STATS 769
Linux and the shell

Paul Murrell

The University of Auckland

August 6, 2019

Overview

- This section of the course (two lectures) explores the Linux operating system and its shell command-line interface.

- We need Linux for several reasons:
  - Linux gives us a command-line interface to the operating system (and the file system), i.e., we can write code.
  - The software and/or the data we need may be on Linux.
  - Big, fast servers are likely to be running Linux.
  - There are some things that are better/faster to do in the shell than in R.
  - Linux is good for you.

- Many different distributions, but very similar user experience; we will mostly work in Ubuntu.
- A desktop Linux GUI looks just like Windows or Mac, just less trendy.
- We will be working with the CLI rather than the GUI.
- Look for icon or menu item called "XTerm" or "Terminal" or "Terminal Emulator".

Linux
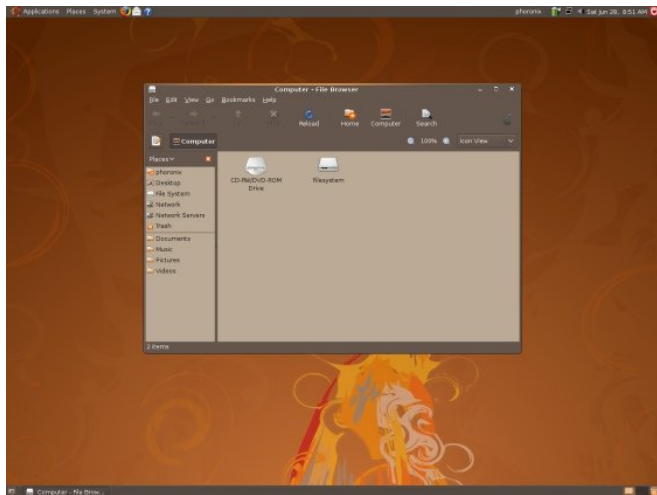
Getting Linux

- Putty from Windows to an existing Linux machine.
  - Centre for eResearch Aspirational Virtual Machine Farm
    upi`@sc-cer00014-04.its.auckland.ac.nz`
    upi`@sc-cer00014-05.its.auckland.ac.nz`
  - Requires VPN from off campus
    `https://www.auckland.ac.nz/en/students/`
    `academic-information/postgraduate-students/`
    `postgraduate-support-and-services/vpn-service.`
    `html`
- Teaching lab (e.g., 302.190) and grad lab machines (303S.175) should be able to boot into Linux.
- Install Linux (possibly as Virtual Machine via VirtualBox)
  - Another option on Windows is to install linux-like software such as Cygwin.

Getting Linux
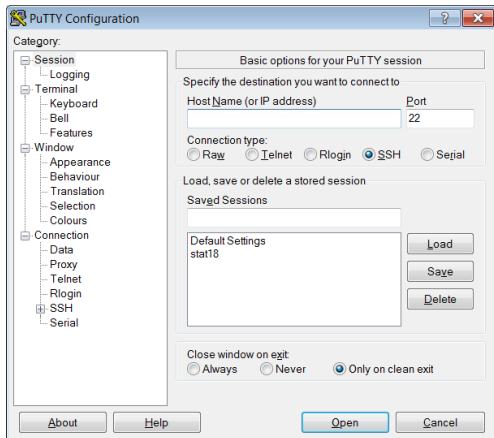
The shell

- A command-line interface to the operating system.
  (Not entirely unlike the command-line in R)
- Communication is text-based: text commands in and text
  output out.
- A "terminal" program allows us to send commands to the shell
  and display the results. It is common to have several terminals
  open at once.
- We can perform all of the tasks that, in a GUI, we would use
  a mouse for, by typing commands at the shell instead.
- Typing commands is writing code: we can record our actions
  (so we can repeat actions, adapt actions, document actions,
  share actions); we can express more complex actions; we can
  programmatically control our actions; we can work without a
  GUI (e.g., on a remote server).
- There are several shell programs; we will use bash (the
  Bourne-Again SHell).

Shell file system commands

`pwd`	print the current (working) directory.
`ls`	list files (in the current directory).
`mkdir`	make a new directory.
`cd`	change directory.
`cp`	copy a file (or directory).
`mv`	move a file (or directory).
`rm`	remove (delete) a file (or directory).
`rmdir`	remove (delete) a directory.
`df`	report file system disk space usage.

WARNING: `rm` does NOT ask for confirmation.

Linux file paths

- "directory" is the same as "folder".
- Can be just a file name (implies a file in the current directory).
- Can be a series of directory names, separated by /.
- Can start with /, in which case the path is absolute (starts from the "root" of the file system).
  Otherwise the path is relative to the current directory.
- . means the current directory.
- .. means the parent directory.

Shell file system commands

Each shell command has options that control its action.

- The `man` command displays help information about a command.
- `ls -latr`
- `cp -ru`
- `rm -r`

Linux help pages

```
LS(1)                           User Commands                          LS(1)

NAME
       ls - list directory contents

SYNOPSIS
       ls [OPTION]... [FILE]...

DESCRIPTION
       List  information  about  the FILEs (the current directory by default).
       Sort entries alphabetically if none of -cftuvSUX nor --sort  is  speci-
       fied.

       Mandatory  arguments  to  long  options are mandatory for short options
       too.

       -a, --all
             do not ignore entries starting with .

       -A, --almost-all
             do not list implied . and ..
```
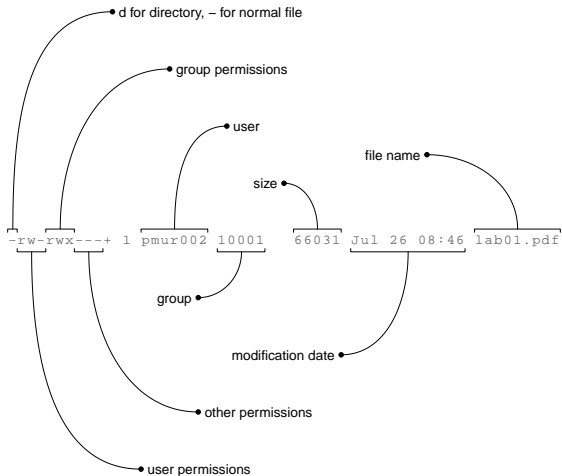
Press the space bar to show the next page, press the Enter key to
show the next line, press the Q key to exit.

Linux file listings

Linux file permissions

- Three sets of `rwx` for user, group, and other.
- `r` means we can read the file.
- `w` means we can write to the file.
- `x` means we can run (execute) the file.
- For a directory, `r` means we can list the directory contents, `w` means we can create and remove files within the directory, `x` means we can enter the directory.
- `chmod` can be used to modify permissions.
- `chmod u+rwx`
- `chmod -w`

Every Linux distribution comes with a basic set of useful programs.

```
echo   Print a value on the screen.
cat    Print a file to the screen.
head   Print the first few lines of a file.
tail   Print the last few lines of a file.
more   Print a file one page/screen at a time.
```
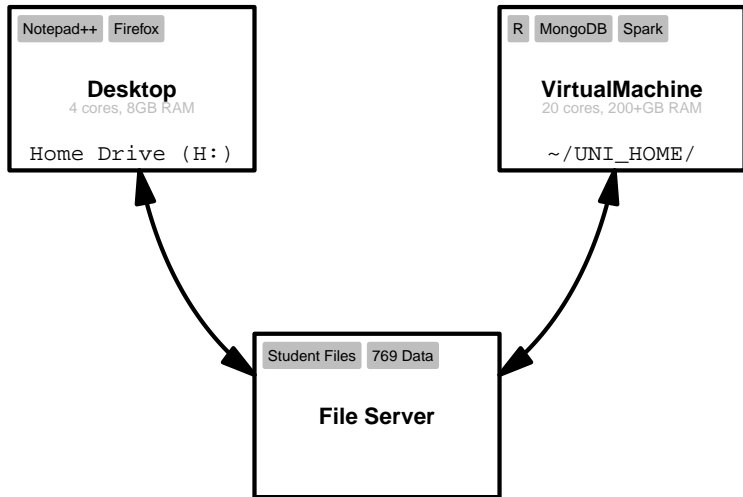
Running programs in the shell

- Shell commands are just small programs, so we have already seen that we run a program by typing its name, plus possibly some options, plus possibly some arguments.
- Use `Ctrl-c` to kill a program.

Compared to a single-user, local storage, Windows PC, Linux provides a …

- Multi-user environment.
  - Use `who` to see who is logged on to the computer.
- Shared-resource environment.
  - Use `top` to see which programs and users are using which resources.

Running programs in the shell

Some differences between running programs in Windows and running programs in the shell …

- programs do not necessarily provide a GUI window, only a CLI, or not even that.
- program input comes from the keyboard or a file and program output goes to a file or to the screen.
- both input and output are often just text.

Running R in the shell

- Just type `R`
- The working directory defaults to the directory in which R started.
- Cutting-and-pasting in Linux can be done with a mouse via select-and-middle-click.
  (When using Putty on Windows, select and right-click?)
- The `source()` function becomes useful for running code in a `.R` file.
- Rscript can be used to run a file in "batch mode."

If we are using Putty (or ssh) to access a remote machine ...

- Work in a directory that is shared by (mounted on) both the local system and the remote system.
- Generate graphics files.

Processing R Markdown files in the shell

- Rscript, rmarkdown::render()
- Run shell commands via system()

Editing files in the shell

The `vi` program is a text editor that works within a terminal.

i	Enter edit mode (typing inserts new text).
o	Insert new line below and enter edit mode.
O	Insert new line above and enter edit mode.
Esc	Exit edit mode (enter normal mode).
	In normal mode
h,j,k,l	Move one left, down, up, or right.
$	Move to the end of the line.
0 (zero)	Move to the start of the line.
Ctrl-f	Page down.
Ctrl-b	Page up.
x	Delete character.
dd	Delete line.
p	Paste line.
u	Undo (only last action).
:w	Save.
:q	Quit.

Resources

- Linux shell introduction
  http://linuxcommand.org/index.php
- Software Carpentry materials for UNIX shell (Sections 1, 2, and 3)
  http://swcarpentry.github.io/shell-novice/
- Putty
  http://www.chiark.greenend.org.uk/~sgtatham/putty/download.html
- 'knitr' Options
  https://yihui.name/knitr/options/