# Operating Systems, Virtual Machine and Cloud computing

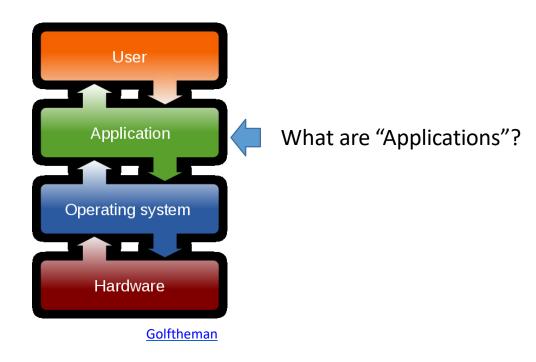
Dumrong Mairiang, PhD

SIRE507: FUNDAMENTAL COMPUTER SCIENCE FOR BIOLOGIST

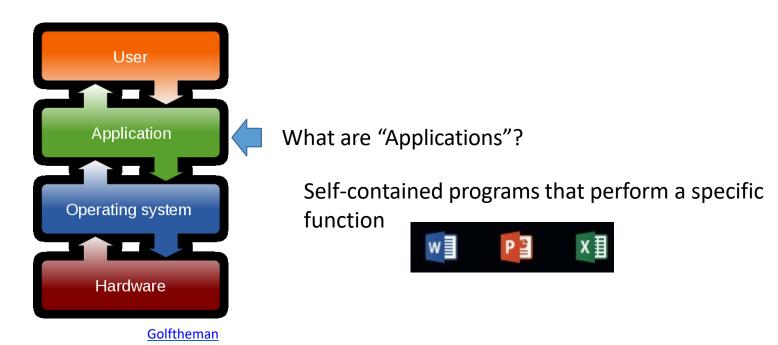
• What is it?

Why is it important?

- What is it?
  - The basic software that manages a computer
- Why is it important?



- What is it?
  - The basic software that manages a computer
- Why is it important?



Do you know any OS?

Do you know any OS?



Do you know any OS?



Mobile OS



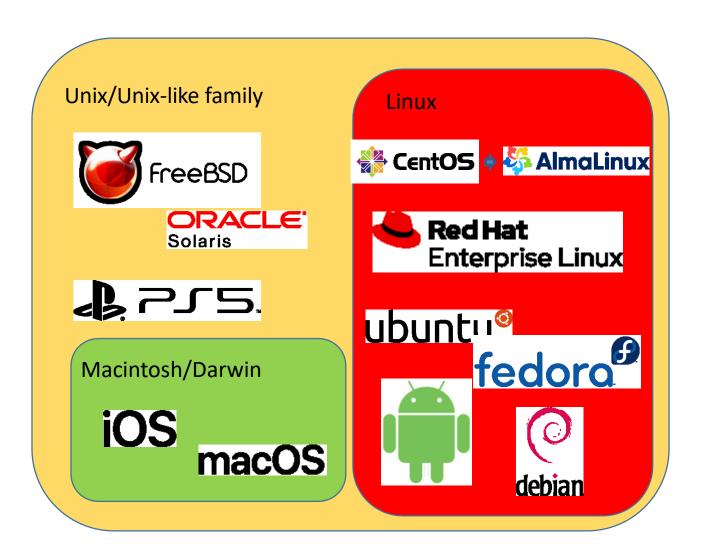


Do you know any OS?



Other





## Why do you need to know about Oses other than MS Windows or MacOS?

- Many of bioinformatic applications were developed for Unix-like or Linux OS
- If you want to develop a web application, web servers usually run with Linux
- Many free and open-source applications must be run on Unix-like or Linux OS

- Old machine, legacy application and IoT:
  - 32-bit vs 64-bit (32-bit apps on Windows?)
  - Lite version/distribution of OS (e.g. Alpine, Lubuntu)
  - AMD/intel vs ARM/Apple silicon
- Users of your application
  - Windows
    - General users with probably no or little bioinformatic background
    - Graphical user interface is likely to be expected
  - Unix-like/Linux
    - Bioinformaticians
    - Command lines are acceptable

- OS-specific file extensions
  - Microsoft Windows
    - File.exe
    - File.msi
  - Mac OS
    - File.dmg
  - Linux (Debian/Ubuntu)
    - File.deb

- File naming rules
  - Microsoft Windows
    - Reserved characters < > : " / \ | ? \*
  - Mac OS
    - Reserved characters / :
  - Linux
    - Reserved characters /
  - Mac OS & Linux
    - Non-reserved characters but can be problematic for a shell command

```
< > : | ? *
```

- As bioinformatician: Please AVOID using "space" in the file name
  - FileName.txt (Camel case)
  - File\_Name.txt (Underscore case)

- File paths to YourFile.txt in your "home" directory
  - Microsoft Windows
    - C:\Users\YourName\YourFile.txt
  - Mac OS
    - /Users/YourName/YourFile.txt
  - Linux
    - /home/YourName/YourFile.txt
- Linux (root access)
  - /root/YourFile.txt

- Application for accessing command line
  - Microsoft Windows
    - Command Prompt
  - Mac OS
    - Terminal
  - Linux
    - Terminal

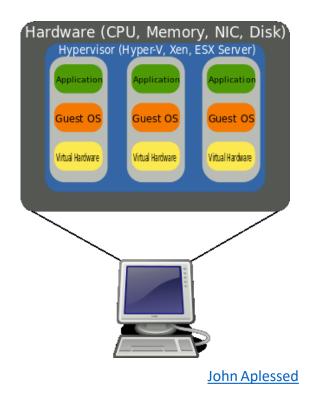
- Application for Back-up
  - Microsoft Windows
    - System Restore and Restore Point
  - Mac OS
    - Time Machine
  - Linux
    - Ubuntu backups
    - Backups application (e.g. cronopete)

- Application for Hardware Management/List
  - Microsoft Windows
    - Device Manager
  - Mac OS
    - System Reports ("About this Mac")
  - Linux (Ubuntu)
    - hardinfo, Ispci, Isusb

- Checking or interrupting programs
  - Microsoft Windows
    - Task Manager
  - Mac OS
    - top (basic) and htop (to be installed)
  - Linux (Ubuntu)
    - top (basic) and htop (to be installed)

## Virtual Machine (VM)

- What is it?
  - A complete environment for a guest operating system to function as though that operating system were installed on its own computer



## Virtual Machine (VM)

- What VM is not:
  - Virtual machine ≠ Emulator
  - Emulator converts commands to and from a host machine to an entirely different platform
  - Emulator: DosBOX (CPU), PuTTY (Terminal), ZSNES (Gaming), PCSX2 (Gaming), N64 Emulator (Gaming), Android Virtual Device (Application Development)

## Why do you need to know about VMs?

- SENARIO 1 (Flexibility): The OS of your machine is not compatible with the application you need to run/test:
  - No spare machine for installing a new OS
  - No space or resources to create a dual boot

## Why do you need to know about VMs?

- SENARIO 2 (Simulation): You want to simulate network connection to your web application in your own machine
  - No spare machine
  - Do not want to deploy in web server yet

## Why do you need to know about VMs?

- SENARIO 3 (Security): Controlled/Quarantined environment for developing, testing or running applications
  - Potentially harmful applications
  - Secured VM

#### Basic information about VM

- Applications for virtualization:
  - VM ware (Commercial)
  - Oracle VM VirtualBox (Free and open-source)
  - Vagrant (Free and open-source, but no GUI)
  - Other...

## Basic information about VM

- Check your machine
  - CPU: at least 2 cores
  - RAM: Host OS minimal requirement + Guest OS minimal requirement
  - Hard disk: Guest OS minimal requirement or External HDD or USB Flash drive

#### Portable VM

- You can save file.vdi to external HDD or USB flash drive
- Connect to any host machine with Virtualbox to quickly create a copy of VM in file.vdi

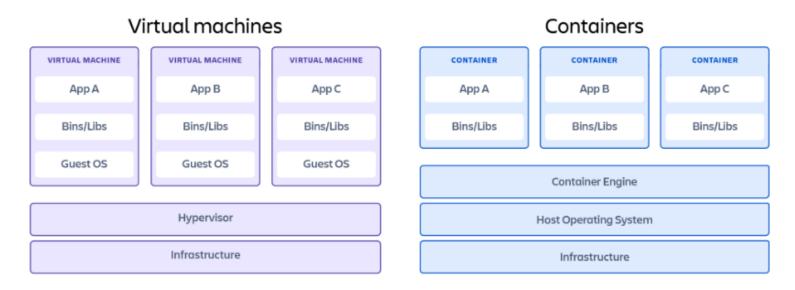
#### • NOTE:

- USB flash drive must be formatted as NTFS
- FAT32 format (default format in some USB flash drives) cannot accommodate a single file larger than 4GB

#### VM versus Container



- VM requires resources (e.g. RAM, CPU, HDD space)
   allocation while a container shares resources with the host.
- VM virtualizes an entire machine while a container virtualizes an "environment" for specific software.



#### VM versus Container





- Container use cases:
  - Run old software in new machine/OS
    - Software needs to run on 32-bit OS with Python 2
  - Deployment with reproducibility
    - Run your software developed on your machine on HPC
  - Microservices
    - 1. Container for database
    - 2. Container for web application
    - 3. Container for backup

#### VM versus Container

- Docker (https://docs.docker.com/)
  - Most popular container (i.e. good documentation)



- Singularity (https://docs.sylabs.io/guides/latest/user-guide/)
  - "Portable" container (Singularity Image Format, SIF)
  - Other containers may be converted to Singularity container (especially Docker)
  - Run on HPC with job scheduler (e.g. Slurm)



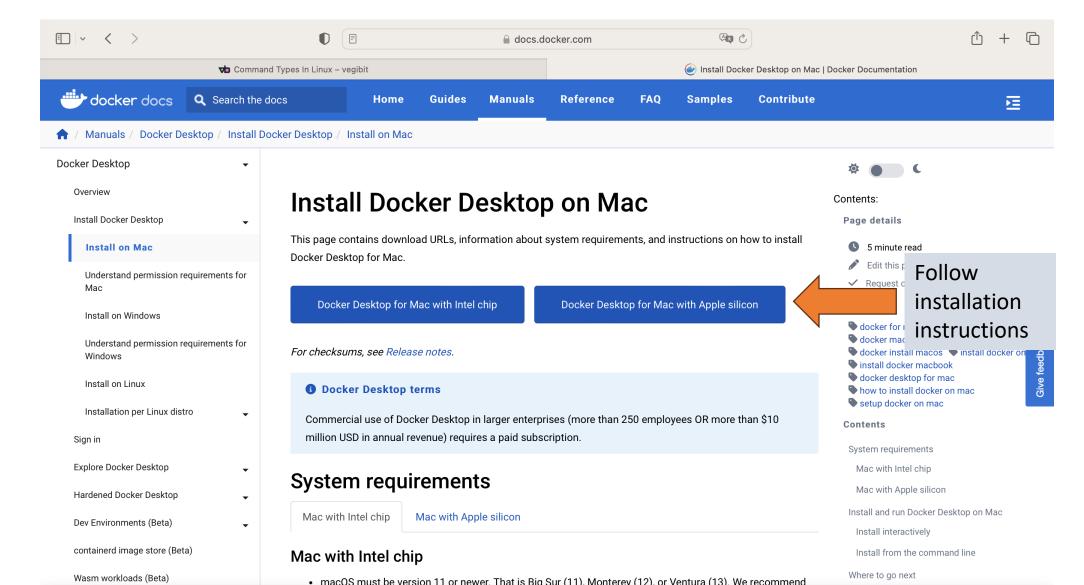


## Running Docker on Windows

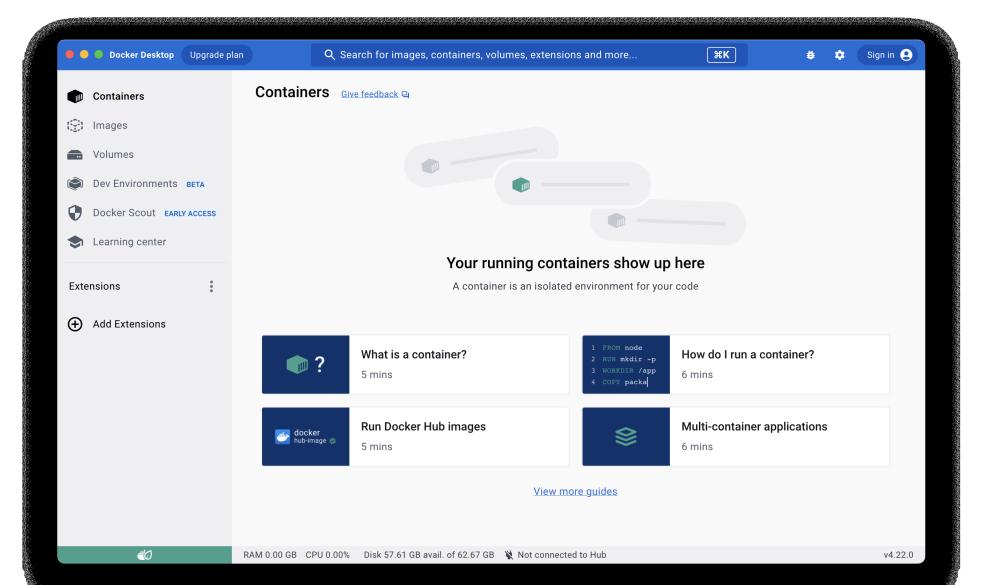
https://docs.docker.com/desktop/install/windows-install/



## Install Docker Desktop for MacOS



## Launch Docker Desktop



#### Run Docker in Terminal

```
kwan — root@855edc88cc25: / — -zsh — 80×24
[(base) kwan@MBPkhxnwanrutai ~ % docker run -it --rm ubuntu:latest /bin/bash
root@855edc88cc25:/# cat etc/os-release
PRETTY_NAME="Ubuntu 22.04.3 LTS"
NAME="Ubuntu"
VERSION_ID="22.04"
VERSION="22.04.3 LTS (Jammy Jellyfish)"
VERSION_CODENAME=jammy
ID=ubuntu
ID_LIKE=debian
HOME_URL="https://www.ubuntu.com/"
SUPPORT_URL="https://help.ubuntu.com/"
BUG_REPORT_URL="https://bugs.launchpad.net/ubuntu/"
PRIVACY_POLICY_URL="https://www.ubuntu.com/legal/terms-and-policies/privacy-poli
су"
UBUNTU_CODENAME=jammy
[root@855edc88cc25:/# exit
exit
(base) kwan@MBPkhxnwanrutai ~ %
```

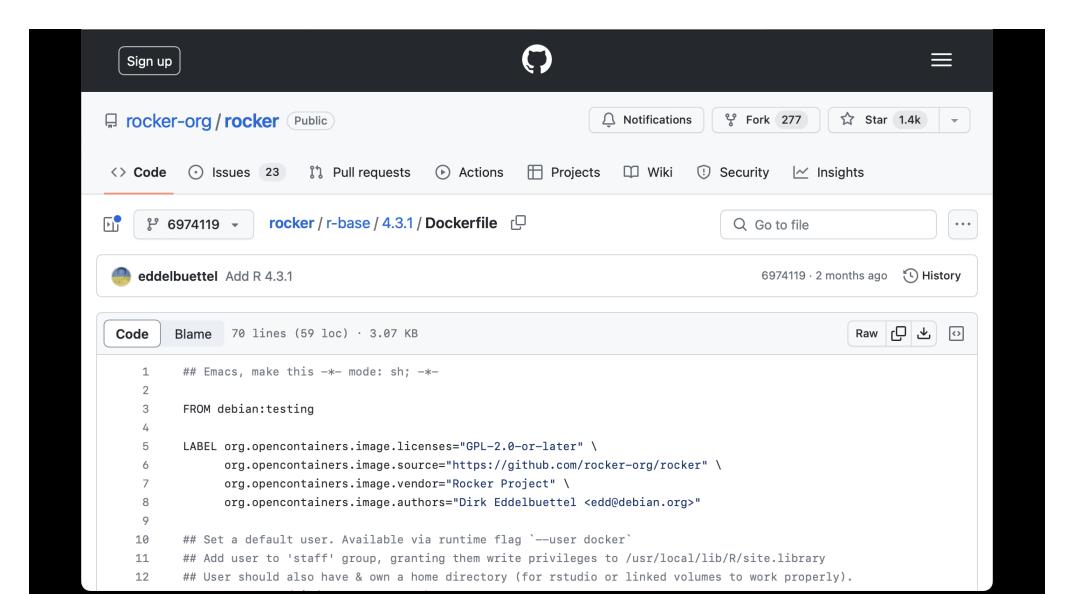
- docker run -it --rm ubuntu:latest /bin/bash
  - docker run → Create and run a new container from an image
  - -it → Interactive session
  - --rm → Automatically remove the container when it exits
  - ubuntu:latest → Image name and tag (i.e. version)
  - /bin/bash → Command to run within the container (/bin/bash = command prompt. Could be omitted.)

- docker run -it --rm -v /my/local/path:/data -w
  /data ubuntu:latest /bin/bash
  - -v /my/local/path:/data → Mount/my/local/path to /data in the container
  - $-w \rightarrow$  Set /data of the container as the working directory

- docker ps → Check all running instances
- docker ps -a → Check all running and exited instances
- docker rm [CONTAINER ID] → Remove an instance
- docker images → Check locally available images
- docker rmi [IMAGE ID] → Remove an image

- docker run -it -v /my/local/path:/data -w /data ubuntu:latest → exit → Exited container (see instance ID with docker ps -a)
- docker start -a -i [instance ID] → Start and access an exited container (-a = attach to access the container; -i = interactive)

### Dockerfile



# Build image from a Dockerfile

- cd /path/to/Dockerfile
- docker build -t image name: tag . (Do not forget '.')

# Simple docker file $\rightarrow$ Test building

```
# Simple Ubuntu container with htop and nano
installed
FROM ubuntu: latest
LABEL AUTHOR=dumrong.mai@mahidol.ac.th
RUN apt-get update && apt-get install -y --no-
install-recommends \
    nano \
    htop
```

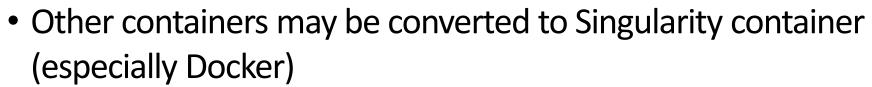
#### VM versus Container

- Docker (https://docs.docker.com/)
  - Most popular container (i.e. good documentation)







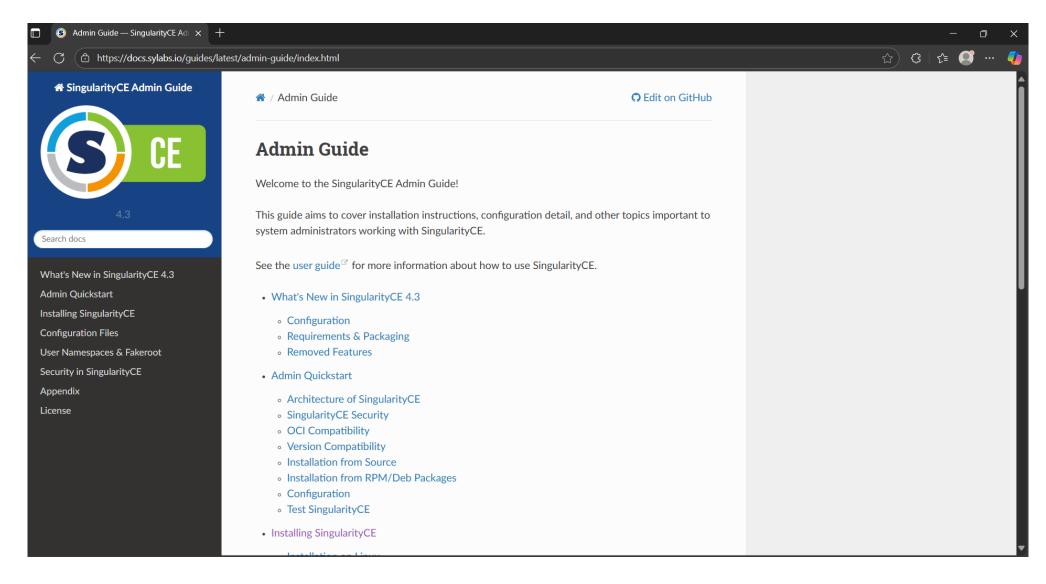


Run on HPC with job scheduler (e.g. Slurm)





# Singularity installation



# Singularity installation (Windows)

Follow the WSL2 installation instructions to enable WSL2 with the default Ubuntu 22.04 environment. On Windows 11 and the most recent builds of Windows 10 this is as easy as opening an administrator command prompt or Powershell window and entering:

wsl --install

Follow the prompts. A restart is required, and when you open the 'Ubuntu' app for the first time you'll be asked to set a username and password for the Linux environment.

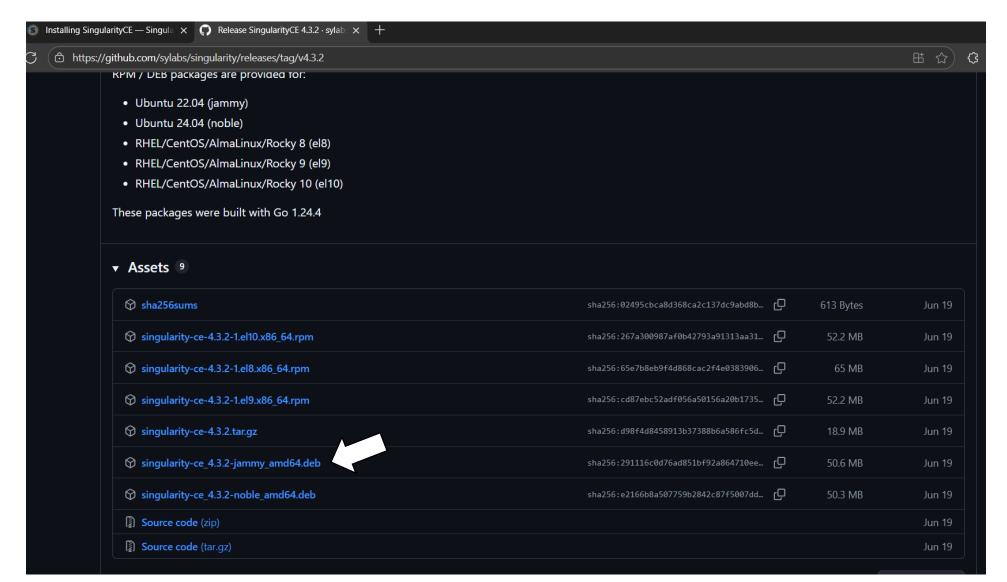
You can install SingularityCE from source, or from the Ubuntu packages at the GitHub releases page. To quickly install the 4.0.0 package use the following commands inside the WSL2 Ubuntu window:

```
$ wget https://github.com/sylabs/singularity/releases/download/v4.0.0/singularity-ce_4.0.0-jamm
$ sudo apt install ./singularity-ce_4.0.0-jammy_amd64.deb
```

The singularity command will now be available in your WSL2 environment:

- Run via WSL2
- Ubuntu on Windows (v20, Recommend v22)

# Singularity installation (Windows)



# Singularity installation (MacOS)

#### Mac %

To install SingularityCE on macOS, we recommend using the lima VM platform, available on Homebrew.

If you don't already have Homebrew installed, you can install it as follows:

```
$ /bin/bash -c "$(curl -fsSL https://raw.githubusercontent.com/Homebrew/install/HEAD/install.sh
```

Follow the instructions at the end of the installation process. In particular, make sure to add the relevant lines to your shell configuration:

```
\ (echo; echo 'eval "$(/home/linuxbrew/.linuxbrew/bin/brew shellenv)"') >> $HOME/.profile $ eval "$(/home/linuxbrew/.linuxbrew/bin/brew shellenv)"
```

Once Homebrew is installed, install lima:

```
$ brew install lima
```

As part of the SingularityCE distribution (starting with version 4), we have provided an example template for using SingularityCE with lima. The example is available under the examples/lima directory in the SingularityCE source bundle, and can also be downloaded directly from the code repository.

- Run via Lima (VM platform)

#### Simple SingularityCE definition file (test.def) -> Test building

```
# Simple Ubuntu container with htop and nano
installed
Bootstrap: docker
From: ubuntu:latest
%post
    apt-get -y update
    apt-get install -y --no-install-recommends \
    nano \
    htop
```

Simple SingularityCE definition file (xxx.def) -> Test building

```
$ sudo singularity build test.sif test.def
$ singularity shell test.sif
Singularity> htop
Singularity> ls
```

### Convert a local docker image to Singularity .sif file

```
$ sudo docker images
$ sudo singularity build docker_test.sif docker-
daemon://docker_test:latest
```

#### Writable sandbox directories

- Anything modified in the sandbox is kept
  - Docker or SIF files do not save any modification
- Sandbox is a directory, not a file
  - Lose portability <-> Gain modification
- Definition file or Dockerfile is recommended for reproducibility

```
$ sudo singularity build --sandbox lolcow/ library://lolcow
$ sudo singularity shell --writable lolcow/
$ sudo singularity build lolcow.sif lolcow/
```

### Basic Singularity commands

# The shell command allows you to spawn a new shell within your container and interact with it as though it were a virtual machine.

```
$ singularity shell lolcow_latest.sif
```

# The exec command allows you to execute a custom command within a container by specifying the image file.

```
$ singularity exec lolcow_latest.sif cowsay moo
```

#SingularityCE containers contain runscripts. These are user-defined scripts that define the actions a container should perform when someone runs it.

```
$ singularity run lolcow_latest.sif
```

# Bind paths in SingularityCE (mount in Docker)

- A Singularity container has an access to host's current working directory unless an admin disables the access.
- You may bind a different directory to the container with a following command:

```
$ singularity shell --bind /data:/mnt
my container.sif
```

• Use cases: HPC, Cloud VM with a persistent storage volume

#### Portable SIF

- You can save file.sif to external HDD or USB flash drive
- Connect to any host machine with Singularity installed to quickly run a Singularity container

#### • NOTE:

- USB flash drive must be formatted as NTFS
- FAT32 format (default format in some USB flash drives) cannot accommodate a single file larger than 4GB

# Cloud computing

- What is it?
- Cloud computing, often referred to as simply "the cloud," is the delivery of on-demand computing resources — everything from applications to data centers — over the internet on a pay-for-use basis.
  - Elastic resources Scale up or down quickly and easily to meet demand
  - Metered service so you only pay for what you use
  - Self service All the IT resources you need with self-service access

# Why do you need to know Cloud Computing?

- Web application deployment
- Requiring intensive but transient computing power
- Backup and archive
- Sharing and hosting (Public cloud)
- Anything else?

# Cloud service providers

Amazon web services



Google cloud platform
 Google Cloud Platform



Microsoft Azure



• IBM cloud



• Digital Ocean



Hands on: Cloud Computing

#### Overview

- Demo for creating a cloud VM
- Practice: remotely access the VM
- Practice: remotely transfer files from/to the VM
- Back-up and clean-up

### Creating a VM

- Platform will be based on "your advisor"
  - Digital ocean (Simple Unix/Linux applications, web application)
  - Google Cloud Platform/AWS (Machine learning, NextFlow)
  - IPGG server/MU cluster/Thai-SC (Intensive calculation with CPUs and RAM)

### Creating a VM

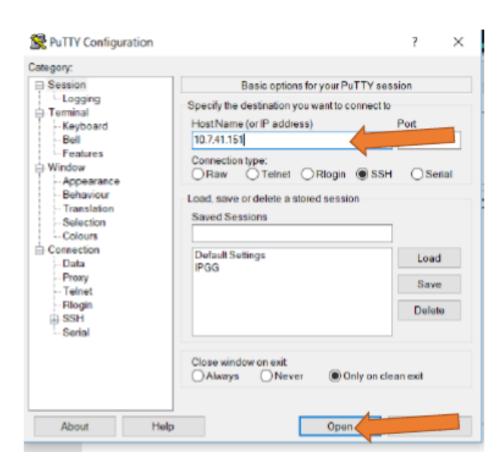
- Calculate/Estimate your resource wisely
  - CPU/RAM/HDD
  - OS
  - Location of your data center
  - Safety features

# Creating a VM on



# Remotely access the VM (or server)

- Windows
  - Require a software
  - Install "putty"
  - Login as 'root'



# Remotely access the VM (or server)

- Windows
  - Alternatively via Ubuntu on Windows (with WSL2)

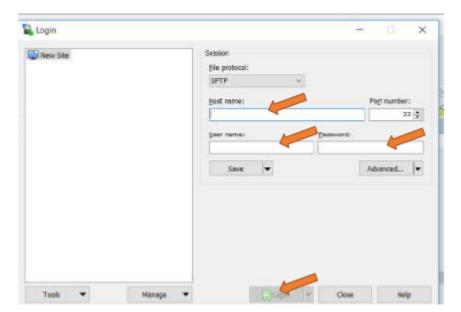
# Remotely access the VM (or server)

- Mac and Linux
  - Use your 'terminal'
  - ssh root@<VM IP address>

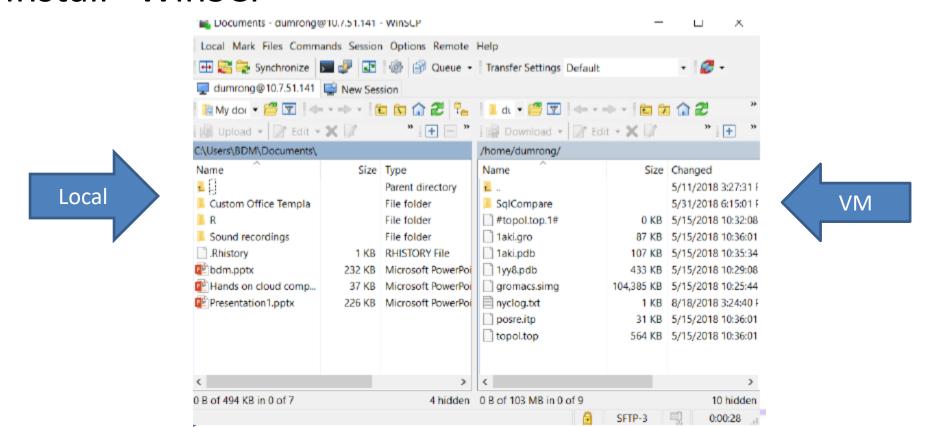
# Test basic Unix/Linux commands

- 1s
- pwd
- touch
- cp
- mv
- rm
- mkdir
- cd
- \*nano\*

- Windows
  - Require a software
  - Install "WinSCP"
    - ullet Windows user: Open WinSCP ullet enter IP Address
      - → Enter username ('root') and password → click "Login"



- Windows
  - Require a software
  - Install "WinSCP"



- Mac and Linux
  - Use your 'terminal'
  - Download:

```
scp root@<VM IP address>:/path/to/file /local/path
```

• Upload:

```
scp /local/path/to/file root@<VM IP address>:/path/
```

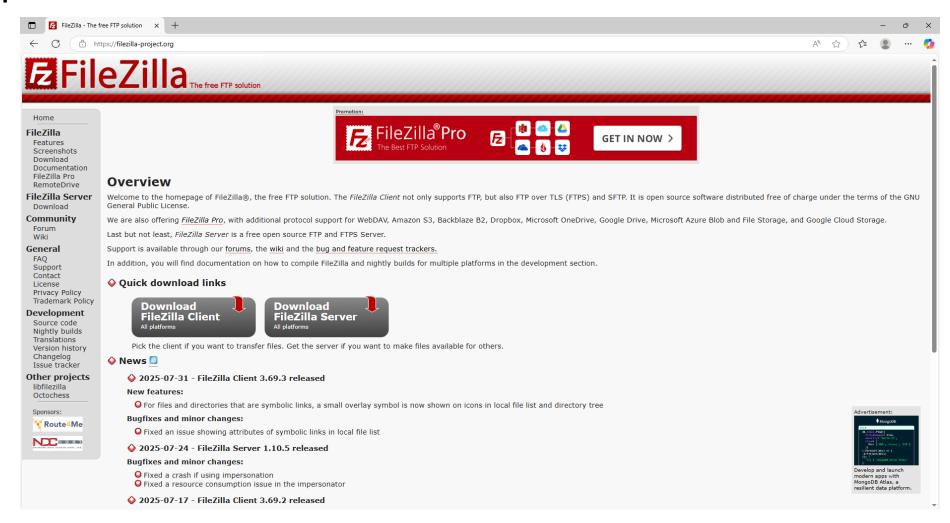
- Windows
  - Alternatively via Ubuntu on Windows (with WSL2)
  - Command lines
  - Download:

```
scp root@<VM IP address>:/path/to/file /local/path
```

• Upload:

```
scp /local/path/to/file root@<VM IP address>:/path/
```

All platforms



# Practice file transfer (to VM)

- 1. Open notepad/text editor in your computer
- 2. Write "Thank you." in a new text file
- 3. Save the file as thank.txt
- 4. Upload this file to your VM (WinSCP or scp)
- 5. Access your VM (PuTTY or ssh)
- 6. Type ls see if you can find your file
- 7. Type cat thank.txt to read the content in the file

# Practice file transfer (from VM)

- 1. Access your droplet (PuTTY or ssh)
- 2.Type echo "you are welcome" > yaw.txt
- 3. Connect to your VM with WinSCP (or use scp)
- 4. Download yaw.txt to you computer
- 5. Use notepad or word to read the file

### Back-up and clean-up

- Back-up with 'snapshot':
  - Freeze everything in your VM at the time of the snapshot.
  - The snapshot is used to instantly restore the VM.
  - Use for:
    - Routine back-up
    - Temporary paused VM
  - Cheaper than let the VM run. (Snapshot is <u>NOT</u> free!)

### Back-up and clean-up

- Clean-up:
  - Running and power-off VMs cost the same!
  - **Everything** must be destroyed to stop the cloud provider from charging you.
  - Snapshots and volumes cost money!
  - Clean-up after:
    - Finish the project
    - Results and scripts are downloaded to local machines
  - After a VM is destroyed, data in the VM are gone for good!