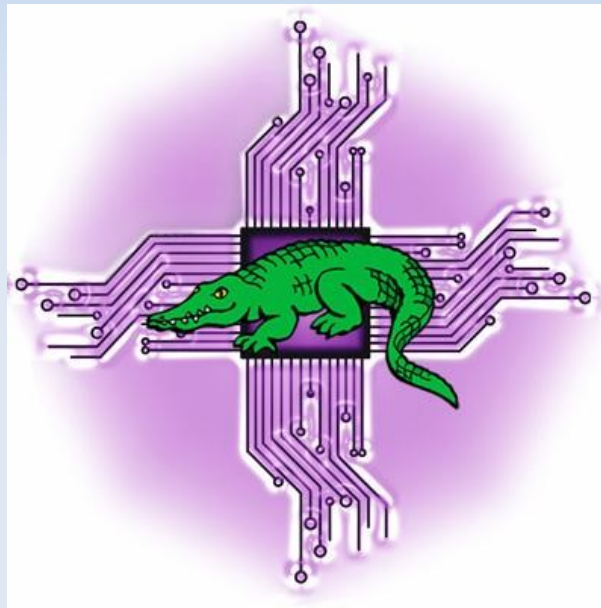# Linux on Spartan

COMP90024 Cluster and Cloud Computing

University of Melbourne, 21ˢᵗ March 2023

lev.lafayette@unimelb.edu.au

Prerequisites and CLI Reasons

An account on Spartan and membership to a project is required.

An SSH (secure shell client) is required. This is typically already installed on Mac and Linux systems. For MS-Windows putty or mobaxterm are options.

The command line interface (CLI) is "close to the metal", and resource efficient. It is also powerful and, with scripting, frees up time and effort for the user.

Knowing the Linux CLI is essential for any serious researcher!

SSH and Logging In

Access the logon node with SSH command; e.g.,

ssh lev@spartan.hpc.unimelb.edu.au

Process can be improved with SSH configuration files and "passwordless SSH" that uses key-exchange for authentication. Agent-forwarding provides additional convenience.

https://dashboard.hpc.unimelb.edu.au/ssh/

File System and Environment

"Everything in the UNIX system is a file" (Kernighan & Pike, 1984, 41).

Login will take the user to /home/$USERID

https://swcarpentry.github.io/shell-novice/fig/standard-filesystem-hierarchy.svg

Core environment commands;
whoami : print user id
pwd : print working directory
ls : List the directory contents
date : print system date and time

Shell Shortcuts

~   Home directory
.   Current working directory
..   One directory level closer to root
/   root directory, directory delimiter
^r Traverse command history or use navigation keys
^L Clear the screen; equivalent of the "clear" command.

Tab completion provides navigation and command options

e.g.,
$ ls /usr/local/common/Int

Interact/   IntroLinux/

Command Options and Manual

Commands are modular and robust and come with a default execution. Alternative execution is invoked with options; <command> -<abbreviation> or <command> --<full-word>

For example:
ls -lart : Directory listing, all files, reverse time of modification
ls -lash : Directory listing, all files, size is in human format

Hundreds of commands, thousands of options! Fortunately, there is a manual :)

man ls or info ls : Summary of "ls" command and options.
whatis ls : Short description.
man -k compiler or apropos compiler : Search man pages.
man man: The manual page for manual pages.

Pipes and Redirects

The pipe is typically used to redirect the output of one command as the input to another for further processing. Multiple commands can be piped in sequence to develop some very specific outcomes.

who -u | less : List of users logged in, piped to less (viewer)
who | wc -l : Count lines of users logged in

Output and input can redirected to/from a file rather than standard output or standard input. Output redirects can be appended.

w > list.txt : Redirect to list.txt
w >> list.txt : Redirect and append to list.txt

Files and Text Editing

Filenames can include any characters except the forward slash (directory delimiter) or null. A good habit to avoid non-printing characters (e.g., spaces)! Consider using snake_case, kebab-case, or CamelCase, instead.

The "touch" command creates a file with no content e.g., touch LIST.TXT

Filenames are case sensitive; list.txt, LIST.TXT are different. File types can be determined by the file command (e.g., file list.txt).

Three main text editors available; nano (easy), emacs (feature-rich), and vim (powerful).

Local Copying

To copy files from a remote system to a local system one can use tools like wget, git, cURL, scp, or rsync. For example:

wget https://swcarpentry.github.io/shell-novice/fig/standard-filesystem-hierarchy.svg
wget -r --no-parent https://swcarpentry.github.io/shell-novice/

To copy files and directories within a system use the cp command. Common options include -r (recursive), -a (archive permissions), -u (update, copy only if source is newer).

cp /usr/local/common/IntroLinux/gattaca.txt .
cp -r /usr/local/common/IntroLinux

Remote Copying

To copy files between systems use scp/stfp or rsync. GUI clients are available (e.g., Filezilla). Again, passwordless SSH and SSH client files are useful!

scp lev@spartan.hpc.unimelb.edu.au:gattaca.txt .
scp -r lev@spartan.hpc.unimelb.edu.au:IntroLinux .

The rsync utility provides a fast way to keep two collections of files "in sync" by tracking changes. This example keeps permissions (-a), is verbose (-v), compresses between source and destination (-z), and only syncs if the source is newer (--update)

rsync -avz --update workfiles/
lev@spartan.hpc.unimelb.edu.au:files/workfiles

Remote Copying

To copy files between systems use scp/stfp or rsync. GUI clients are available (e.g., Filezilla). Again, passwordless SSH and SSH client files are useful!

scp lev@spartan.hpc.unimelb.edu.au:gattaca.txt .
scp -r lev@spartan.hpc.unimelb.edu.au:IntroLinux .

The rsync utility provides a fast way to keep two collections of files "in sync" by tracking changes. This example keeps permissions (-a), is verbose (-v), compresses between source and destination (-z), and only syncs if the source is newer (--update)

rsync -avz --update workfiles/
lev@spartan.hpc.unimelb.edu.au:files/workfiles

Creating Directories, Moving Files

Directories can be created with the mkdir (make directory) command. Multiple directories can be created at the same time with space delimiting the separate directories.

The mv (move) command will move (and rename) a file. There are options for -f (force), -i (interactive), and -n (no clobber). If more than one of these is selected only the last one is applied.

Example:
mkdir BRAF
nano gattaca.txt (make a single character change)
mv gattaca.txt BRAF/
scp -r BRAF lev@spartan.hpc.unimelb.edu.au:

File Differences

The -l option is ls gives a timestamp for modification. Differences in content can be determined by the diff command. Or, with a side-by-side representation with sdiff. Further, there is the comm command, which compares two files sorted line by line.

ls -l gattaca.txt BRAF/gattaca.txt
diff gattaca.txt BRAF/gattaca.txt
sdiff gattaca.txt BRAF/gattaca.txt
comm gattaca.txt BRAF/gattaca.txt

Note that the diff command also gives the sequence of commands for the very old UNIX editor, ed. This is still used for invoking patch files!

File Searches, Globbing

On many Linux systems one can use the locate command to find files; but often disabled a large shared system like Spartan!

Instead, use the "find" command; very powerful, many options. The following examples also use wildcards (*, all files and ?, single character) and ranges.

find . -name "*.log"
find . -name "comsol?.log"
find . -name "comsol[1-3].log"

The term "glob" is short for "global" in the UNIX function glob().

Searches Within Files

Searching within a file for a string of text is achieved with grep ("global search for regular expression and print"). Often used with "strong" quotes to protect from command substitution or variables.

The following command is grep, with the option -i for "case-insensitive", followed by the regular expression 'ATEK'.

```
grep -i 'ATEK' BRAF/*
touch braf/file1
grep -i 'ATEK' BRAF/*
```

Deletions

Deleting files actually means deleting files on the CLI! There is no "trashcan" or "recycle bin" or anything like that.

Files can be deleted with the "rm" (remove) command. Empty directories can be deleted with "rmdir" (remove directory), however the directory must be empty.

A very powerful shortcut is rm -rf (remove, recursive, force).

rm gattaca.txt
rmdir braf/
rm -rf braf/

Think before running rm -rf – and then think again!

This Matters

BRAF is a human gene that makes a protein (imaginatively) named B-Raf. This protein is involved in sending signals inside cells, which are involved in directing cell growth.

In 2002, it was shown to be faulty (mutated) in human cancers. In particular, the difference between the two files "ATVKSRWSGS" and "ATEKSRWSGS" is the difference that leads to susceptibility to metastatic melanoma.

THANKS FOR WATCHING & LISTENING PATIENTLY