Introduction to Linux

Francisco Fuster Tormo 13 - 03 - 18 ffuster@carrerasresearch.org

What is Linux?

Operating System (OS) based on UNIX

- Multiuser → Allows more than one user at the same time
- Multitasking → Allows more than one program executed at the same time
- Networking → Allows network access
- Various user interfaces → Different "desktops" allowed

Other UNIX-based OS?

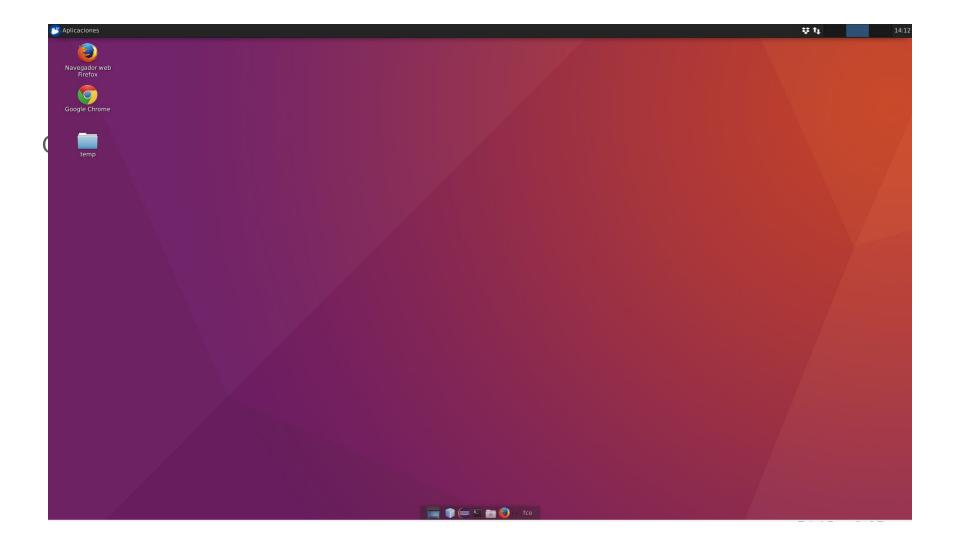
- Mac OS
- Android

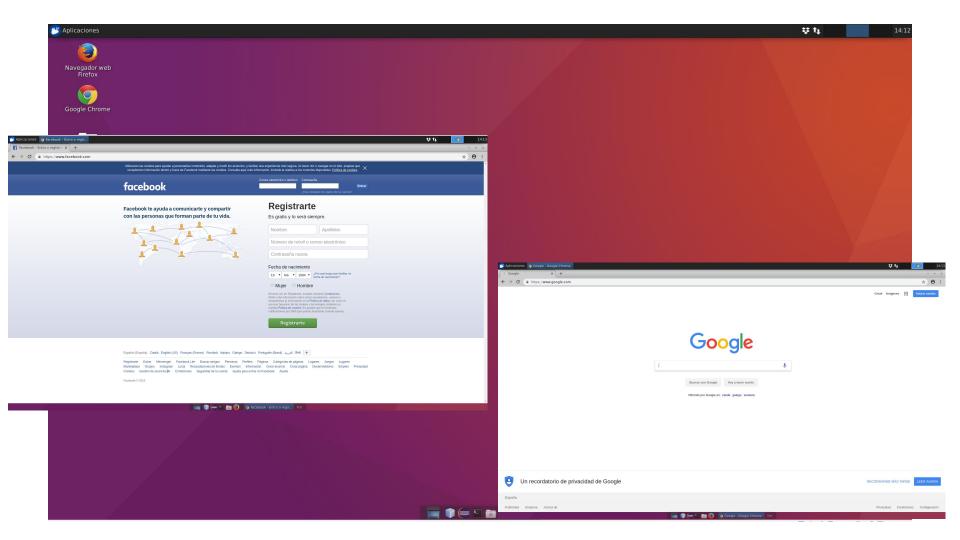
How many Linux are?

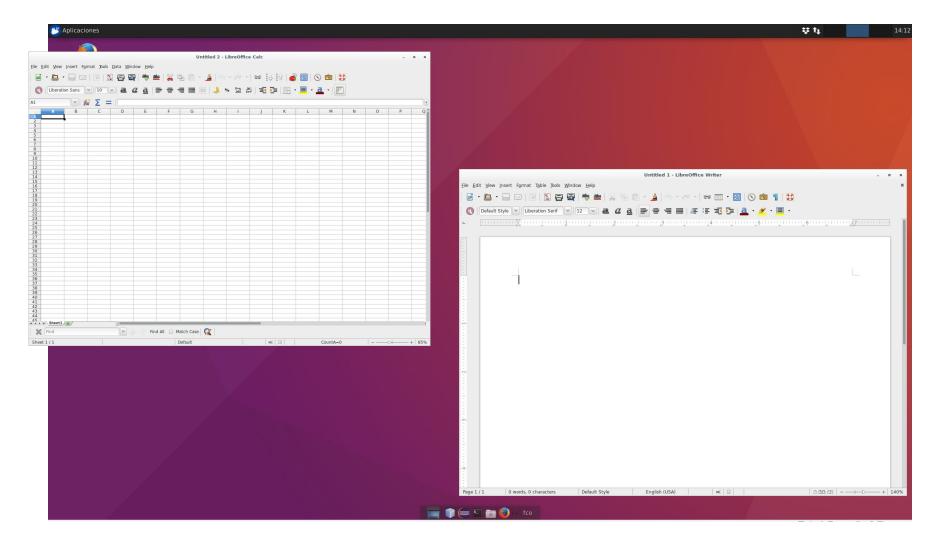
- Debian
- Ubuntu
- Linux Mint
- Red Had
- Fedora
- OpenSuse
- CentOS
- Gentoo
- ...

Is Linux difficult to use?

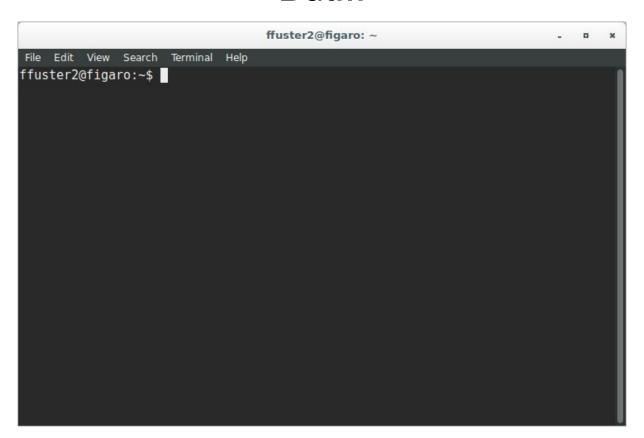
Graphic User Interface (Desktop)







But...

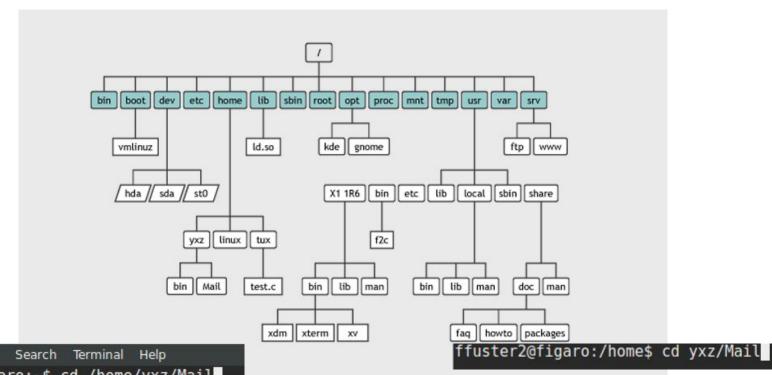


Differences between Windows & UNIX

```
fco@fco-VirtualBox: ~
Archivo Editar Ver Buscar Terminal Ayuda
fco@fco-VirtualBox:~$ ls
                                  Imágenes NetBeansProjects Público Vídeos
          Documentos eclipse
                      Escritorio Música
                                           Plantillas
Descargas Dropbox
fco@fco-VirtualBox:~5
```

```
CX Command Prompt
                                                                                                            - | U ×
Microsoft Windows XP [Version 5.1.2600]
(C) Copyright 1985-2001 Microsoft Corp.
C:\Documents and Settings\Fco>ls
'ls' is not recognized as an internal or external command,
operable program or batch file.
C:\Documents and Settings\Fco>
```

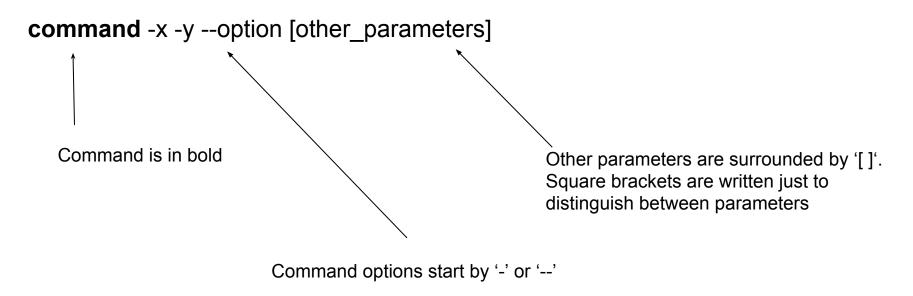
The Linux filesystem



File Edit View Search Terminal Help

ffuster2@figaro:~\$ cd /home/yxz/Mail

Special considerations



Linux basic commands (1). Folder navigation

- cd [folder] → Navigation in directories
- **Is** → List the files/directories in a path
- pwd → Get absolute path of current directory
- mkdir [folder_name] → creates the folder [folder_name] in the current path
- **find** [path] -name [pattern] → Searches files in the specified position

```
fco@fco-VirtualBox:~$ ls
Descargas Dropbox Escritorio Música Plantillas R
Documentos eclipse Imágenes NetBeansProjects Público Vídeos
fco@fco-VirtualBox:~$ cd Escritorio/
fco@fco-VirtualBox:~/Escritorio$ ls
firefox.desktop google-chrome.desktop temp
```

```
fco@fco-VirtualBox:~/Escritorio$ pwd
/home/fco/Escritorio
```

```
fco@fco-VirtualBox:~/Escritorio$ find . -name new*
/new_dir
```

```
fco@fco-VirtualBox:~/Escritorio$ mkdir new_dir
fco@fco-VirtualBox:~/Escritorio$ ls
firefox.desktop google-chrome.desktop new_dir temp
```

Exercise (1)

Go to your home folder /home/user

Create the directory 'course' in you home folder

List all the elements you have in your home folder

Get absolute path of your folder

Exercise (1)

Go to your home folder /home/user

```
fco@fco-VirtualBox:~$ cd /home/fco/
```

Create the directory 'course' in you home folder

```
fco@fco-VirtualBox:~$ mkdir course
```

List all the elements you have in your home folder

```
fco@fco-VirtualBox:~$ ls
course Documentos eclipse Imágenes NetBeansProjects Público Vídeos
Descargas Dropbox Escritorio Música Plantillas R
```

Get absolute path of your folder

```
fco@fco-VirtualBox:~$ pwd
/home/fco
```

Linux basic commands (2). Copy, move, remove files

- **' → touch** [file_name] → creates an empty file called [file_name]
- cp [file_origin] [file_destination] → copy [file_origin] in [file_destination]
- rsync [origin_folder] [destination_folder] → Synchronize both folders
- mv [file_origin] [file_destination] → move [file_origin] to [file_destination]
- rm [file] → removes [file]

Linux basic commands (2). Copy, move, remove files

- **touch** [file name] → creates an empty file called [file name]
- **cp** [file origin] [file destination] \rightarrow copy [file origin] in [file destination]
- **rsync** [origin folder] [destination folder] → Synchronize both folders

```
mv [file_origin] [file_fco@fco-VirtualBox:~$ cd course/fco@fco-VirtualBox:~/course$ ls
rm [file] → removes fco@fco-VirtualBox:~/course$ touch file_name.txt
                          file name.txt
                          fco@fco-VirtualBox:~/course$ cp file_name.txt file_name_copy.txt
                          fco@fco-VirtualBox:~/course$ ls
                          file_name_copy.txt file_name.txt
                          fco@fco-VirtualBox:~/course$ mv file name copy.txt file name new.txt
                          fco@fco-VirtualBox:~/course$ ls
                          file_name_new.txt file_name.txt
                          fco@fco-VirtualBox:~/course$ rm file_name.txt
                          fco@fco-VirtualBox:~/course$ ls
                          file name new.txt
                          fco@fco-VirtualBox:~/course$
```

Exercise (2)

Create a new file called ex2.sh in your home folder

Copy this file to your recently created folder

Remove the file from your home folder

Exercise (2)

Create a new file called ex2.sh in your home folder

```
fco@fco-VirtualBox:~$ touch ex2.sh
```

Copy this file to your recently created folder

```
fco@fco-VirtualBox:~$ cp ex2.sh course/
```

Remove the file from your home folder

```
fco@fco-VirtualBox:~$ rm ex2.sh
```

Exercise (2)

Create a new file called ex2.sh in your home folder

fco@fco-VirtualBox:~\$ touch ex2.sh

Copy this file to your recently created folder

Remove the file from your home folder

fco@fco-VirtualBox:~\$ rm ex2.sh

fco@fco-VirtualBox:~\$ mv ex2.sh course/

Linux commands (3). Working with files

- head -n [number_of_lines] [file] → Show first [number_of_lines] of the [file]
- tail -n [number_of_lines] [file] → Show last [number_of_lines] of the [file]
- cat [file] → Print all the file
- less [fille] → Print all the file
- grep [word] [file] → Find [word] in [file]
- **cut** -f [columns] [file] → Show only [columns] of [file]
- sort -k [column] [file] → Sort the lines of a file by [column]
- **diff** [file1] [file2] → Compare [file1] and [file2] line by line
- uniq [file] → Report or omit repeated lines in file
- **tr** → Translate or delete characters

Examples (2)

cat & less fco@fco-VirtualBox:~\$ less example.txt fco@fco-VirtualBox:~\$ cat example.txt Hi humans! This is just a text example With a lot of different lines I'm planning to have, at least 20 But I think I'm not going to have enough ideas I will try No, sure I haven't How many lines do we have? 8 Bullshit, not enough. We need at least 10 more So . . . What I'm going to write? Idea! just count last lines and that's it Line 13 Line 14 Line 15 Line 16 Line 17 Almost done. This is 18 Second last: 19 And line 20!! Got it! Congratulations! fco@fco-VirtualBox:~\$

Hi humans! This is just a text example With a lot of different lines I'm planning to have, at least 20 But I think I'm not going to have enough ideas I will try No. sure I haven't How many lines do we have? 8 Bullshit, not enough. We need at least 10 more So... What I'm going to write? Idea! just count last lines and that's it Line 13 Line 14 Line 15 Line 16 Line 17 Almost done. This is 18 Second last: 19 And line 20!! Got it! Congratulations!

head & tail

```
fco@fco-VirtualBox:~$ head -n 5 example.txt
Hi humans!
This is just a text example
With a lot of different lines
I'm planning to have, at least 20
But I think I'm not going to have enough ideas
fco@fco-VirtualBox:~$
```

```
fco@fco-VirtualBox:~$ tail -n 8 example.txt
Line 14
Line 15
Line 16
Line 17
Almost done. This is 18
Second last: 19
And line 20!! Got it!
Congratulations!
fco@fco-VirtualBox:~$
```

Examples (3)

diff

```
fco@fco-VirtualBox:~$ cp example.txt example2.txt
fco@fco-VirtualBox:~$ echo "new line attached" >> example2.txt
fco@fco-VirtualBox:~$ tail example2.txt
Line 13
Line 14
Line 15
Line 16
Line 17
Almost done. This is 18
Second last: 19
And line 20!! Got it!
Congratulations!
new line attached
fco@fco-VirtualBox:~$ diff example.txt example2.txt
21a22
 new line attached
fco@fco-VirtualBox:~$ diff example2.txt example.txt
22d21
 new line attached
fco@fco-VirtualBox:~$
```

Examples (4)

sort, uniq & cut

```
fco@fco-VirtualBox:~$ cat example2.txt
cell1-1 cell1-2
cell2-1 cell2-2
cell3-1 cell3-2
cell3-1 cell3-2
cell3-1 cell3-2
cell4-1 cell4-2
fco@fco-VirtualBox:~$ cut -f 1 example2.txt
cell1-1
cell2-1
cell3-1
cel13-1
cell3-1
cell4-1
fco@fco-VirtualBox:~$ uniq example2.txt
cell1-1 cell1-2
cell2-1 cell2-2
cell3-1 cell3-2
cell4-1 cell4-2
fco@fco-VirtualBox:~$
```

```
fco@fco-VirtualBox:~$ sort example.txt
Almost done. This is 18
And line 20!! Got it!
Bullshit, not enough. We need at least 10 more
But I think I'm not going to have enough ideas
Congratulations!
Hi humans!
How many lines do we have? 8
Idea! just count last lines and that's it
I'm planning to have, at least 20
 will trv
line 13
ine 14
Line 15
Line 16
ine 17
No. sure I haven't
Second last: 19
So...
This is just a text example
What I'm going to write?
With a lot of different lines
fco@fco-VirtualBox:~$
```

Examples (5)

tr

```
fco@fco-VirtualBox:~$ cat example2.txt
cell1-1 cell1-2
cell2-1 cell2-2
cell3-1 cell3-2
cell1-1 cell1-2
cell5-1 cell5-2
cell4-1 cell4-2
fco@fco-VirtualBox:~$ tr '\t' ',' < example2.txt
cell1-1,cell1-2
cell2-1,cell2-2
cell3-1,cell3-2
cell1-1,cell1-2
cell5-1,cell5-2
cell4-1,cell4-2
fco@fco-VirtualBox:~$</pre>
```

Exercise (3)

Copy exercise.tsv in course folder

Print last 10 lines of the file

Print 6th column of the file

Convert all the ':' to '-' in the file

Exercise (3)

Copy exercise.tsv in course folder ffuster2@figaro:~/course\$ cp /media/ffuster2/USB\ DISK/exercise.tsv

Print last 10 lines of the file

```
ffuster2@figaro:~/course$ tail -n 10 exercise.tsv
chr3 128199429 T 649 568 T:555:2
chr3 128199662 G 797 786 G:422:2
chr3 128204951 C 562 540 C:261:2
chr3 128205860 G 302 297 G:0:0:0
chr3 128206618 C 251 240 C:123:2
chr3 128206710 T 173 167 T:0:0:0
chr3 136056728 G 723 719 G:697:2
chr3 136056861 T 775 771 T:1:1:6
chr3 136191272 C 447 434 C:430:2
chr3 168801495 C 743 734 C:1:1:5
```

Print 6th column of the file

```
ffuster2@figaro:~/course$ cut -f 6 exercise.tsv
base:reads:strands:avg_qual:map_qual:plus_reads:minus_reads
T:451:2:42:1:308:143:183
T:635:2:44:1:409:226:0
G:650:2:45:1:472:178:13
A:307:2:37:1:154:153:0
A:0:0:0:0:0:0:0
T:567:2:36:1:518:49:0
A:545:2:32:1:509:36:2
```

Convert all the ':' to '-' in the file

```
ffuster2@figaro:~/course$ tr ':' '-' < exercise.tsv
chrom position ref_base depth q20_depth base-reads-
chr1 35649754 T 647 634 T-451-2-42-1-308-143-183
-3-AAA-1-1-13-1-1-0 INS-1-A-54-2-41-1-32-22 INS-2-AA-7-2-32-1-5-2
chr1 35649768 T 669 645 T-635-2-44-1-409-226-0 T
chr1 35649862 G 692 663 G-650-2-45-1-472-178-13 GA
chr1 35658369 A 347 308 A-307-2-37-1-154-153-0 A
```

Special characters

"

All characters

- '>' [file] → Store the output in [file]
- '>>' [file] → Append output to [file]
- 2> [file] → Store error output to [file]
- 2>> [file] → Append output error to [file]
- '<' [file] → Use [file] as input
- ';' → Separator between commands
- '|' \rightarrow Send the output to other program (pipe)

```
fco@fco-VirtualBox:~$ cat example2.txt
cell1-1 cell1-2
cell2-1 cell2-2
cell3-1 cell3-2
cell1-1 cell1-2
cell5-1 cell5-2
cell4-1 cell4-2
fco@fco-VirtualBox:~$ tr '\t' ',' < example2.txt
cell1-1,cell1-2
cell2-1,cell2-2
cell3-1,cell3-2
cell1-1,cell1-2
cell5-1,cell5-2
cell4-1,cell4-2
fco@fco-VirtualBox:~$ tr '\t' ',' < example2.txt > example3.txt
fco@fco-VirtualBox:~$ cat example3.txt
cell1-1,cell1-2
cell2-1,cell2-2
cell3-1,cell3-2
cell1-1,cell1-2
cell5-1,cell5-2
cell4-1,cell4-2
fco@fco-VirtualBox:~$
```

```
fco@fco-VirtualBox:~$ cat example2.txt
cell1-1 cell1-2
cell2-1 cell2-2
cell3-1 cell3-2
cell1-1 cell1-2
cell5-1 cell5-2
cell4-1 cell4-2
fco@fco-VirtualBox:~$ cat example2.txt >> example3.txt
fco@fco-VirtualBox:~$ cat example3.txt
cell1-1,cell1-2
cell2-1,cell2-2
cell3-1,cell3-2
cell1-1,cell1-2
cell5-1,cell5-2
cell4-1,cell4-2
cell1-1 cell1-2
cell2-1 cell2-2
cell3-1 cell3-2
cell1-1 cell1-2
cell5-1 cell5-2
cell4-1 cell4-2
fco@fco-VirtualBox:~$
```

```
fco@fco-VirtualBox:~/course$ grep "!" example.txt
Hi humans!
Idea! just count last lines and that's it
And line 20!! Got it!
Congratulations!
fco@fco-VirtualBox:~/course$ grep "!" example.txt | tr "!" "?"
Hi humans?
Idea? just count last lines and that's it
And line 20?? Got it?
Congratulations?
fco@fco-VirtualBox:~/course$ grep "!" example.txt | tr "!" "?" | sort
And line 20?? Got it?
Congratulations?
Hi humans?
Idea? just count last lines and that's it
fco@fco-VirtualBox:~/course$ grep "!" example.txt | tr "!" "?" | sort | tr "?" "."
And line 20., Got it.
Congratulations.
Hi humans.
Idea. just count last lines and that's it
fco@fco-VirtualBox:~/course$
```

Exercise (4)

Print 1st column of exercise.tsv and store the output in a new file called exercise2.tsv

Add to exercise2.tsv, the first 5 lines of exercise.tsv

Find ':' in exercise2.tsv. Convert '.' to ','. Store the output in new file called exercise3.tsv

Exercise (4)

Print 1st column of exercise.tsv and store the output in a new file called exercise2.tsv

[CONFIGN | CONFIGN | CONFI

Add to exercise 2.tsv, the first 5 Lines of exercise tsv > exercise tsv > exercise tsv

```
fco@fco-VirtualBox:~/course$ head -n 5 exercise.tsv >> exercise2.tsv
fco@fco-VirtualBox:~/course$ tail -n 10 exercise2.tsv
chr3
chr3
chr3
chr3
chr3
chrom position ref_base depth q20_depth base:reads:strands:avg_qual:map_qual:plus_read
chr1 35649754 T 647 634 T:451:2:42:1:308:143:183 TA 0 0 0
-3-AAA:1:1:13:1:10 INS-1-A:54:2:41:1:32:22 INS-2-AA:7:2:32:1:5:2
chr1 35649768 T 669 645 T:635:2:44:1:409:226:0 T 635 2 44 1
chr1 3564962 G 692 663 G:650:2:45:1:472:178:13 GA 0 0 0 0
chr1 35658369 A 347 308 A:307:2:37:1:154:153:0 A 307 2 37 1
fco@fco-VirtualBox:~/course$
```

Find ':' in exercise2.tsv. Convert ':' to ','. Store the output in new file called

```
fco@fco-VirtualBox:-/course$ grep ': 'exercise2.tsv | tr ':' ',' > exercise3.tsv | fco@fco-VirtualBox:-/course$ cat exercise3.tsv | fco@fco-VirtualBox:-/course$ | fco@fco-Virtu
```

Linux commands (4). Other useful commands

- man [command] → Get the manual of [command]
- wc [file] → Print newline, word, and byte counts of [file]
- history → Print all the commands used
- **clear** → "Cleans" the terminal
- echo ["Some words"] → Prints ["some words"] in the terminal

Exercise (5)

What is the parameter in 'wc' command to print only the number of lines that file has?

Print all the 'cd' command you have used during the course

Exercise (5)

What is the parameter in 'wc' command to print only the number of lines that file has?

```
ffuster2@figaro:~/course$ man wc
ffuster2@figaro:~/course$
ffuster2@figaro:~/course$ wc -l exercise.tsv
40 exercise.tsv
```

Print all the 'cd' command you have used during the course

```
ffuster2@figaro:~/course$ history | grep cd
```

Merging all

We can combine different instructions if needed

```
#!/bin/bash

#Do all exercises at once
cd /home/fco #Go to home folder
mkdir course #Create new folder
cd course #Go to folder
touch ex2.sh #Create new blank file
cp /media/ffuster/USB/exercise.tsv . #Copy example exercise from the USB.
tr ':' '-' < exercise.tsv > exercise2.tsv #Change ':' to '-' in the file and store the ouput in a new file
sort exercise2.tsv > exercise3.tsv #Sort the file and store it in a new file
```

Merging all

We can combine different instructions if needed

```
#!/bin/bash

#Do all exercises at once
cd /home/fco #Go to home folder
mkdir course #Create new folder
cd course #Go to folder
touch ex2.sh #Create new blank file
cp /media/ffuster/USB/exercise.tsv . #Copy example exercise from the USB.
tr ':' '-' < exercise.tsv > exercise2.tsv #Change ':' to '-' in the file and store the ouput in a new file
sort exercise2.tsv > exercise3.tsv #Sort the file and store it in a new file
```

My first bash script!!

Bash scripts. Variables

Pieces of code that can get whatever value

```
fco@fco-VirtualBox:~/course$ cat script1.sh
#! /bin/bash

variable="My first variable"

echo $variable
fco@fco-VirtualBox:~/course$ ./script1.sh
My first variable
fco@fco-VirtualBox:~/course$
```

Special variables:

- \$1 \$9 → Get the value of first ninth option
- ...

Bash Scripting. If statement

Piece of code that will be executed if a condition is true

```
#!/bin/bash

#My first if statement

if [ condition ]

then

[Commands that will be executed if condition is true]

else

[Commands that will be executed if the condition is false]

fi
```

Bash Scripting. If statement

Piece of code that will be executed if a condition is true

```
#!/bin/bash

#My first if statement

if [ condition ]

then

[Commands that will be executed if condition is true]

else

[Commands that will be executed if the condition is false]

fi
```

Bash scripting. For loop

Piece of code that will be executed several times

```
#!/bin/bash

#My first for loop

for ()

do

[Commands that will be executed several times]

done
```

Bash scripting. For loop

Piece of code that will be executed several times

```
#!/bin/bash

#My first for loop

for ()

do

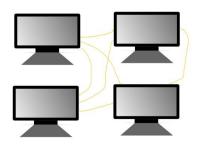
[Commands that will be executed several times]

done
```

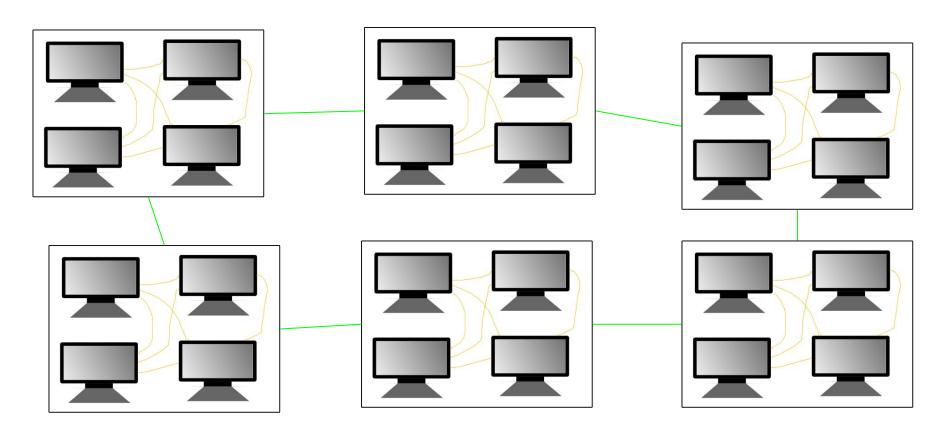
Cluster



Cluster



Cluster



Queuing system

Orders the jobs sent to cluster according to priorities

- SGE
- SLURM
- ...

SGE useful commands

qsub [bash_script] → Sends [bash_script] to cluster

qstat →Check cluster status

qdel → Delete a job from cluster

Exercise (6)

Copy cluster_template.sh to your course folder

Modify cluster_template.sh, with your name

Send the job to cluster

Check the cluster status

Exercise (6)

Copy cluster_template.sh to your course folder

Modify cluster_template.sh, with your name

Send the job to cluster

Check the cluster status

Additional exercises (1)

Print last 10 lines of exercise.tsv, but excluding last one

Print 2nd column of exercise.tsv, sort the output and store it in a new file called exercise2.tsv

Find the files in course folder that starts with 'ex' and has '2' in the name

Print 6th column of exercise.tsv, change ':' to ',' and sort the output. Append the output to exercise2.tsv

Additional exercises (1)

Print last 10 lines of exercise.tsv, but excluding last one

```
ffuster2@figaro:~/course$ tail -n 10 exercise.tsv
                                   649
                                            568
chr3
                                   797
                                            786
chr3
                                   562
                                            540
chr3
                                           297
chr3
                                   251
                                           240
chr3
                                   173
                                           167
chr3
                                           771
```

Print 2nd column of exercise.tsv, sort the output and store it in a new file called exercise2.tsv

ffuster2@figaro:~/course\$ cut -f 2 exercise.tsv | sort > exercise2.tsv

Additional exercises (1)

Find the files in course folder that starts with 'ex' and has '2' in the name

ffuster2@figaro:~/course\$ find . -name ex*2*

Print 6th column of exercise.tsv, change ':' to ',' and sort the output. Append the output to exercise2.tsv

ffuster2@figaro:~/course\$ cut -f 6 exercise.tsv | tr ':' ',' >> exercise2.tsv