Computing Basics for Bioinformatics on the Cloud

Session 2 Yaoyu Wang





Overview

- Introduction to cloud computing
- Get on to the cloud
- Get around on the cloud with linux
- Free bioinformatics resources on the cloud

Download slide at:

http://tinyurl.com/cccbtrain





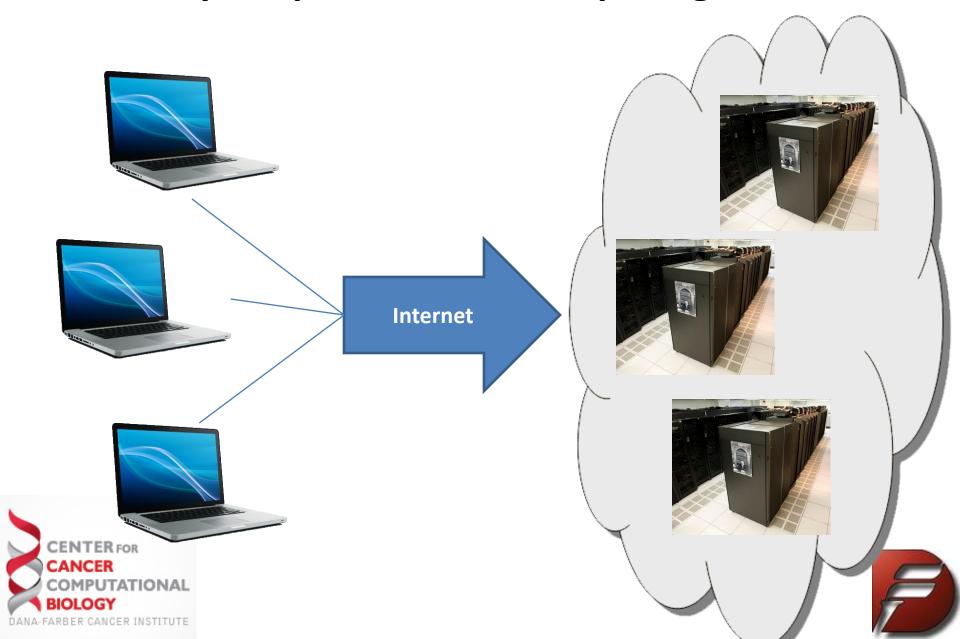
What is Cloud Computing?

Cloud Computing is a general term used to describe a network based computing that takes place over the Internet





Overly Simplistic Cloud Computing Schema



Basic Cloud Characteristics

- The "flexibility and elasticity" allows these systems
 to scale up and down at will utilising the resources
 of all kinds (CPU, storage, server capacity, load
 balancing, and databases).
- The "pay as much as used and needed" type of utility computing and the "always on!, anywhere and any place" type of network-based computing.
- The "no-need-to-know" in terms of the underlying details of infrastructure, applications interface with the infrastructure via the APIs.





Some Commercial Cloud Offerings





























Amazon Web Services

- Elastic Compute Cloud (EC2)
 - Rent computing resources by the hour
 - Basic unit of accounting = instance-hour
 - Additional costs for bandwidth
- Elastic Block Storage (EBS)
 - Persistent storage
 - Charge by provision block
 - Additional costs for bandwidth





Cloud Computing

Quick-start an Amazon EC2 instance...





Cloud Computing

Running a server on Amazon Web Service (AWS) requires you to go through the following steps:

- Create your account on Amazon (if you do not already have one) and setup your credentials
- 2. Install Amazon EC2 API Tools
- 3. Instantiate your image (i.e. your account)
- Configure your <u>instance</u> (i.e. your virtual machine)



Cloud Computing: Start the Machine

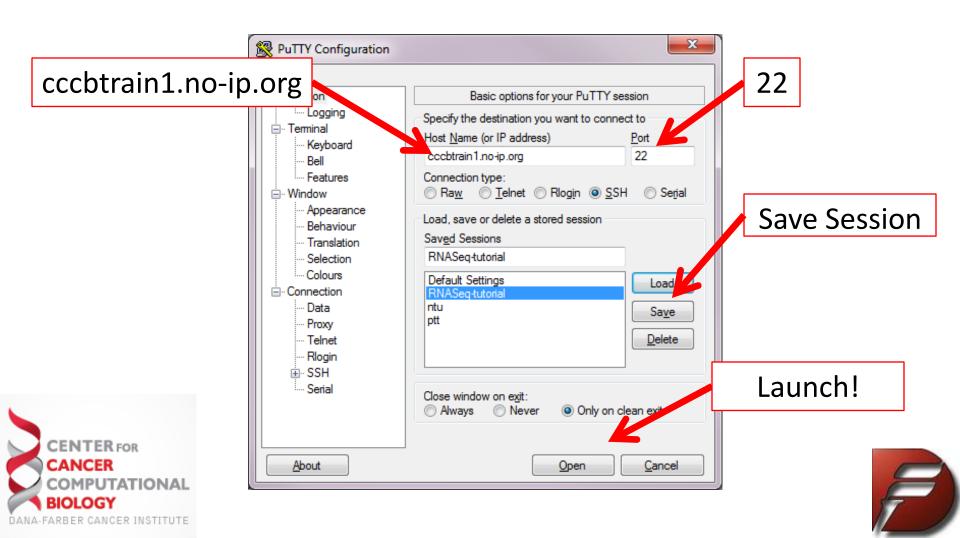
http://aws.amazon.com/

a, Amazon.com Sign In		2 X
← → C ♠ Attps://www.amazor	n.com/ap/signin?openid.assoc_handle=aws&openid.return_to=https%3A% 😭	4 3
**		Î
amazon		
amazon webservices™		
Amazon Web Servic	ces Sign In	
Please enter the AWS Identity & Access	Management (IAM) User name and password assigned by your system ac	dminis
AWS Account:	сссь	=
User Name:		
Password:		
	Sign in using our secure server	
	e contact your system administrator if	
	ave forgotten your user credentials.	
Sign	in using AWS Account credentials	
	Terms of Use Privacy Policy © 1996-2010, Amazon.com, Inc. or its affil	liates.
		,



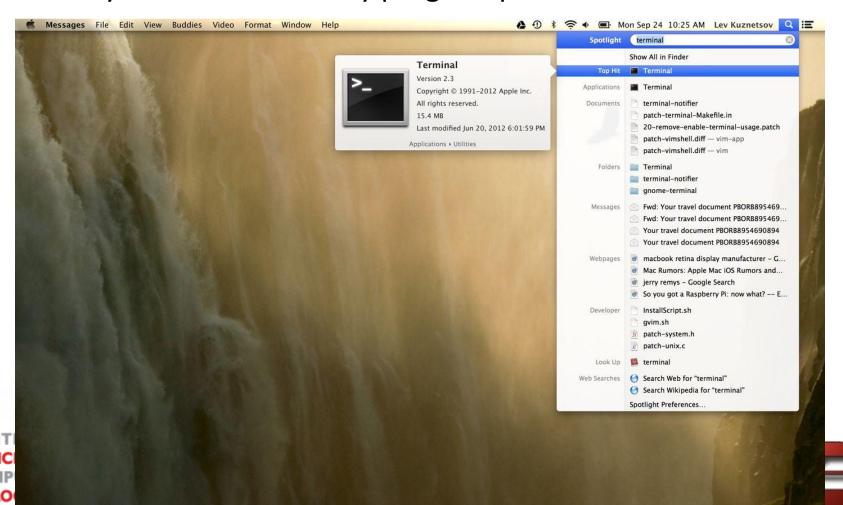
For Windows

Download PuTTY at: http://tinyurl.com/putty-d



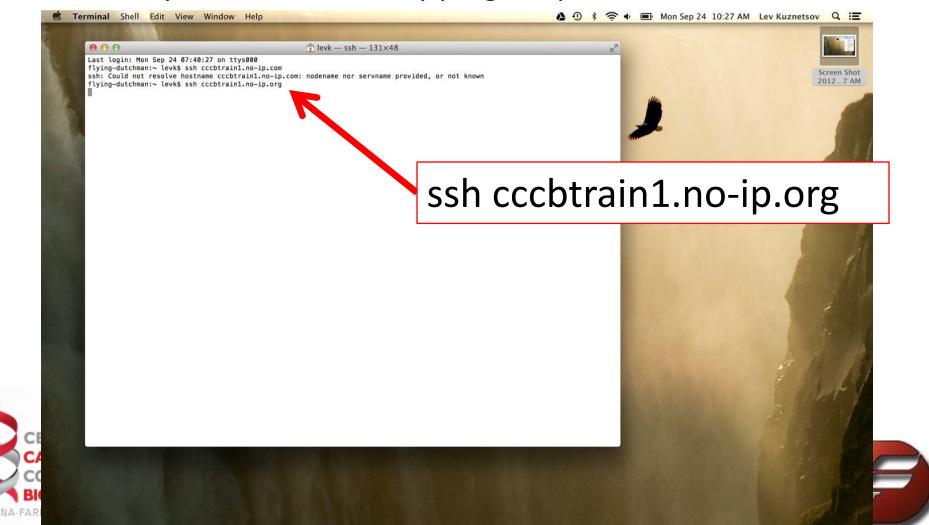
For Mac

You already have the necessary program pre-installed

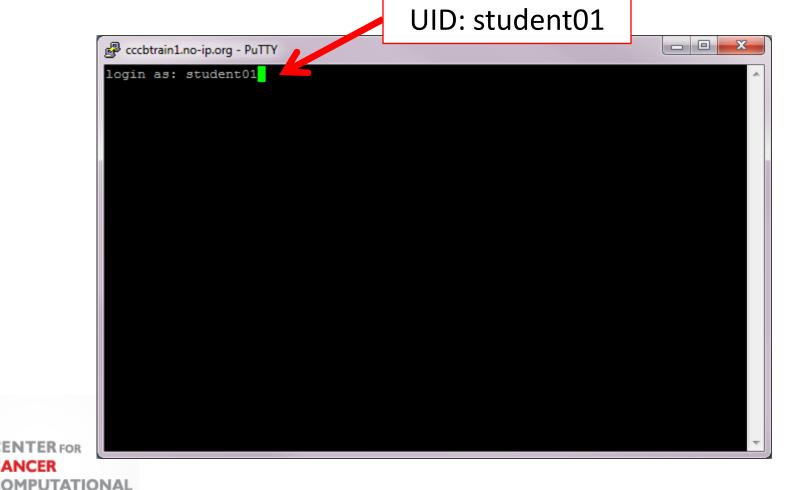


For Mac

You already have the necessary program pre-installed



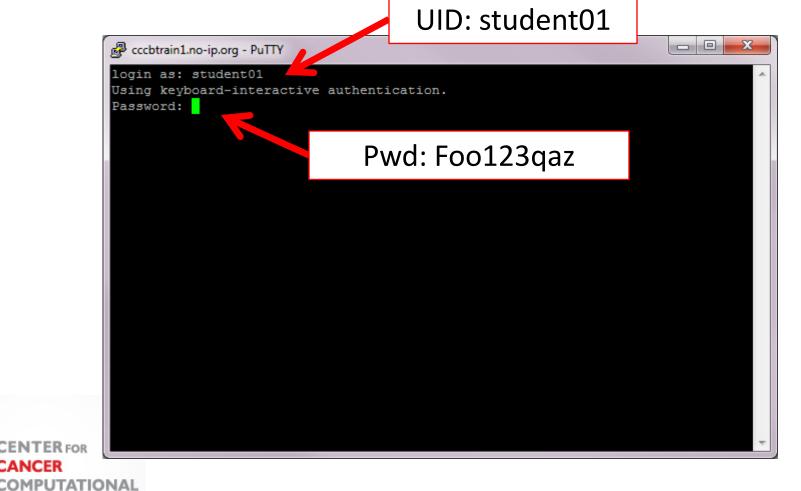
Now you are on THE Cloud...



ARBER CANCER INSTITUTE



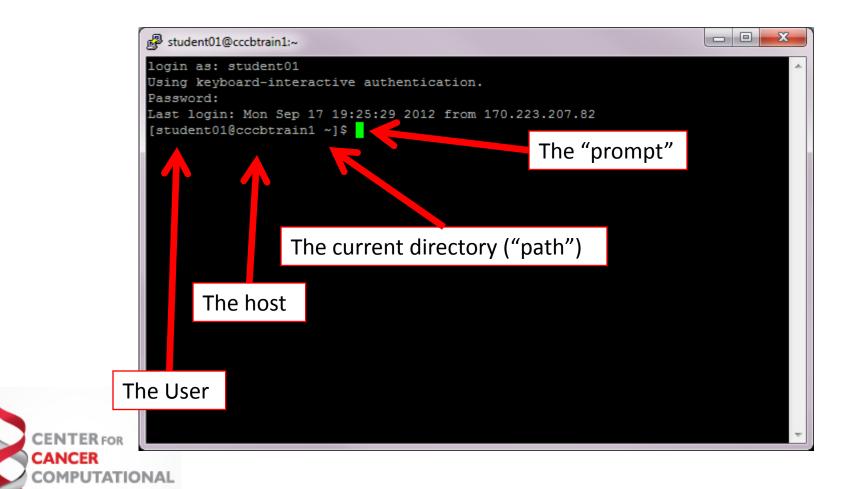
Now you are on THE Cloud...



FARBER CANCER INSTITUTE



Now you are on THE Cloud...



FARBER CANCER INSTITUTE



Cloud Computing

How to navigate around the Cloud?

(a.k.a Linux Basics)





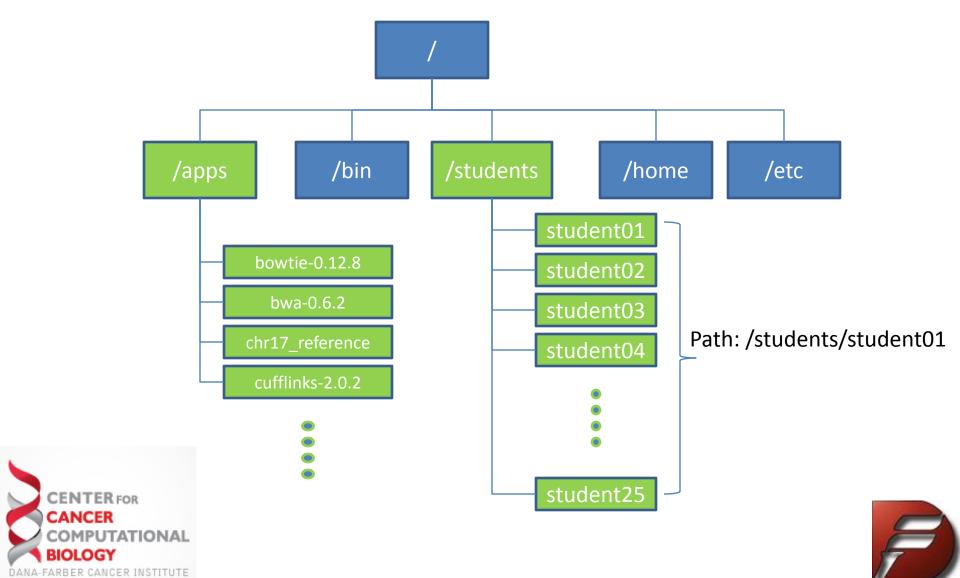
What exactly is a "shell"?

- After logging in, Linux/Unix starts another program called the shell
- The shell interprets commands the user types and manages their execution
 - The shell communicates with the internal part of the operating system called the kernel
 - The most popular shells are: tcsh, csh, korn, and bash
 - The differences are most times subtle
 - For this tutorial, we are using bash
- Shell commands are CASE SENSITIVE!





File System Organization



```
student01@cccbtrain1:~
[student01@cccbtrain1 ~]$ pwd
students/student01
[student01@cccbtrain1 ~]$
```



pwd - print name of current/working directory



```
student01@cccbtrain1:~
[student01@cccbtrain1 ~]$ pwd
/students/student01
[student01@cccbtrain1 ~]$ ls
TP53.fasta
[student01@cccbtrain1 ~]$
```



Is - list directory contents



```
student01@cccbtrain1:~
[student01@cccbtrain1 ~]$ pwd
/students/student01
[student01@cccbtrain1 ~]$ ls
TP53.fasta
[student01@cccbtrain1 ~]$ ls -1
-rw-r--r-. 1 student01 students 19510 Sep 18 22:31 TP53.fasta
[student01@cccbtrain1 ~]$
```



Is —I use a long listing format



```
student01@cccbtrain1:~
[student01@cccbtrain1 ~]$ pwd
/students/student01
[student01@cccbtrain1 ~]$ ls
TP53.fasta
[student01@cccbtrain1 ~]$ ls -1
total 20
-rw-r--r-. 1 student01 students 19510 Sep 18 22:31 TP53.fasta
[student01@cccbtrain1 ~]$ less TP53.fasta
```



less – quick file viewer













less – allows for simple text searching







less – allows for simple text searching



```
student01@cccbtrain1:~
[student01@cccbtrain1 ~]$ pwd
/students/student01
[student01@cccbtrain1 ~]$ ls
TP53.fasta
[student01@cccbtrain1 ~]$ ls -1
total 20
-rw-r--r-. 1 student01 students 19510 Sep 18 22:31 TP53.fasta
[student01@cccbtrain1 ~]$ less TP53.fasta
[student01@cccbtrain1 ~]$
```





```
student01@cccbtrain1:~
[student01@cccbtrain1 ~]$ pwd
/students/student01
[student01@cccbtrain1 ~]$ ls
TP53.fasta
[student01@cccbtrain1 ~]$ ls -1
total 20
-rw-r--r-. 1 student01 students 19510 Sep 18 22:31 TP53.fasta
[student01@cccbtrain1 ~]$ less TP53.fasta
[student01@cccbtrain1 ~]$ mkdir FASTA
```



mkdir – make directories



```
student01@cccbtrain1:~
[student01@cccbtrain1 ~]$ pwd
/students/student01
[student01@cccbtrain1 ~]$ ls
TP53.fasta
[student01@cccbtrain1 ~]$ ls -1
total 20
-rw-r--r-. 1 student01 students 19510 Sep 18 22:31 TP53.fasta
[student01@cccbtrain1 ~]$ less TP53.fasta
[student01@cccbtrain1 ~]$ mkdir FASTA
[student01@cccbtrain1 ~]$ ls
FASTA TP53.fasta
[student01@cccbtrain1 ~]$ ls -1
total 24
drwxr-xr-x. 2 student01 students 4096 Sep 18 23:11 FASTA
-rw-r--r-. 1 student01 students 19510 Sep 18 22:31 TP53.fasta
[student01@cccbtrain1 ~]$
```



mkdir – make directories



```
student01@cccbtrain1:~
[student01@cccbtrain1 ~]$ pwd
/students/student01
[student01@cccbtrain1 ~]$ ls
TP53.fasta
[student01@cccbtrain1 ~]$ ls -1
total 20
-rw-r--r-. 1 student01 students 19510 Sep 18 22:31 TP53.fasta
[student01@cccbtrain1 ~]$ less TP53.fasta
[student01@cccbtrain1 ~]$ mkdir FASTA
[student01@cccbtrain1 ~]$ ls
FASTA TP53.fasta
[student01@cccbtrain1 ~]$ ls -1
total 24
drwxr-xr-x. 2 student01 students 4096 Sep 18 23:11 FASTA
-rw-r--r-. 1 student01 students 19510 Sep 18 22:31 TP53.fasta
[student01@cccbtrain1 ~]$ mv TP53.fasta FASTA/
```





```
student01@cccbtrain1:~
[student01@cccbtrain1 ~]$ pwd
/students/student01
[student01@cccbtrain1 ~]$ ls
TP53.fasta
[student01@cccbtrain1 ~]$ ls -1
total 20
-rw-r--r-. 1 student01 students 19510 Sep 18 22:31 TP53.fasta
[student01@cccbtrain1 ~]$ less TP53.fasta
[student01@cccbtrain1 ~]$ mkdir FASTA
[student01@cccbtrain1 ~]$ ls
FASTA TP53.fasta
[student01@cccbtrain1 ~]$ ls -1
total 24
drwxr-xr-x. 2 student01 students 4096 Sep 18 23:11 FASTA
-rw-r--r-. 1 student01 students 19510 Sep 18 22:31 TP53.fasta
[student01@cccbtrain1 ~]$ mv TP53.fasta FASTA/
[student01@cccbtrain1 ~]$ ls
[student01@cccbtrain1 ~]$
```





```
student01@cccbtrain1:~
[student01@cccbtrain1 ~]$ pwd
/students/student01
[student01@cccbtrain1 ~]$ ls
TP53.fasta
[student01@cccbtrain1 ~]$ ls -1
total 20
-rw-r--r-. 1 student01 students 19510 Sep 18 22:31 TP53.fasta
[student01@cccbtrain1 ~]$ less TP53.fasta
[student01@cccbtrain1 ~]$ mkdir FASTA
[student01@cccbtrain1 ~]$ ls
FASTA TP53.fasta
[student01@cccbtrain1 ~]$ ls -1
total 24
drwxr-xr-x. 2 student01 students 4096 Sep 18 23:11 FASTA
-rw-r--r-. 1 student01 students 19510 Sep 18 22:31 TP53.fasta
[student01@cccbtrain1 ~]$ mv TP53.fasta FASTA/
[student01@cccbtrain1 ~]$ ls
[student01@cccbtrain1 ~]$ cd FASTA/
```





```
student01@cccbtrain1:~/FASTA
[student01@cccbtrain1 ~]$ pwd
/students/student01
[student01@cccbtrain1 ~1$ ls
TP53.fasta
[student01@cccbtrain1 ~]$ ls -1
total 20
-rw-r--r-. 1 student01 students 19510 Sep 18 22:31 TP53.fasta
[student01@cccbtrain1 ~]$ less TP53.fasta
[student01@cccbtrain1 ~]$ mkdir FASTA
[student01@cccbtrain1 ~1$ 1s
FASTA TP53.fasta
[student01@cccbtrain1 ~]$ ls -1
total 24
drwxr-xr-x. 2 student01 students 4096 Sep 18 23:11 FASTA
-rw-r--r-. 1 student01 students 19510 Sep 18 22:31 TP53.fasta
[student01@cccbtrain1 ~]$ mv TP53.fasta FASTA/
[student01@cccbtrain1 ~]$ ls
[student01@cccbtrain1 ~]$ cd FASTA/
[student01@cccbtrain1 FASTA]$
```





```
student01@cccbtrain1:~/FASTA
[student01@cccbtrain1 ~]$ pwd
/students/student01
[student01@cccbtrain1 ~1$ ls
TP53.fasta
[student01@cccbtrain1 ~]$ ls -1
total 20
-rw-r--r-. 1 student01 students 19510 Sep 18 22:31 TP53.fasta
[student01@cccbtrain1 ~]$ less TP53.fasta
[student01@cccbtrain1 ~]$ mkdir FASTA
[student01@cccbtrain1 ~]$ 1s
FASTA TP53.fasta
[student01@cccbtrain1 ~]$ ls -1
total 24
drwxr-xr-x. 2 student01 students 4096 Sep 18 23:11 FASTA
-rw-r--r-. 1 student01 students 19510 Sep 18 22:31 TP53.fasta
[student01@cccbtrain1 ~]$ mv TP53.fasta FASTA/
[student01@cccbtrain1 ~]$ ls
FASTA
[student01@cccbtrain1 ~]$ cd FASTA/
[student01@cccbtrain1 FASTA]$ ls
TP53.fasta
[student01@cccbtrain1 FASTA]$
```





```
student01@cccbtrain1:~/FASTA
[student01@cccbtrain1 ~]$ pwd
/students/student01
[student01@cccbtrain1 ~1$ ls
TP53.fasta
[student01@cccbtrain1 ~]$ ls -1
total 20
-rw-r--r-. 1 student01 students 19510 Sep 18 22:31 TP53.fasta
[student01@cccbtrain1 ~]$ less TP53.fasta
[student01@cccbtrain1 ~]$ mkdir FASTA
[student01@cccbtrain1 ~1$ 1s
FASTA TP53.fasta
[student01@cccbtrain1 ~]$ ls -1
total 24
drwxr-xr-x. 2 student01 students 4096 Sep 18 23:11 FASTA
-rw-r--r-. 1 student01 students 19510 Sep 18 22:31 TP53.fasta
[student01@cccbtrain1 ~]$ mv TP53.fasta FASTA/
[student01@cccbtrain1 ~]$ 1s
FASTA
[student01@cccbtrain1 ~]$ cd FASTA/
[student01@cccbtrain1 FASTA]$ ls
TP53.fasta
[student01@cccbtrain1 FASTA]$ pwd
/students/student01/FASTA
student01@cccbtrain1 FASTA1$
```





```
student01@cccbtrain1:~/FASTA
[student01@cccbtrain1 ~]$ pwd
/students/student01
[student01@cccbtrain1 ~1$ ls
TP53.fasta
[student01@cccbtrain1 ~]$ ls -1
total 20
-rw-r--r-. 1 student01 students 19510 Sep 18 22:31 TP53.fasta
[student01@cccbtrain1 ~]$ less TP53.fasta
[student01@cccbtrain1 ~]$ mkdir FASTA
[student01@cccbtrain1 ~1$ 1s
FASTA TP53.fasta
[student01@cccbtrain1 ~]$ ls -1
total 24
drwxr-xr-x. 2 student01 students 4096 Sep 18 23:11 FASTA
-rw-r--r-. 1 student01 students 19510 Sep 18 22:31 TP53.fasta
[student01@cccbtrain1 ~]$ mv TP53.fasta FASTA/
[student01@cccbtrain1 ~]$ 1s
FASTA
[student01@cccbtrain1 ~]$ cd FASTA/
[student01@cccbtrain1 FASTA]$ ls
TP53.fasta
[student01@cccbtrain1 FASTA]$ pwd
/students/student01/FASTA
[student01@cccbtrain1 FASTA]$ cd ..
```





```
student01@cccbtrain1:~
/students/student01
[student01@cccbtrain1 ~]$ ls
TP53.fasta
[student01@cccbtrain1 ~]$ ls -1
total 20
-rw-r--r-. 1 student01 students 19510 Sep 18 22:31 TP53.fasta
[student01@cccbtrain1 ~]$ less TP53.fasta
[student01@cccbtrain1 ~]$ mkdir FASTA
[student01@cccbtrain1 ~1$ ls
FASTA TP53.fasta
[student01@cccbtrain1 ~]$ ls -1
total 24
drwxr-xr-x. 2 student01 students 4096 Sep 18 23:11 FASTA
-rw-r--r-. 1 student01 students 19510 Sep 18 22:31 TP53.fasta
[student01@cccbtrain1 ~]$ mv TP53.fasta FASTA/
[student01@cccbtrain1 ~]$ 1s
FASTA
[student01@cccbtrain1 ~]$ cd FASTA/
[student01@cccbtrain1 FASTA]$ ls
TP53.fasta
[student01@cccbtrain1 FASTA]$ pwd
/students/student01/FASTA
[student01@cccbtrain1 FASTA]$ cd ..
[student01@cccbtrain1 ~]$
```





```
student01@cccbtrain1:~
/students/student01
[student01@cccbtrain1 ~]$ ls
TP53.fasta
[student01@cccbtrain1 ~]$ ls -1
total 20
-rw-r--r-. 1 student01 students 19510 Sep 18 22:31 TP53.fasta
[student01@cccbtrain1 ~]$ less TP53.fasta
[student01@cccbtrain1 ~]$ mkdir FASTA
[student01@cccbtrain1 ~1$ ls
FASTA TP53.fasta
[student01@cccbtrain1 ~]$ ls -1
total 24
drwxr-xr-x. 2 student01 students 4096 Sep 18 23:11 FASTA
-rw-r--r-. 1 student01 students 19510 Sep 18 22:31 TP53.fasta
[student01@cccbtrain1 ~]$ mv TP53.fasta FASTA/
[student01@cccbtrain1 ~]$ 1s
FASTA
[student01@cccbtrain1 ~]$ cd FASTA/
[student01@cccbtrain1 FASTA]$ 1s
TP53.fasta
[student01@cccbtrain1 FASTA]$ pwd
/students/student01/FASTA
[student01@cccbtrain1 FASTA]$ cd ..
[student01@cccbtrain1 ~]$ man pwd
```





```
student01@cccbtrain1:~
PWD(1)
                                 User Commands
                                                                        PWD(1)
NAME
      pwd - print name of current/working directory
SYNOPSIS
       pwd [OPTION]...
DESCRIPTION
       Print the full filename of the current working directory.
       -L, --logical
             use PWD from environment, even if it contains symlinks
       -P, --physical
              avoid all symlinks
       --help display this help and exit
       --version
              output version information and exit
       NOTE: your shell may have its own version of pwd, which usually super-
```

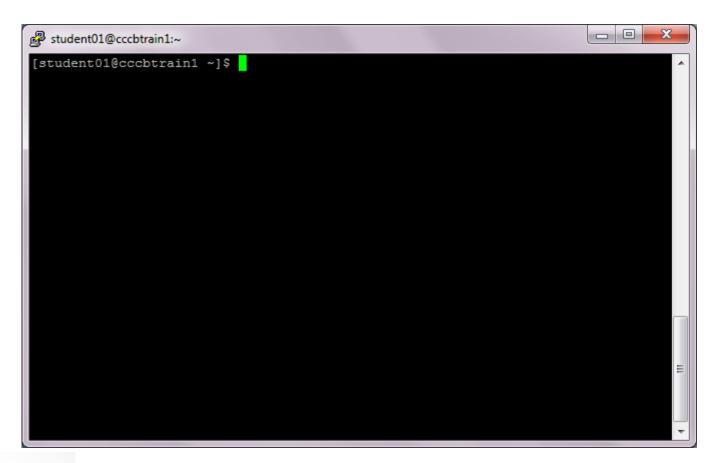




```
student01@cccbtrain1:~
[student01@cccbtrain1 ~]$ 1s
TP53.fasta
[student01@cccbtrain1 ~]$ ls -1
total 20
-rw-r--r-. 1 student01 students 19510 Sep 18 22:31 TP53.fasta
[student01@cccbtrain1 ~]$ less TP53.fasta
[student01@cccbtrain1 ~]$ mkdir FASTA
[student01@cccbtrain1 ~]$ ls
FASTA TP53.fasta
[student01@cccbtrain1 ~]$ 1s -1
total 24
drwxr-xr-x. 2 student01 students 4096 Sep 18 23:11 FASTA
-rw-r--r-. 1 student01 students 19510 Sep 18 22:31 TP53.fasta
[student01@cccbtrain1 ~]$ mv TP53.fasta FASTA/
[student01@cccbtrain1 ~]$ ls
FASTA
[student01@cccbtrain1 ~]$ cd FASTA/
[student01@cccbtrain1 FASTA]$ 1s
TP53.fasta
[student01@cccbtrain1 FASTA]$ pwd
/students/student01/FASTA
[student01@cccbtrain1 FASTA]$ cd ..
[student01@cccbtrain1 ~]$ man pwd
[student01@cccbtrain1 ~]$ clear
```











Cloud Computing

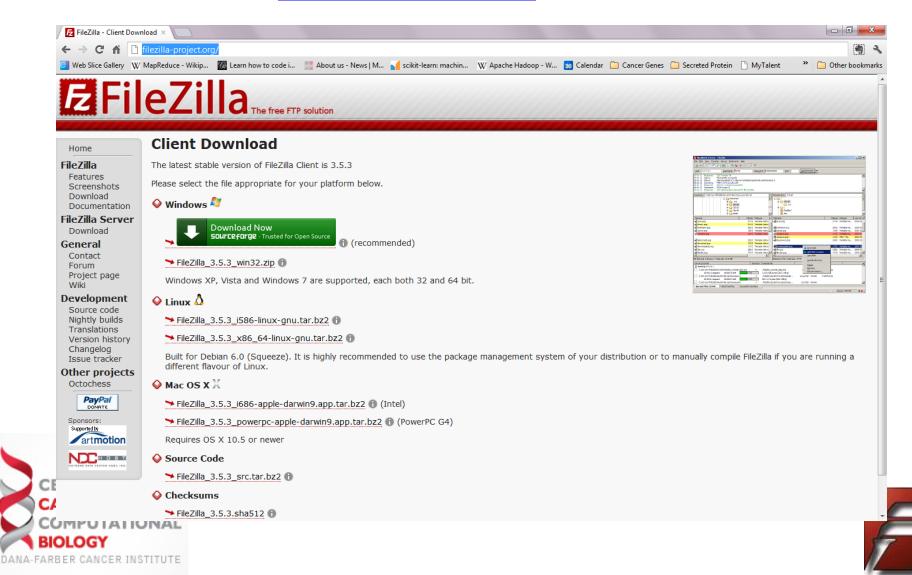
How to get files up and down the Cloud?





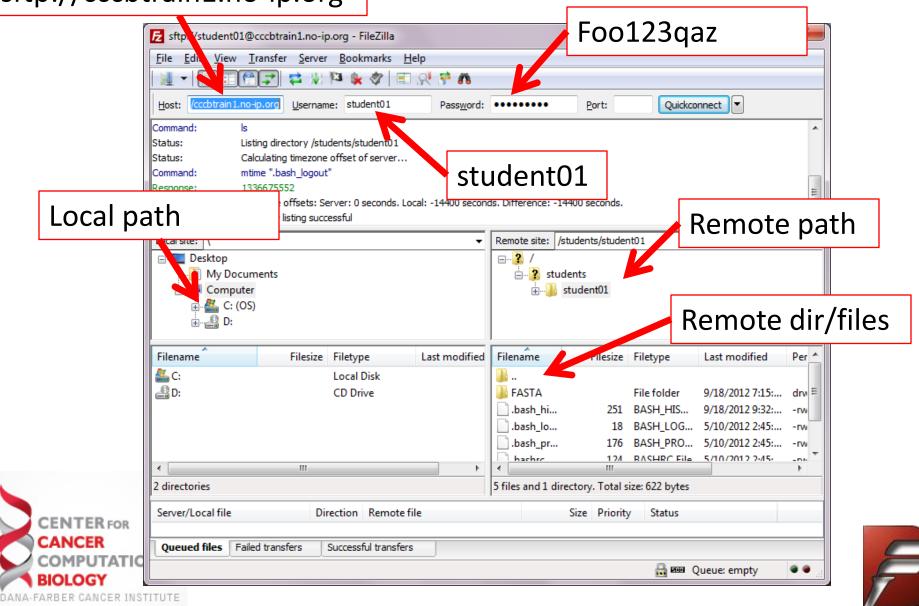
How to get files up and down the Cloud?

Download FileZilla at: http://filezilla-project.org/

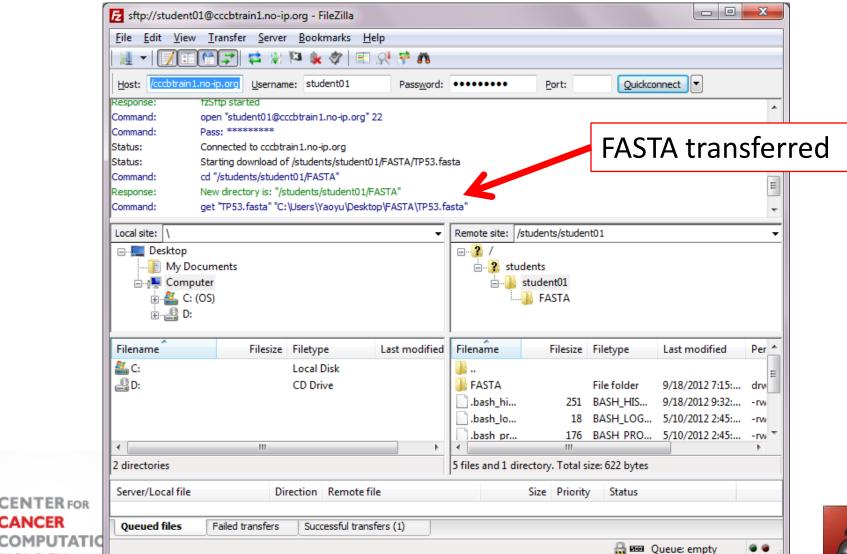


File Transfer with FileZilla

sftp://cccbtrain1.no-ip.org



File Transfer with FileZilla



CENTER FOR CANCER

BIOLOGY DANA-FARBER CANCER INSTITUTE



Session Summary

- Could Computing are remotely hosted computational resources and ubiquitous accessible by internet
- Amazon Web Service (AWS) and Elastic Compute Cloud (EC2) provides cloud computing service and storage by demand
- Linux-based Cloud servers can be accessed through SSH (Secure SHell) by <u>PuTTy</u>
- Files can be transferred up and down the Cloud Server by SFTP using programs such as <u>FileZilla</u>





Download R at: /apps/R-installation



