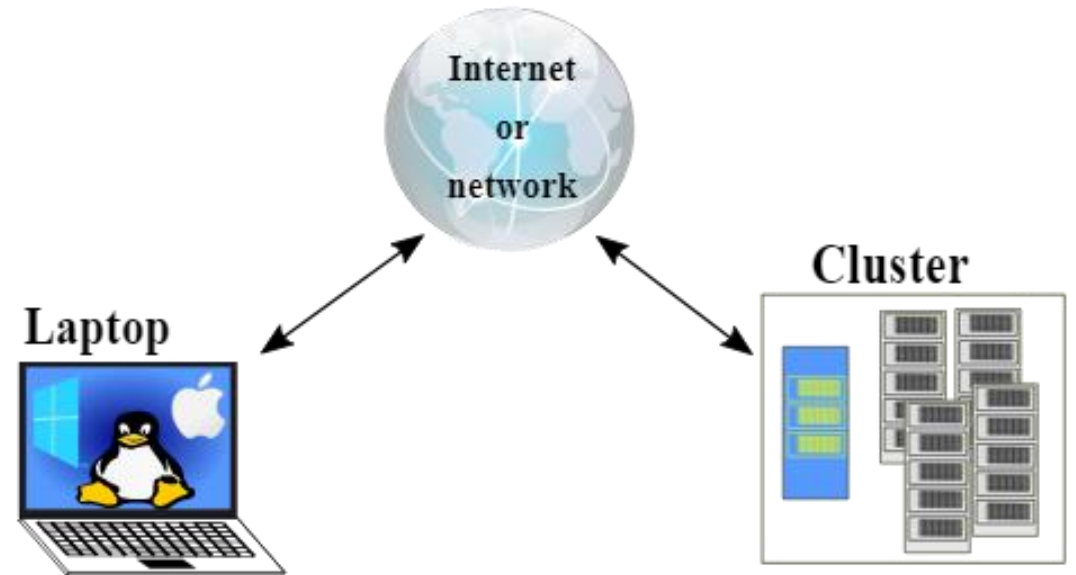
Some estimations

My forward model only needs 10 seconds to run. Why do I need to run it on cluster?

- Resolution study – refine grid and timestep until converge
 - In 3D problem, each grid resolution multiply run-time by 8 every time you resolve your grid twice
 - 3 levels of resolution makes single run longer than 1 hour
- Optimization based on forward simulation
 - Number of runs scaled by the number of optimization variables if gradient information is required
 - With 20 optimization parameters, 50 iterations of optimization will be run in ~ 3 hours
- Uncertainty quantification based on numerical model
 - Depends on complexity of the model, Monte-Carlo method requires from several thousands to millions of runs
 - Time spend on forward solution starts from day and can spend up to months

Running your program on cluster

- The majority of clusters using Linux as an operation system
- To connect to cluster, you need terminal and internet
- Your account should be created on cluster



Start with Linux

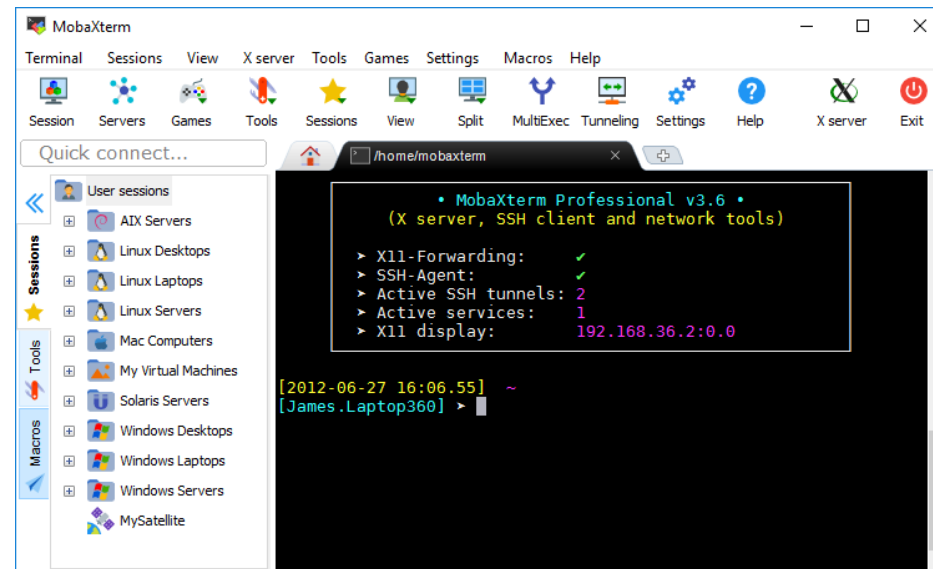
Several ways to start with Linux:

- Install Linux as a prime or second OS on your PC
- Connect to other Linux machine via ssh protocol. [mobaXterm](#)
- Cygwin (through Git Bash for example)
- Run Linux through Windows Subsystem for Linux (WSL2) - Windows 10 and higher
- Run terminal in Jupyter environment
- Install virtual machine, then install and run Linux as a guest OS. [VirtualBox](#)
- Install Docker, download preferred Linux image, run. [Docker Desktop](#)

Run inside your
Windows OS

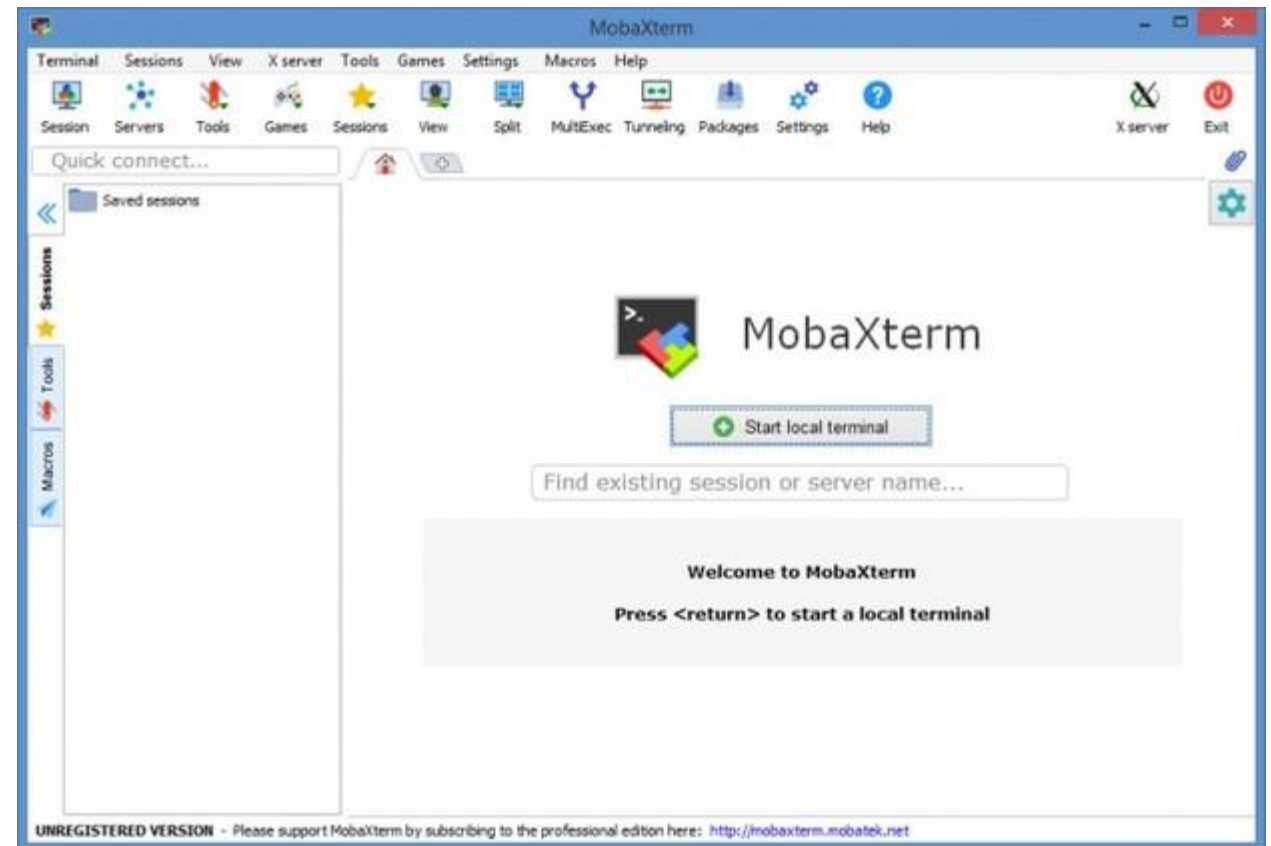
Use terminal application MobaXterm on Windows

- MobaXterm is your ultimate toolbox for remote computing.
- In a single Windows application, it provides loads of functions that are tailored for programmers, webmasters, IT administrators and pretty much all users who need to handle their remote jobs in a more simple fashion.
- MobaXterm provides all the important remote network tools (SSH, RDP, X11, SFTP, FTP, Telnet, Rlogin, ...) to Windows desktop, in a single portable exe file which works out of the box.



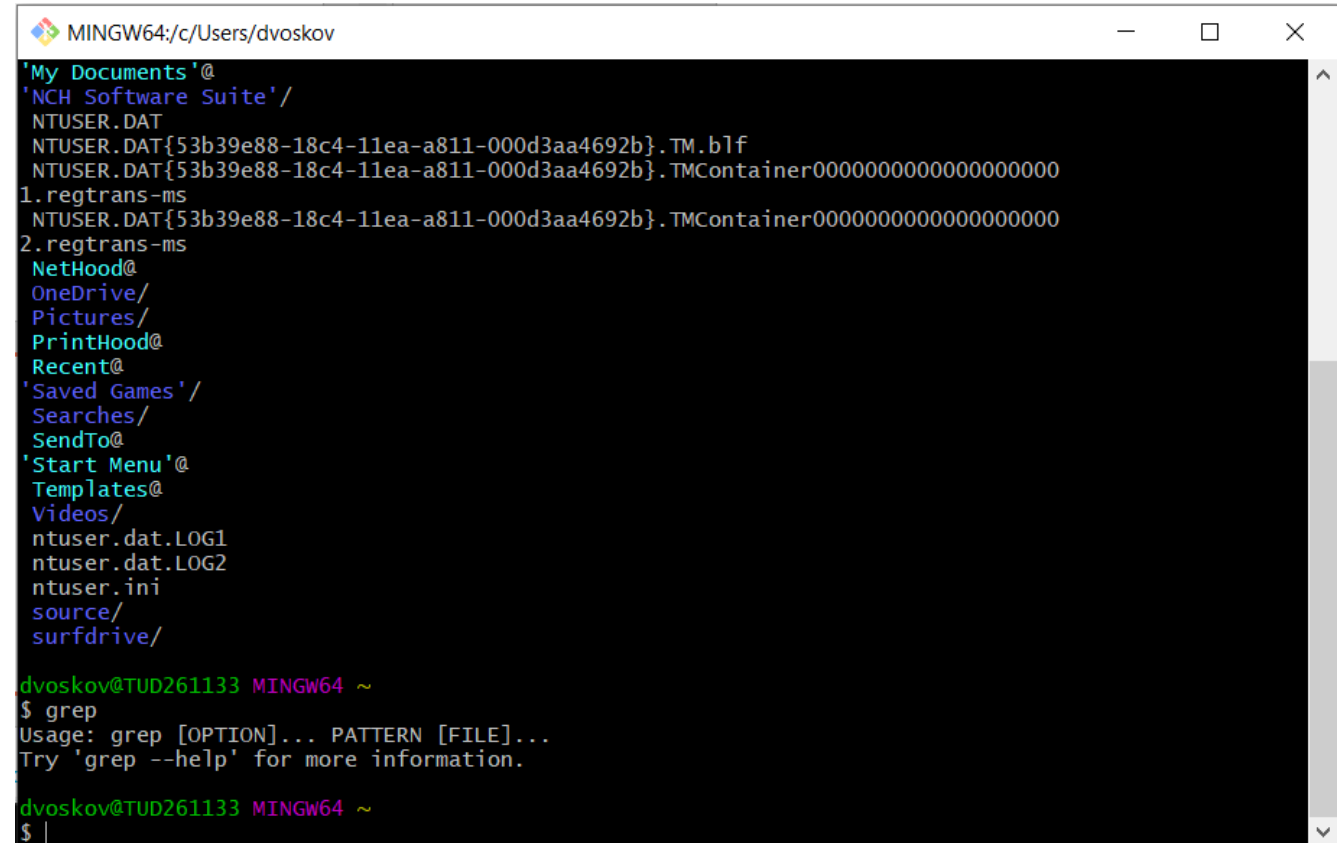
MobaXterm installation

- [Download the Home Edition of MobaXterm.](#) Make sure to select the Installer edition rather than the Portable edition.
- Click on the downloaded zip file to open it, and then click on the MobaXterm installer msi file to begin the installation.
- Once the install has finished, open the MobaXterm app. In the main window, you will see a list of saved sessions in the left-hand column. If you have previously installed and used PuTTY, any saved PuTTY sessions will also be listed.
- From here, you will want to start a local terminal. You can start a local terminal by clicking on the “Start local terminal” button in the MobaXterm main window.



Linux on Windows, option 1: Using Git Bash (local)

- Git Bash is a part of Git software which is already installed.
- It allows to execute Linux commands on Windows (uses MINGW).



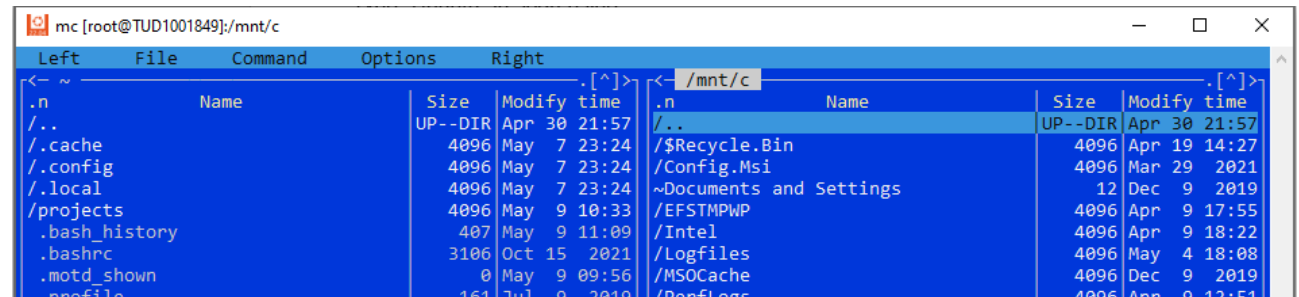
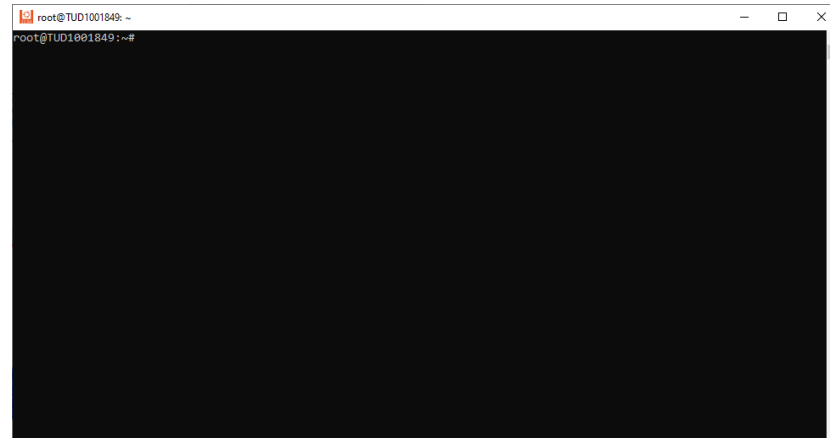
```
MINGW64:/c/Users/dvoskov
'My Documents'@
'NCH Software Suite' /
NTUSER.DAT
NTUSER.DAT{53b39e88-18c4-11ea-a811-000d3aa4692b}.TM.blf
NTUSER.DAT{53b39e88-18c4-11ea-a811-000d3aa4692b}.TMContainer00000000000000000000
1.regtrans-ms
NTUSER.DAT{53b39e88-18c4-11ea-a811-000d3aa4692b}.TMContainer00000000000000000000
2.regtrans-ms
NetHood@
OneDrive/
Pictures/
PrintHood@
Recent@
'Saved Games' /
Searches/
SendTo@
'Start Menu'@
Templates@
Videos/
ntuser.dat.LOG1
ntuser.dat.LOG2
ntuser.ini
source/
surfdrive/

dvoskov@TUD261133 MINGW64 ~
$ grep
Usage: grep [OPTION]... PATTERN [FILE]...
Try 'grep --help' for more information.

dvoskov@TUD261133 MINGW64 ~
$
```

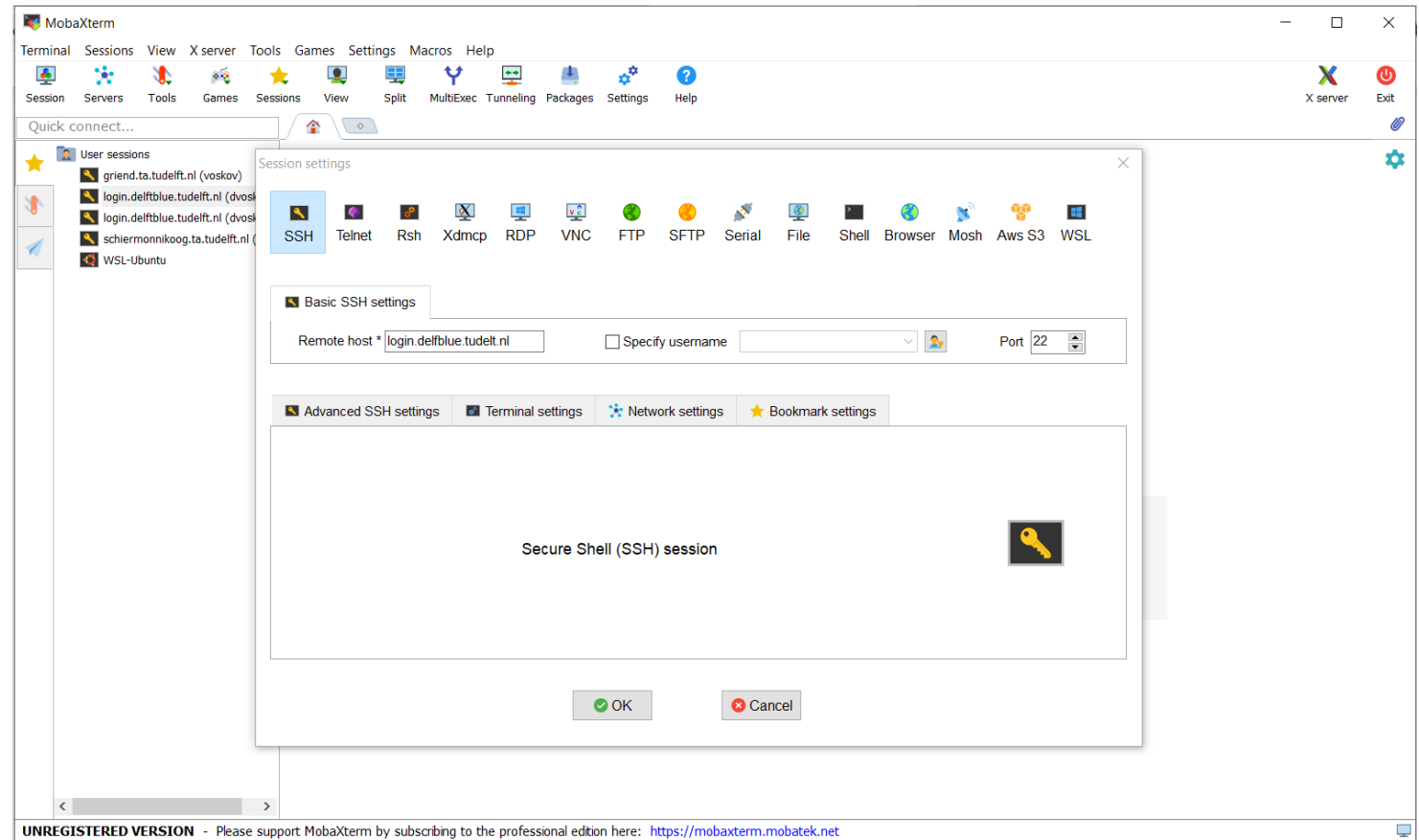

Linux on Windows, option 2: 10 PC through WSL2 (local)

- Run PowerShell from Administrator
- Run `wsl --install`
- Reboot Windows
- Open the Microsoft Store
- Type “Ubuntu” in search line
- Select “Ubuntu 22.04 LTS”
- Click “Get”
- Click “Run” after installation
- You will see the window with command line



Linux on Windows, option 3: connect to cluster (requires internet)

- Open MobaXTerm or another terminal
- In the menu, choose Session
- Click on SSH
- Write login.delfblue.tudelt.nl in 'Remote host'
- Put your netid in 'Specify username'
- Enter your password



Linux on Windows, option 4: Jupyter Hub (requires internet)

- Open Jupyter Hub in the browser
- Click on New and pick Terminal
- New tab will be opened with Linux terminal
- This is where we will perform our practical exercises

The screenshot displays the Jupyter Hub web interface. At the top, there's a header with the 'jupyterhub' logo and 'Logout' and 'Control Panel' buttons. Below the header, there are tabs for 'Files', 'Running', and 'Clusters'. A message says 'Select items to perform actions on them.' Below this, there's a list of folders: '0', '/', 'Week_1_1', 'Week_1_2', 'Week_1_3', 'Week_1_4', and 'Week_1_5'. A 'New' button is visible, and a dropdown menu is open, showing options: 'Notebook: Python 3 (ipykernel)', 'R', and 'Other: Text File', 'Folder', 'Terminal'. The 'Terminal' option is highlighted. Below the dropdown, there's a table with columns for 'Name', 'Created', and 'Size'. The table shows several entries, including 'a month ago', '21 days ago', and '2 months ago'.

Below the main interface, there's a terminal window with the prompt 'lcur1922@r11n24:~\$'.

Some of Linux differences to Windows

- Multi-user OS (several users can simultaneously work in Linux but only few and only in Server editions of Windows)
- User/admin permissions (sudo, su)
- Package manager (same as Microsoft Store, but one can add additional sources for custom packages)
- Directory tree, path separator
- Case in file names is important
- Encoding, line ending
- No need in file extensions
- File and folder permissions
- Windows executable and libraries will not work in Linux
- Environment variables. LD_LIBRARY_PATH – place where to search dynamic libraries at runtime
- No bin (be careful with deleting files and folders)
- Different file system (ext4) with symbolic links