

**Programmazione Avanzata per il Calcolo Scientifico**

**Advanced Programming for Scientific Computing**

**Lab 01 - 13/03/2015**

**Carlo de Falco**

# Preliminaries

Stuff you'd be better off reading ASAP

- Some useful preliminary documents are available on the BeeP page for this course.
- If you did not attend the seminars, you should have a look at:
  - `unix2015.pdf` : Intro to the UNIX environment
  - `bash_intro.pdf` : “A Gentle Introduction to Bash”
  - `moduli.pdf` : using Lmod environment modules
  - `git_intro.pdf` : “An introduction to git”

# Set-up Working Environment

Remote Login to the Maths Department



Good software should be portable to different platforms  
And build with a variety of standard conforming toolchains

Well ... at least it should build and run on the platform where  
it is expected to be deployed ... that is,  
should work on **MY** machine, **NOT YOURS**

```
guglielmo.local $ ssh -CY carlo@tosca.mate.polimi.it  
Last login: Fri Mar 13 02:17:32 2015 from 93-35-0-22.ip52.fastwebnet.it
```

ssh (secure shell), sftp (secure file transfer protocol)  
remote login/file transfer programs.

ssh/sftp are installed by default on OSX and GNU/Linux, on Windows  
use, e.g., PuTTY/FileZilla

common options : -C compress data,  
-Y enable trusted X forwarding (allow using remote GUI programs)

```
guglielmo.local $ ssh -CY carlo@tosca.mate.polimi.it
```

```
Last login: Fri Mar 13 02:17:32 2015 from 93-35-0-22.ip52.fastwebnet.it
```

```
guglielmo.local $ ssh -CY carlo@tosca.mate.polimi.it  
Last login: Fri Mar 13 02:17:32 2015 from 93-35-0-22.ip52.fastwebnet.it  
tosca.mate.polimi.it > ssh -CY mactaris  
The authenticity of host 'mactaris (10.48.139.6)' can't be established.  
RSA key fingerprint is af:5b:94:9b:cc:01:e8:d2:6b:a1:fe:f2:f0:c1:7d:a9.  
Are you sure you want to continue connecting (yes/no)? yes  
Warning: Permanently added 'mactaris' (RSA) to the list of known hosts.  
Last login: Fri Mar 13 02:18:43 2015 from toska.mate.polimi.it  
mactaris.mate.polimi.it >
```



```
guglielmo.local $ ssh -CY carlo@tosca.mate.polimi.it  
Last login: Fri Mar 13 02:17:32 2015 from 93-35-0-22.ip52.fastwebnet.it  
tosca.mate.polimi.it > ssh -CY mactaris  
The authenticity of host 'mactaris (10.48.139.6)' can't be established.  
RSA key fingerprint is af:5b:94:9b:cc:01:e8:d2:6b:a1:fe:f2:f0:c1:7d:a9.  
Are you sure you want to continue connecting (yes/no)? yes  
Warning: Permanently added 'mactaris' (RSA) to the list of known hosts.  
Last login: Fri Mar 13 02:18:43 2015 from toska.mate.polimi.it  
mactaris.mate.polimi.it >
```

Do not run any jobs on 'tosca', move on to a different machine

Once you are logged in to the correct host, set up the environment

```
guglielmo.local $ ssh -CY carlo@tosca.mate.polimi.it
Last login: Fri Mar 13 02:17:32 2015 from 93-35-0-22.ip52.fastwebnet.it
tosca.mate.polimi.it > ssh -CY mactaris
The authenticity of host 'mactaris (10.48.139.6)' can't be established.
RSA key fingerprint is af:5b:94:9b:cc:01:e8:d2:6b:a1:fe:f2:f0:c1:7d:a9.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'mactaris' (RSA) to the list of known hosts.
Last login: Fri Mar 13 02:18:43 2015 from toska.mate.polimi.it
mactaris.mate.polimi.it >
mactaris.mate.polimi.it > export mkPrefix=/u/geo2/sw/
mactaris.mate.polimi.it > source /u/geo2/sw/System/mkConfig.sh
carlo@mactaris ~ $
```

Do not run any jobs on 'tosca', move on to a different machine

Once you are logged in to the correct host, set up the environment

```
carlo@mactaris ~ $ emacs -nw ~/.bashrc
carlo@mactaris ~ $ vi ~/.bashrc
carlo@mactaris ~ $ gedit ~/.bashrc
```

You can write commands you run often in you ~/.bashrc file  
Use a 'Text Editor' to edit the file

- emacs / vi can run in their own GUI or in text mode in the terminal emulator (good for slow connections)
- both emacs and vi are extremely powerfull but have a steep learning curve, if you are not comfortable with them use gedit (GUI mode only) or nano (text mode only)

```
carlo@mactaris ~ $ ssh idra
```

To run parallel jobs connect to the cluster frontend

```
Politecnico di Milano

Clustered computational resources

Active queues :
  name      # nodes  # max nodes p/job  # max time p/job  # max job
  gigat      5        2          24             1
  idra       16       16          24             2
  idralong    8        4           -             1
  cerbero    4        1           -             2

HW resources
  gigat : 160 cores / 5 nodes Xeon E4610-v2 / 1.2 TB RAM

  idra : 128 cores / 16 nodes Xeon 5560 / 426 GB RAM

  cerbero: 24 cores / 4 nodes i7 / 64 GB RAM

idra.mate.polimi.it >
```

DO NOT try running jobs on the frontend, use torque (qsub) to submit jobs to the scheduler, more details on this later on in the course

Various “queues”, the one usable for this course is “cerbero”

## Environment modules

```
carlo@mactaris ~ $ module avail
Rebuilding cache, please wait ... (not written to file) done
----- /u/geo2/sw/Modules/compilers -----
clang+system/3.4    gcc+system/4.6    gcc+system/4.8    gcc+system/4.9 (D)
Where:
(D):  Default Module

carlo@mactaris ~ $ module load gcc+system
```

A toolchain is a collection of tools used for the development of software.

A toolchain consists of:

- C/C++ compiler transforms source code in object code
- linker combines object codes in a single file
- C/C++ standard library provides a interface to the system
- loader resolves the undefined symbols at runtime

```
carlo@mactaris ~ $ module avail
----- /u/geo2/sw/Toolchains/gcc+system/4.9/Modules/langs -----
R/3.0    octave/3.6    octave/3.8 (D)
----- /u/geo2/sw/Toolchains/gcc+system/4.9/Modules/libs -----
arpack/96      fftw/3.3      netcdf/4.3      scotch/5.1
blacs/1        gdal/1.10     netcdf-mpi/4.3  scotch-mpi/5.1
blas/9999     getfem/4.3    numpy/1.8       suitesparse/4.2
```

Once you select a toolchain, more modules will become available all built with that toolchain

Loading a module will automatically bring in its dependencies

```
carlo@mox33 ~ $ module load mumps  
carlo@mox33 ~ $ module list
```

Currently Loaded Modules:

1) gcc+system/4.9    2) openblas/9999    3) blas/9999    4) scotch/5.1    5) mumps/4.10

Loading a module will set up environment variables describing where the headers and object files for a library are located

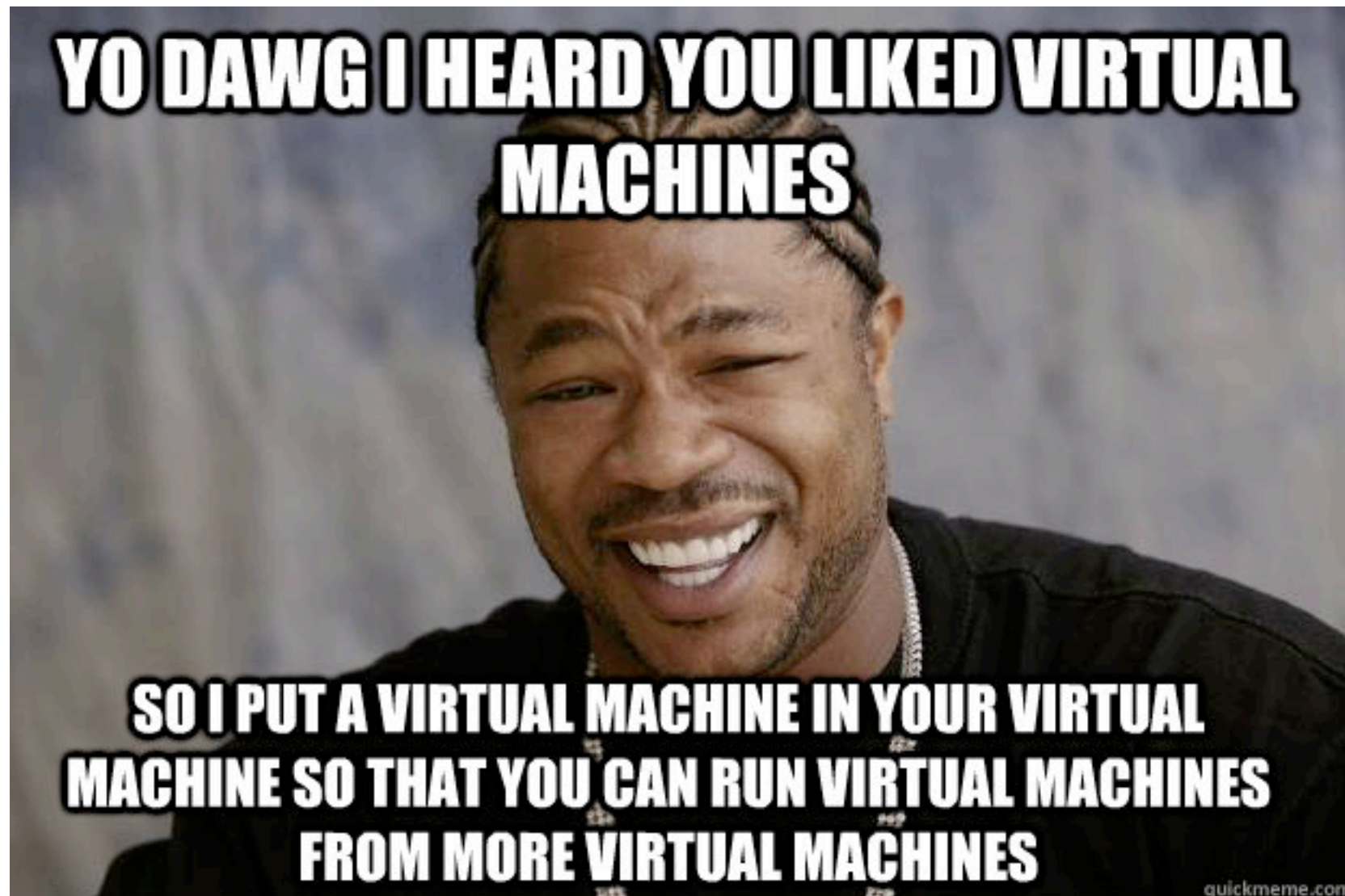
```
carlo@mox33 ~ $ printenv mkMumpsInc  
/u/geo2/sw/Toolchains/gcc+system/4.9/Packages/libs/mumps/4.10/include  
carlo@mox33 ~ $ printenv mkMumpsLib  
/u/geo2/sw/Toolchains/gcc+system/4.9/Packages/libs/mumps/4.10/lib  
carlo@mox33 ~ $ printenv mkMumpsHome  
/u/geo2/sw/Toolchains/gcc+system/4.9/Packages/libs/mumps/4.10
```

Such variables come handy when compiling and linking a program (or writing a Makefile)

```
carlo@mox33 ~ $ module load boost  
carlo@mox33 ~ $ g++ -I${mkBoostInc} -c hydrogen.cpp  
carlo@mox33 ~ $ g++ -I${mkBoostInc} -c main.cpp
```

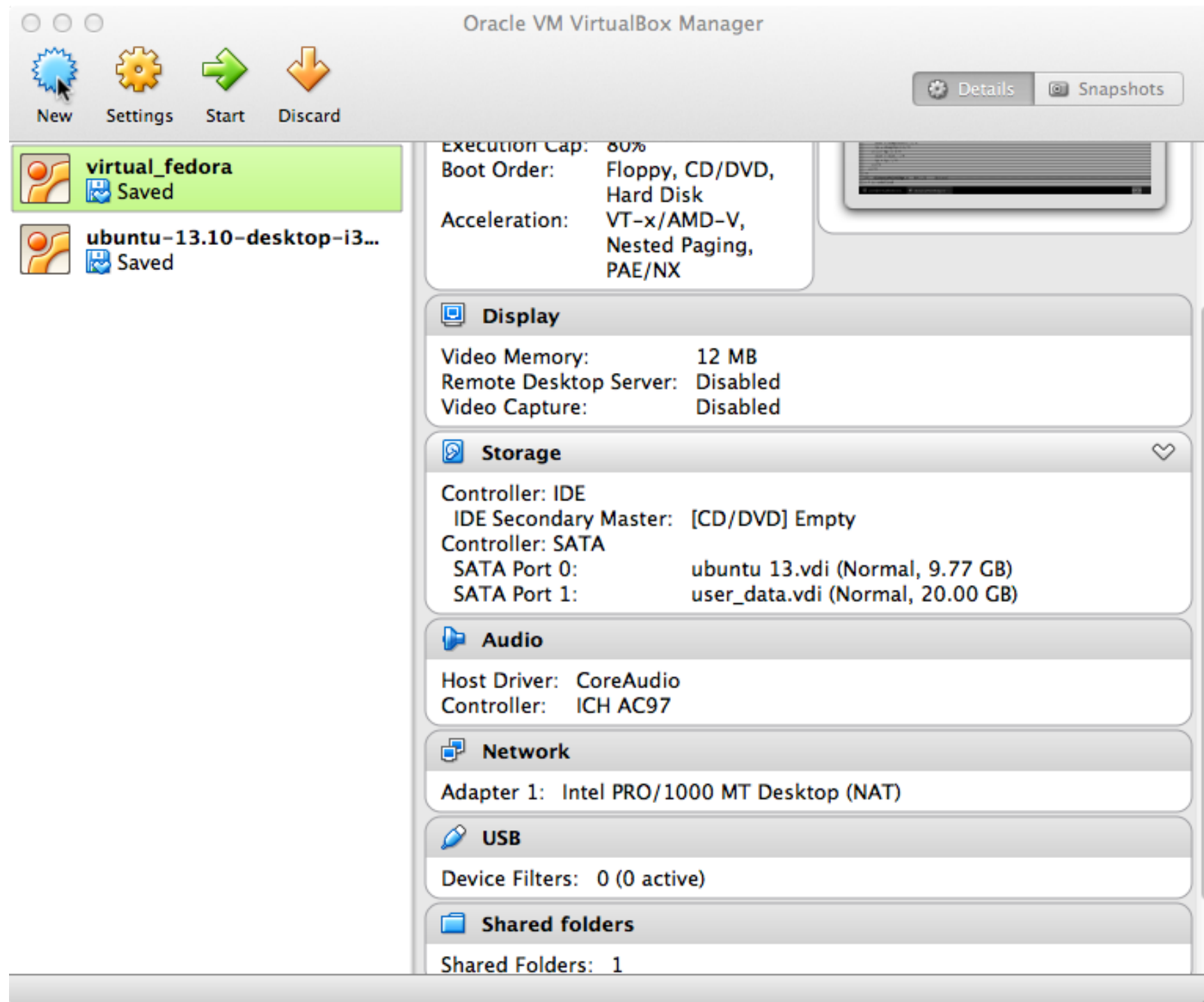
# Set-up Working Environment

Virtual Machine

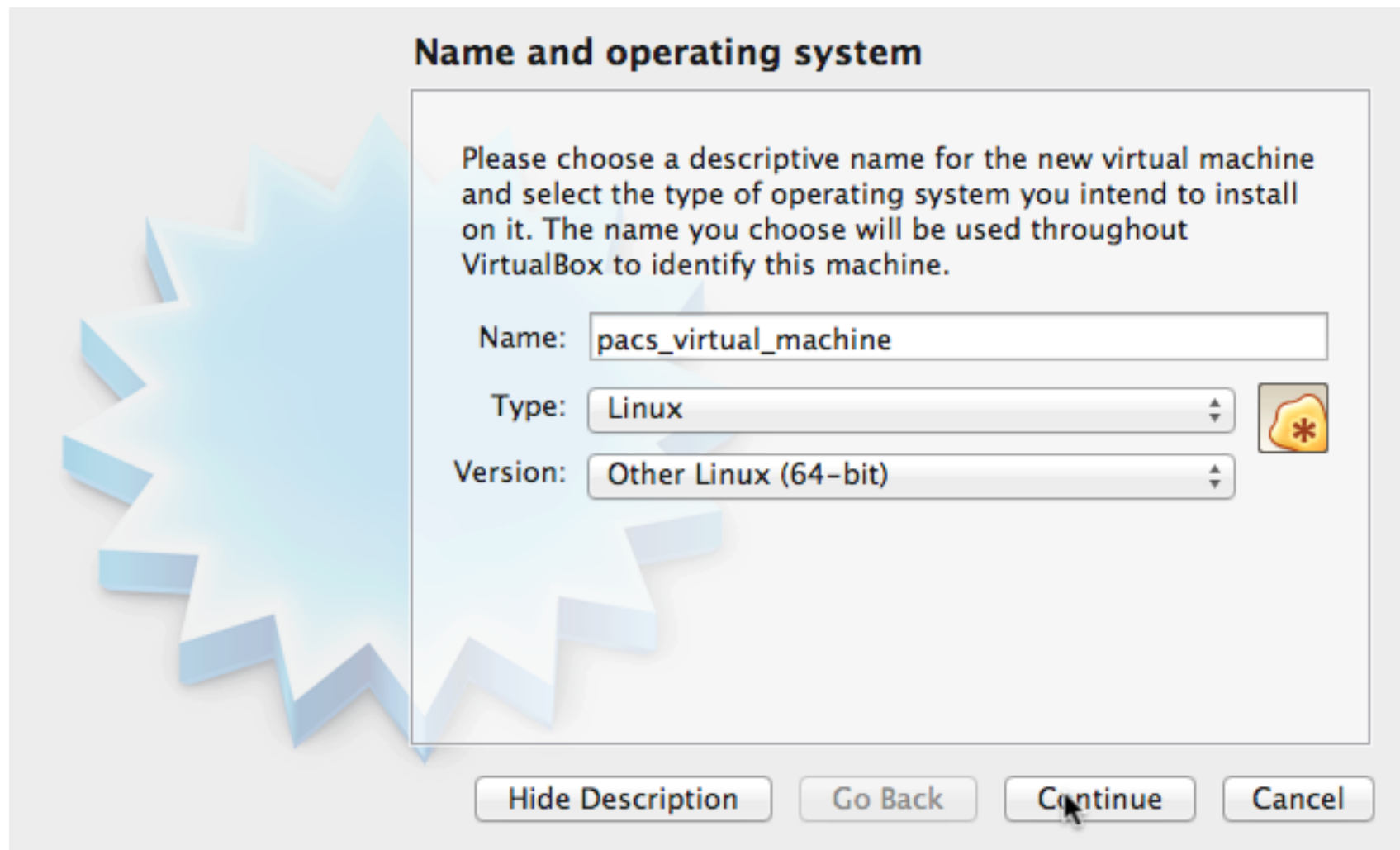


enjoy!



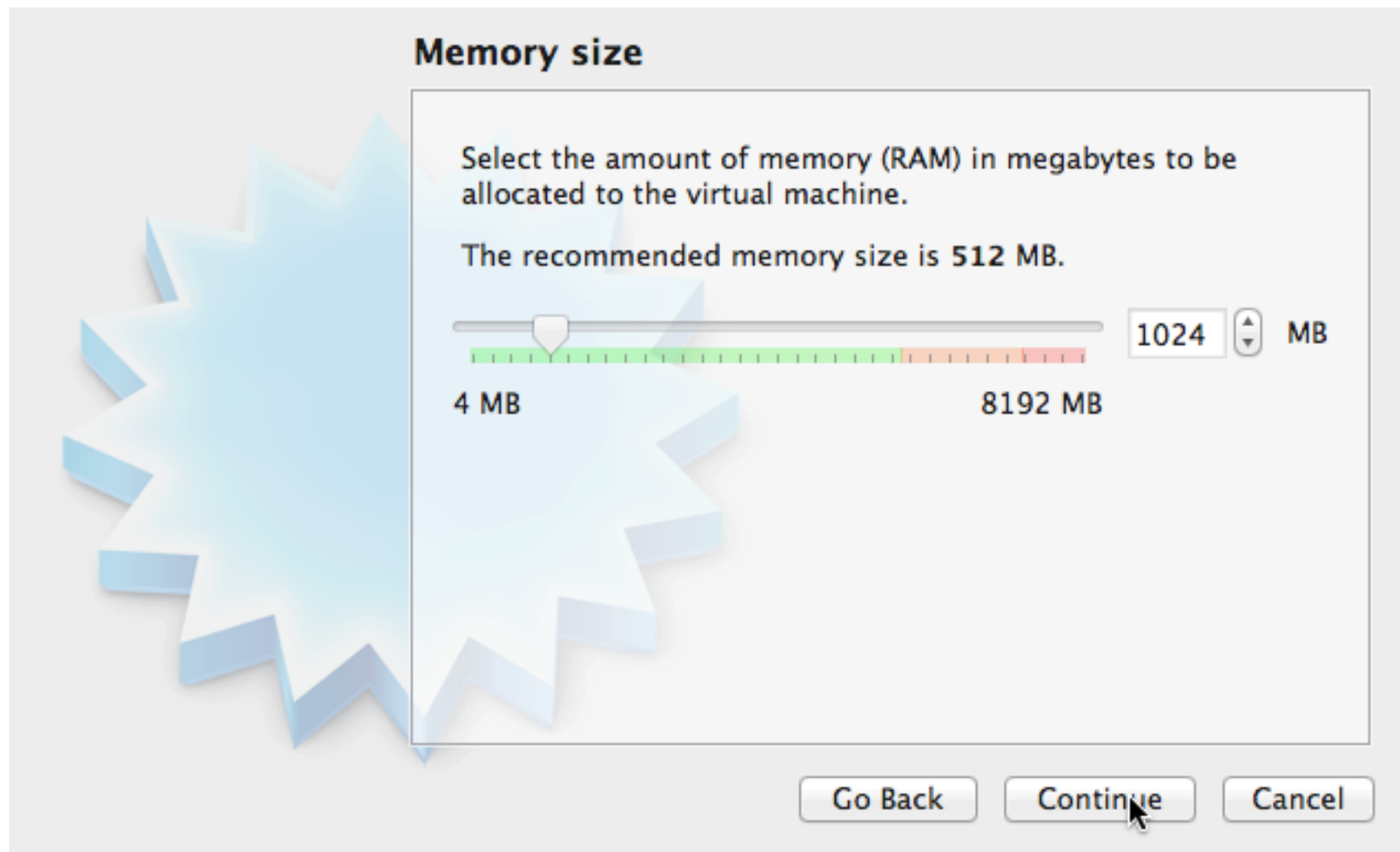


Create a new virtual machine

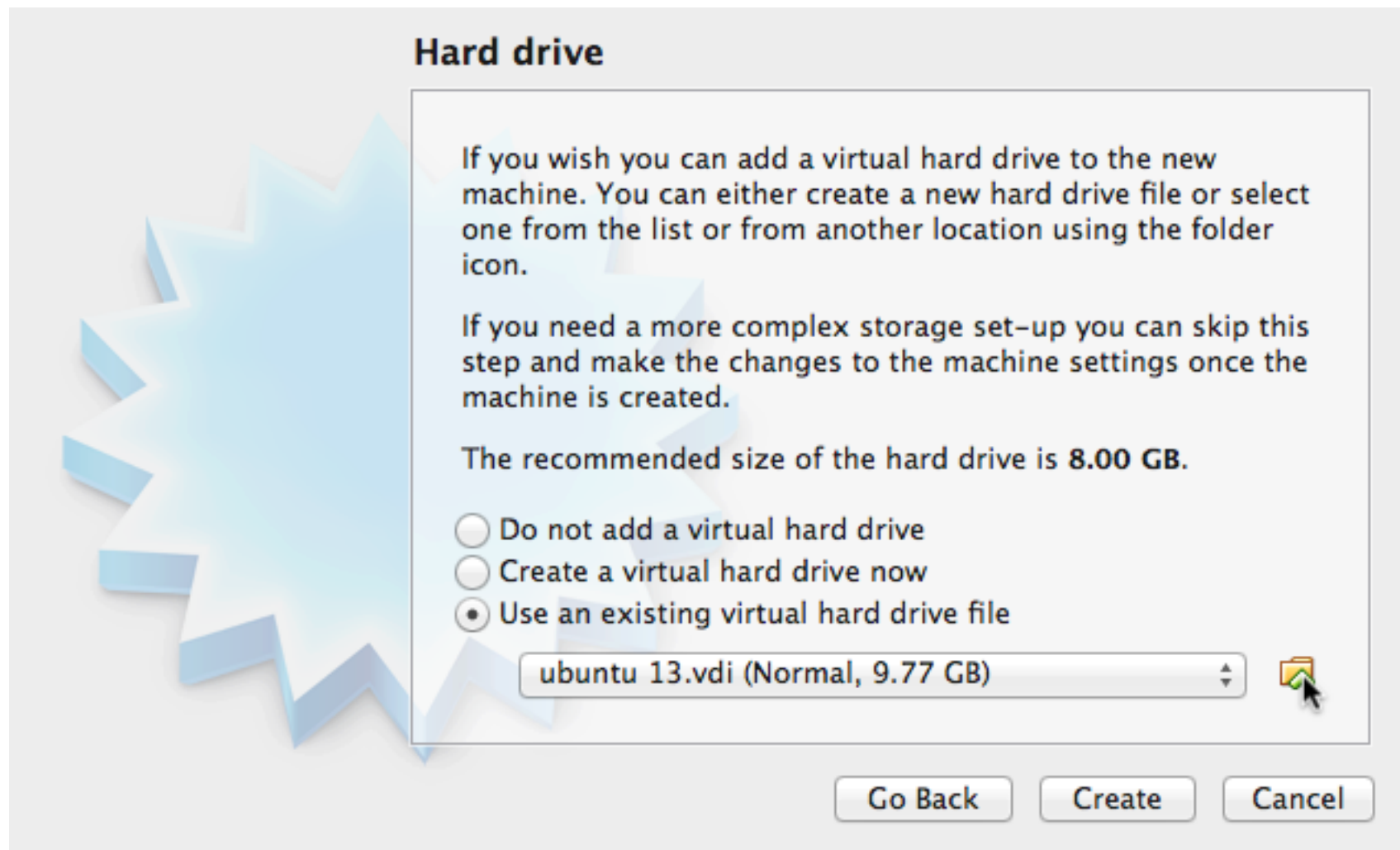


## Choose Operating System

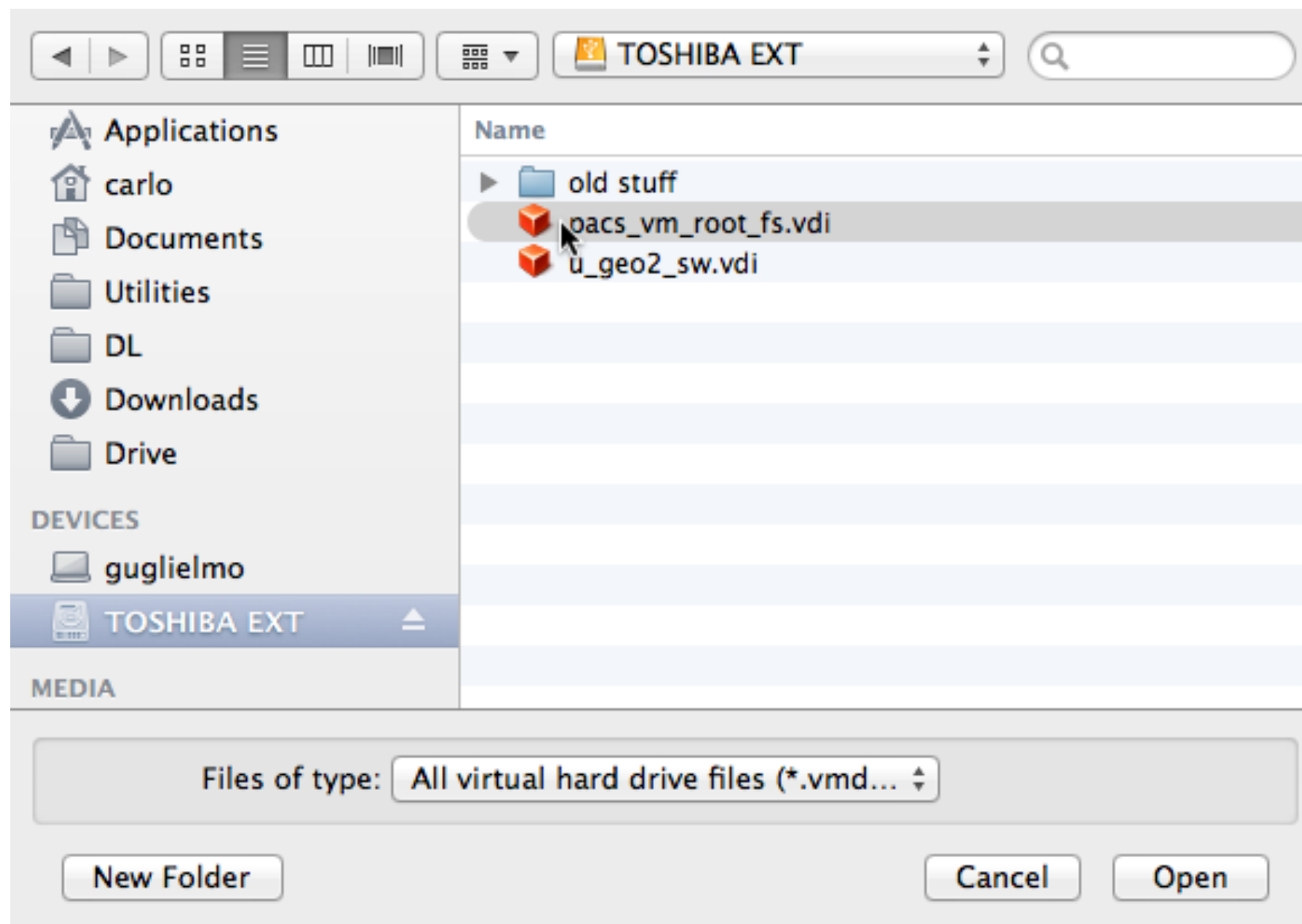
**TIP :** If you don't see 64-bit options, check whether virtualization is enabled in your BIOS



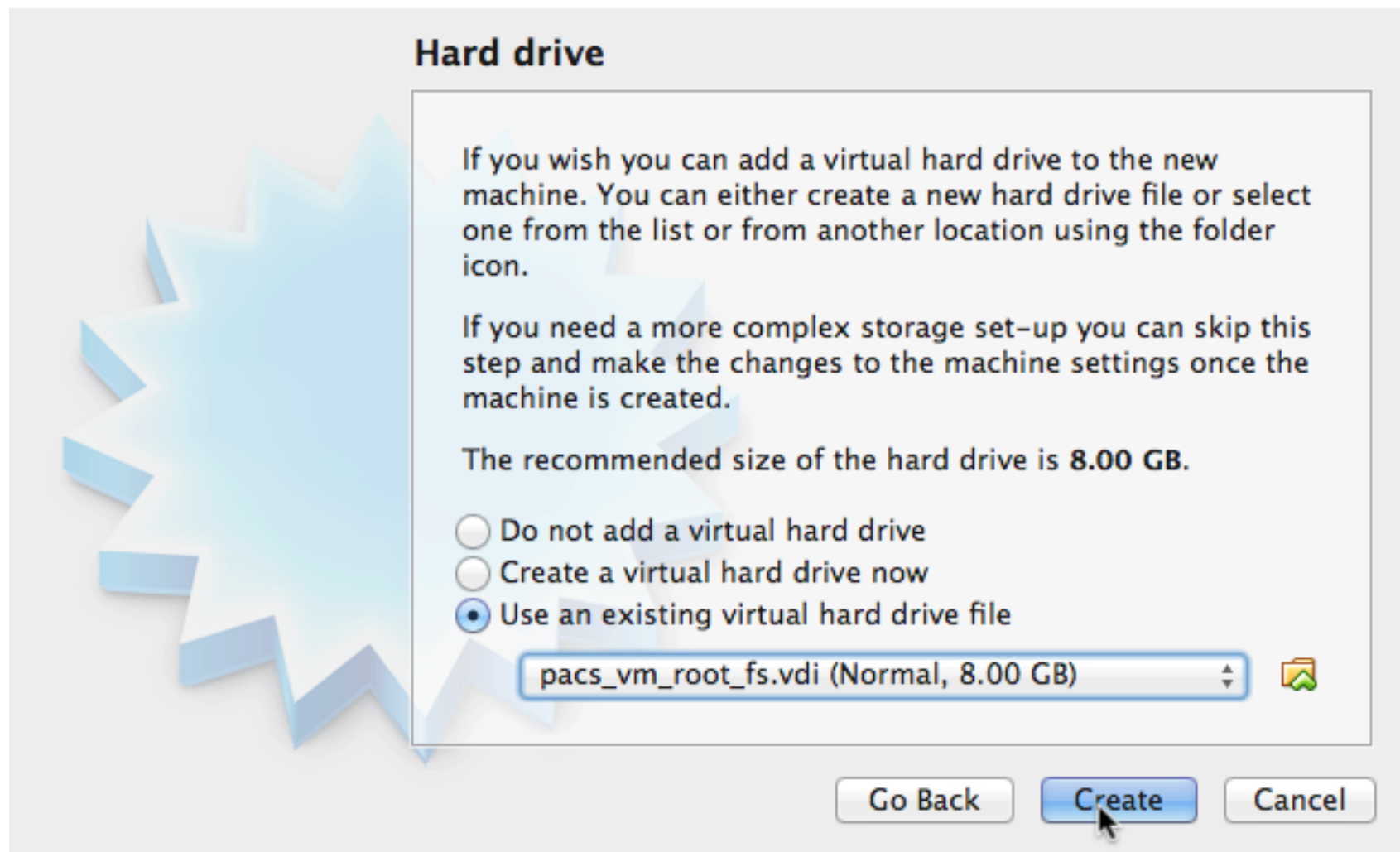
**At least 1024MB RAM Recommended**



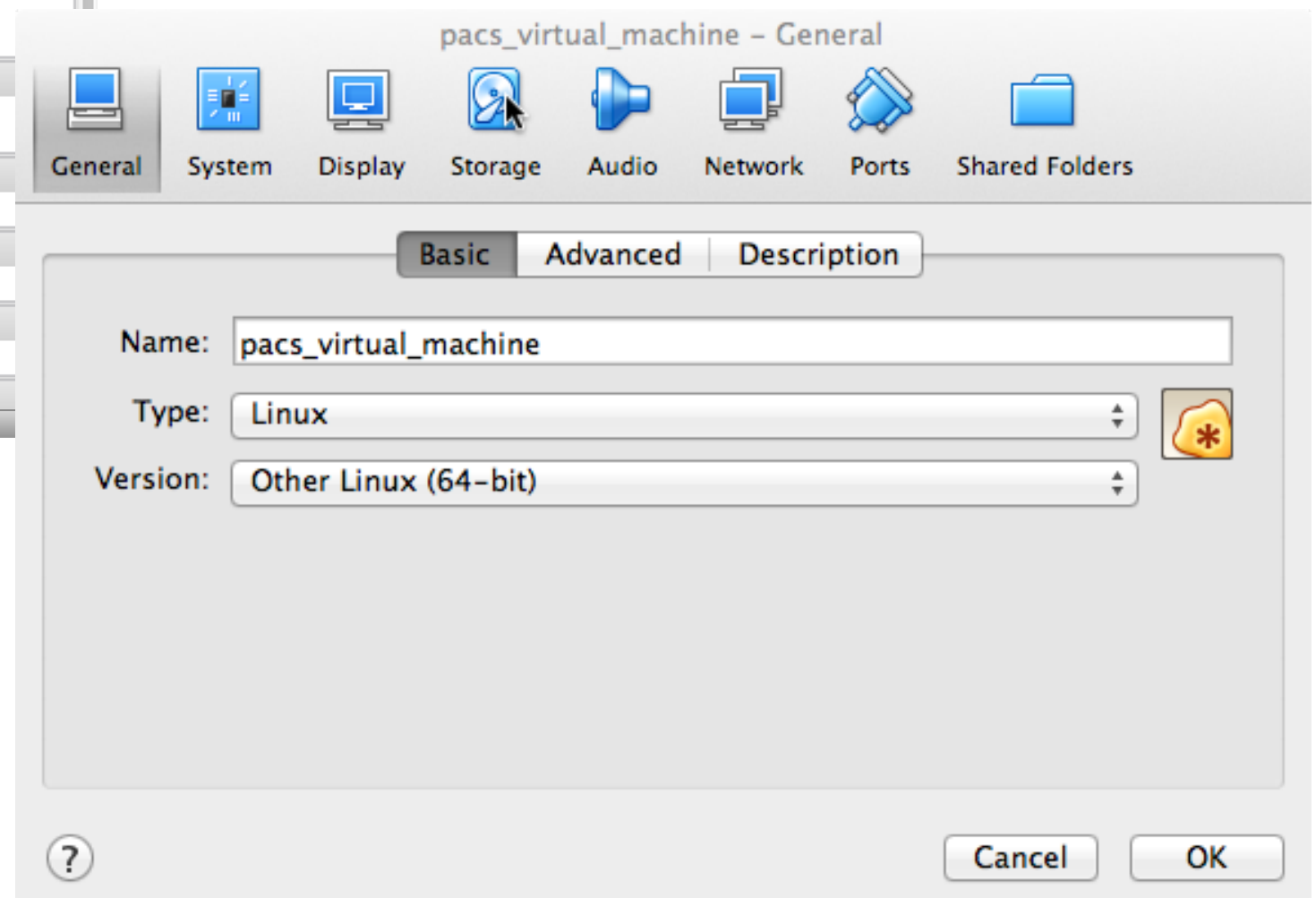
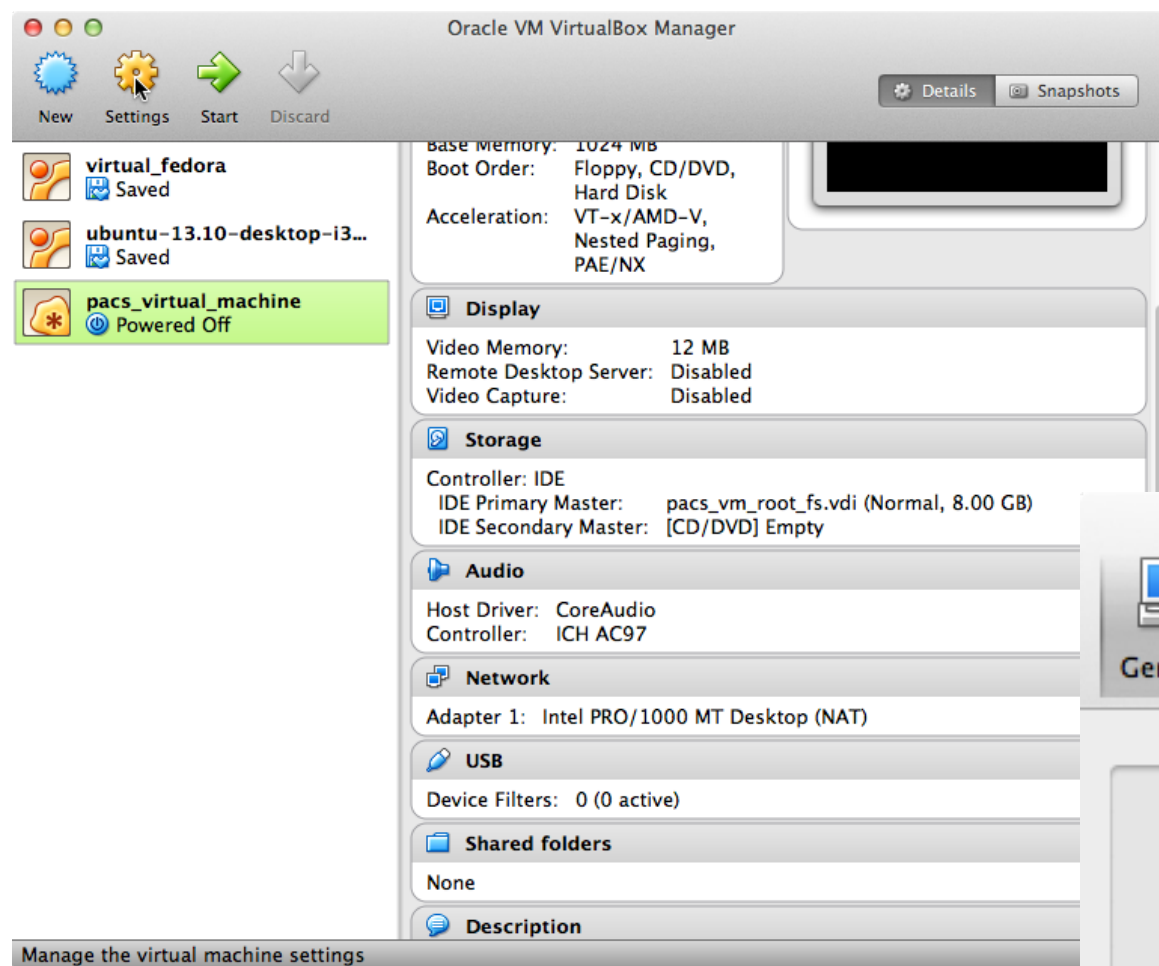
Use an existing Hard Drive



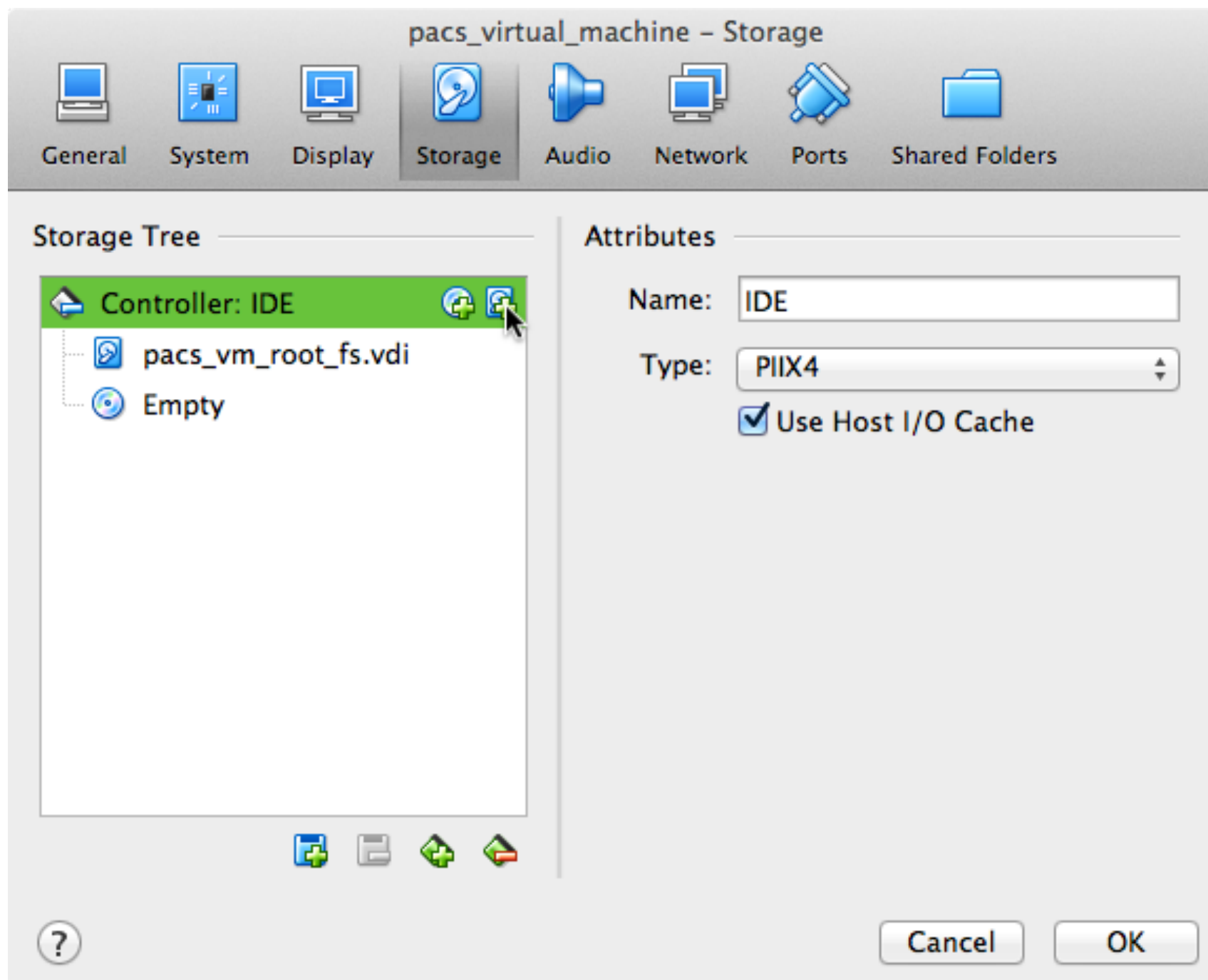
Choose 'pacs\_vm\_root\_fs.vdi'  
as main Hard Drive image



Create the virtual machine



Change virtual machine settings

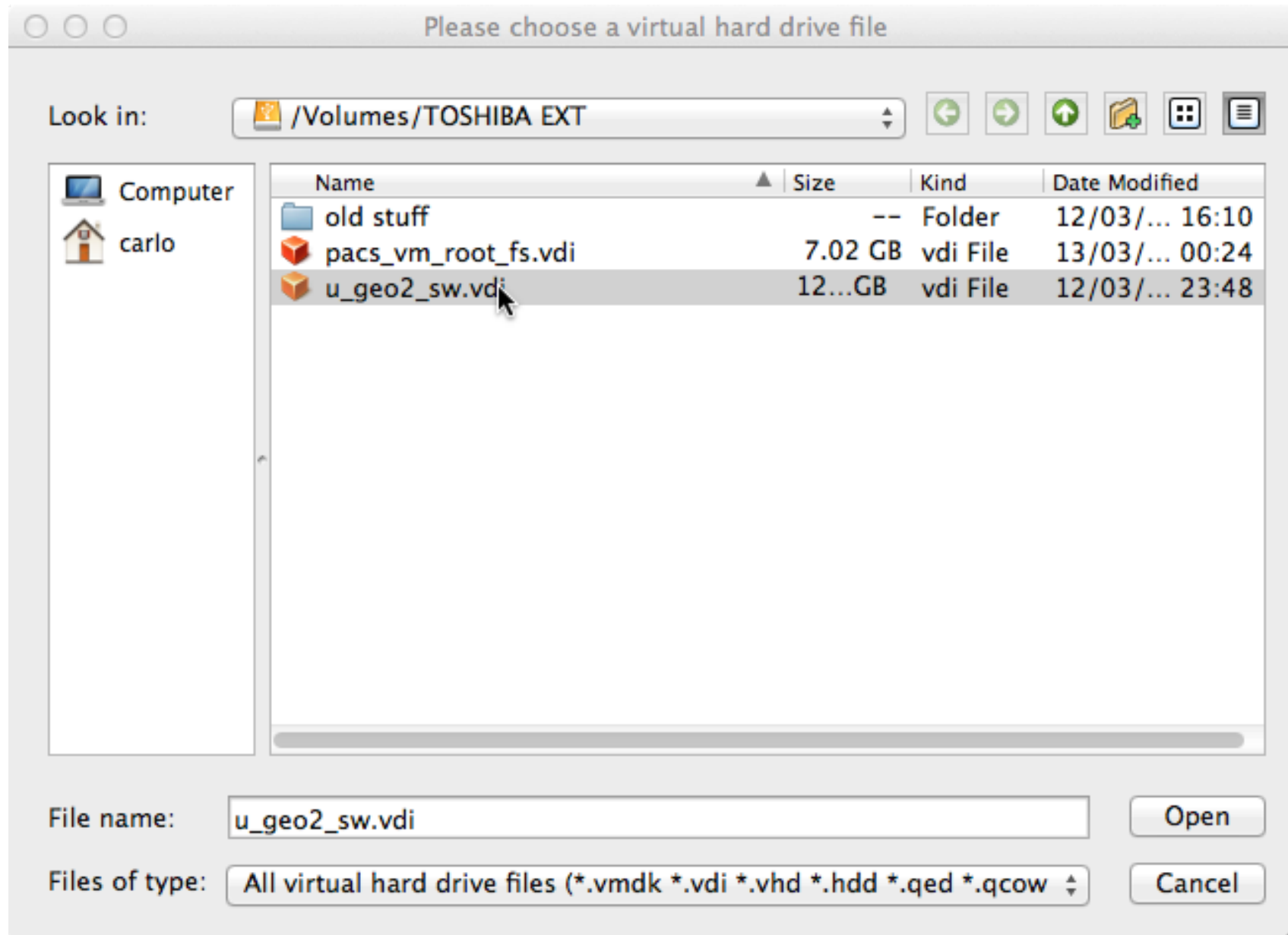


Add a drive to the IDE controller

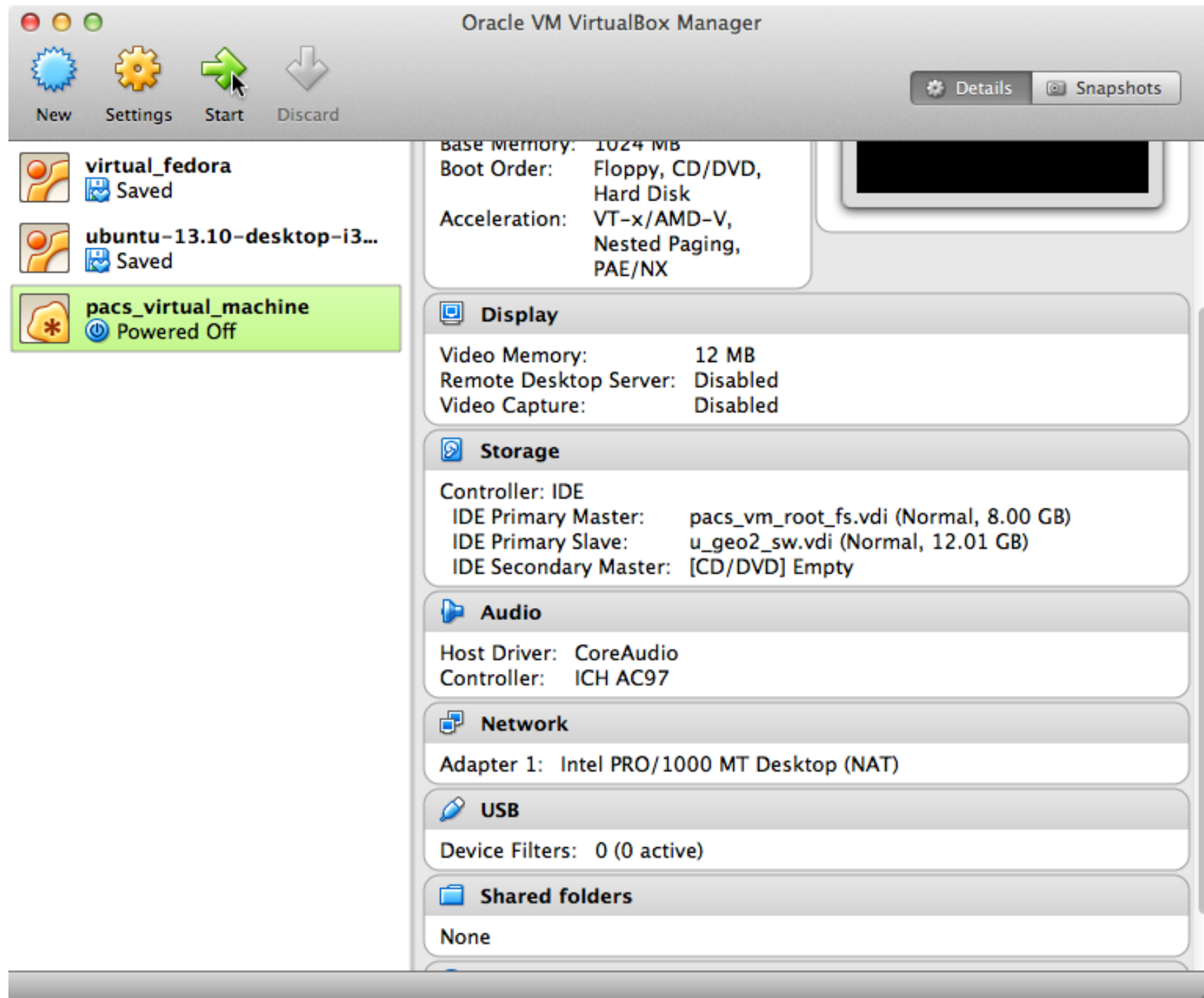




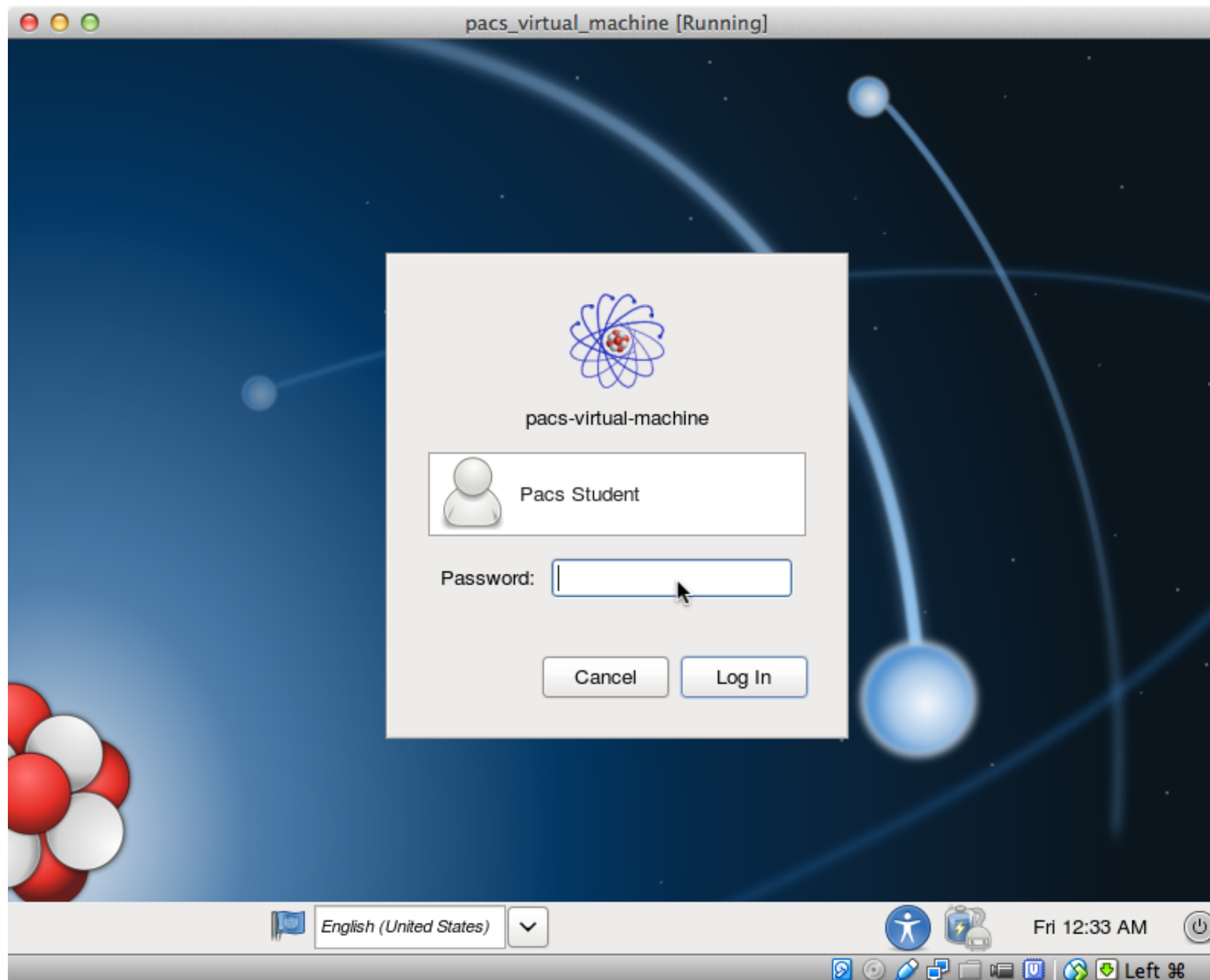
Use existing disk image



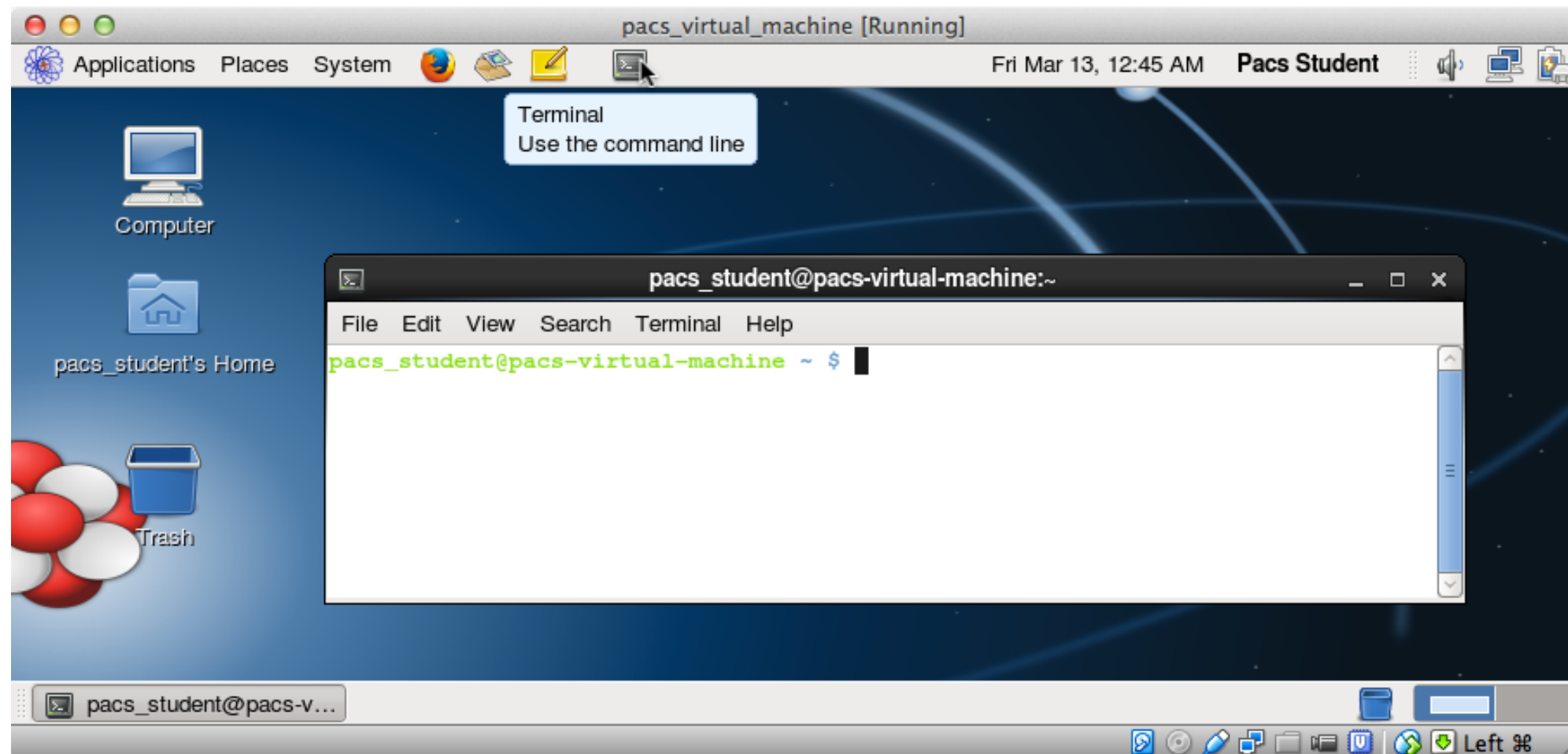
select 'u\_geo2\_sw.vdi' as secondary drive



start the virtual machine



user name: pacs\_student  
user password: pacs\_2015  
superuser password: pacs\_2015



start terminal emulator

# Collaboration Tools

(Distributed) Revision Control Systems

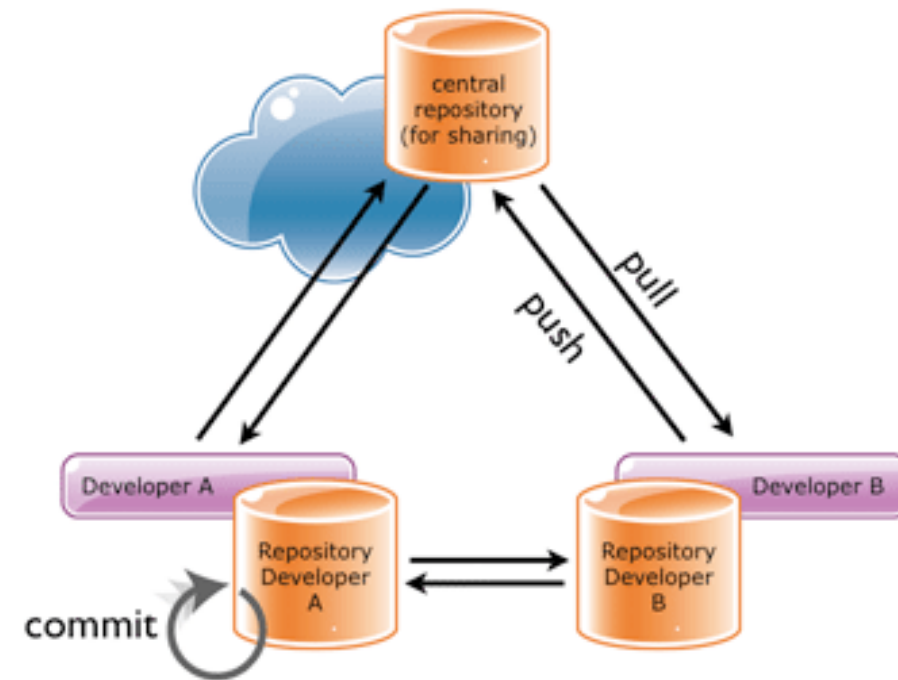
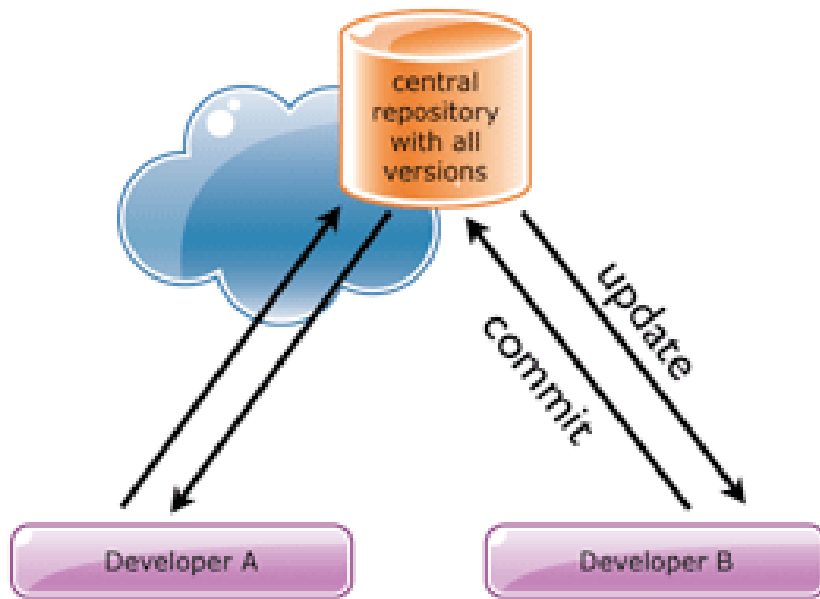
Software needs to be **maintained, debugged, improved...**

while keeping track of the changes and being able to recover old revisions if needed.

Those task can be done automatically using a **Revision Control System**.

In addition, one may want to keep a copy of the code (and of all its changes) in a remote repository, for safety and for sharing with others.

Most common RCS: CVS, RCS, git, mercurial (Hg)....



Centralised RCS compares with a distributed RCS as an **hub** compares with a **peer-to-peer** model

Each user has a local copy of the repository: you have full control of all the history even if you are not connected to the web

- Most commands operate on the local copy (minimize communications).
- You may have multiple remote repositories: better distributed development

see “an introduction to git” by A. Cervone and L. Formaggia in Beep for more



# Collaboration Tools

On-line hosting and collaborative development tools

GitHub is a web-based Git repository hosting service,  
We will use github for the examples of the PACS course.  
<https://github.com/pacs-course/pacs>

Using GitHub or another similar hosting service (BitBucket, Launchpad, Savannah, ~~Google Code~~) for your exam project is encouraged

Get your own account on GitHub in order to participate in “challenges” and homework, choose an account name containing your real name or your student ID number to make yourself recognizable.

GitHub tutorials are available here:  
<https://guides.github.com>

Using Git through GitHub usually requires a different “workflow”  
<https://guides.github.com/introduction/flow/>

We will use “pull requests” to manage answers to homework questions