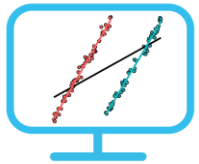
Imperial College London



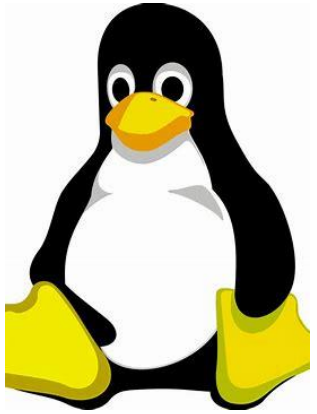
DS Helper Team 2020

Linux and Command Line

Zhuoyu (Joy) Li

2nd June 2021

What is Linux?

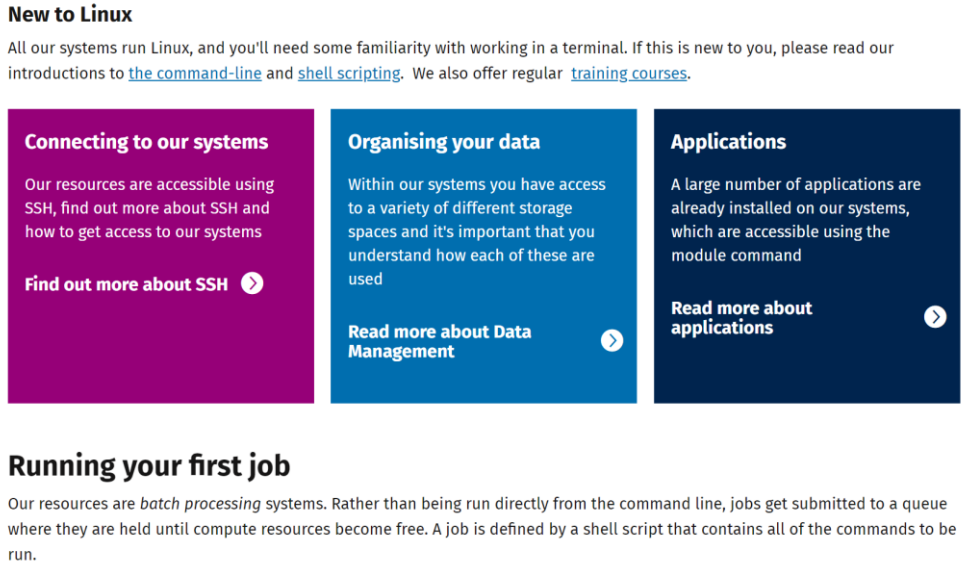


- Just like Windows, iOS, and Mac OS, Linux is an operating system. It is free and open-source.
- An operating system is software that manages all of the hardware resources associated with your desktop or laptop.
- Many enterprise-level projects (c/c++/python/java) are deployed to Linux system. One of the most popular platforms, Android, is powered by the Linux operating system.

- Linux has a number of different versions to suit any type of user.
- Distributions (distros): specific group of software and conventions
- Major distributions: Arch, Debian (Linux Mint, Ubuntu, Kali, etc.), Red Hat (CentOS, Fedora, etc.), Slackware

Why Linux?

- For us … HPC!
- Much of scientific experimentation such as modeling, simulations or data management and mining is made possible by computing -> **high performance computing (HPC) cluster.**
- HPC environment: Linux operating system & Bash shell
- Available teaching courses https://www.imperial.ac.uk/admin-services/ict/self-service/research-support/rcs/support/getting-started/

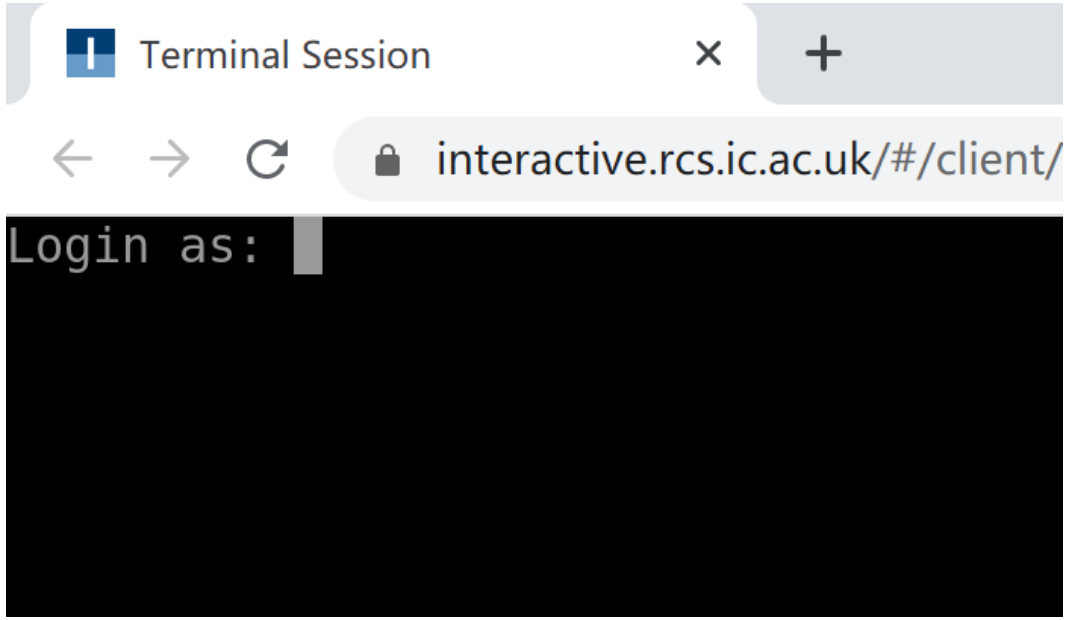


- Bioinformatics is highly dependent on Linux-based computers and software.

1. With the advent of new technologies (e.g. sequencing), analysis tend to rely more and more on powerful computing clusters.
2. Linux systems come with pre-installed versions of most popular programing languages and thus tend to become the natural platform for newly developed software(some programs are exclusively developed for Linux-based systems).

Environment

- Terminal
- Mac – easier – Terminal
- Windows – harder – “run Linux on Windows”
  possible: Ubuntu, CentOS, Fedora…



- Login via the browser http://login.rcs.ic.ac.uk
  Without VPN? Try this: https://copy.sh/v86/?profile=linux26

Basic Bash commands

- **ls:**

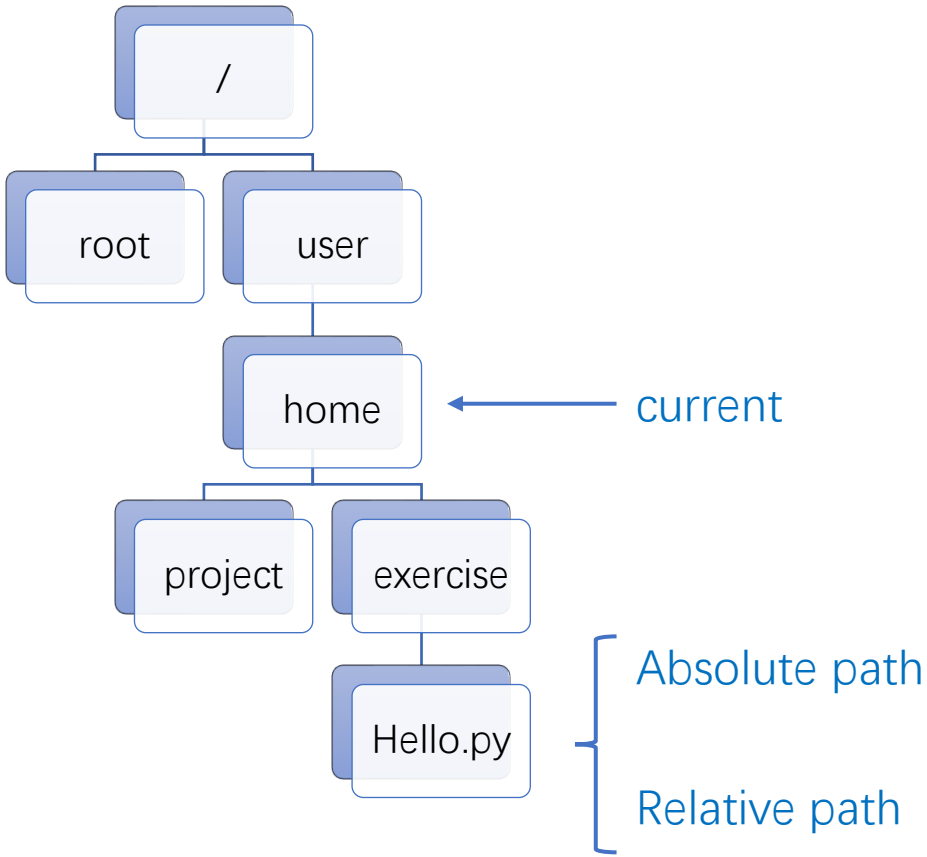
Description: lists contents of current directory.
Common options:
-a lists all files of the current directory including the hidden files.
-l lists contents of the current directory with extra details.
Examples: list all the contents in a specific directory.

- **man:** Displays the manual page of a given command.
- **help:** Displays information about built-in commands.

- **pwd:**

Description: prints working directory (the full path), or your location on the filesystem)
Absolute path (full path) and relative path

- Absolute path and relative path



- **cd:**

Description: change directory

Examples:

change directory to

1. child directory
2. parent directory
3. a specific directory using absolute path
4. a specific directory using relative path
5. your home directory
6. the previous directory

Folders

- **mkdir:**

Description: makes a directory
Examples:
1. Specify the path to where you want to create the new directory.
2. Create multiple directories.
3. Create nested directories.

- **rmdir:**

Description: removes a directory, which should be empty.

Files

- **cp:**

Syntax: cp source_file destination_file
Description: copy files and folders from one place to another.
Examples:
1. Copy files to a different path.
2. Copy multiple files.

- **rm:**

Description: remove files.
Options:
-r Remove directories and their contents recursively.
-d Remove empty directories. Equivalent to using rmdir.
Examples: remove multiple files with similar names.

- **mv:**

Examples:

1. move a file to different directory
2. move multiple files and the use of *
3. rename a file
4. move a directory
5. rename a directory

”

cp command = copy and paste

mv command = cut and paste

- **cat:**

Description: concatenate files and displays the output to the shell.

Common options:

-b number all the non-empty lines

-n number all output lines

Examples: Create a new file and check. (**>** redirects the output from a command.)

Editors: vim, Sublime, Eclipse, etc.

- **wget:**

Description: a network downloader which is used to download files from the server. It supports HTTP, HTTPS, and FTP protocols, as well as retrieval through HTTP proxies.

Common options:

-c resume a partially downloaded file

-o save the downloaded file under a different name

-P save the file to a specific location

- **|:** redirect the output of the first command as the input to next command.
- **more:** shows the contents of one or more files, one page at a time.
- **less:** displays the content of the file or the output of a command.

- **head:**

Description: print the top N numbers of data of the given input.
Common options: -n num: prints the first ‘num’ lines (default: 10).

- **tail:**

Description: print the last N numbers of data of the given input.
Common options: -n num: prints the last ‘num’ lines (default: 10).

- **wc:**

Description: count the number of lines, words, characters, and bytes of each given file or standard input and print the result.
Common options:
-l --lines - Print the number of lines.
-w --words - Print the number of words.
-m --chars - Print the number of characters.
-c --bytes - Print the number of bytes.
-L --max-line-length - Print the length of the longest line.