

Data Science (ciencia de los datos):

Aplicaciones a la biología y a la medicina con Python y R

**Curso de Experto Universitario
Universidad de Barcelona (UB)**

<http://www.ub.edu>

Con el soporte de:



GNU/Linux OS

Introduction

Administration

Xavier de Pedro

Thursday, May 3rd. 2018
Course «*Data Science*» UB



Image source: <http://tekopsglobal.com>

Xavier de Pedro Puente, Ph.D.

xdepedro@bcn.cat

Academics:

- Degree in Biology
(University of Barcelona - UB)
- Ph.D. in Ecology
(University of Barcelona - UB)
- Postgraduate in Bioinformatics
(Open University of Catalonia – UOC)

Current Work:

- Technician at Oficina Municipal de Dades (Barcelona City Council)

Past (related) Work:

- Bioinformatics technician (UEB, VHIR)
- Systems administrator (UEB, VHIR)



Session Outline

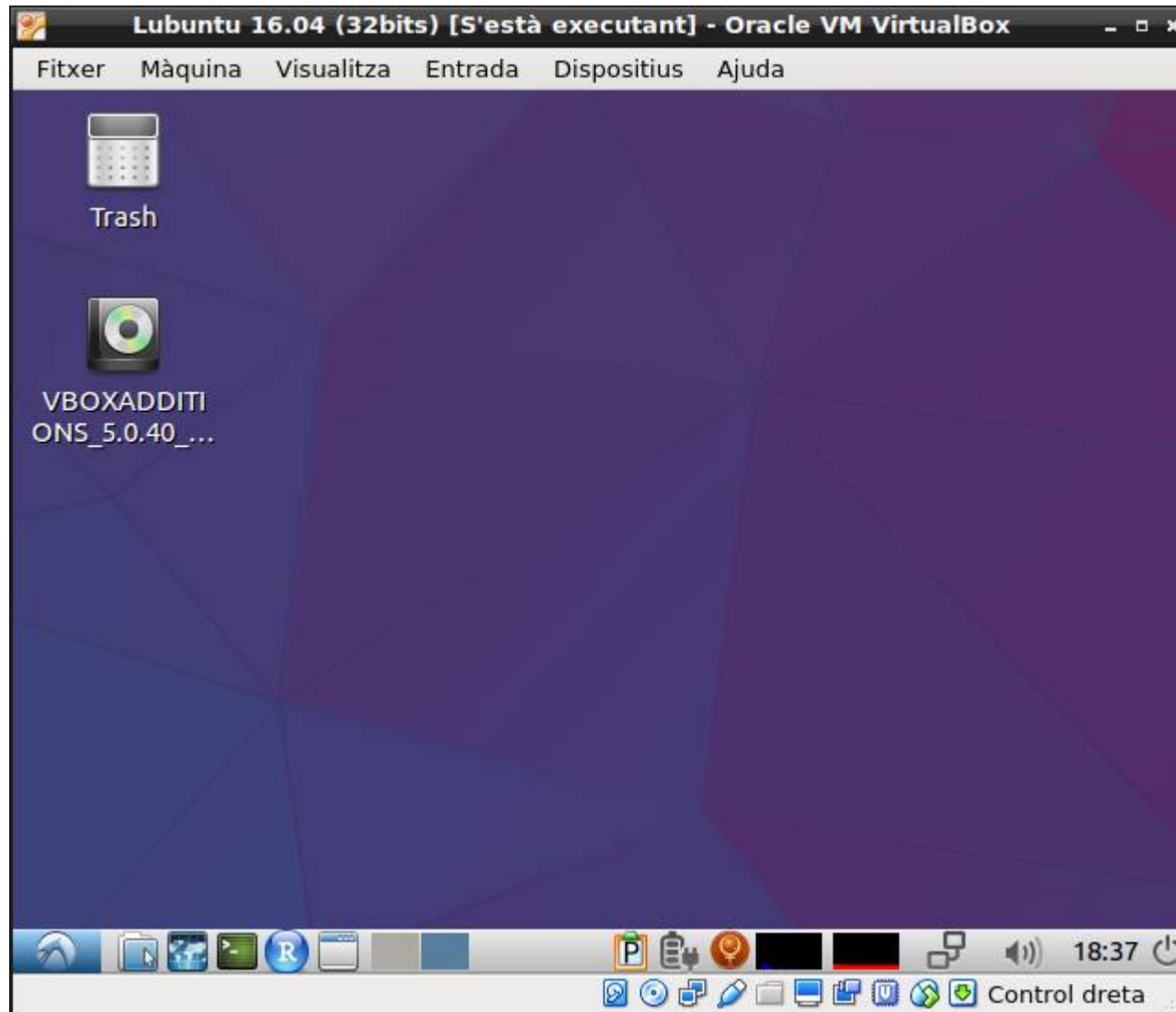
(1) GNU/Linux Introduction

- Concepts, History, "Distributions", Ubuntu
- Basic Differences respect to:
 - Windows
 - MacOSX
- Command line runs

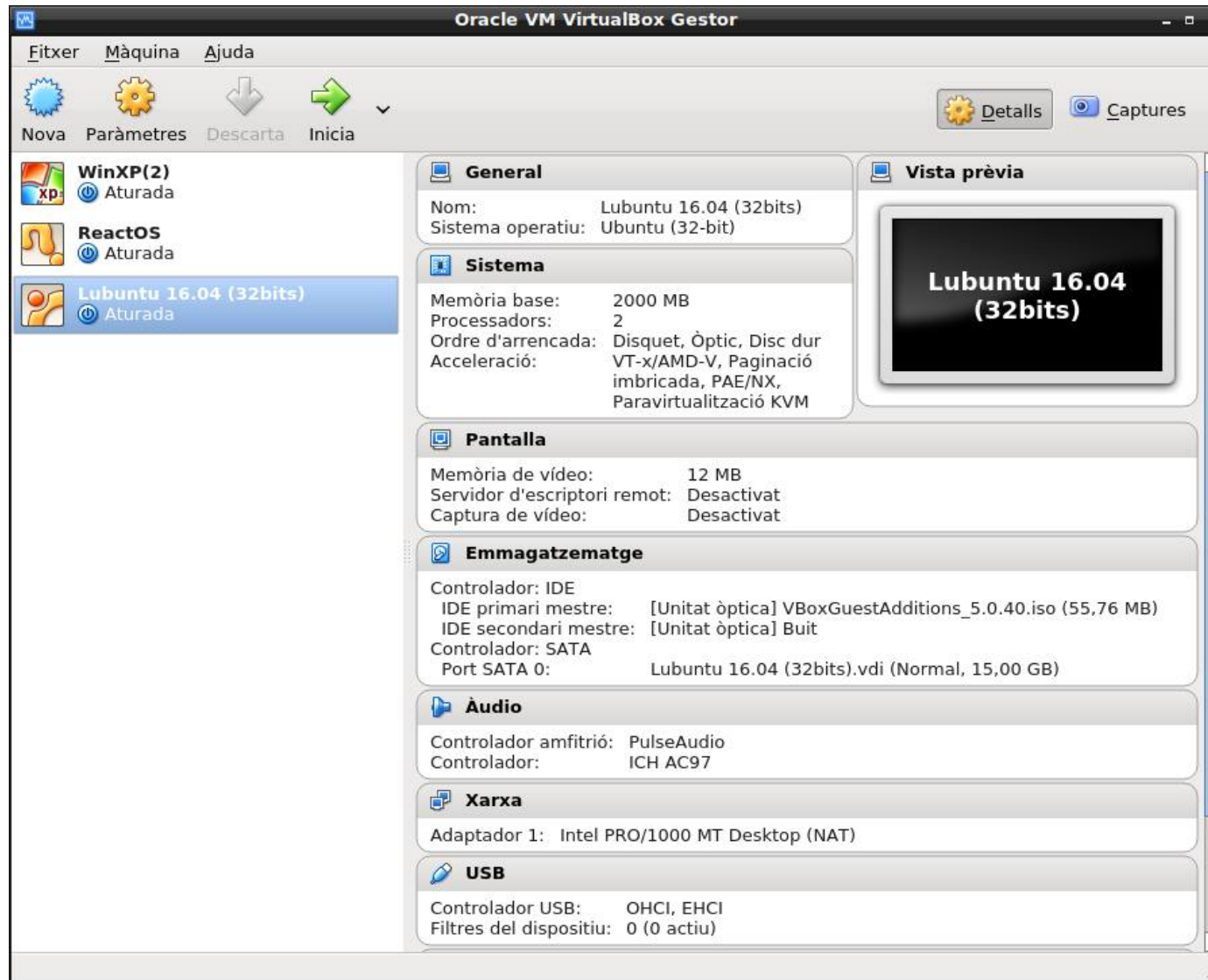
(2) GNU/Linux Administration

- Installation
- Package Management
- User Management
- Permission Management
- Device Management
- Backup Management
- Security
- Computer Client Management

Tool: VirtualBox VM with Lubuntu GNU/Linux

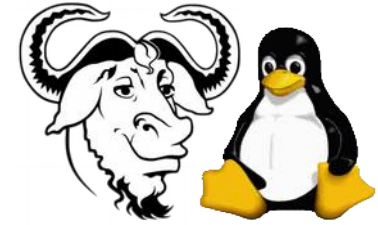


Tool: VirtualBox VM with Lubuntu GNU/Linux



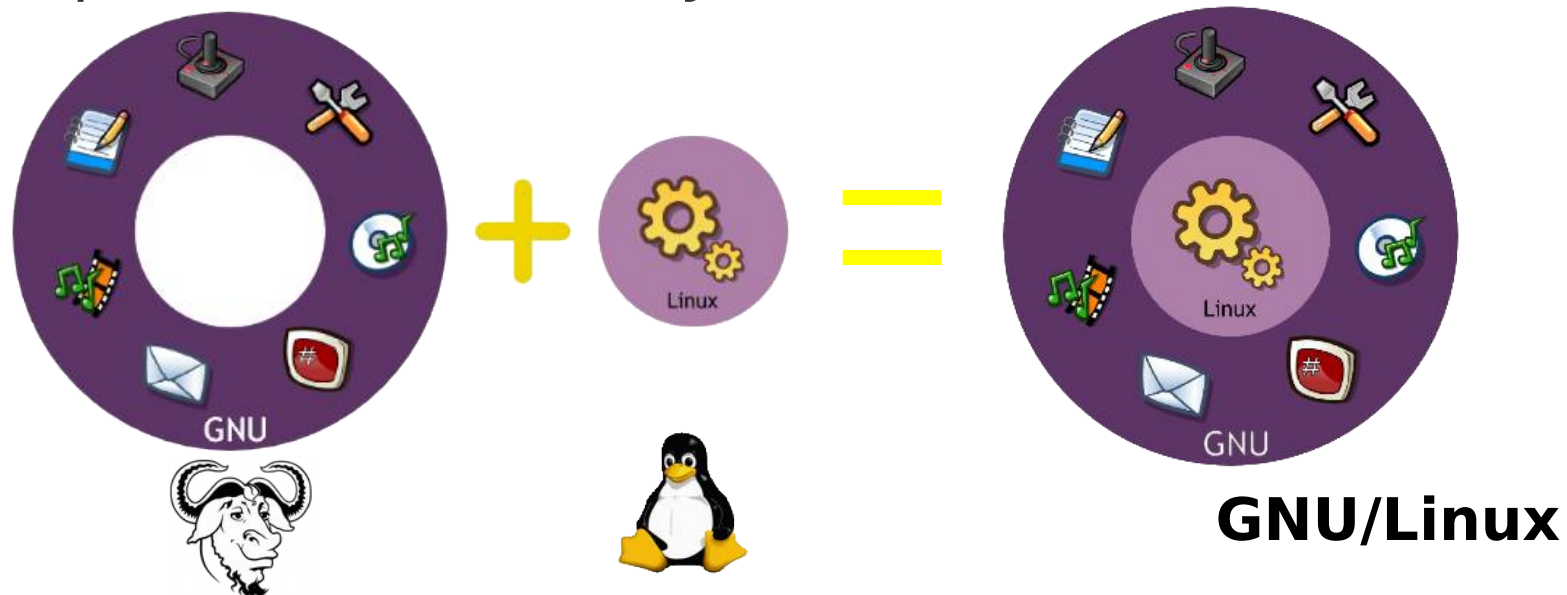
What is GNU/Linux?

(a.k.a «Linux»)



⚙️ It is ...

- An Open Source (Free/Libre) operating system
- Sum of GNU surroundings and the Linux kernel/core
- Compatible with UNIX systems



Operating System *(in context)*

End User Applications



Desktop Environments

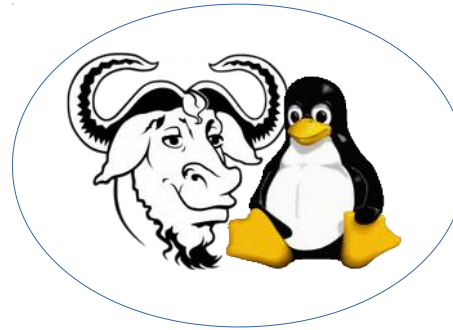


Aqua



Luna (XP)
Aero (Vista +)

Operating System



Hardware



GNU/Linux

Some basic concepts

- GNU Project
- Hardware vs software
- Open Source & Free software (as in «free» seats)
 - **FLOSS**: Free/Libre Open Source Software
- Free Software Foundation (FSF)
- Free Operating Systems
 - GNU/Linux, GNU/Hurd, [Open,Free,Net]-BSD, OpenSolaris, ReactOS...
- Open Source Hardware

Historical Evolution

In the origins of software....

- Software was born free.
- In the decade of the '60, when buying **HW**, access was granted to the manufacturer's **SW** catalogue
- All software distributed together with source code
- At the end of the '70, **IBM** announced their intention to sell parts of their **SW** separately.
- From then on, proprietary **SW** (no free) became common

Historical Evolution

⚙ GNU/Linux History

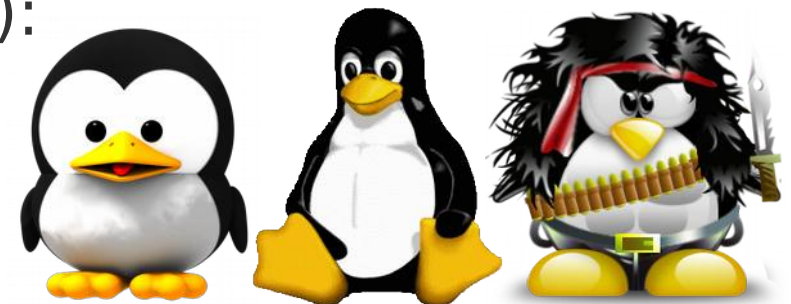
- 1983. Richard M. Stallman (RMS): **GNU** project
- 1984. Free Software Foundation (FSF).
- First components of the GNU system, all written by RMS:
 - a C compiler (gcc)
 - an text editor (emacs)
 - and a debugger (gdb)



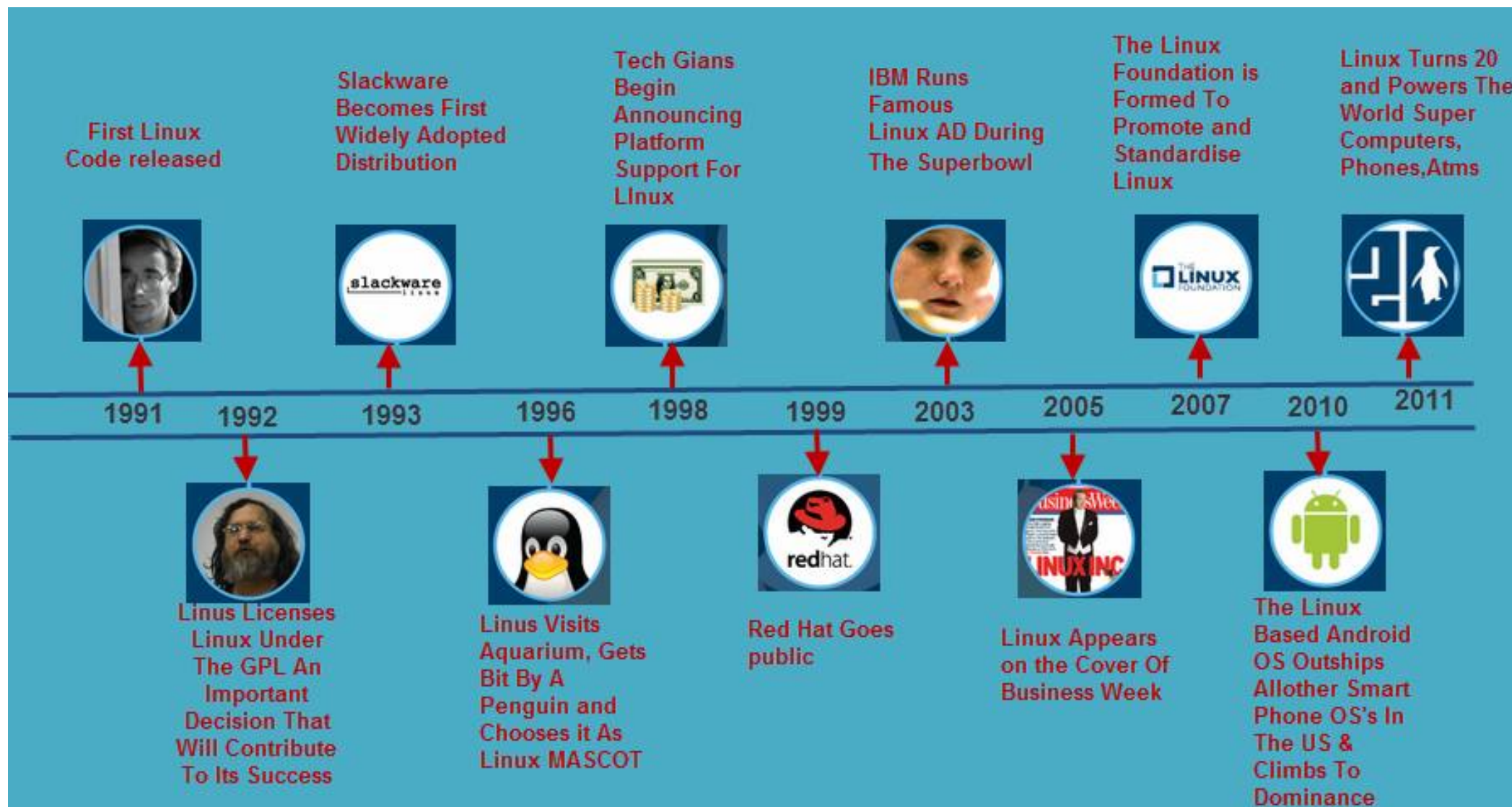
Historical Evolution

⚙ GNU/Linux History

- To guarantee the four freedoms, RMS invented the concept of copy-left (reverse of copy-right).
- 1990: GNU system was almost complete, only missing to finish an ambitious kernel (core).
- 1991: Linus Torvalds wrote a monolithic kernel
- GNU + Linux kernel (by Linus T.): GNU/Linux
- 1996: A penguin bit to Linus. Logo: penguin Tux



GNU/Linux History



Desktop Environments *(in context)*

Applications

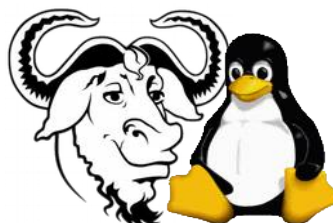


Desktop Environments



Aqua  Luna (XP)
Aero (Vista +)

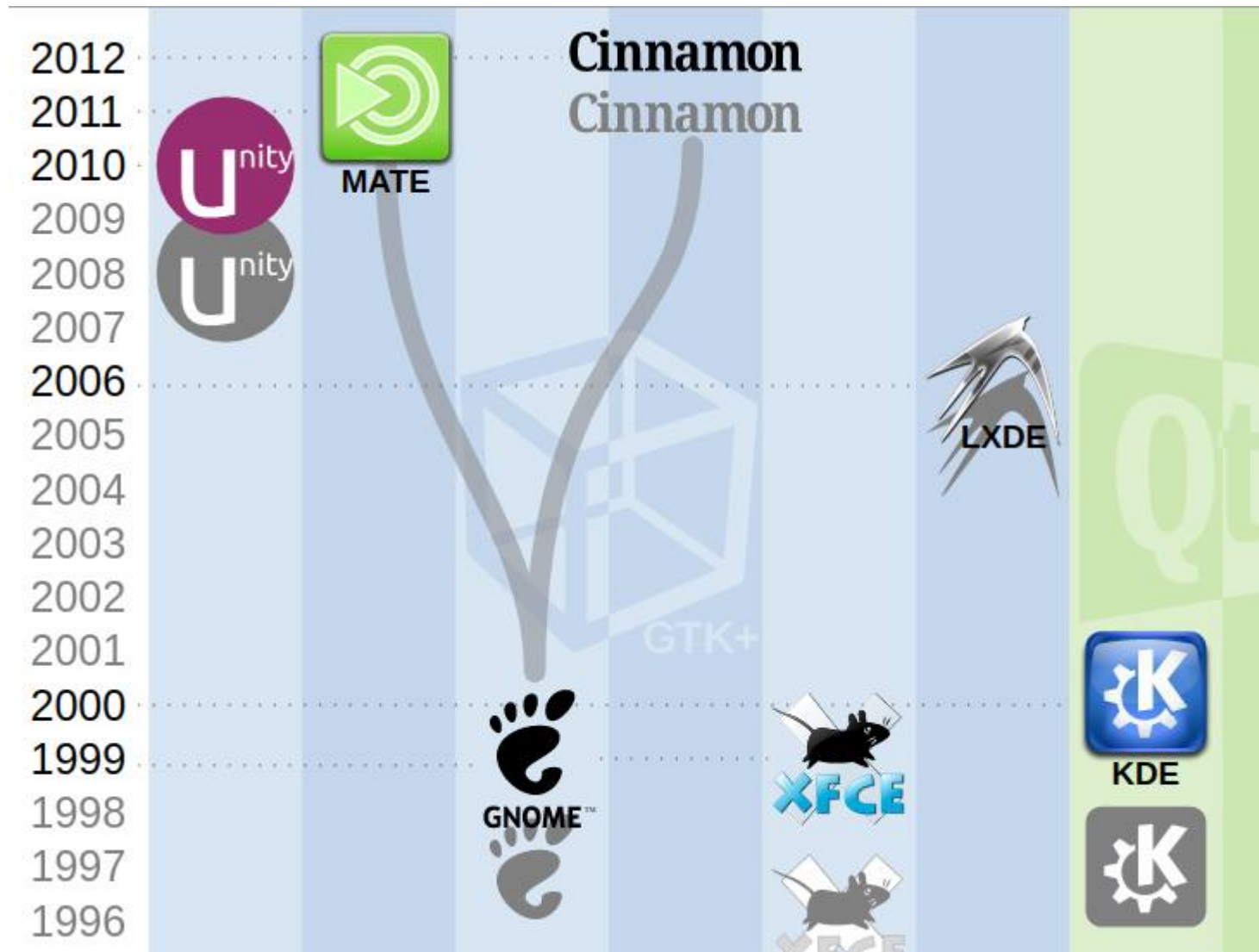
Operating System



Hardware



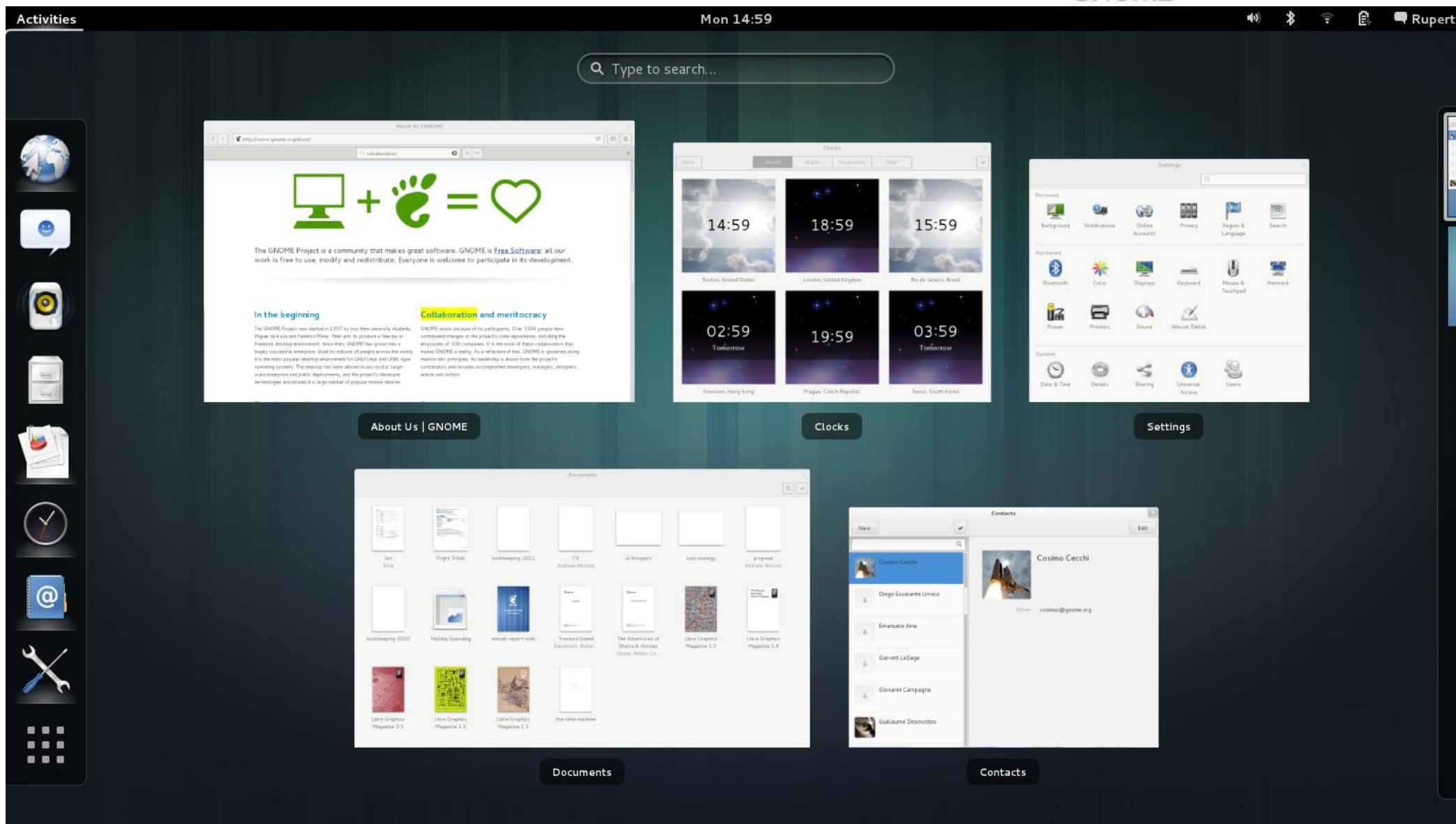
History of Desktop Environments



Desktop Environments (DE)



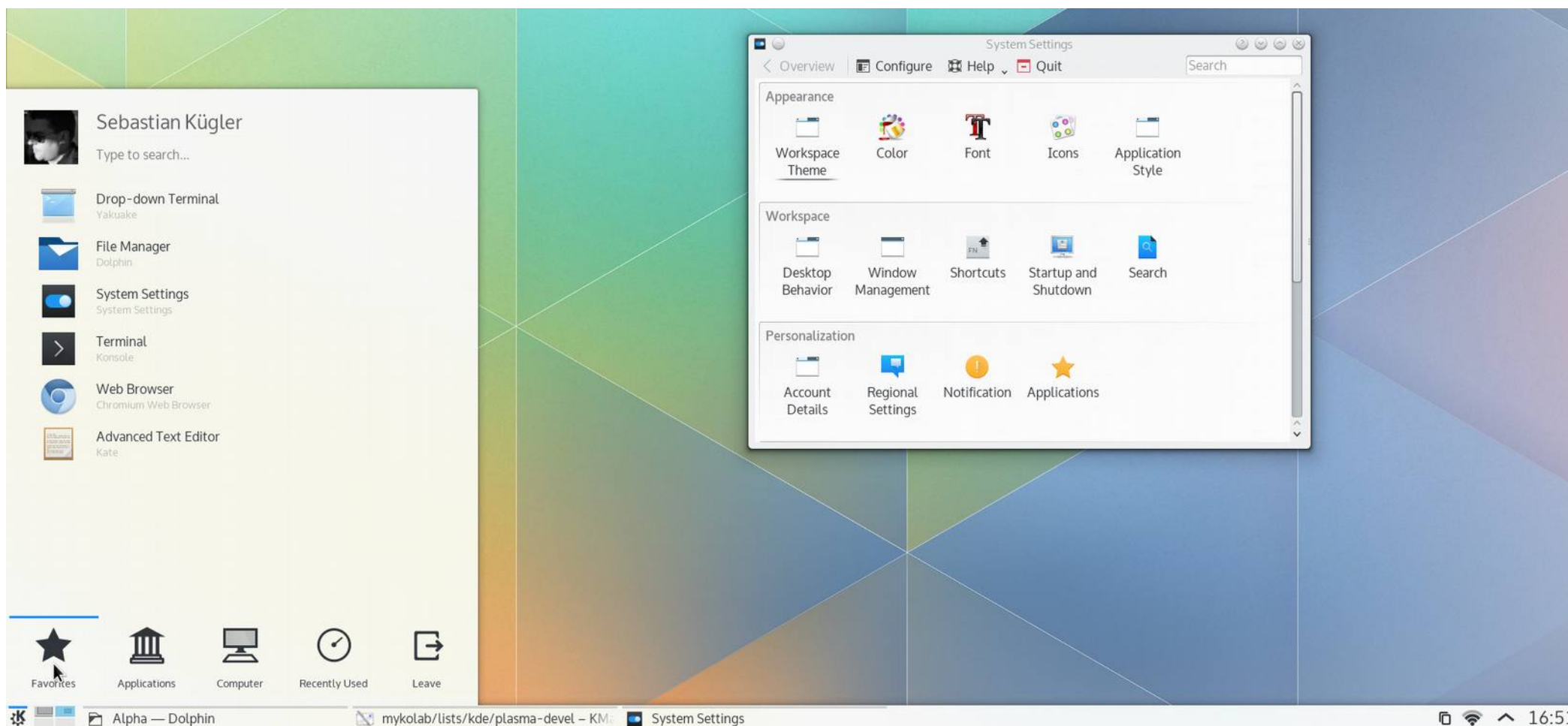
GNU/Linux DE:



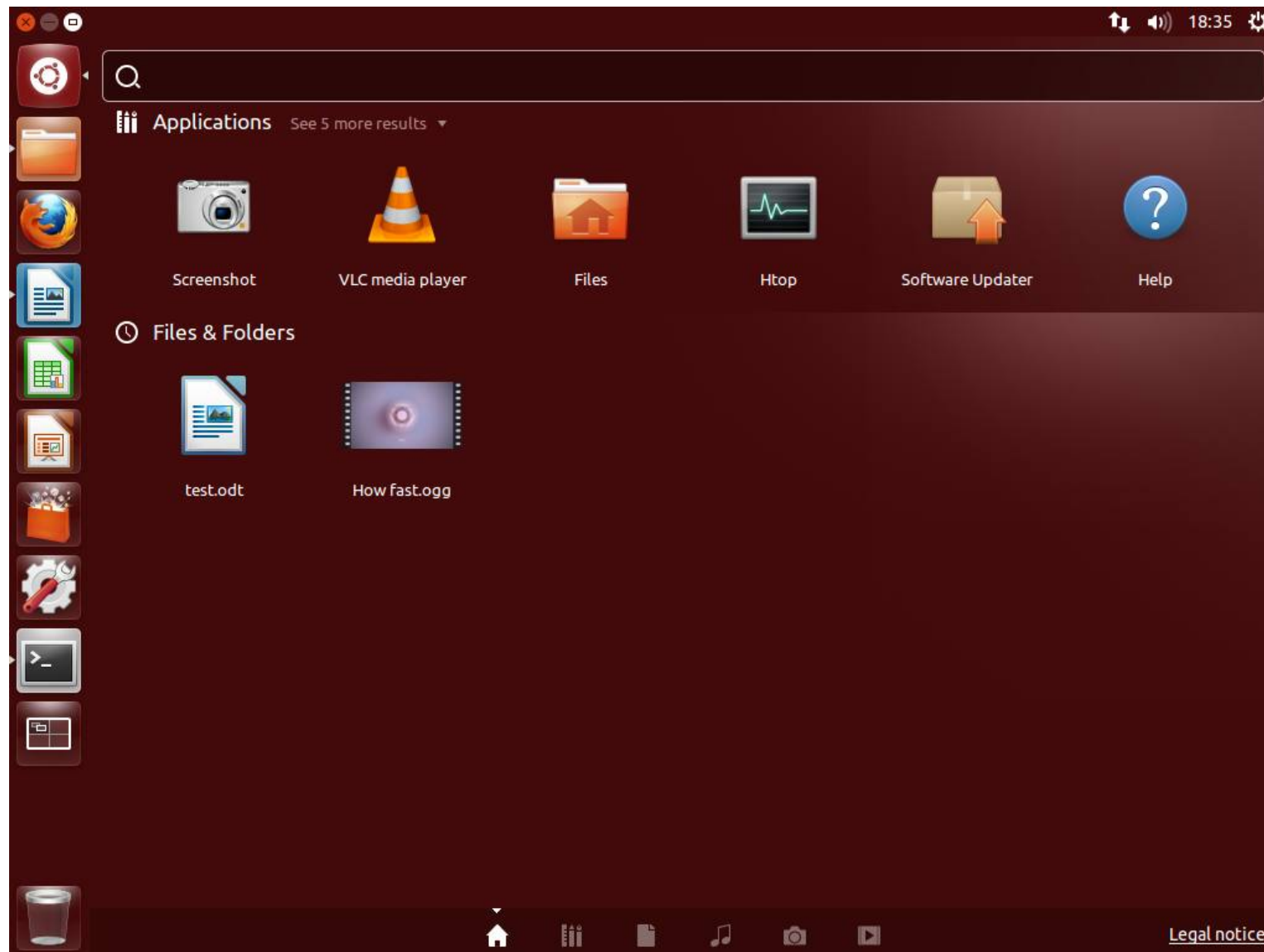
GNU/Linux DE:



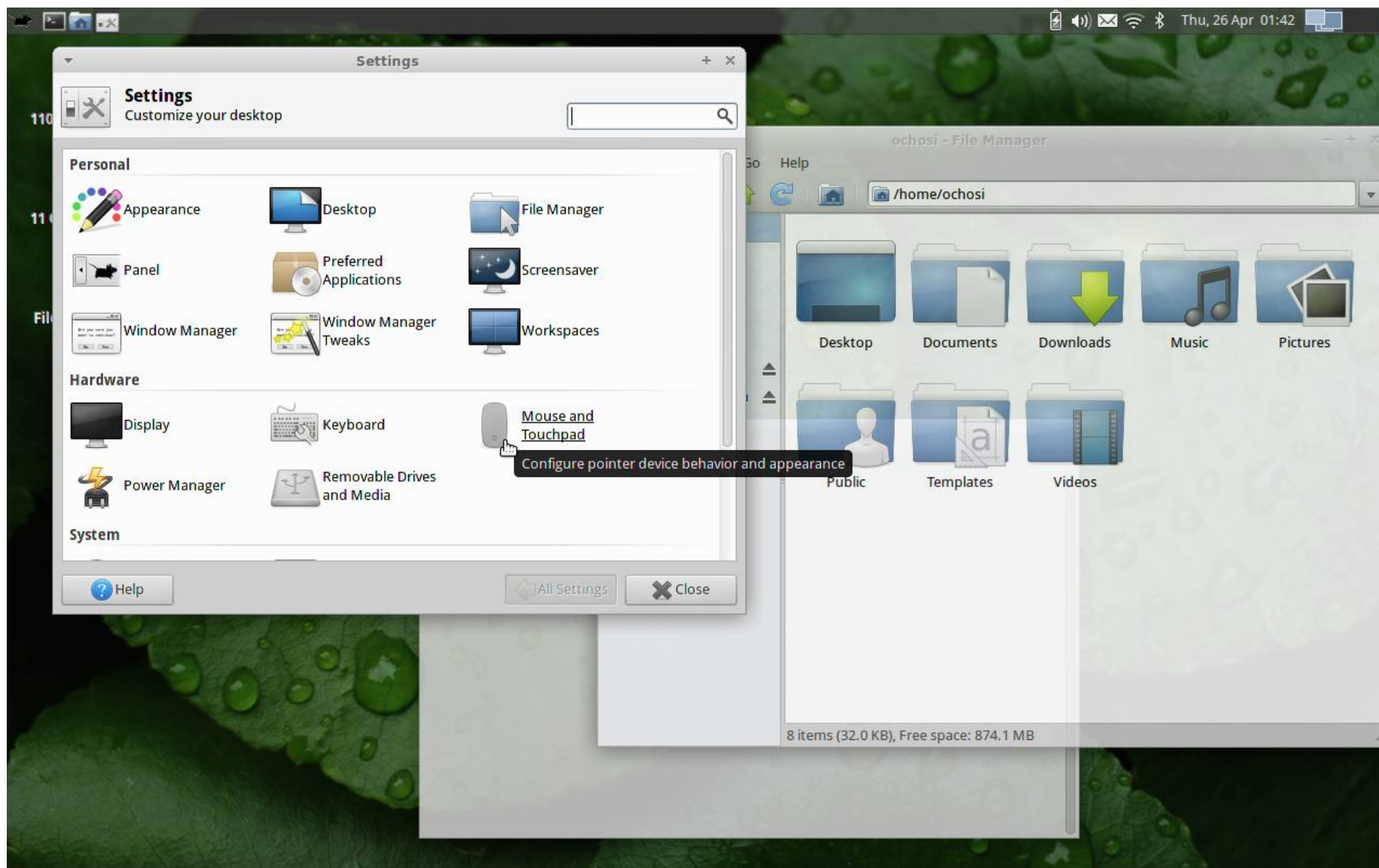
KDE



GNU/Linux DE:



GNU/Linux DE:



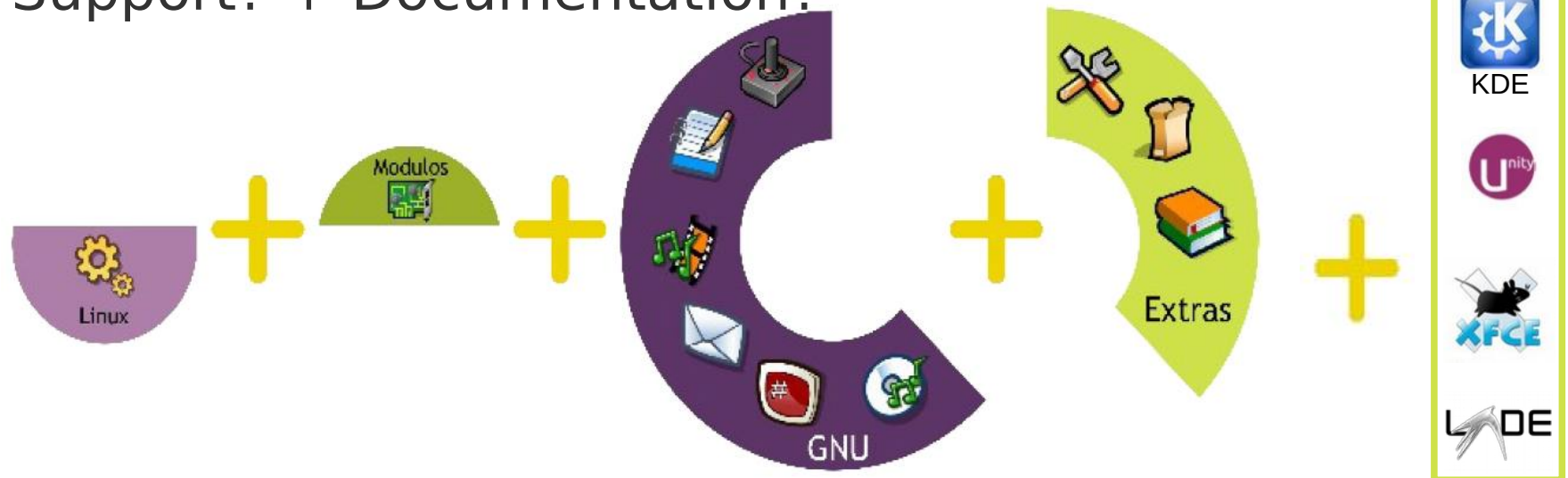
GNU/Linux DE: LXDE



What is a GNU/Linux «distribution»?

⚙️ A collection of free software

- Core + drivers (modules)
- Desktop Environment
- Extra + programs + utilities
- Support? + Documentation?



Distributions of GNU/Linux



Debian



Ubuntu



Slackware



Gentoo



Fedora



Redhat



Mandriva



Suse

Ubuntu (GNU/Linux)



⚙ Main characteristic

- Distribution based in Debian
- Developed by Canonical Ltd. (South Africa)
- Ubuntu:
 - Philosophy zulú: *"Mankind to others", "I am because we are"*
 - Slogan Ubuntu: *"Linux for human beings"* (or "beans"?)
- Definitely, it is the easiest to install/use



Ubuntu (GNU/Linux)

Variants

	1st Release
Ubuntu	4.10
Kubuntu	5.04
Edubuntu	5.10
Xubuntu	6.06 LTS
Mythbuntu	7.10
UbuntuStudio	7.10
Lubuntu	11.10
Ubuntu GNOME	13.04
Ubuntu Kylin	13.04
Ubuntu MATE	15.04
Ubuntu Budgie	17.04



Source: <https://itsfoss.com/which-ubuntu-install/>



Kubuntu

Kubuntu offers the KDE Plasma Workspace experience, a good-looking system for home and office use.



Lubuntu

Lubuntu is a fast, energy saving and lightweight variant of Ubuntu using LXDE. It is popular with PC and laptop users running on low-spec hardware.



Ubuntu Budgie

Ubuntu Budgie provides the Budgie desktop environment which focuses on simplicity and elegance. It provides a traditional desktop metaphor based interface utilising a customisable panel based menu driven system.



Ubuntu Kylin

The Ubuntu Kylin project is tuned to the needs of Chinese users, providing a thoughtful and elegant Chinese experience out-of-the-box.



Ubuntu MATE

Ubuntu MATE expresses the simplicity of a classic desktop environment. Ubuntu MATE is the continuation of the GNOME 2 desktop which was Ubuntu's default desktop until October 2010.



Ubuntu Studio

Ubuntu Studio is a multimedia content creation flavor of Ubuntu, aimed at the audio, video and graphic enthusiast or professional.



Xubuntu

Xubuntu is an elegant and easy to use operating system. Xubuntu comes with Xfce, which is a stable, light and configurable desktop environment.

GNU/Linux vs. Windows

- **Technical** differences
 - User: **root** vs. administrator
 - Case **sensitive** vs. Case insensitive
 - "MyFile != myfile" vs. "MyFile = myfile"
 - "**Symbolic or Hard Links**" vs. "Shortcuts"
 - Paths: **Slash ("/")** vs. Backslash ("\")
 - Main harddrive/partition: **"/"** vs. "C:\"
 - Partition formats (file systems): **ext2, ext3, ext4, ...** vs. fat16/fat32/ntfs
 - User default folder: **/home/username** vs. "C:\Documents and Settings\" or "C:\Users\username\MyDocuments"
 - USBdisk default folder: **/media/username/usbdiskname** vs. "X:\usbdiskname"
 - **Secure (Viruses???)** vs. Insecure (viruses, bots, worms, trojans, backdoors, ...)
 - **It can extend computer useful life** vs. Planned & perceived obsolescence
 - Performance, with same hardware: **Faster** vs. Slower (Antivirus, antispyware,...)
 - Run as admin: **"sudo program"** in console vs. "Run program as administrator"

GNU/Linux vs. Windows

- **Philosophical** differences
 - Freedom: "**Free/Libre Open Source Software (FLOSS)**" vs. "Closed Source/Proprietary software"
 - Software Sustainability: **High** vs. **Low**
 - Usually: **product given for free (at no Cost)** vs. **product for (excessive?) profit**
 - Money comes through: **customizations and training** vs. **Selling Software**
 - **FLOSS Fosters local economies as well as big companies** vs. **Big & remote Corporations growth.**

GNU/Linux vs. Mac OS X

- **Technical** differences

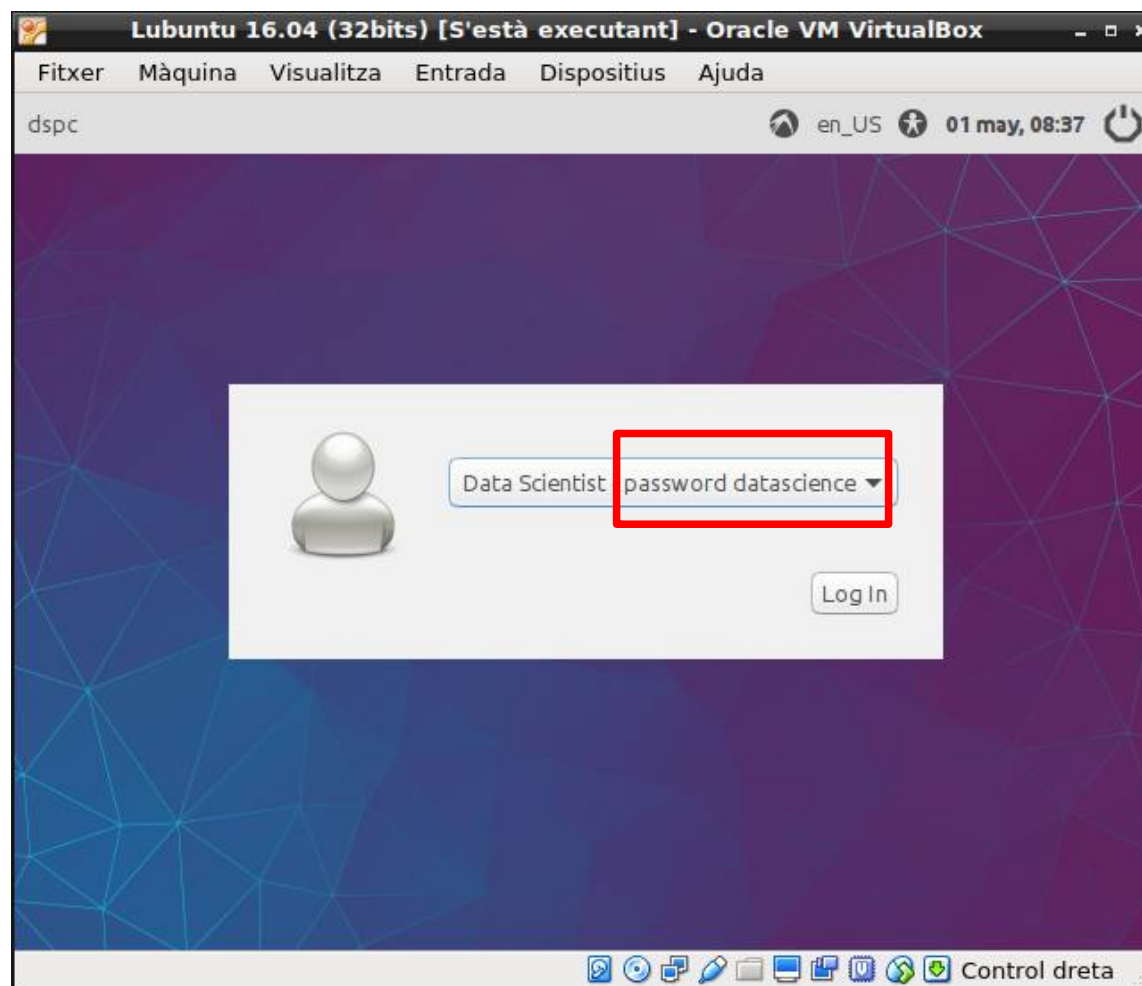
- Users:
- Case:
- Links:
- Paths:
- Main harddrive/partition
- Partition formats: **ext2, ext3, ext4, ...** vs. **HFS+**
- User default folder: **/home/username** vs. **"/Users/Username"**
- USBdisk default folder:
- **Secure (Viruses???)** vs. **Secure (Viruses?)**
- **It can extend computer useful life** vs. **Planned & perceived obsolescence**
- Performance, with same hardware: **Faster???** vs. **Extremely Fast**
- Run as admin:

GNU/Linux vs. Mac OS X

- **Philosophical** differences
 - Freedom: "**Free/Libre Open Source Software (FLOSS)**" vs. "Closed Source/Proprietary software"
 - Software Sustainability: **High** vs. **Low**
 - Usually: **product given for free (at no Cost)** vs. **product for (excessive?) profit**
 - Money comes through: **customizations and training** vs. **Selling Software & Hardware**
 - **FLOSS Fosters local economies as well as big companies** vs. **Big & remote Corporations growth.**

Ubuntu (GNU/Linux)

See it in action ...



Command line runs (i)

- Shell = Terminal = Console ("Black" text screen to run commands)s
 - Secure Shell = **ssh** = a safe way to connect to remote computers or servers (encrypted)
 - **ftp**: File Transfer Protocol
 - to transfer files between computers or servers
 - you can NOT run commands (other than listings)
 - **sftp**: like FTP but "secure" (best: using ssh libraries)
- Local terminal window
 - > **whoami**

Command line runs (ii)

- **Example of simple local commands:**

- > `ps -e`

- > `ls -l`

- > `df -h`

- > `top`

- > `tree . | head`

- > `du . -h | grep G`

- **Example of simple editors:**

- > `nano` (simple editor: press "Ctrl + X" to eXit)

Command line runs (iii)

- **Example of network-related commands:**
 - > `ping google.com`
 - > `ifconfig`
- **Example of system administration (sysadmin) commands:**
 - > `sudo apt update & sudo apt install tree`
 - > `sudo adduser foo`
 - > `sudo passwd foo`
 - > `sudo service apache restart`
 - > `sudo kill -9 java`
- **Example of simple editors:**
 - > `nano` (simple editor: press "Ctrl + X" to eXit)

Command line runs (iv)

- **Example of other shell-based programs:**

- > `htop` (type "q" to quit/exit)

- > `mc` (Midnight Commander).

- Click on F10 (with mouse or trackpad) to quit

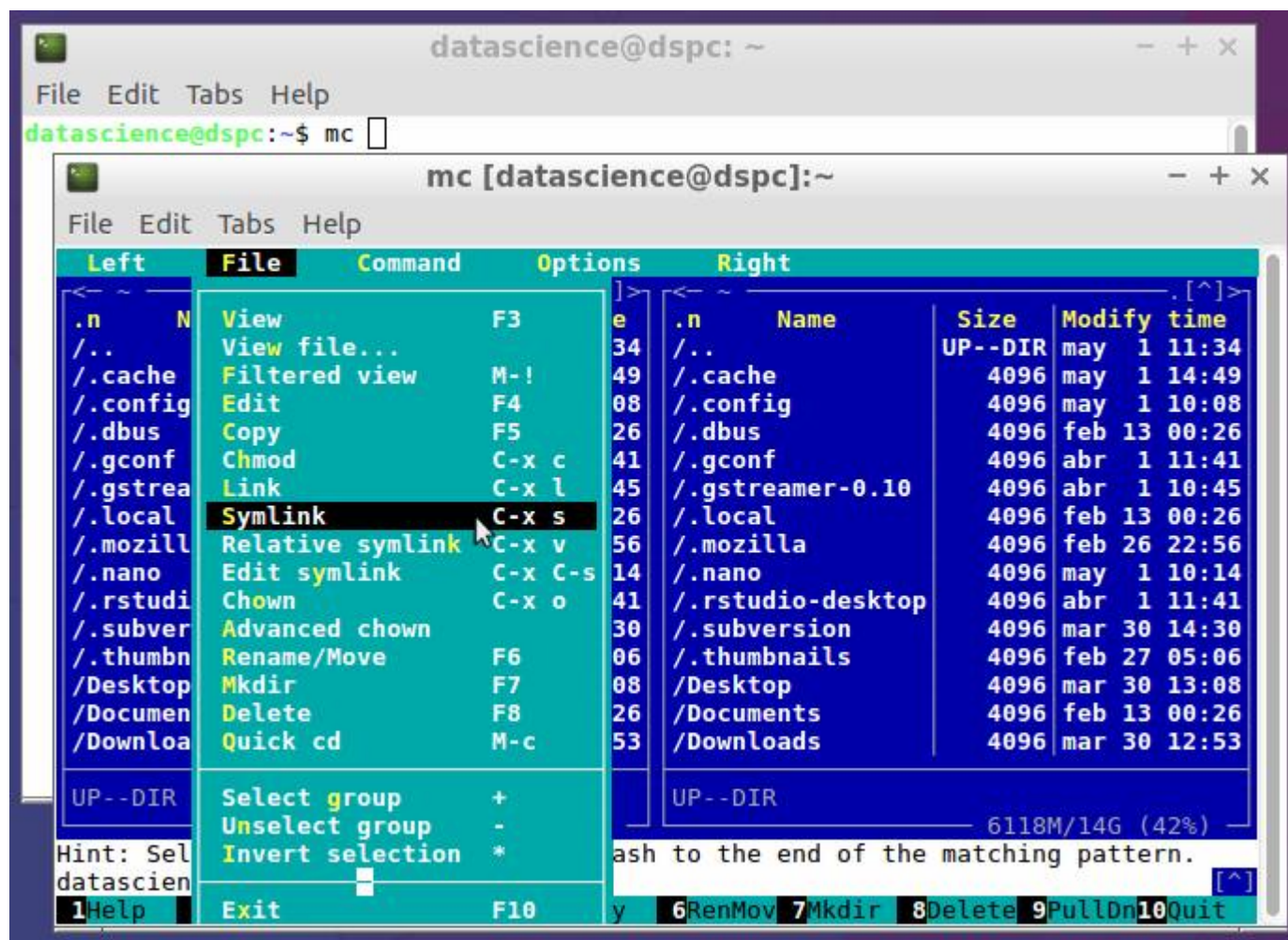
- > `R` (Type "q()" to quit)

- **Serial vs. Parallel tasks**

- <https://www.datascienceatthecommandline.com/chapter-8-parallel-pipelines.html>

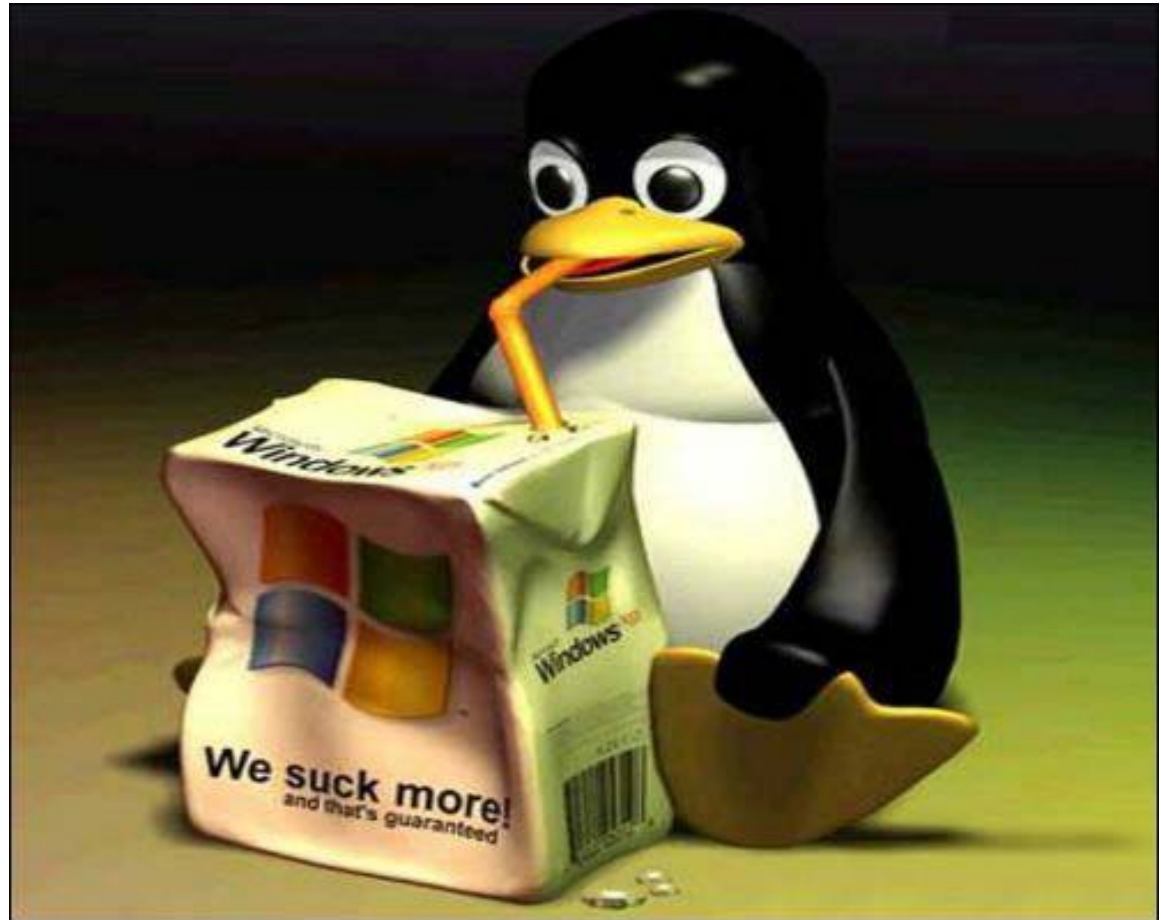
Midnight Commander (mc)

- **Powerful dual-pane file manager in terminals. A Life-saver for human beans when no X windows in servers**



Pause...

- **Coffee / soda time?**



Source: <http://wallpaperpicture--photo.blogspot.com.es/2015/01/funny-ads-pictures-0.html>

(2) GNU/Linux Administration

- ⚙ Installation
- ⚙ Package Management
- ⚙ User Management
- ⚙ Permission Management
- ⚙ Device Management
- ⚙ Backup Management
- ⚙ Security
- ⚙ Computer Client Management
- ⚙ Computer Clusters



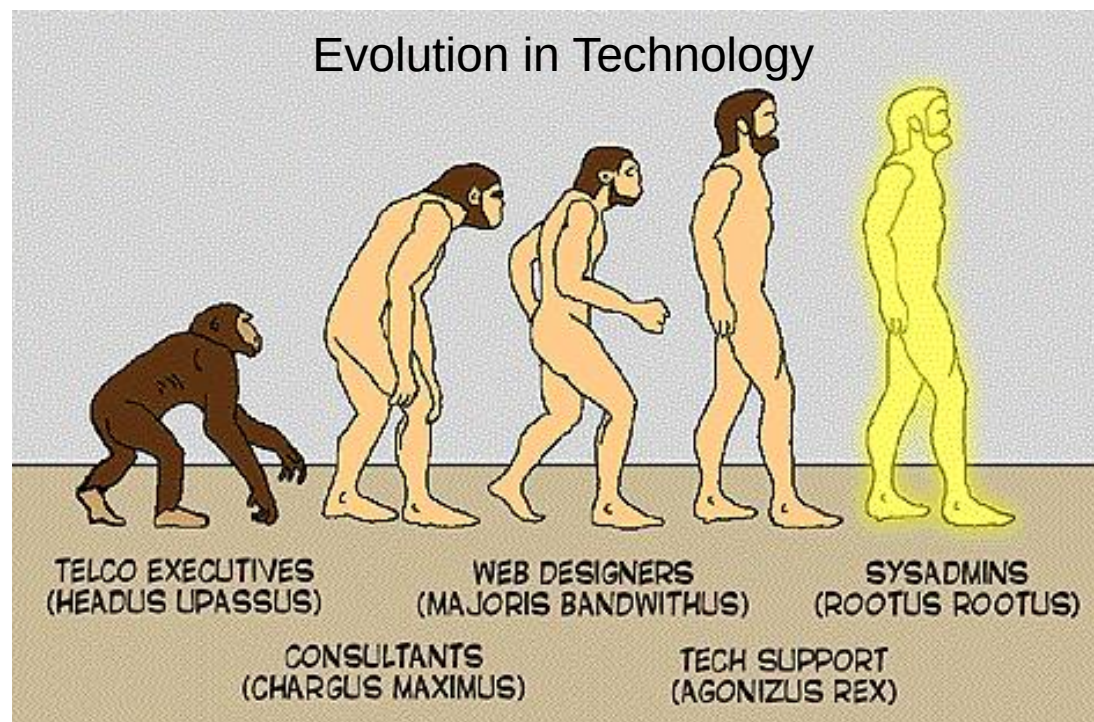


Image source: <http://ars.userfriendly.org/cartoons/?id=19990718>



Image Source: <https://www.linux.com/blog/sysadmin-ebook/2017/9/future-proof-your-sysadmin-career-advancing-open-source>

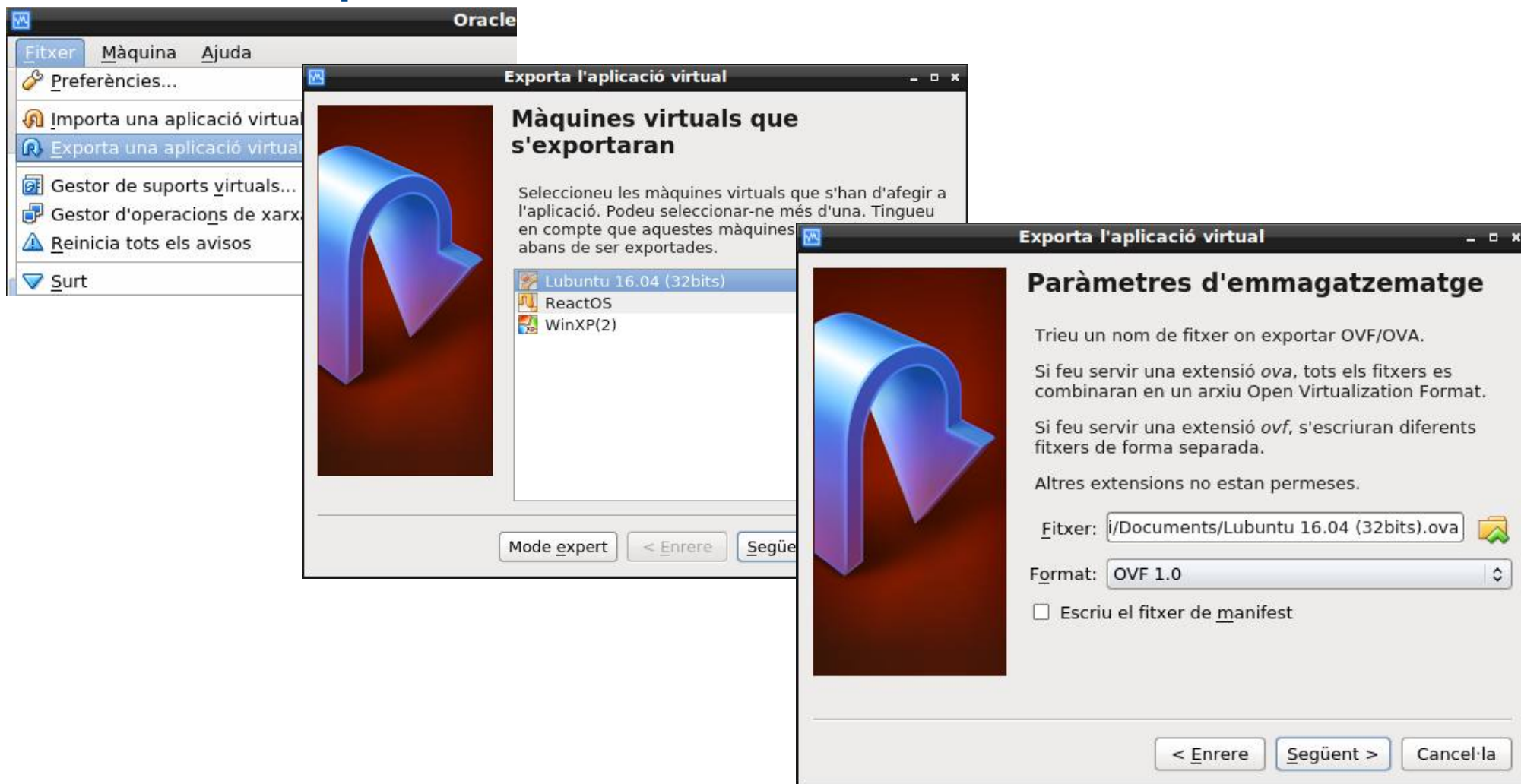
Installation

How to use or install GNU/Linux?

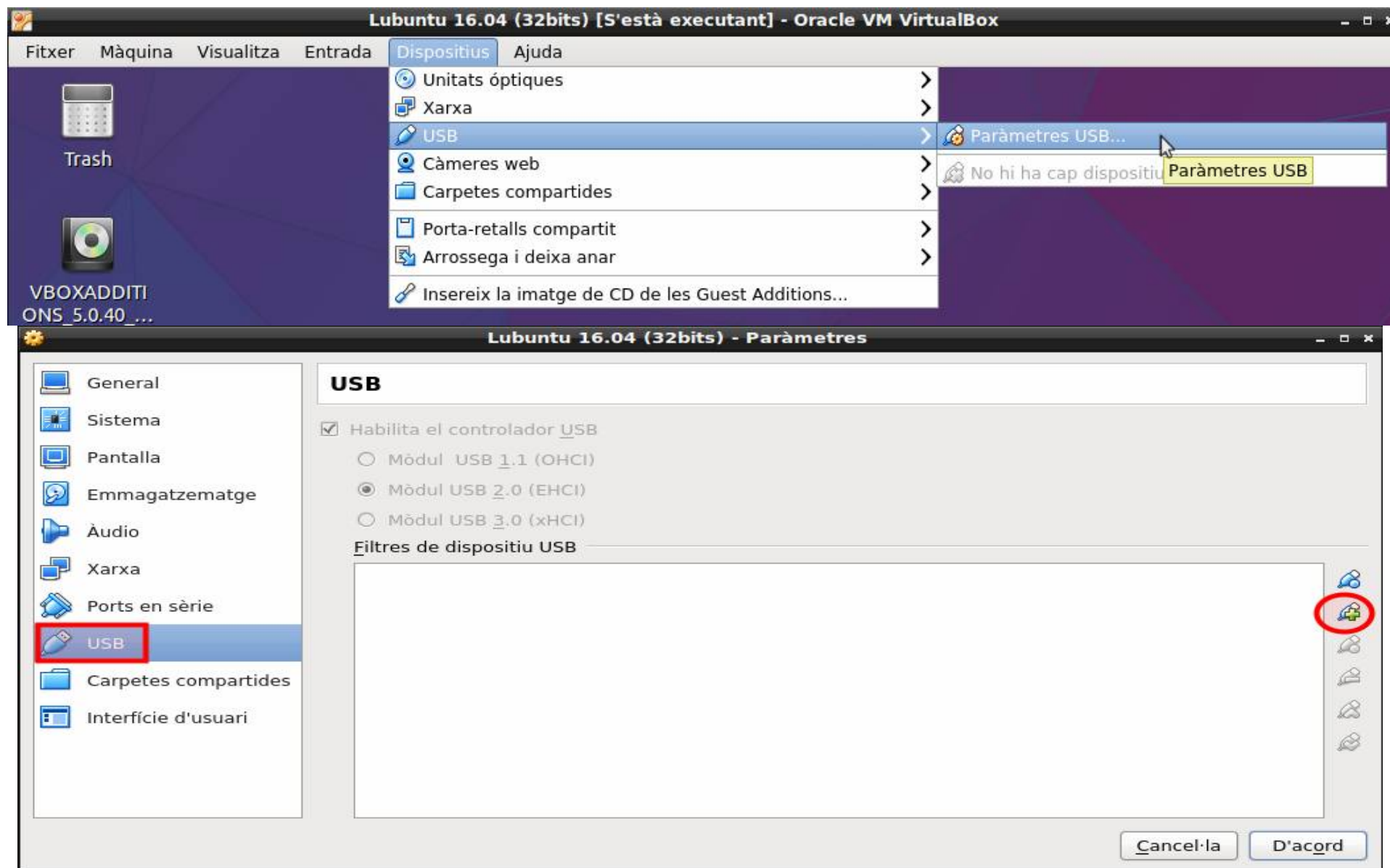
- LiveCD/LiveUSB: Without installing it in the hard disk
- Install in the hard disk
 - More efficient using your hardware, and you can install next to your other installed OS. You choose OS at booting time.
- Install in a disk USB or pendrive
- Install in a local virtual machine (or in the cloud)
 - From scratch (from .iso file), or from a previously exported virtual machine (.ova file or equivalent)
- Windows + andLinux/CoLinux

Exporting VirtualBox (.OVA)

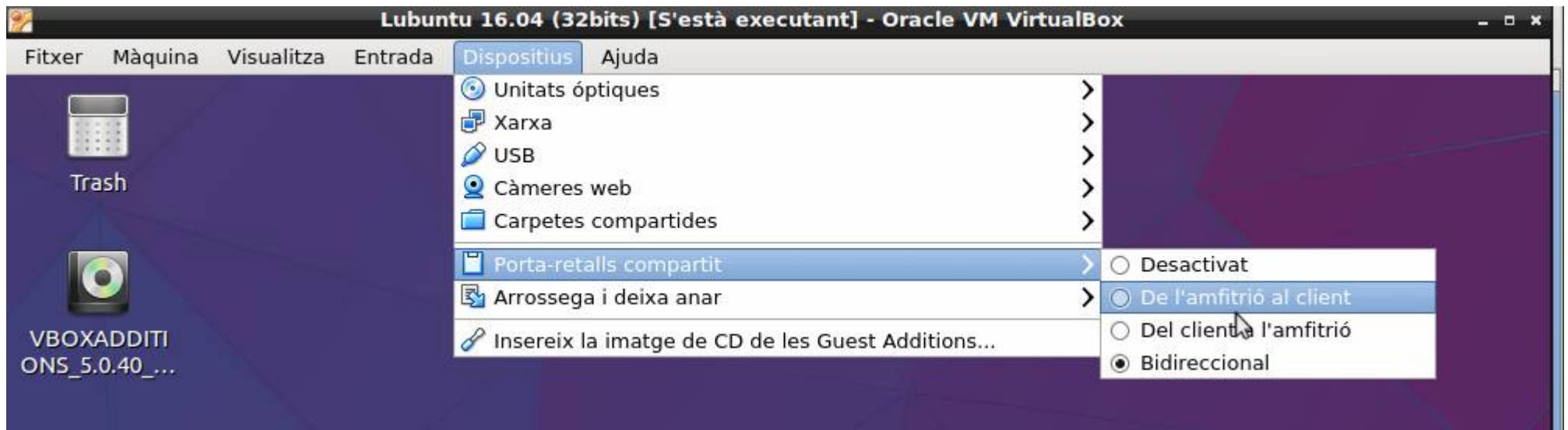
With current parameters



Willing to export to USB? Set it up

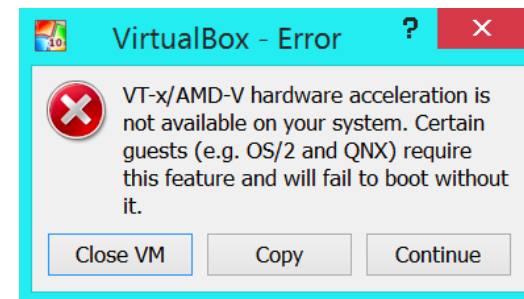


Willing to copy&paste to and/or from Host/Client?



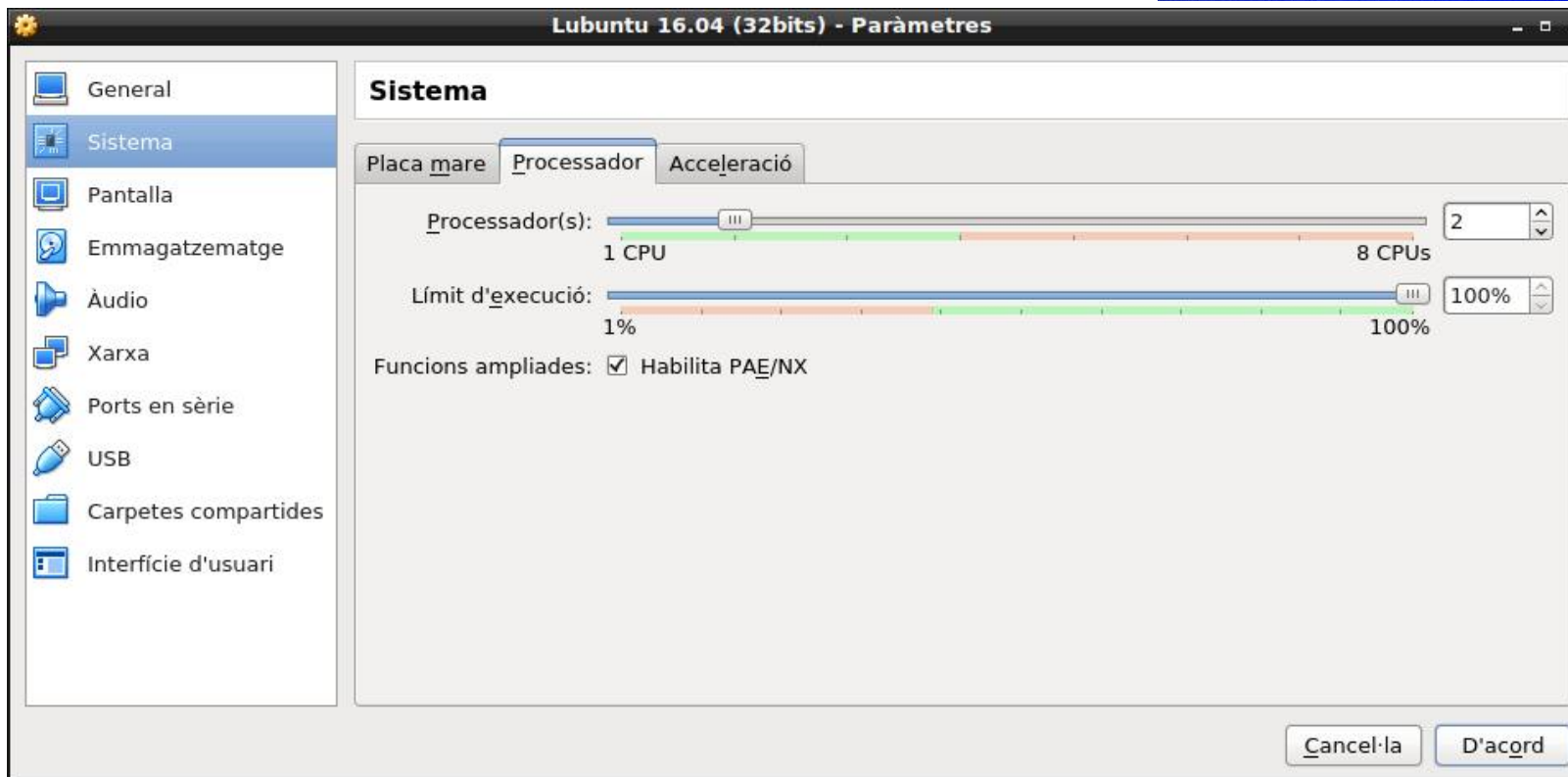
Importing VirtualBox (.OVA)

VT-X option disabled in BIOS allows only 1 cpu



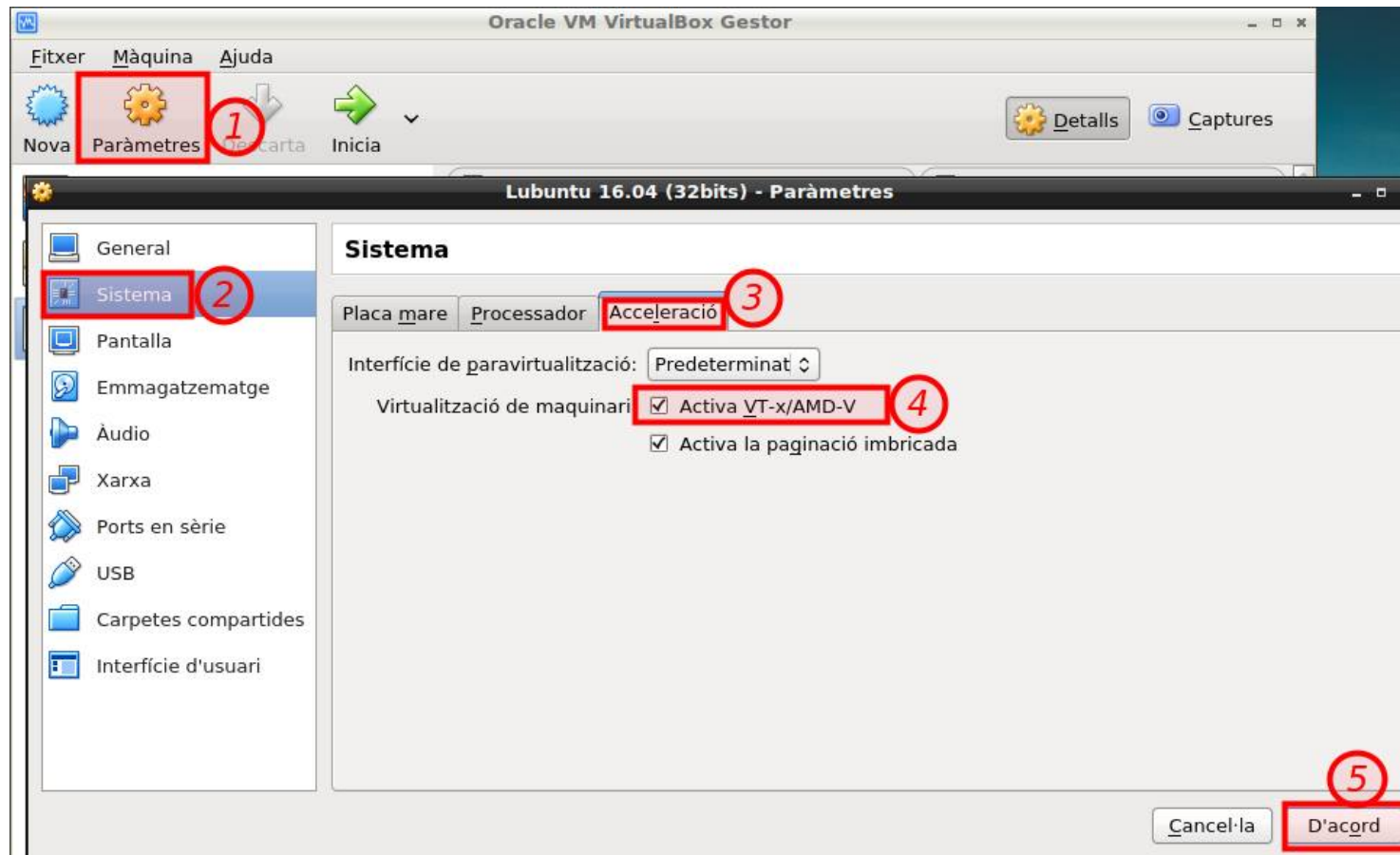
Importing VirtualBox (.OVA)

VT-X option enabled in BIOS
allows 2+ cpu



Where to enable/disable it?

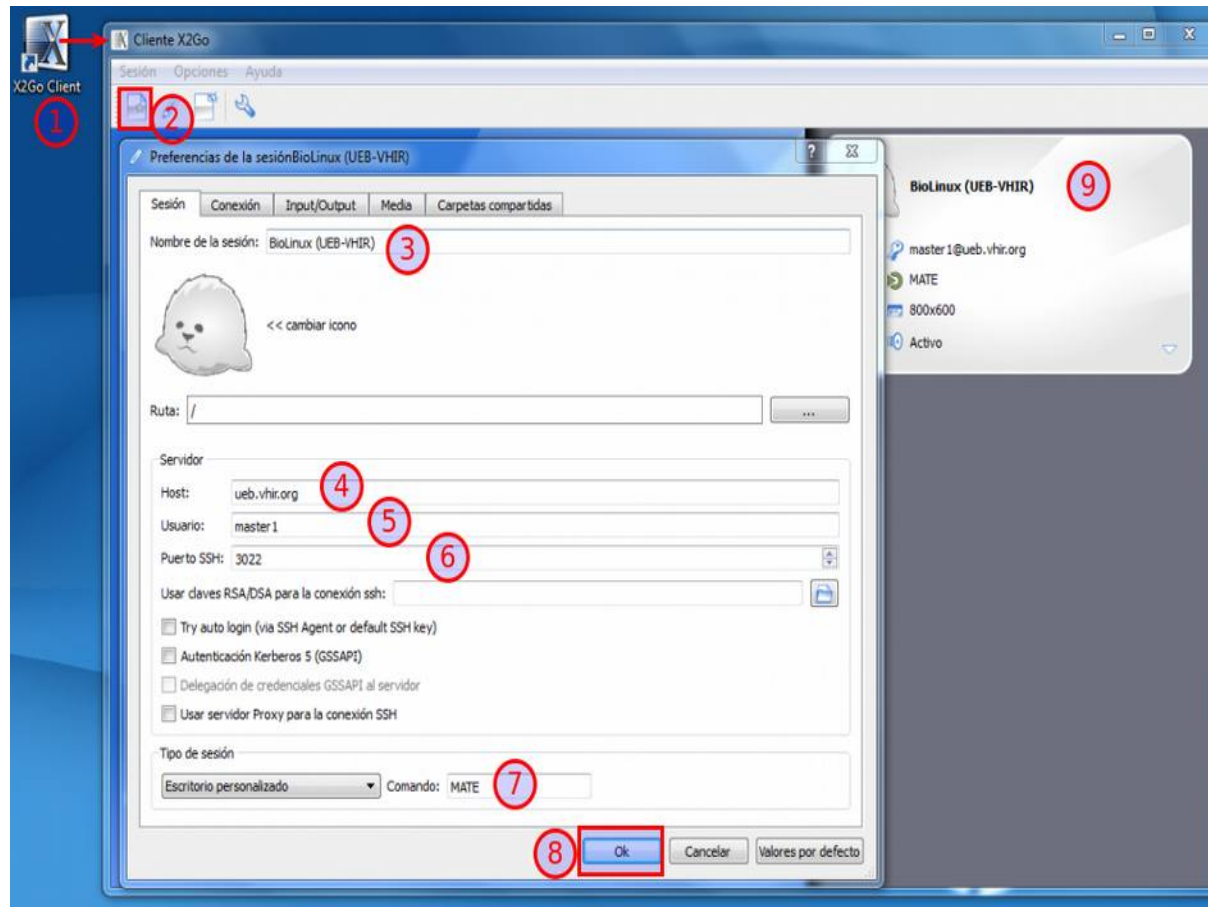
VirtualBox VM config for VT-X setting in BIOS



Lubuntu: Lightweight Desktop on a Ubuntu GNU/Linux Distribution. <http://lubuntu.net>

Other Options to use VirtualBox VM

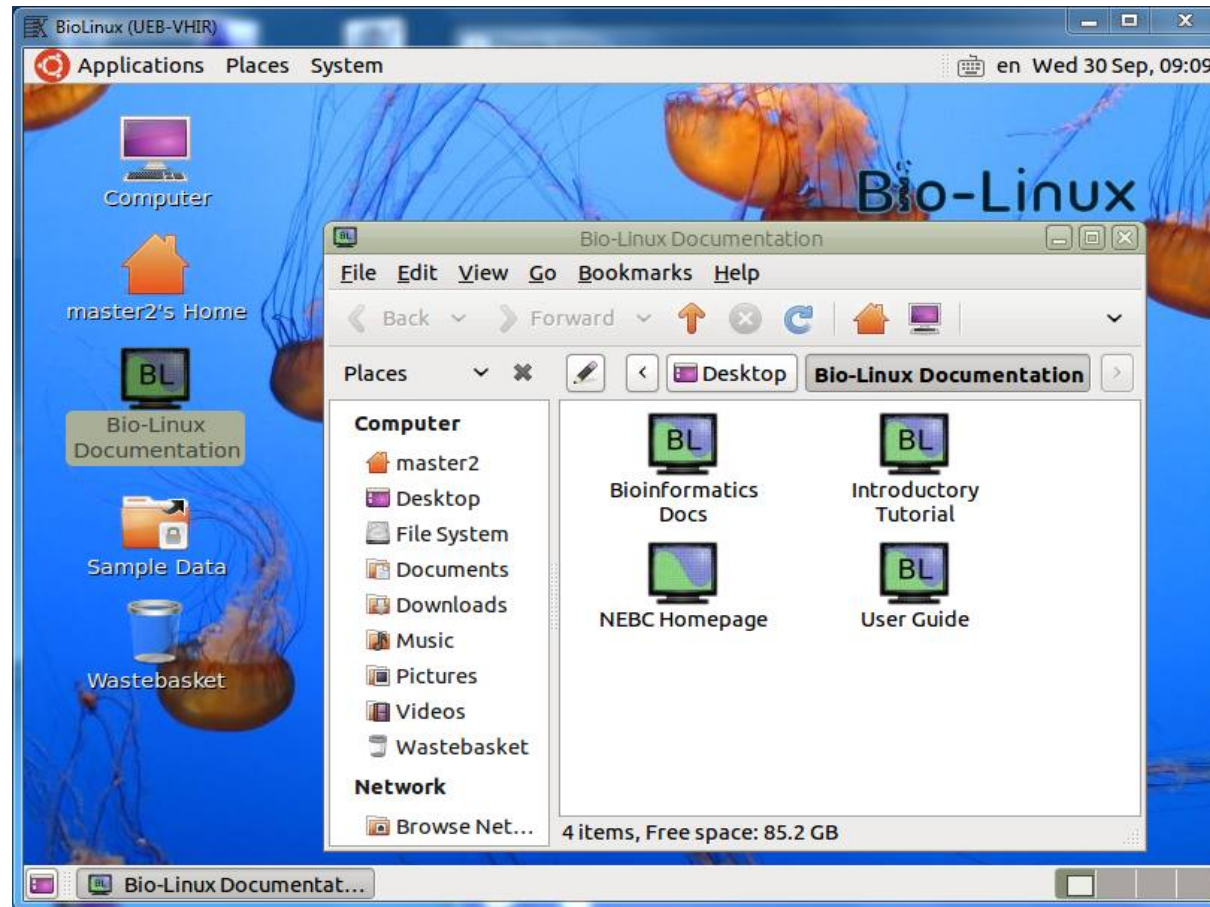
Connection to remote VBox using X2Go (ssh)



X2Go: Program to connect to a remote computer through GUI (Graphical User Interface)

Other Options to use VirtualBox VM

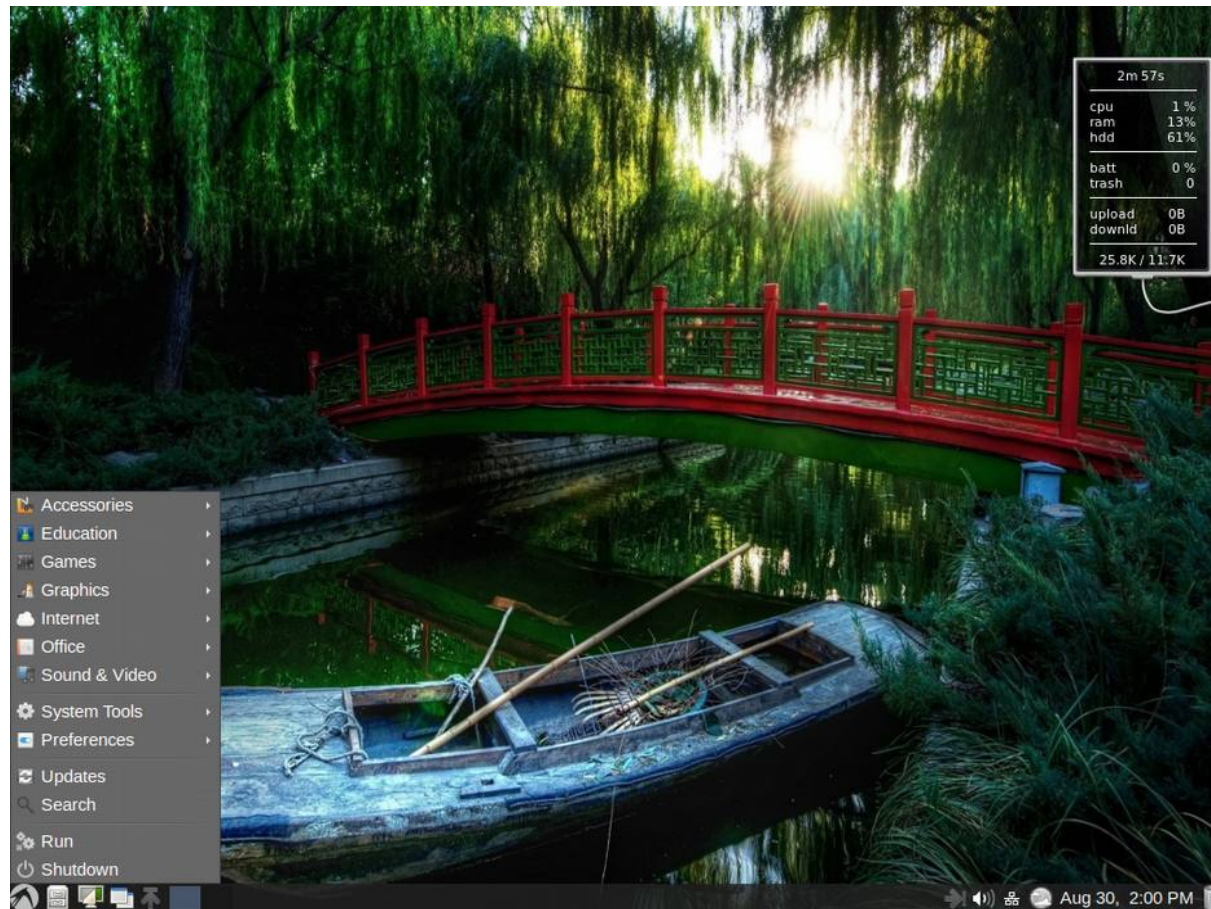
Connection to remote VBox using X2Go (ssh)



X2Go: <http://wiki.x2go.org>

Other Options

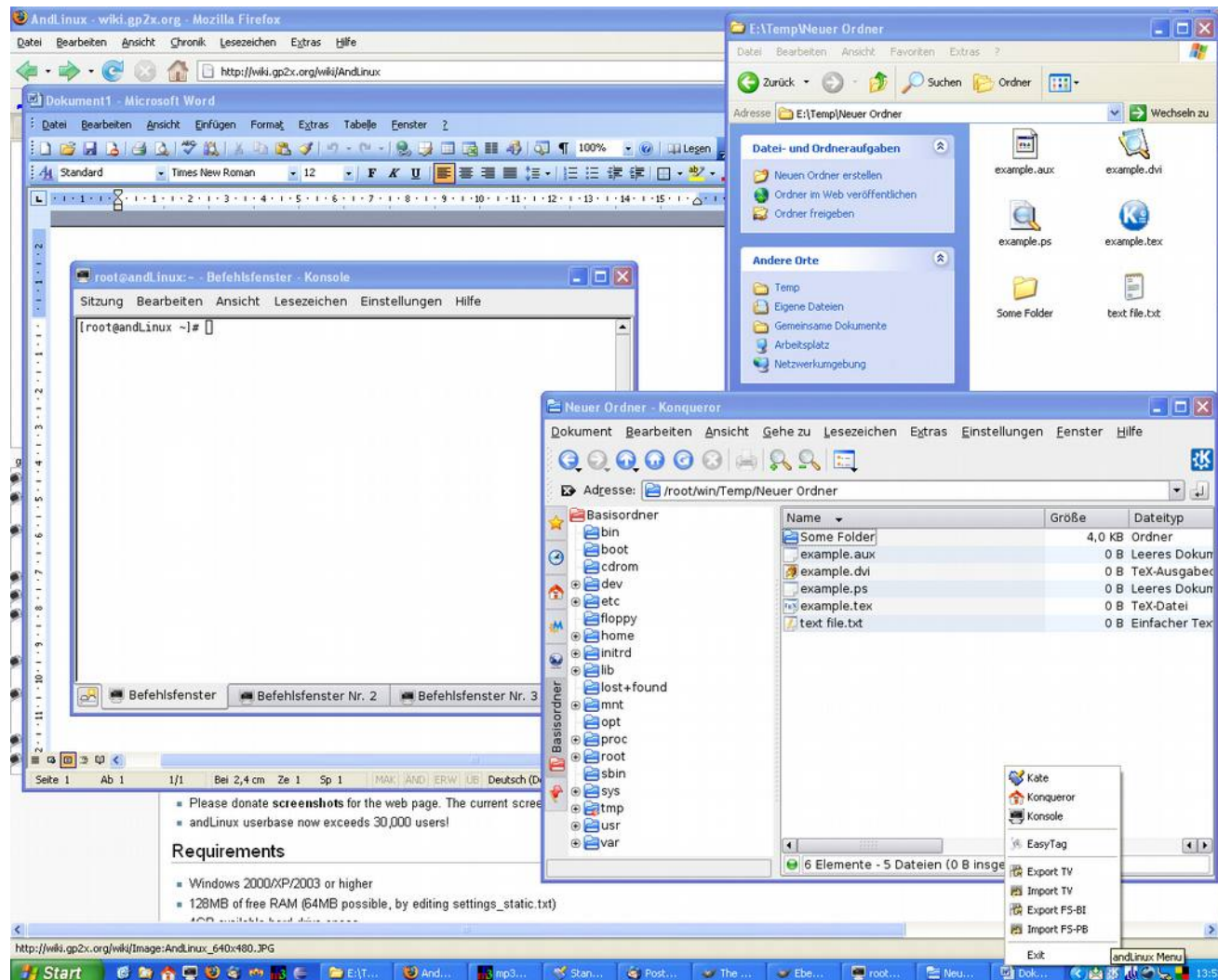
LiveUSB (example: LXLE Lubuntu-based; BIOS to boot from USB)



LXLE: A Lubuntu-based GNU/Distribution with improved desktop <http://www.lxle.net>

Other Options

andLinux (side to side to Windows - <http://andlinux.sf.net>)



Installation on Hard Disk

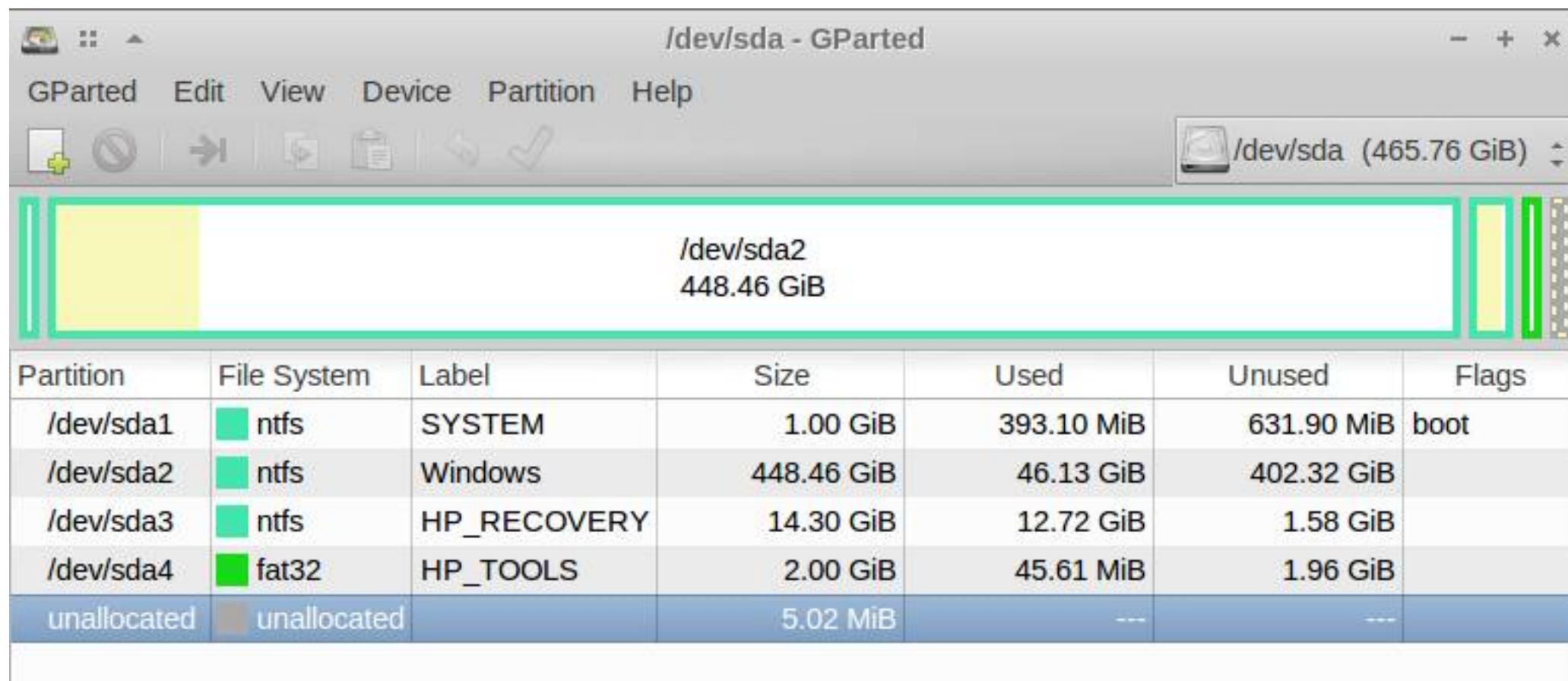
- GNU/Linux Installation
 - Dual/Multi-boot
 - Ubuntu (GNU/Linux) + Windows (or Mac OSX, ...)
 - From USB or CD/DVD
 - Requirements
 - Defrag Hard Drive (in Windows) – if needed
 - Backup Data in external device (usb, network drive, ...)
 - Identify Partitions in your hard drive
 - You may have **4** primary **partitions** (maximum)
 - You need **1 free partition** (minimum)
 - To make (preferably) 3 logic partitions inside

Using Live USB/CD/DVD

- Start from a Live USB or CD/DVD for inspecting & safer partition management
- Required: setup computer BIOS to allow booting from USB or CD/DVD
 - Press a key to enter configuration
 - Usually: DEL, F2, ESC,
 - Once there, look for the start sequence (Boot, Boot device, ...)

Hard Drive Partition Management before Installing GNU/Linux.
Challenging example:

- HP Laptop from 2015 with M\$ Win7



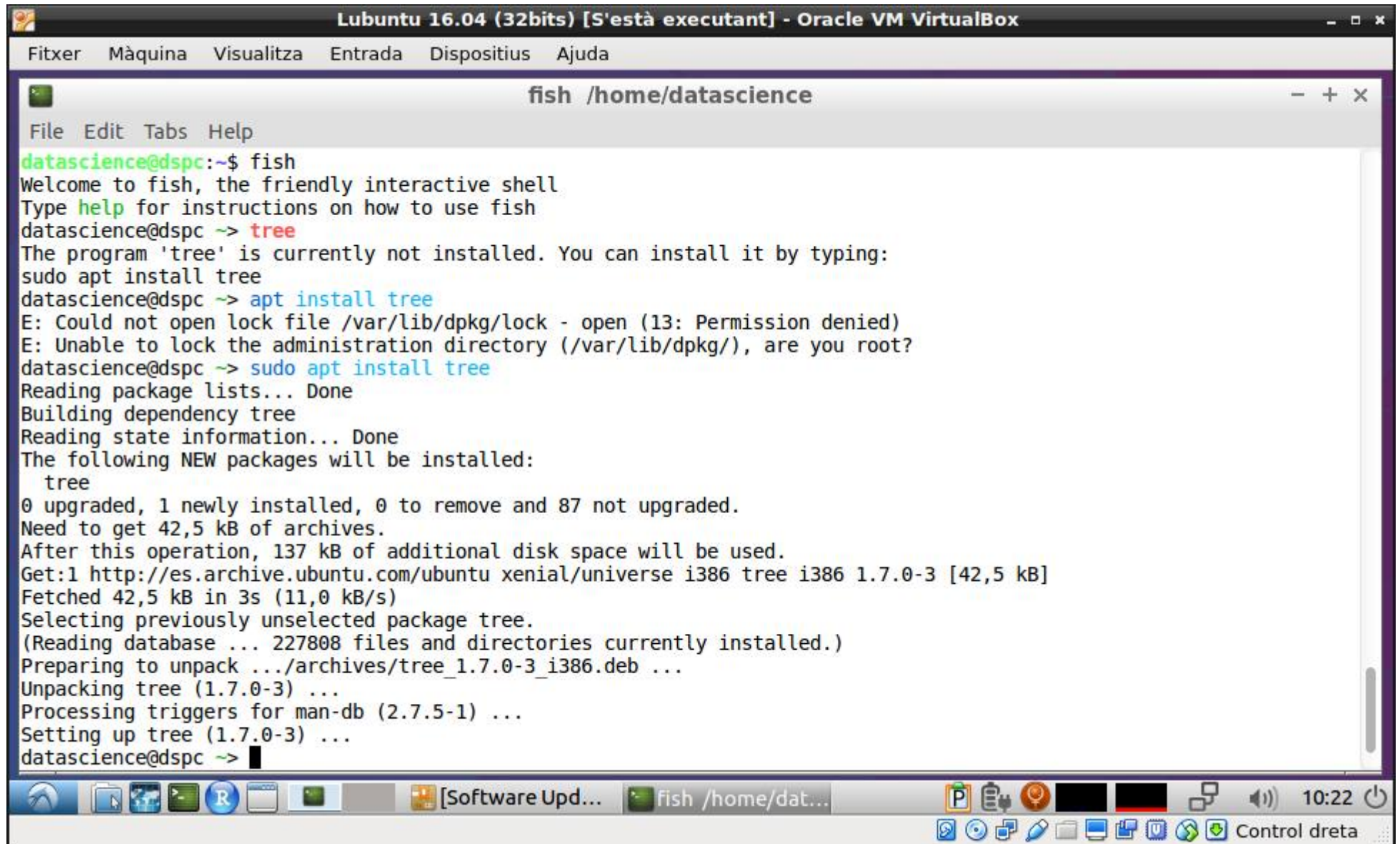
Partition	File System	Label	Size	Used	Unused	Flags
/dev/sda1	ntfs	SYSTEM	1.00 GiB	393.10 MiB	631.90 MiB	boot
/dev/sda2	ntfs	Windows	448.46 GiB	46.13 GiB	402.32 GiB	
/dev/sda3	ntfs	HP_RECOVERY	14.30 GiB	12.72 GiB	1.58 GiB	
/dev/sda4	fat32	HP_TOOLS	2.00 GiB	45.61 MiB	1.96 GiB	
unallocated	unallocated		5.02 MiB	---	---	

Package Management (installing «programs»)

Standard Package Management

- Console based:
 - APT (Debian/Ubuntu/Mint/LXLE)
 - YUM (Redhat/CentOS/Fedora Core)
 - YAST2 (SUSE)
 - Portage (Gentoo)
- GUI based
 - Synaptic (Debian-based)
 - Many others (distro-specific)

sudo apt install foo



The screenshot shows a terminal window titled "fish /home/datascience" running inside a "Lubuntu 16.04 (32bits) [S'està executant] - Oracle VM VirtualBox". The user enters the command `fish` to start a new shell. Then, they enter `tree`, which fails because the package is not installed. They then enter `sudo apt install tree`, which also fails due to a permission error. Finally, they enter `sudo apt install tree` with the 'sudo' command, which successfully installs the 'tree' package. The terminal output shows the progress of the installation, including reading package lists, building a dependency tree, and downloading the package from the Ubuntu repository.

```
Lubuntu 16.04 (32bits) [S'està executant] - Oracle VM VirtualBox
Fitxer Màquina Visualitza Entrada Dispositius Ajuda

fish /home/datascience
File Edit Tabs Help

datascience@dspc:~$ fish
Welcome to fish, the friendly interactive shell
Type help for instructions on how to use fish
datascience@dspc -> tree
The program 'tree' is currently not installed. You can install it by typing:
sudo apt install tree
datascience@dspc -> apt install tree
E: Could not open lock file /var/lib/dpkg/lock - open (13: Permission denied)
E: Unable to lock the administration directory (/var/lib/dpkg/), are you root?
datascience@dspc -> sudo apt install tree
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following NEW packages will be installed:
  tree
0 upgraded, 1 newly installed, 0 to remove and 87 not upgraded.
Need to get 42,5 kB of archives.
After this operation, 137 kB of additional disk space will be used.
Get:1 http://es.archive.ubuntu.com/ubuntu xenial/universe i386 tree i386 1.7.0-3 [42,5 kB]
Fetched 42,5 kB in 3s (11,0 kB/s)
Selecting previously unselected package tree.
(Reading database ... 227808 files and directories currently installed.)
Preparing to unpack .../archives/tree_1.7.0-3_i386.deb ...
Unpacking tree (1.7.0-3) ...
Processing triggers for man-db (2.7.5-1) ...
Setting up tree (1.7.0-3) ...
datascience@dspc -> 
```

Package Management

From extra repositories

add-apt-repository & apt update & apt install

- Oracle Java
 - `sudo add-apt-repository -y ppa:webupd8team/java`
- R (updated version)
 - `sudo add-apt-repository -y ppa:marutter/rrutter`
- Dependencies for R GIS packages (such as tmap)
 - `sudo add-apt-repository -y ppa:ubuntugis/ubuntugis-unstable`
 - `sudo add-apt-repository -y ppa:opencpu/jq`

```
fish /home/datascience
File Edit Tabs Help
datascience@dspc -> sudo add-apt-repository ppa:webupd8team/java
Oracle Java (JDK) Installer (automatically downloads and installs Oracle JDK8). There are no actual Java files in this PPA.

Important -> Why Oracle Java 7 And 6 Installers No Longer Work: http://www.webupd8.org/2017/06/why-oracle-java-7-and-6-installers-no.html

Update: Oracle Java 9 has reached end of life: http://www.oracle.com/technetwork/java/javase/downloads/jdk9-downloads-3848520.html

The PPA supports Ubuntu 18.04, 17.10, 16.04, 14.04 and 12.04.

More info (and Ubuntu installation instructions):
- for Oracle Java 8: http://www.webupd8.org/2012/09/install-oracle-java-8-in-ubuntu-via-ppa.html

Debian installation instructions:
- Oracle Java 8: http://www.webupd8.org/2014/03/how-to-install-oracle-java-8-in-debian.html

For Oracle Java 10, see a different PPA: https://www.linuxuprising.com/2018/04/install-oracle-java-10-in-ubuntu-or.html
More info: https://launchpad.net/~webupd8team/+archive/ubuntu/java
Press [ENTER] to continue or ctrl-c to cancel adding it

gpg: keyring `/tmp/tmpjn0qsu8o/secring.gpg' created
gpg: keyring `/tmp/tmpjn0qsu8o/pubring.gpg' created
gpg: requesting key EEA14886 from hkp server keyserver.ubuntu.com
gpg: /tmp/tmpjn0qsu8o/trustdb.gpg: trustdb created
gpg: key EEA14886: public key "Launchpad VLC" imported
gpg: no ultimately trusted keys found
gpg: Total number processed: 1
gpg:         imported: 1 (RSA: 1)
OK
datascience@dspc -> sudo apt update
Hit:1 http://ppa.launchpad.net/marutter/rrutter/ubuntu xenial InRelease
Hit:2 http://es.archive.ubuntu.com/ubuntu xenial InRelease
```

Add missing gpg repo keys

 With a helper: **launchpad-getkeys**

```
datascience@dspc ~> sudo add-apt-repository -y ppa:nilarimogard/webupd8  
datascience@dspc ~> sudo apt install launchpad-getkeys  
datascience@dspc ~> sudo launchpad-getkeys  
datascience@dspc ~> sudo apt update
```

Add all required ubuntu packages (dependencies) & dev helpers for some R packages

⚙ For Ubuntu 16.06 LTS with some extra repos:

```
datascience@dspc ~> sudo add-apt-repository -y ppa:marutter/rrutter
datascience@dspc ~> sudo add-apt-repository -y ppa:ubuntugis/ubuntugis-unstable
datascience@dspc ~> sudo add-apt-repository -y ppa:opencpu/jq
datascience@dspc ~> sudo apt update
datascience@dspc ~> sudo apt install -y r-recommended r-cran-xml libgraphviz-dev
libcairo2-dev r-cran-cairodevice freeglut3 freeglut3-dev r-cran-rgl r-cran-rgl r-cran-misc3d
libx11-dev libxt-dev libcurl4-gnutls-dev libxml2-dev r-cran-xml libgraphviz-dev libcairo2-dev
bwidget tk-table libv8-dev r-cran-rjava libmpfr-dev libc6 libssl-dev texlive-latex-extra texlive-
lang-spanish libx11-dev libxml2-dev libxml2:i386 libxt-dev r-cran-misc3d subversion git tk-
dev unaccent xvfb libgdal1-dev libproj-dev r-cran-rmysql libmagick++-dev r-cran-
rcolorbrewer r-cran-doparallel libssh2-1-dev libudunits2-dev libgdal-dev libgeos-dev libproj-
dev libv8-3.14-dev libjq-dev libprotobuf-dev protobuf-compiler libssl-dev libcairo2-dev
```

⚙ Fix permissions

```
sudo chmod 777 /usr/lib/R/site-library /usr/lib/R/site-library/* -R
sudo chmod 777 /usr/lib/R/library /usr/lib/R/library/* -R
sudo chmod 777 /usr/share/R/doc/html/* -R
```

Package Management

From External sources

- Example: Adobe Acrobat Reader 9.x

```
wget http://ardownload.adobe.com/pub/adobe/reader/unix/9.x/9.5.5/enu/AdbeRdr9.5.5-1_i386linux_enu.deb  
sudo dpkg -i AdbeRdr9.5.5-1_i386linux_enu.deb; sudo apt-get -f install #For 32 bits  
sudo dpkg -i --force-architecture AdbeRdr9.5.5-1_i386linux_enu.deb; sudo apt-get -f  
install #For 64 bits
```

- Example: Portable Signer:

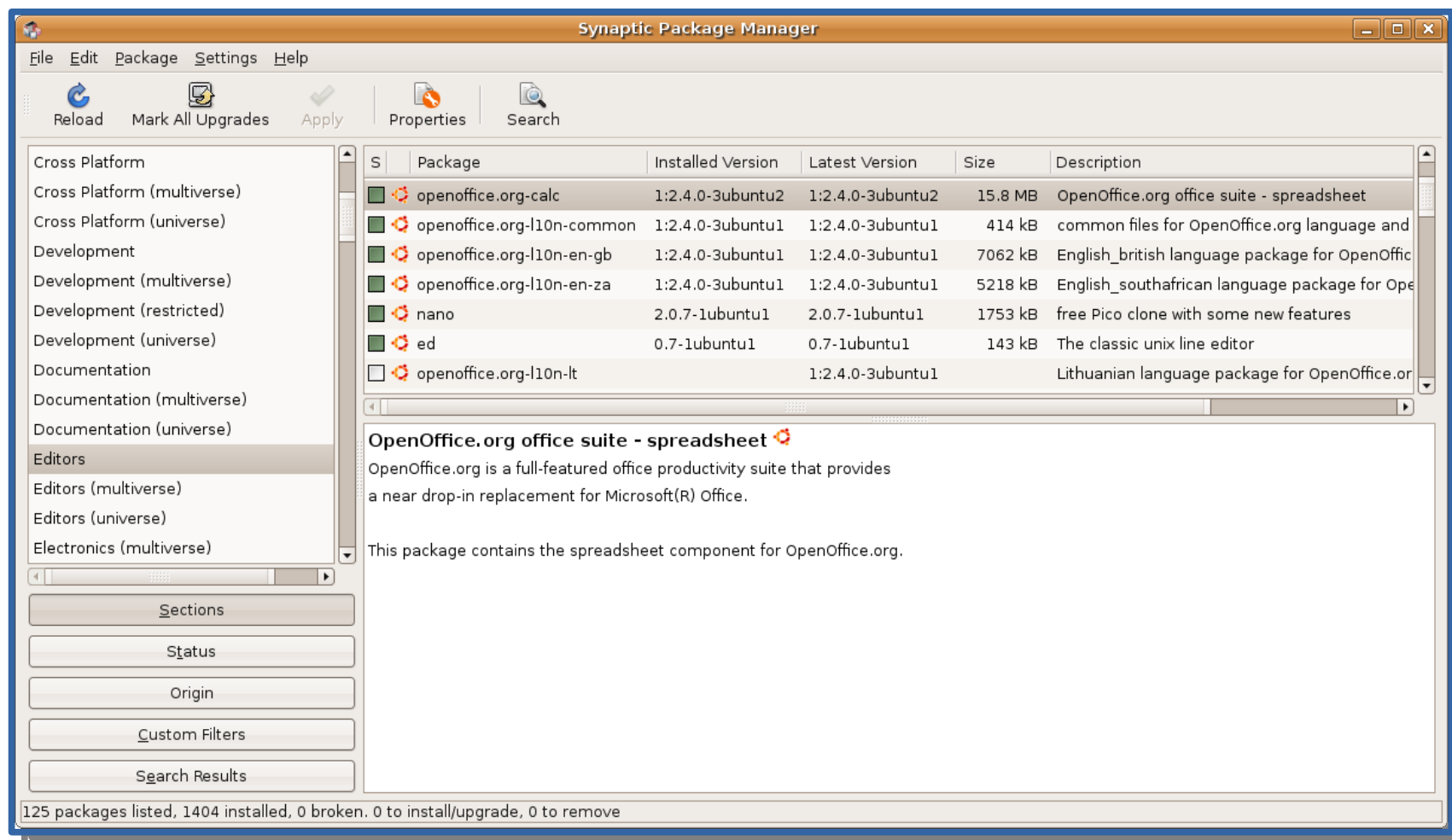
To digitally sign PDF with certificates like FNMT (.p12)

<http://portablesigner.sf.net>

- download, uncompress, make executable and run (java app)

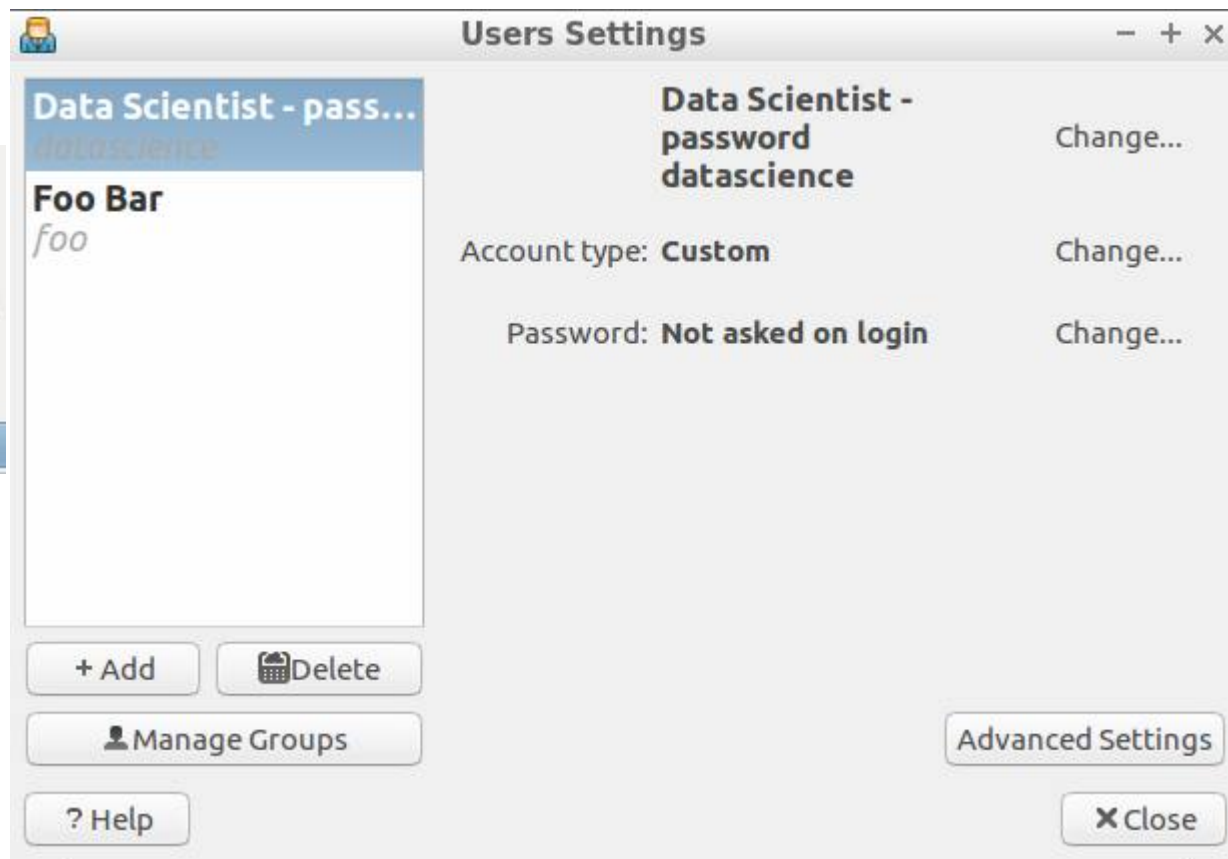
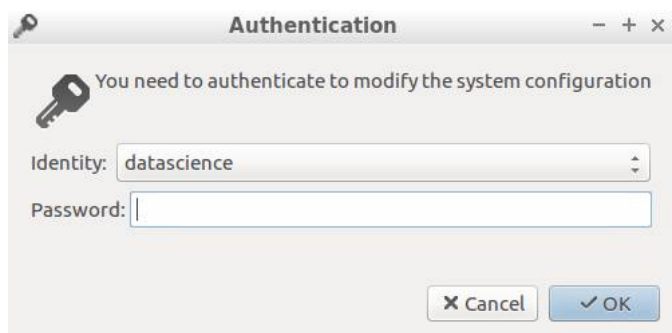
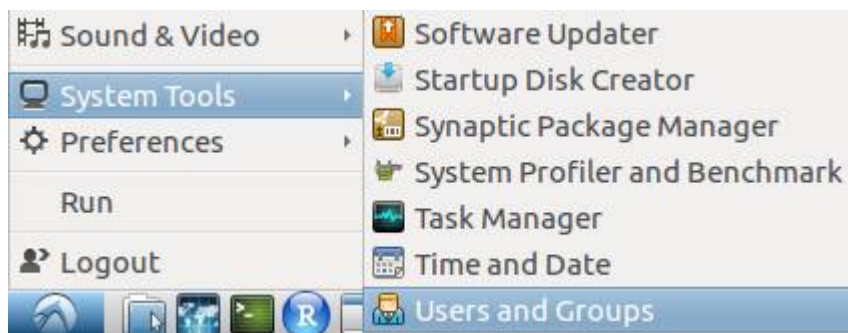
Package Management (GUI)

Synaptic



User (& Group) Management

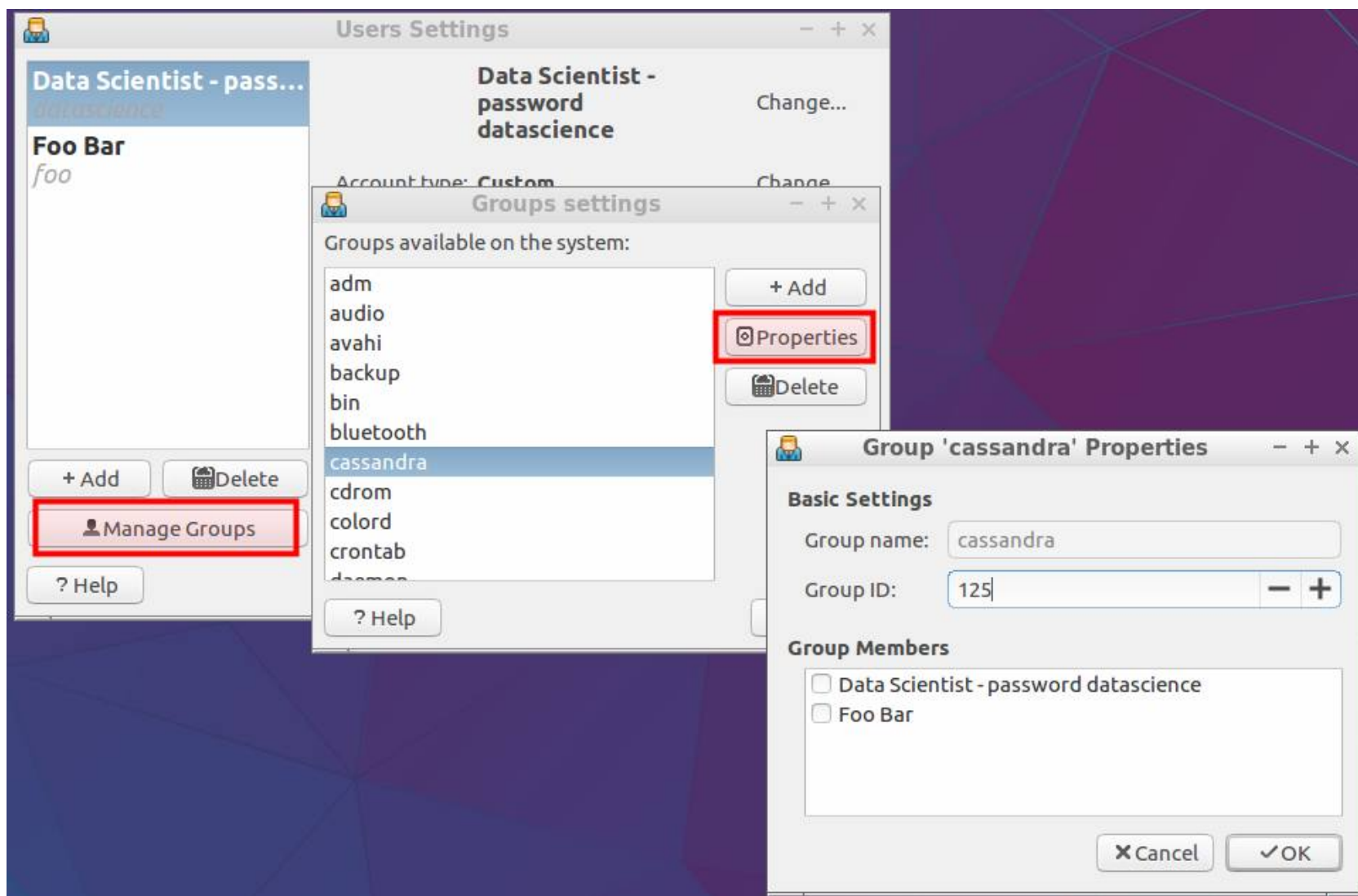
GUI based



Console based: adduser, passwd, ...

```
datascience@dspc ~> sudo su #become root user («Switch User»)
[sudo] password for datascience:
root@dspc:/home/datascience# adduser foo
Adding user `foo' ...
Adding new group `foo' (1001) ...
Adding new user `foo' (1001) with group `foo' ...
Creating home directory `/home/foo' ...
Copying files from `/etc/skel' ...
Enter new UNIX password: (new-user password typed here)
Retype new UNIX password: (new-user password typed again)
passwd: password updated successfully
Changing the user information for foo
Enter the new value, or press ENTER for the default
  Full Name []: Foo Bar
  Room Number []:
  Work Phone []:
  Home Phone []:
  Other []:
Is the information correct? [Y/n] Y
root@dspc:/home/datascience# exit
exit
datascience@dspc ~>
```

User (& Group) Management



Permission Management

Console based

- **chmod** – change permissions. **chown** – change ownership

u – user g – group o – other

r – read w – write x – execute

sudo chmod -R ug+rw /DATA/SHARE

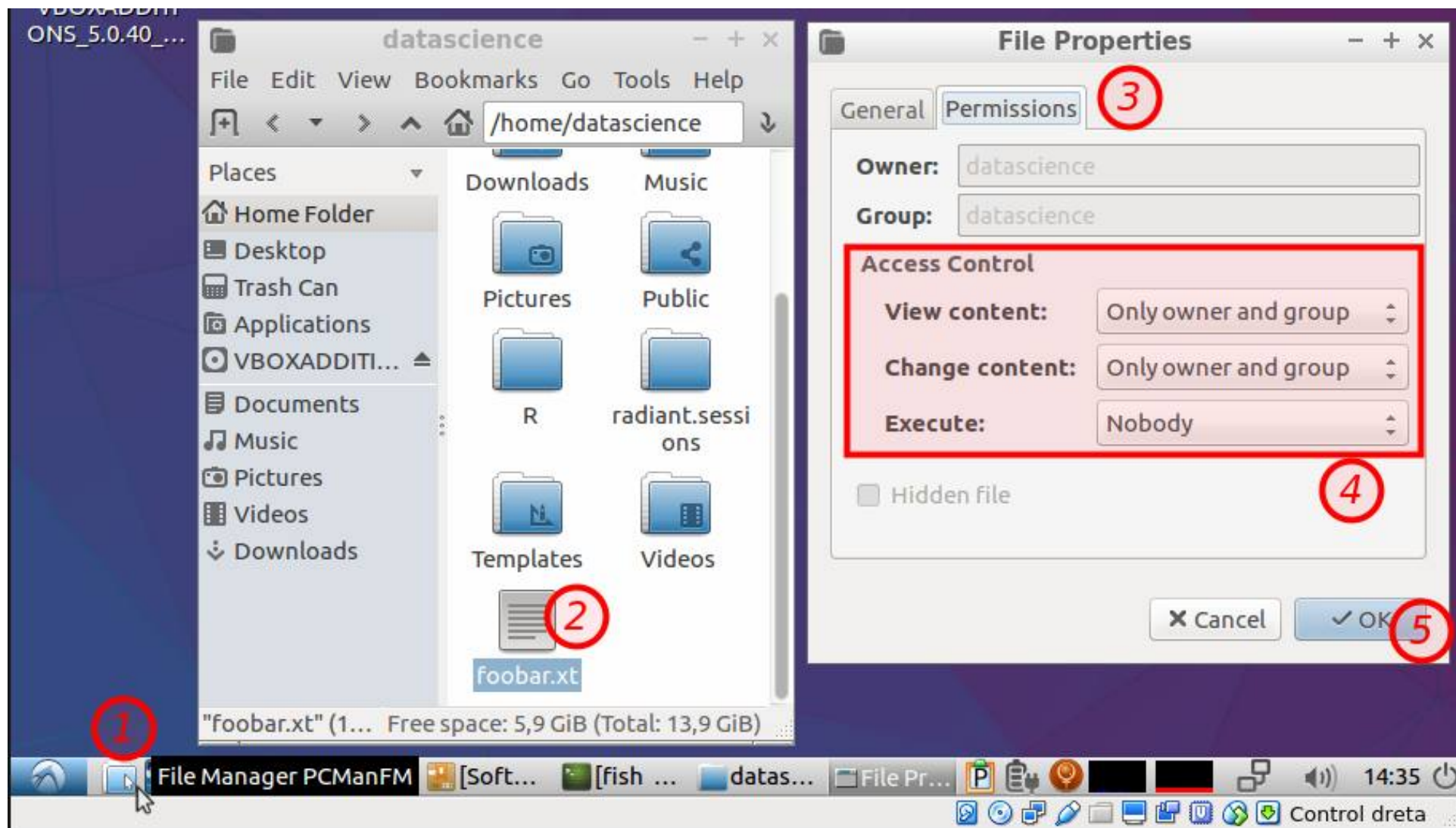
sudo chmod -R 660+rw /DATA/SHARE

- **-R** → it modifies permission of parent folder & child objects within
- **ug+rw** (= **660**) → it gives *User & Group (but not Others)* read and write access (but not execute access).
- See:
<https://www.linux.com/learn/understanding-linux-file-permissions> &
<https://www.linux.com/learn/how-manage-file-and-folder-permissions-linux>

Permission Management

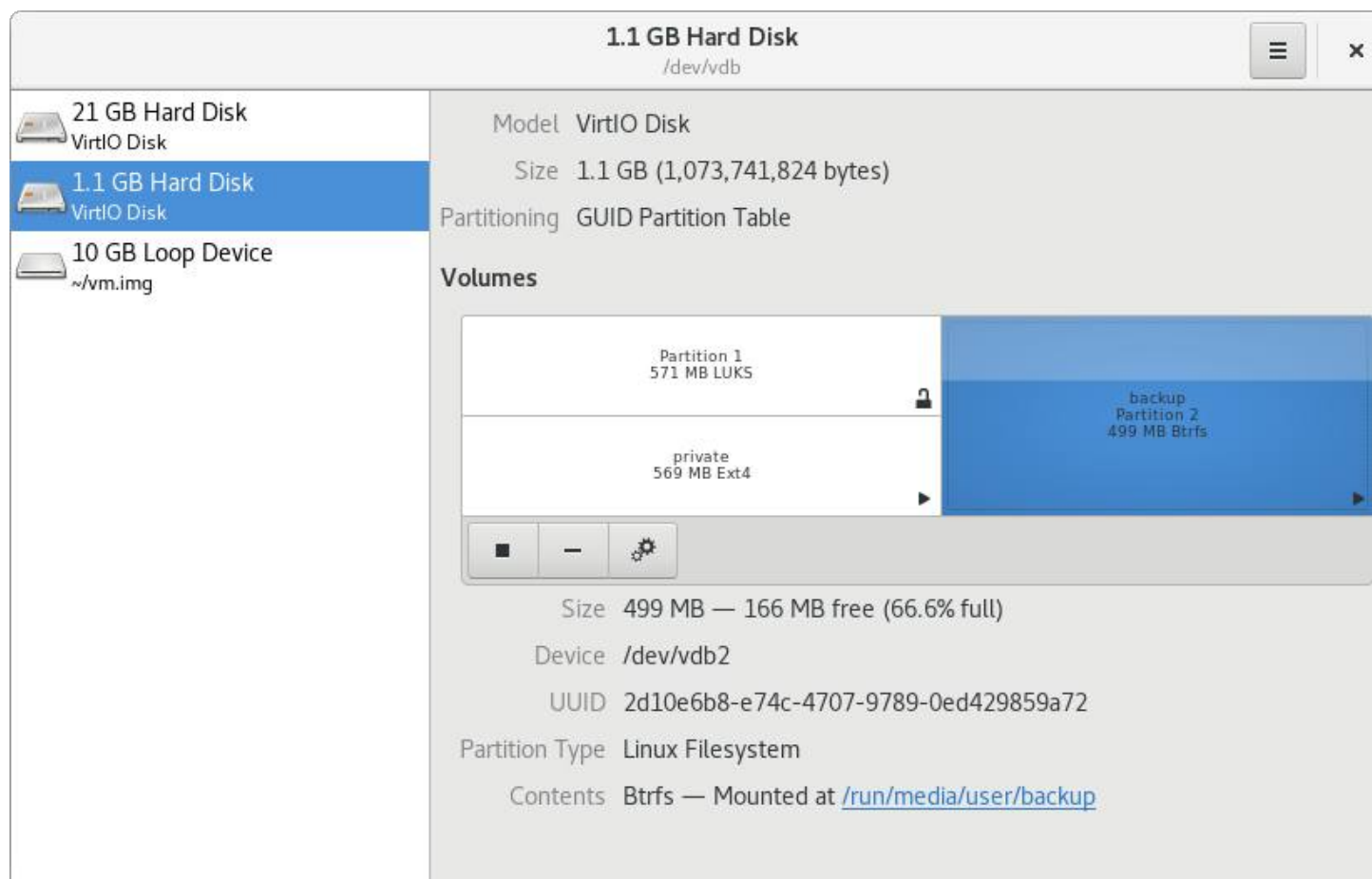
```
fish /home/datascience
File Edit Tabs Help
datascience@dspc ~-> nano foobar.txt
datascience@dspc ~-> ls -l
total 44
drwxr-xr-x 2 datascience datascience 4096 mar 30 13:08 Desktop/
drwxr-xr-x 2 datascience datascience 4096 feb 13 00:26 Documents/
drwxr-xr-x 2 datascience datascience 4096 mar 30 12:53 Downloads/
-rw-rw-r-- 1 datascience datascience 16 may 1 14:29 foobar.txt
drwxr-xr-x 2 datascience datascience 4096 feb 13 00:26 Music/
drwxr-xr-x 2 datascience datascience 4096 feb 13 00:26 Pictures/
drwxr-xr-x 2 datascience datascience 4096 feb 13 00:26 Public/
drwxrwxr-x 3 datascience datascience 4096 feb 27 07:26 R/
drwxrwxr-x 2 datascience datascience 4096 mar 31 04:09 radiant.sessions/
drwxr-xr-x 2 datascience datascience 4096 feb 13 00:26 Templates/
drwxr-xr-x 2 datascience datascience 4096 feb 13 00:26 Videos/
datascience@dspc ~-> chmod ug+rw foobar.txt
datascience@dspc ~-> ls -l | grep foobar
-rw-rw-r-- 1 datascience datascience 16 may 1 14:29 foobar.txt
datascience@dspc ~-> chmod o-x foobar.txt
datascience@dspc ~-> ls -l | grep foobar
-rw-rw-r-- 1 datascience datascience 16 may 1 14:29 foobar.txt
datascience@dspc ~-> chmod o-r foobar.txt
datascience@dspc ~-> ls -l | grep foobar
-rw-rw---- 1 datascience datascience 16 may 1 14:29 foobar.txt
datascience@dspc ~-> █
```

Permission Management (GUI)



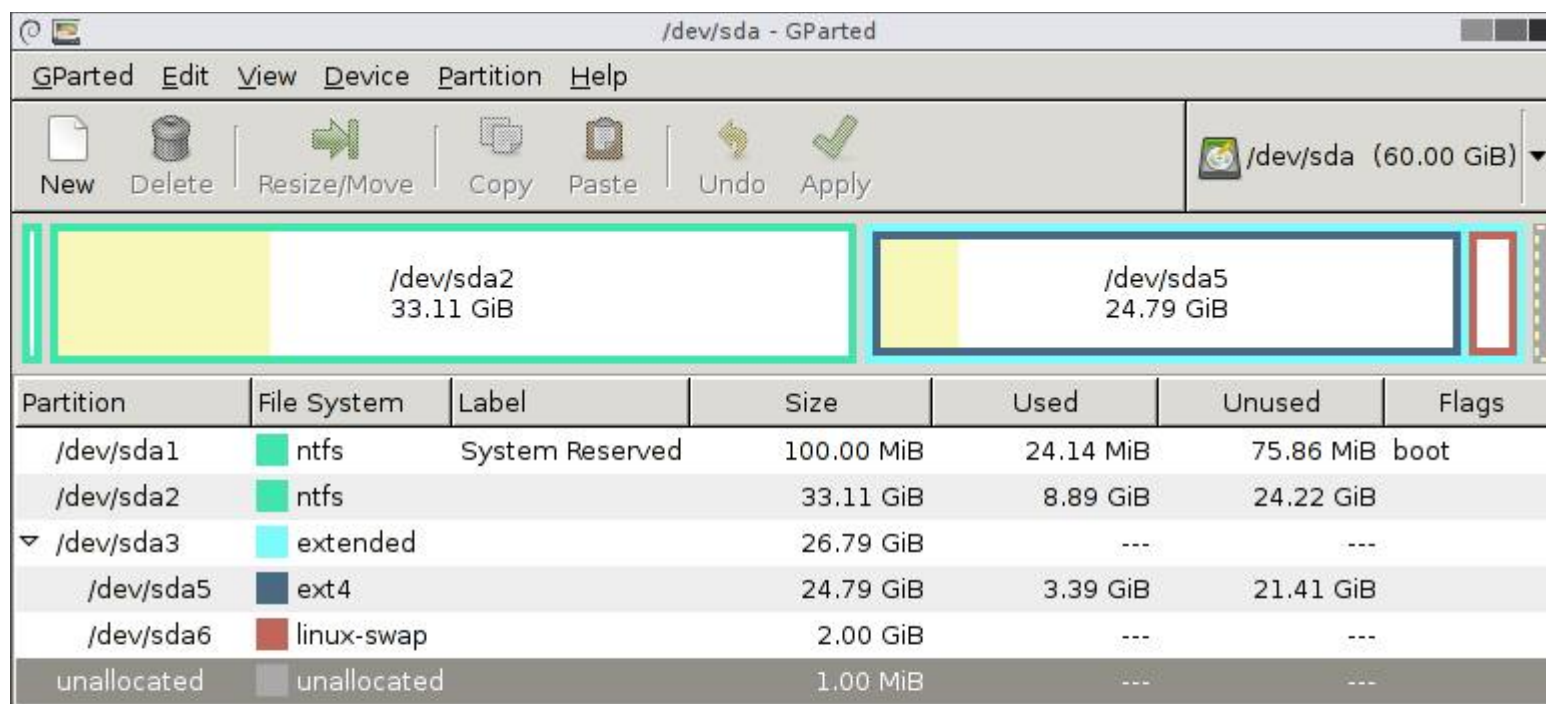
Device Management

🔧 **Gnome Disks:** <https://wiki.gnome.org/Apps/Disks>



Device Management

🔧 Gparted: <https://gparted.org>



The screenshot shows the GParted application window titled "/dev/sda - GParted". The menu bar includes GParted, Edit, View, Device, Partition, and Help. The toolbar contains icons for New, Delete, Resize/Move, Copy, Paste, Undo, and Apply. A dropdown menu on the right shows "/dev/sda (60.00 GiB)". The main display area shows a visual representation of the disk with two partitions highlighted: /dev/sda2 (33.11 GiB) and /dev/sda5 (24.79 GiB). Below this is a table with the following data:

Partition	File System	Label	Size	Used	Unused	Flags
/dev/sda1	ntfs	System Reserved	100.00 MiB	24.14 MiB	75.86 MiB	boot
/dev/sda2	ntfs		33.11 GiB	8.89 GiB	24.22 GiB	
▼ /dev/sda3	extended		26.79 GiB	---	---	
/dev/sda5	ext4		24.79 GiB	3.39 GiB	21.41 GiB	
/dev/sda6	linux-swap		2.00 GiB	---	---	
unallocated	unallocated		1.00 MiB	---	---	

Backup Management: tips

Keep copies elsewhere than same computer

- «Elsewhere»: different hardware, different device, different room, different building, ...

Automatic (regular) backups

«Smart remove» even more important

Efficiency vs. resilience

- Simple enough so that some team mates can restore them?
- Complex RAID disk setups vs redundant external hard drives elsewhere?
- Encrypted vs unencrypted?
- Advice: Use your team crowd wisdom; stay away from single-man freaks with too-techie solutions

Backup Management: tools

Déjà Dup (Duplicity GUI)

- <https://wiki.gnome.org/Apps/DejaDup/>
<https://wiki.gnome.org/Apps/DejaDup/>

Backintime (GUI & console; rsync & hard link based)

- <https://github.com/bit-team/backintime>

Luckybackup (rsync-based GU)

- <http://luckybackup.sourceforge.net/>

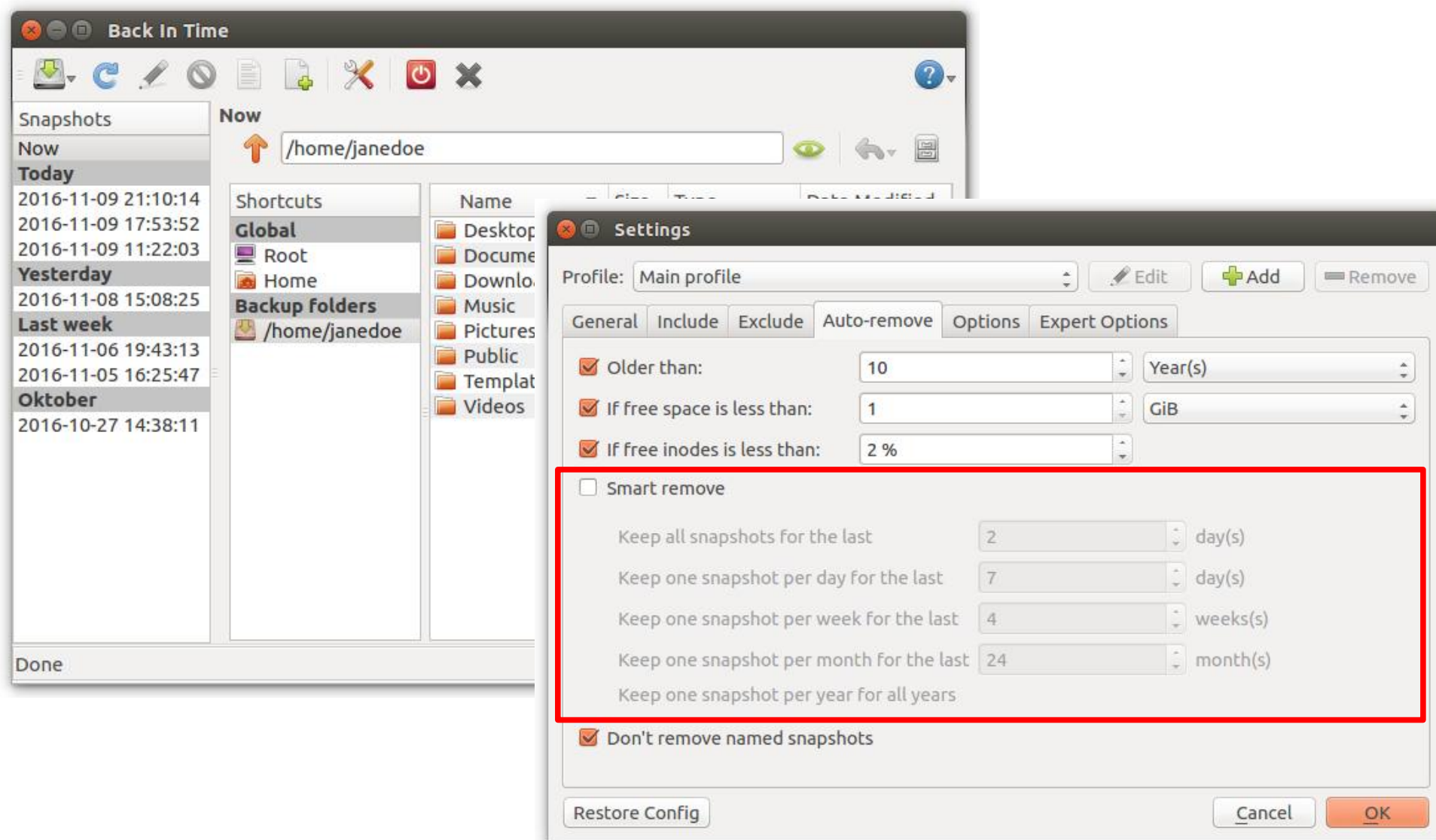
Backup Ninja (console-based)

- <https://0xacab.org/riseuplabs/backupninja>

Custom Bash Scripts

Some examples at: <https://github.com/xavidp/bashscripts>

Backintime & Smart remove

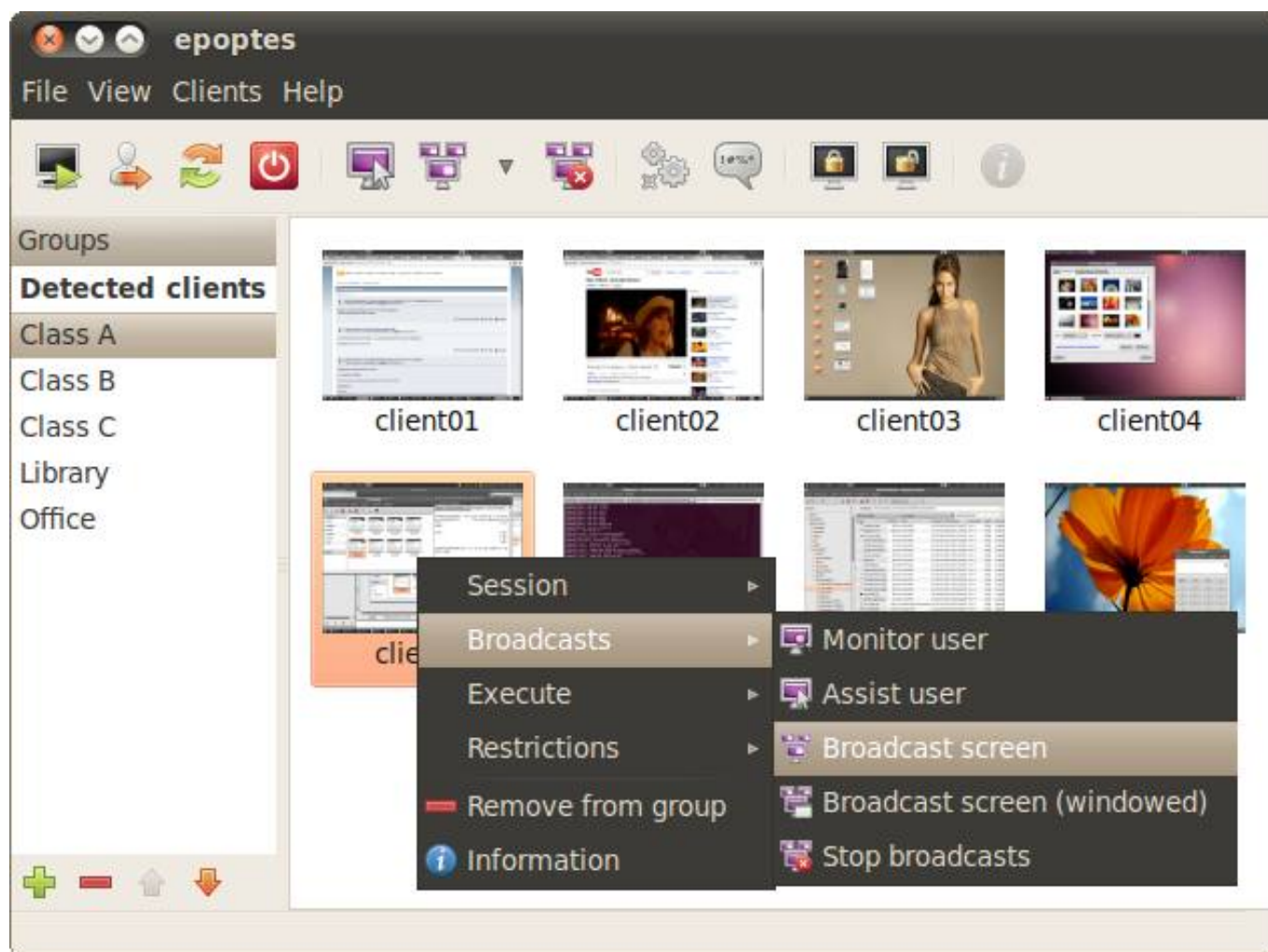


Security

- ⚙ ~~http~~ > https, ~~ftp~~ > sftp, ~~root~~ > sudo, ~~root ssh~~, jailkit, rootkit hunter, fail2ban, AppArmor/SELinux
- ⚙ Firewalls. UFW (Uncomplicated Firewall)
 - <https://help.ubuntu.com/community/Gufw>
- ⚙ Mail tweaks (postfix) & Spamassassin ...
- ⚙ Antivirus (dual boots)?: amavis + GUI (clamTk, Klamav)
- ⚙ Monitoring: logwatch, swatch, munin, monit, nagios, zabbix, ...
- ⚙ Password management: keepass2, ...
- ⚙ sudo apt-get dist-upgrade

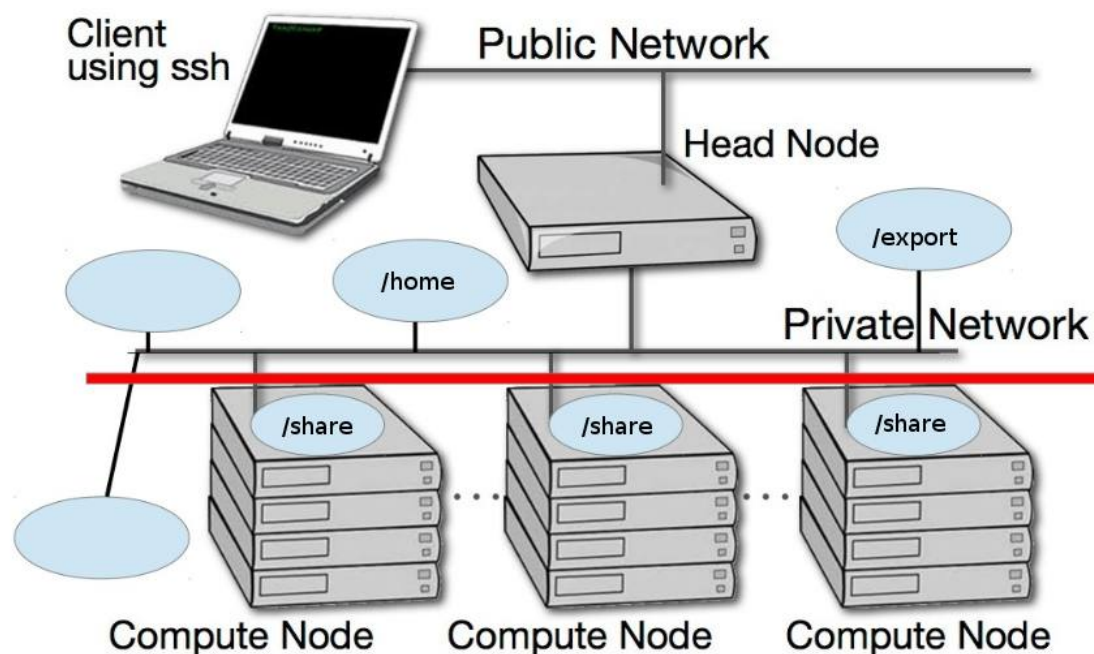
Computer Client Management

⚙️ **Epoptes** - <http://www.epoptes.org>



Computer Clusters, ...

🔧 Grid, program modules, job queue management ...



- For newbies: Rocks Clusters distro (CentOS based)



Open-Source Toolkit
for Real and Virtual Clusters

<http://www.rocksclusters.org>

Computer Clusters

SgeMonitor - Phoenix

File Edit View Go Bookmarks Tools Help

Fonts Colors Images JavaScript Clear Cache Real UA PrefBar Help What's New Customize

SGEMonitor

Rows with a red background have load average near 0 meaning they are not doing much so you may want to check the status of those jobs.

[Get per user stats](#)

Job Id	Pri	Job Description	Userid	State	Time in Queue or Runtime	Queue	Resources requested	Load Average
9361	-50	test_25.200	john	r	4 day(s), 20 hour(s), 6 minute(s)	acoma.q	vcs=1, qe_regress_long=1	1.07
10097	0	make	dennis	r	3 hour(s), 58 minute(s)	isleta.q	vcs=1	1.26
9831	-50	t22_3.002	john	S	1 day(s), 5 hour(s), 26 minute(s)	isleta.r	vcs=1, qe_regress_long=1, regress=TRUE	1.26
9819	-50	t22_320.0	john	r	1 day(s), 5 hour(s), 26 minute(s)	jemez.r	vcs=1, qe_regress_long=1, regress=TRUE	1.19
10171	0	test_11.721	jim	r	8 minute(s)	keresan.q	vcs=1	0.91
10152	0	sleep	josh	r	1 hour(s), 14 minute(s)	lab_sleep.	system15=1	
10168	0	wrap	chris	r	30 minute(s)	lago.q		1.06
10073	0	simulate	jack	r	4 hour(s), 36 minute(s)	pueblo.q	vcs=1	1.59
9821	-50	t22_19.41	john	S	1 day(s), 5 hour(s), 26 minute(s)	pueblo.r	vcs=1, qe_regress_long=1, regress=TRUE	1.59
10173	0	t22_19.16	john	r	8 minute(s)	taos.r	vcs=1, regress=TRUE, qe_regress_long=1	0.99
10158	0	make	dennis	r	1 hour(s), 7 minute(s)	tesuque.q	vcs=1	1.13
10166	0	r5_7.030	john	r	44 minute(s)	zia.r	vcs=1, regress=TRUE, qe_regress_long=1	1.18

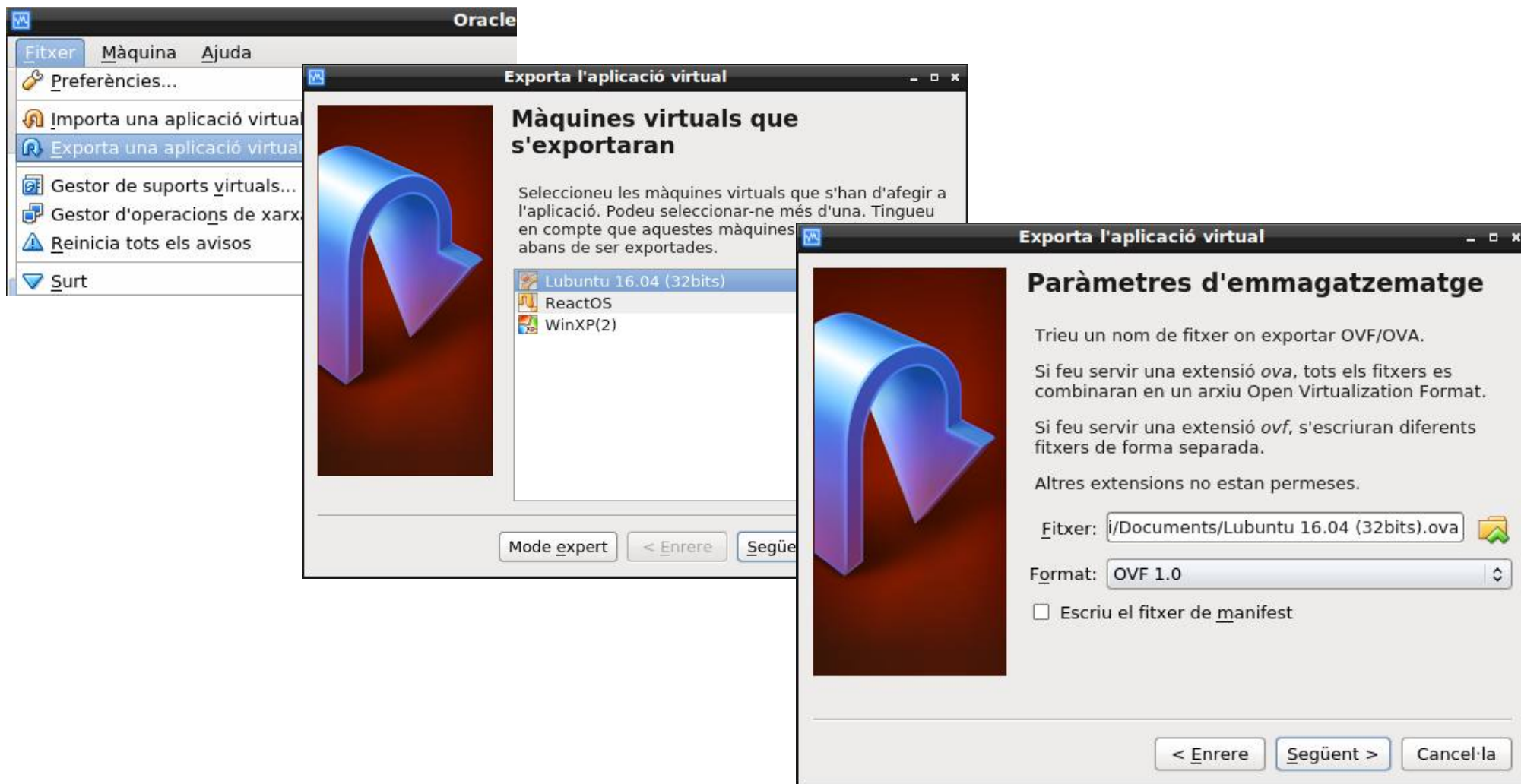
User	Total jobs	Running jobs	Queued/Suspended jobs
ryan	1	1	
eric	1	1	
john	16	14	2

scheduling info: queue "acoma.r" dropped because it is temporarily not available

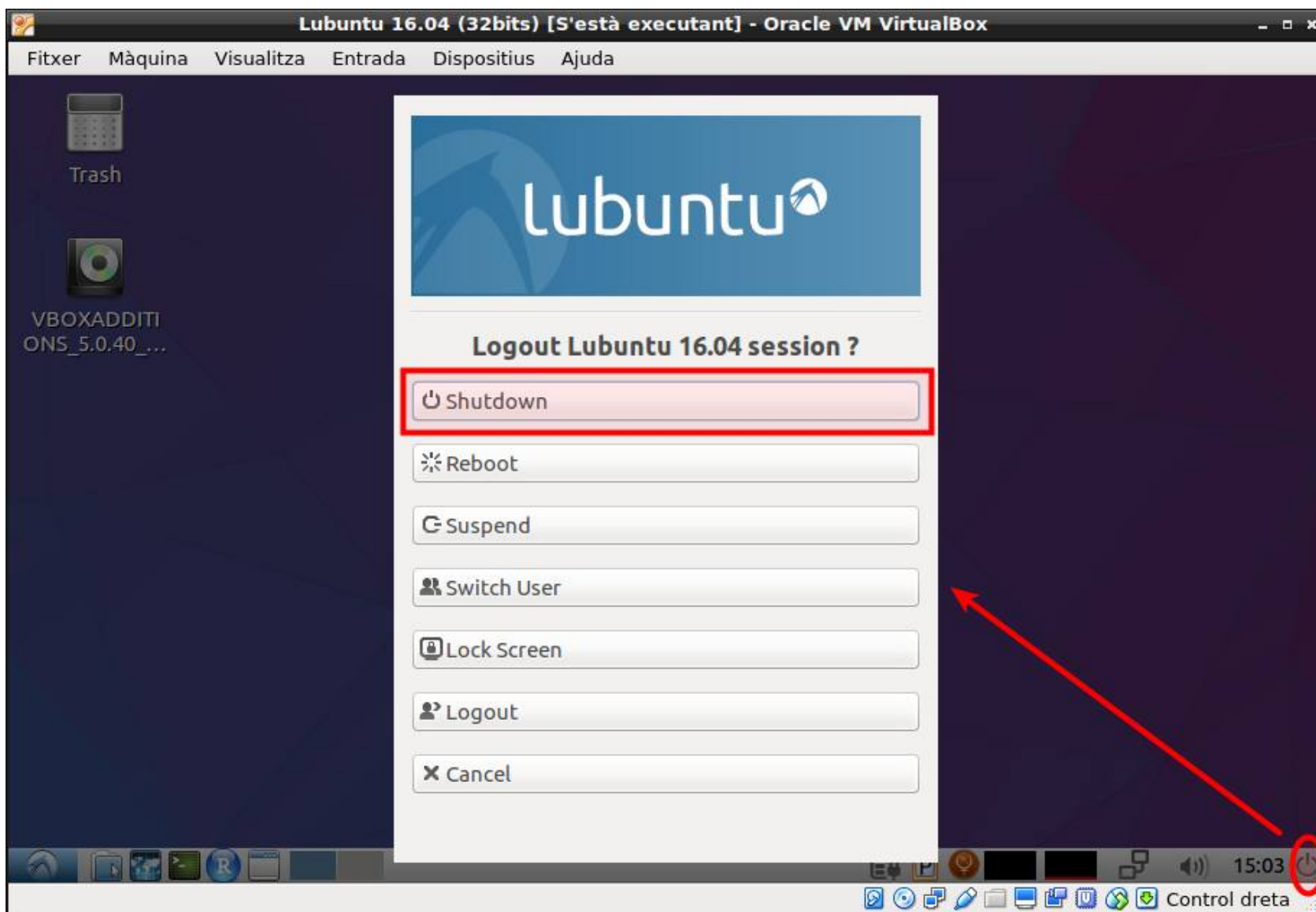
Done

Did you Export your VM (.OVA) to your USB?

Otherwise, you'll loose your changes when you reboot this computer in UB-Geologia S1 computer classroom (frozen images at work)



Shutdown Lubuntu VM within VirtualBox



More information

- **Ubuntu GNU/Linux:**
 - <http://www.ubuntu.com>
- **Data Science Virtual Machine for Linux (Ubuntu)**
 - <https://azuremarketplace.microsoft.com/en-us/marketplace/apps/microsoft-ads.linux-data-science-vm-ubuntu>
- **Linkat (Ubuntu):**
 - Manual Installation with custom partitions (advanced, or for servers):
http://linkat.xtec.cat/portal_linkat/wikilinkat/index.php/Wiki_Linkat_edu_14.04
- **Forums for help and support:**
 - Ubuntu-es Forums: <https://www.ubuntu-es.org/forum>
 - Ubuntu (Catalan LoCo Team): <http://ubuntuforums.org/forumdisplay.php?f=206>
- **Data Science Toolbox (to run locally or in the cloud with AWS):**
 - <http://datasciencetoolbox.org/>
 - <https://www.datascienceatthecommandline.com>