



# Canadian Bioinformatics Workshops

[www.bioinformatics.ca](http://www.bioinformatics.ca)

This page is available in the following languages:

Afrikaans Български Català Dansk Deutsch Ελληνικά English English (CA) English (GB) English (US) Esperanto Castellano Castellano (AR) Español (CL) Castellano (CO) Español (Ecuador) Castellano (MX) Castellano (PE) Euskara Suomeksi français français (CA) Galego ລາວ hrvatski Magyar Italiano 日本語 한국어 Macedonian Melayu Nederlands Norsk Sesotho sa Leboa polski Português română slovenščina srpski (latinica) Sotho svenska 中文 華語 (台灣) IsiZulu



## Attribution-Share Alike 2.5 Canada

### You are free:



to Share — to copy, distribute and transmit the work

to Remix — to adapt the work



### Under the following conditions:



**Attribution.** You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).



**Share Alike.** If you alter, transform, or build upon this work, you may distribute the resulting work only under the same or similar licence to this one.

- For any reuse or distribution, you must make clear to others the licence terms of this work.
- Any of the above conditions can be waived if you get permission from the copyright holder.
- The author's moral rights are retained in this licence.

[Disclaimer](#)

Your fair dealing and other rights are in no way affected by the above.

This is a human-readable summary of the Legal Code (the full licence) available in the following languages:

[English](#) [French](#)

# RNA-Seq Module 0

## Introduction to Cloud Computing

Malachi Griffith and Obi Griffith

Modification of slides from Francis Ouellette

Informatics for RNA-seq Analysis

May 28-30, 2018



# Learning objectives of the course

- **Module 0: Introduction to Cloud Computing**
- Module 1: Introduction to RNA Sequencing
- Module 2: Alignment and Visualization
- Module 3: Expression and Differential Expression
- Module 4: Alignment Free Expression Estimation
- Module 5: Isoform Discovery and Alternative Expression
- Tutorials
  - Provide a working example of an RNA-seq analysis pipeline
  - Run in a ‘reasonable’ amount of time with modest computer resources
  - Self contained, self explanatory, portable

# Learning Objectives

- Introduction to cloud computing concepts
- Introduction to cloud computing providers
- Use the Amazon EC2 console to create an instance for each student
  - Will be used for many hands-on tutorials throughout the course
- How to log into your cloud instance

# Disk Capacity vs Sequencing Capacity, 1990-2012

Disk Storage  
(Mbytes/\$)

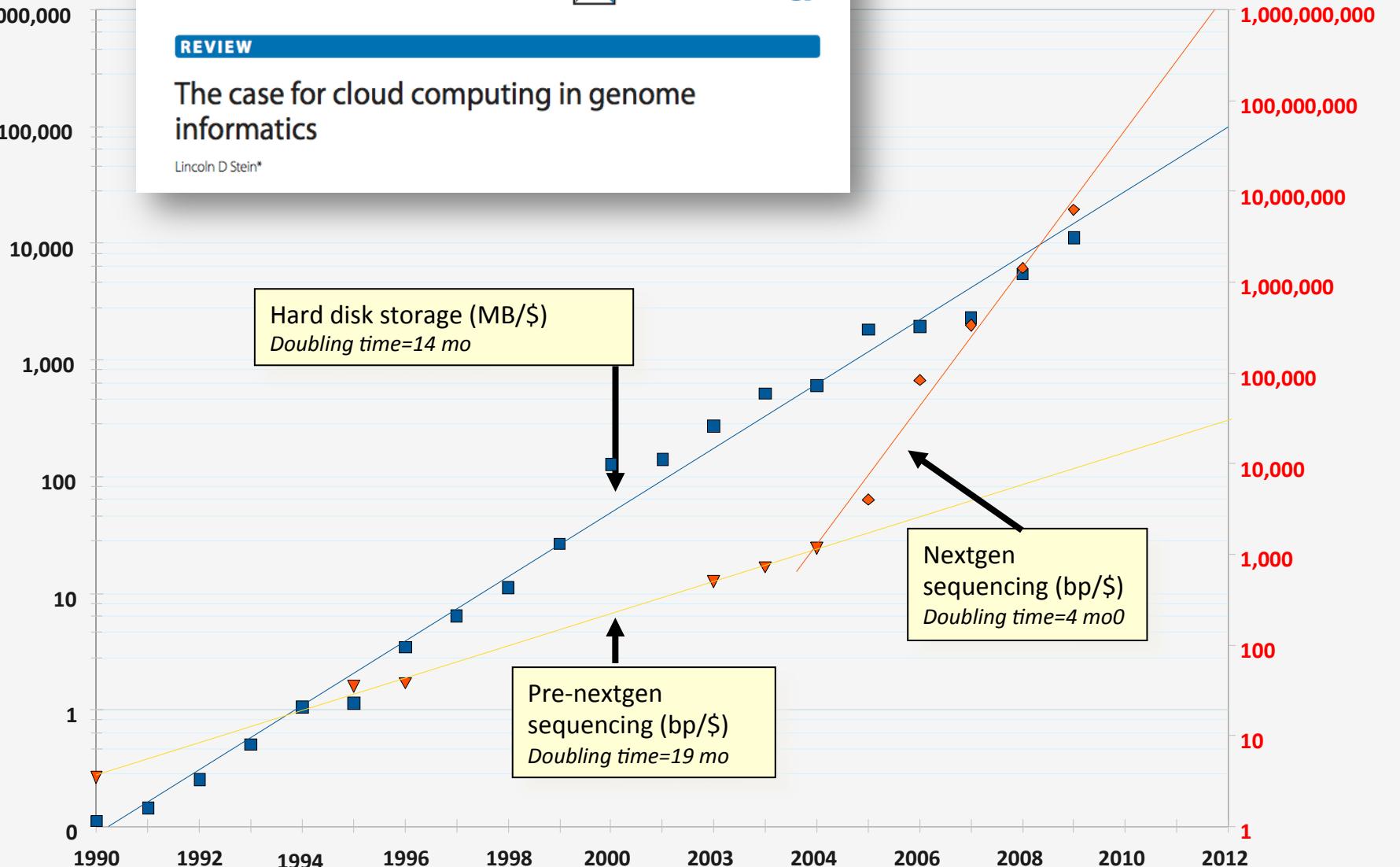
Stein *Genome Biology* 2010, 11:207  
<http://genomebiology.com/2010/11/5/207>



REVIEW

The case for cloud computing in genome informatics

Lincoln D Stein\*



# About DNA and computers

- We hit the \$1000 genome\* in ~2016
  - Need to think about the \$100 genome
- The doubling time of sequencing has been ~5-6 months.
- The doubling time of storage and network bandwidth is ~12 months.
- The doubling time of CPU speed is ~18 months.
- The cost of sequencing a base pair will eventually equal the cost of storing a base pair

# What is the general biomedical scientist to do?

- Lots of data
- Poor IT infrastructure in many labs
- Where do they go?
- Write more grants?
- Get bigger hardware?

# Cloud computing providers

- Amazon AWS
  - <https://aws.amazon.com/>
- Google cloud
  - <https://cloud.google.com/>
- Digital ocean
  - <https://www.digitalocean.com/>
- Microsoft Azure
  - <https://azure.microsoft.com/en-us/>
- More...

# Amazon Web Services (AWS)

- Infinite storage (scalable): S3 (simple storage service)
- Compute per hour: EC2 (elastic cloud computing)
- Ready when you are High Performance Computing
- Multiple football fields of HPC throughout the world
- HPC are expanded at one container at a time:



# Some of the challenges of cloud computing:

- Not cheap!
- Getting files to and from there
- Not the best solution for everybody
- Standardization
- PHI: personal health information & security concerns
- In the USA: HIPAA act, PSQIA act, HITECH act, Patriot act, CLIA and CAP programs, etc.
  - <http://www.biostars.org/p/70204/>

# Some of the advantages of cloud computing:

- We received a grant from Amazon, so supported by ‘AWS in Education grant award’.
- There are better ways of transferring large files, and now AWS makes it free to upload files.
- A number of datasets exist on AWS (e.g. 1000 genome data).
- Many useful bioinformatics AMI’s (Amazon Machine Images) exist on AWS: e.g. cloudbiolinux & CloudMan (Galaxy) – now one for this course!
- Many flavors of cloud available, not just AWS

# Key AWS concepts and terminology

- AWS - Amazon Web Services. A collection of cloud computing services provided by Amazon.
- EC2 - Elastic Compute. An AWS service that allows you to configure and rent computers to meet your compute needs on an as needed basis.
- EBS - Elastic Block Storage. A data storage solution that allows you to rent disk storage and associate that storage with your compute resources. EBS volumes are generally backed by SSD devices.

# Key AWS concepts and terminology

- S3 - Simple storage service. Cheaper than EBS and allows for storage of larger amounts of data with some drawbacks compared to EBS. S3 volumes store data as objects that are accessed by an API or command line interface or other application designed to work with S3. EBS volumes on the other hand can be mounted as if they were a local disk drive associated with the Instance.
- SSD - Solid state drive. A particular type of storage hardware that is generally faster and more expensive than traditional hard drives.

# Key AWS concepts and terminology

- HDD - Hard disk drive. A particular type of storage hardware that is generally cheaper and larger but slower than SSD. HDD drives are traditional hard drives that access data on a spinning magnetic disk.
- Ephemeral storage - Also known as Instance Store storage. Data storage associated with an EC2 instance that is local to the host computer. This storage does not persist when the instance is stopped or terminated. In other words, anything you store in this way will be lost if the system is stopped or terminated. Instance store volumes may be backed by SSD or HDD devices.

# What is a Region?

- An AWS Region is set of compute resources that Amazon maintains (like the Data Center image shown before)
- Each Region corresponds to a physical warehouse of compute hardware (computers, storage, networking, etc.).
- At the time of writing there are 14 regions: (US East (N.Virginia), US East (Ohio), US West (Oregon), US West (N. California), Canada (Central), EU (Ireland), EU (Frankfurt), EU (London), Asia Pacific (Singapore), Asia Pacific (Sydney), Asia Pacific (Seoul), Asia Pacific (Tokyo), Asia Pacific (Mumbai) and South America (Sao Paulo)).
- When you are logged into the AWS EC2 console you are always operating in one of these regions.

# What is a Region?

- Current region shown in the upper right corner of console
- It is important to pay attention to what region you are using for several reasons.
  - When you create an EC2 instance (EBS volume, etc) in one region you won't see it in another region.
  - The cost to use many AWS resources varies by region.
  - The region may influence network performance when you are accessing the instance, especially if you need to transfer large amounts of data in or out.
  - Billing is tracked separately for each region
  - Generally you should choose a region that is close to you or your users. But cost is also a consideration.

# What is difference between the 'Start', 'Stop', 'Reboot', and 'Terminate' (Instance States)?

- Start – turn on an EC2 instance that you have previously created
- Stop – turn off an EC2 instance that you have previously created
- Reboot – restart an EC2 instance
- Terminate – permanently stop and destroy an EC2 instance. Any associated EBS volumes may also be destroyed at this time depending on configuration

# What is an AMI/snapshot?

- AMI (Amazon Machine Image) – a template that specifies how to launch EC2 instances
  - Root volume with operating system (OS), pre-installed applications, etc
  - Launch permissions determine who can use the AMI
  - Specification of (data) volumes to attach when launched
- You can create an AMI for any instance you have created/configured
- AMI can be made public for sharing (region-specific)
- Creating an AMI involves creating a snapshot of the root and any attached volumes. You will be charged to store this snapshot.

# I can not log into my EC2 instance, what might have gone wrong?

- Is your instance running?
- Are you providing the correct path to your key file?
- Is it the correct key file?
- Have you set the permissions for your key file correctly?
- Did you specify a valid user for your AMI (e.g., ubuntu)?
- Did you specify the correct IP address?
- Does the Security Group for the instance allow access for your connection protocol (e.g., SSH) and location?

# How much does it cost to use AWS EC2 resources?

Linux	RHEL	SLES	Windows	Windows with SQL Standard	Windows with SQL Web
Windows with SQL Enterprise					
Region: US West (Oregon)					
vCPU	ECU	Memory (GiB)	Instance Storage (GB)	Linux/UNIX Usage	
General Purpose - Current Generation					
t2.nano	1	Variable	0.5	EBS Only	\$0.0058 per Hour
t2.micro	1	Variable	1	EBS Only	\$0.0116 per Hour
t2.small	1	Variable	2	EBS Only	\$0.023 per Hour
t2.medium	2	Variable	4	EBS Only	\$0.0464 per Hour
t2.large	2	Variable	8	EBS Only	\$0.0928 per Hour
t2.xlarge	4	Variable	16	EBS Only	\$0.1856 per Hour
t2.2xlarge	8	Variable	32	EBS Only	\$0.3712 per Hour
m4.large	2	6.5	8	EBS Only	\$0.1 per Hour
m4.xlarge	4	13	16	EBS Only	\$0.2 per Hour
m4.2xlarge	8	26	32	EBS Only	\$0.4 per Hour

Data transfer (GB): In: free or \$0.01; Out: free, \$0.01 or \$0.02

EBS storage (GB/Month): \$0.10

S3 storage (GB/Month): \$0.023 standard, \$0.0125 infrequent access, or  
\$0.004 glacier

# Why am I still getting a monthly bill?

- Generally you get an accounting of usage and cost on a 30 day cycle
  - Pricing is per instance-hour (now instance-second!) consumed for each instance type.
  - Also charges for storage, transfers, etc
- Be aware of regions!
- Even when an instance is stopped, storage for root or other EBS volumes persist
- Creating AMIs/snapshots requires storage
- Explore the billing and cost management tools of AWS to track your spending, set warnings, etc

# Amazon AWS documentation

[https://github.com/griffithlab/rnaseq\\_tutorial/wiki/Intro-to-AWS-Cloud-Computing](https://github.com/griffithlab/rnaseq_tutorial/wiki/Intro-to-AWS-Cloud-Computing)

<http://aws.amazon.com/console/>

# In this workshop:

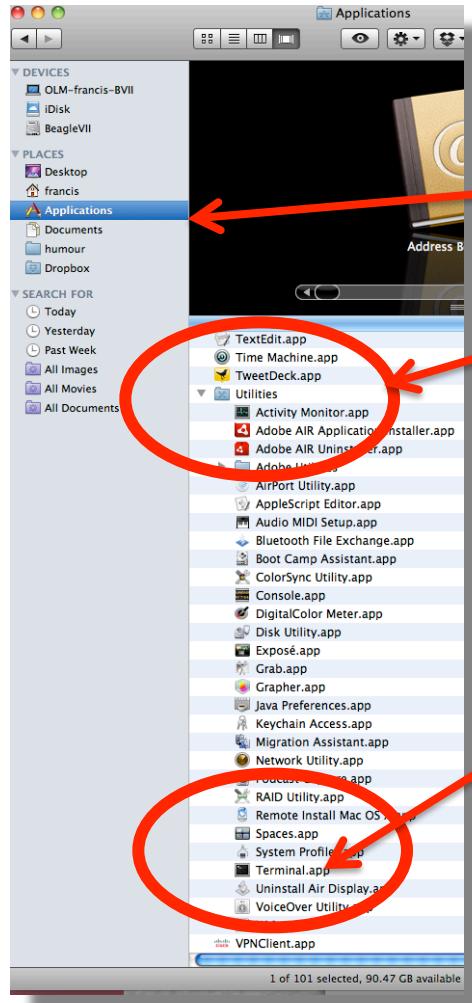
- Some tools (data) are
  - on your computer
  - on the web
  - on the cloud.
- You will become efficient at traversing these various spaces, and finding resources you need, and using what is best for you.
- There are different ways of using the cloud:
  1. Command line (like your own very powerful Unix box)
  2. With a web-browser (e.g. Galaxy): not in this workshop

# Things we have set up:

- Loaded data files to a web server
- We brought up an Ubuntu (Linux) instance, and loaded a whole bunch of software for NGS analysis.
- We will clone this and create separate instances for everybody in the class.
- We've simplified the security: you basically all have the same login and file access, and opened ports. In your own world you would be more secure.

# **Logging into Amazon AWS**

# Opening a ‘terminal session’ on a Mac

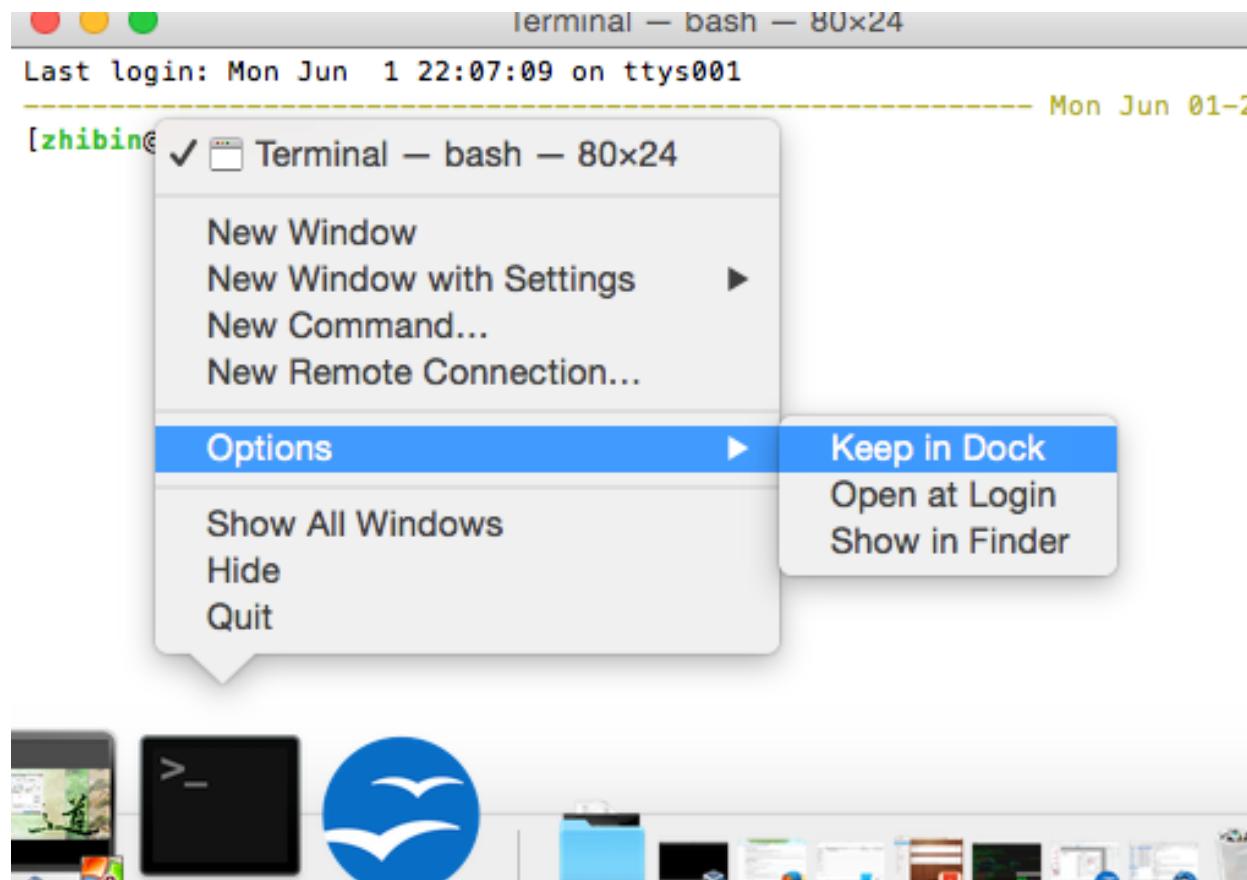


In a Finder window  
‘Applications’ -> ‘Utilities’ -> ‘Terminal’



Or on your dock

# Add the terminal App to your dock



# Creating a working directory on your Mac called ‘cbw’

```
obis-air:~ ogriffit$ pwd
/Users/ogriffit
obis-air:~ ogriffit$ ls
Applications   Desktop          Dropbox        Movies       Public      gittemp     temp
Attachments    Documents         Google Drive   Music        bin         igv
Box Sync       Downloads        Library        Pictures    git         ncbi
obis-air:~ ogriffit$ mkdir cshl
obis-air:~ ogriffit$ cd cshl
obis-air:cshl ogriffit$ ls -la
total 0
drwxr-xr-x  2 ogriffit  staff   68 Nov 13 22:18 .
drwxr-xr-x+ 58 ogriffit  staff  1972 Nov 13 22:18 ..
obis-air:cshl ogriffit$ █
```

mkdir cbw  
cd cbw

# Go to course wiki, “Accessing the cloud” page

## Download the certificate

Secure | [https://bioinformaticsdotca.github.io/rnaseq\\_2018](https://bioinformaticsdotca.github.io/rnaseq_2018)

Welcome

*Ann Meyer*

Introduction to Cloud Computing

*Obi Griffith*

Lecture

- We have set up 30 instances on the Amazon cloud - one for each student. In order to log in to your instance, you will need a security certificate. If you plan on using Linux or Mac OS X, please download [this certificate](#). Otherwise if you plan on using Windows (with Putty and Winscp), please download [this certificate](#).
- Detail instructions can be found [here](#).

Module 1: Introduction to RNA Sequencing Analysis

[https://bioinformaticsdotca.github.io/rnaseq\\_2018](https://bioinformaticsdotca.github.io/rnaseq_2018)

# Go to course wiki, “Accessing the cloud” page

The screenshot shows a web-based wiki interface for the C-SEQTEC 2017 workshop. The left sidebar contains links to Home, Wiki (which is selected), Resources, Roster, Announcements, Drop Box, Polls, Forums, Chat Room, Email Archive, Meetings & Courses, CSHL RSS, Statistics, Site Info, and Help. The main content area has a header "C-SEQTEC 2017: Wiki" and a sub-header "Accessing The Cloud". It includes a search bar and navigation buttons for Home, View, Edit, Info, History, Watch, and a recent changes link. The page content starts with a section titled "Instructions for students to access their AWC EC2 cloud instance". It explains that most hands-on components will be performed on Amazon AWS EC2 instances and provides instructions for logging in via the AWS Console. It also shows terminal commands for connecting to the AWS instance via SSH.

Accessing The Cloud last modified by Obi Griffith on November 10, 2017 11:43:40 AM EST

## Instructions for students to access their AWC EC2 cloud instance

Most hands on components of this workshop will be performed on Amazon AWS EC2 instances. Each student is assigned their own instance (instance type) running the Ubuntu operating system. Use the following instructions to log in.

Visit AWS Console to create instance:

- AWS Console Signin: <https://workshops.signin.aws.amazon.com/console>
- User Name: cshl.student
- Password: seqtec2017

Connect to AWS instance via Terminal:

```
wget genomedata.org/seq-tec-workshop/cshl_2017.pem
chmod 400 cshl_2017.pem
ssh -i cshl_2017.pem ubuntu@YOUR_IP_ADDRESS
```

Be the first to comment

# Login to AWS console



**i Coming Soon: Changes to Multi-Factor Authentication (MFA)**

Entry of an MFA security code for IAM users will move from this sign-in page to a subsequent page

Account: 364840684323

User Name: cshl.student

Password: .....  
 I have an MFA Token (more info)

**Sign In**

[Sign-in using root account credentials](#)



<https://workshops.signin.aws.amazon.com/console>

# Select "EC2" service

The screenshot shows the AWS Management Console homepage. At the top, there is a navigation bar with icons for AWS, Services, Edit, and user information (cshl.student @ 3648-4068-4323, Oregon, Support). Below the navigation bar is a section titled "Shortcuts and Recently Viewed Services" featuring two boxes: one for IAM (with a key icon) and one for EC2 (with a server icon). A large red arrow points to the EC2 box. To the right of this section is a "Service Health" status bar indicating all services are operating normally (updated Nov 10 2016 15:34:00 GMT-0600). A large red arrow also points to the "Oregon" region selection in the top navigation bar. The main content area contains sections for "Quick Starts" (Build a web app, Launch a Virtual Machine (EC2 Instance), Back up your files, Build a back end for your mobile app, Host a static website, Analyze big data) and "AWS Services" (Compute, Storage & Content Delivery, Database, Developer Tools, Management Tools, Security & Identity, Internet of Things, Game Development, Mobile Services, Application Services). A search bar labeled "Search services" is located above the service lists. On the far right, large text reads "Make sure you are in Oregon region". Below this, links for the Amazon Appstore, Google Play, and iTunes are provided. Further down are links for AWS Marketplace and Feedback.

Shortcuts and Recently Viewed Services

AWS Services

Compute

Storage & Content Delivery

Database

Developer Tools

Management Tools

Security & Identity

Internet of Things

Game Development

Mobile Services

Application Services

Service Health

All services are operating normally.  
Updated Nov 10 2016 15:34:00 GMT-0600

View Dashboard

Amazon Appstore, Google Play, or iTunes.

AWS Marketplace

Feedback

Make sure you are in Oregon region

# Launch a new Instance

Screenshot of the AWS EC2 Dashboard showing the 'Create Instance' section. A red arrow points to the 'Launch Instance' button.

The dashboard shows the following resource counts:

Category	Count
Running Instances	4
Volumes	7
Key Pairs	3
Placement Groups	0
Elastic IPs	0
Snapshots	9
Load Balancers	0
Security Groups	2

A message box suggests using Chef recipes and OpsWorks.

**Create Instance**

To start using Amazon EC2 you will want to launch a virtual server, known as an Amazon EC2 instance.

**Launch Instance** 

Note: Your instances will launch in the US West (Oregon) region

**Service Health**

**Service Status:**

- US West (Oregon): This service is operating normally

**Availability Zone Status:**

- us-west-2a: Availability zone is operating normally
- us-west-2b: Availability zone is operating normally
- us-west-2c: Availability zone is operating normally

**Scheduled Events**

US West (Oregon): No events

**AWS Marketplace**

Find free software trial products in the AWS Marketplace from the [EC2 Launch Wizard](#). Or try these popular AMIs:

- Tableau Server (10 users)
- Provided by Tableau
- Rating ★★★★☆
- Pay by the hour for Tableau software and AWS usage
- [View all Business Intelligence](#)
- SAP HANA One 244GiB
- Provided by SAP America, Inc
- Rating ★★★★☆

# Choose an AMI – Find the CSHL SEQTEC 2017 AMI in the Community AMIs

Step 1: Choose an Amazon Machine Image (AMI)

An AMI is a template that contains the software configuration (operating system, application server, and applications) required to launch your instance. You can select an AMI provided by AWS, our user community, or the AWS Marketplace; or you can select one of your own AMIs.

Quick Start

My AMIs

AWS Marketplace

**Community AMIs**

Operating system

- Amazon Linux
- Cent OS
- Debian
- Fedora
- Gentoo
- OpenSUSE
- Other Linux
- Red Hat
- SUSE Linux
- Ubuntu
- Windows

cshl\_seqtec\_2015\_v2 - ami-28130249

Root device type: ebs Virtualization type: hvm

cshl\_seqtec\_2015\_v2\_noworkspace - ami-e9100188

Root device type: ebs Virtualization type: hvm

Select 64-bit

Select 64-bit

Search for: "cshl\_seqtec\_2017\_v1" (US West - Oregon)

# Choose “m4.2xlarge” instance type, then “Next: Configure Instance Details”.

Screenshot of the AWS EC2 instance creation wizard, Step 2: Choose an Instance Type. The instance type **m4.2xlarge** is selected. A red arrow points to the **m4.2xlarge** row in the table, and a red box highlights the **Next: Configure Instance Details** button.

Step 2: Choose an Instance Type

Amazon EC2 provides a wide selection of instance types optimized to fit different use cases. Instances are virtual servers that can run applications. They have varying combinations of CPU, memory, storage, and networking capacity, and give you the flexibility to choose the appropriate mix of resources for your applications. [Learn more](#) about instance types and how they can meet your computing needs.

Filter by: All instance types Current generation Show/Hide Columns

Currently selected: m4.2xlarge (26 ECUs, 8 vCPUs, 2.4 GHz, Intel Xeon E5-2676v3, 32 GiB memory, EBS only)

	Family	Type	vCPUs	Memory (GiB)	Instance Storage (GB)	EBS-Optimized Available	Network Performance
<input type="checkbox"/>	General purpose	t2.micro <small>Free tier eligible</small>	1	1	EBS only	-	Low to Moderate
<input type="checkbox"/>	General purpose	t2.small	1	2	EBS only	-	Low to Moderate
<input type="checkbox"/>	General purpose	t2.medium	2	4	EBS only	-	Low to Moderate
<input type="checkbox"/>	General purpose	t2.large	2	8	EBS only	-	Low to Moderate
<input type="checkbox"/>	General purpose	m4.large	2	8	EBS only	Yes	Moderate
<input type="checkbox"/>	General purpose	m4.xlarge	4	16	EBS only	Yes	High
<input checked="" type="checkbox"/>	General purpose	<b>m4.2xlarge</b>	<b>8</b>	<b>32</b>	EBS only	<b>Yes</b>	<b>High</b>
<input type="checkbox"/>	General purpose	m4.4xlarge	16	64	EBS only	Yes	High

Cancel Previous Review and Launch **Next: Configure Instance Details**

# Select "Protect against accidental termination", then "Next: Add Storage".

The screenshot shows the AWS Launch Wizard interface for launching an Amazon Linux 2 AMI instance. The top navigation bar includes 'AWS', 'Services', 'Edit', and user information 'cshl.student @ 3648-4068-4323' for the Oregon region.

The current step is '3. Configure Instance'. Below it, the steps are listed: 1. Choose AMI, 2. Choose Instance Type, 3. Configure Instance (highlighted in orange), 4. Add Storage, 5. Tag Instance, 6. Configure Security Group, and 7. Review.

**Step 3: Configure Instance Details**

Configure the instance to suit your requirements. You can launch multiple instances from the same AMI, request Spot instances to take advantage of the lower pricing, assign an access management role to the instance, and more.

Number of instances: 1

Purchasing option: Request Spot instances

Network: vpc-ebcc188e (172.31.0.0/16) (default)

Subnet: No preference (default subnet in any Availability Zone)

Auto-assign Public IP: Use subnet setting (Enable)

IAM role: None

Shutdown behavior: Stop

Enable termination protection:  Protect against accidental termination (highlighted with a red arrow)

Monitoring:  Enable CloudWatch detailed monitoring  
Additional charges apply.

Tenancy: Shared tenancy (multi-tenant hardware)  
Additional charges will apply for dedicated tenancy.

Advanced Details

Cancel Previous Review and Launch **Next: Add Storage** (button highlighted with a red border)

You should see "snap-xxxxxxx" (32GB) and "snap-xxxxxxx" (250GB) as the two storage volumes selected. Then, "Next: Tag Instance"

Screenshot of the AWS EC2 instance creation wizard, Step 4: Add Storage.

The screenshot shows the "Add Storage" section where two EBS volumes are selected:

Type	Device	Snapshot	Size (GiB)	Volume Type	IOPS	Delete on Termination	Encrypted
Root	/dev/sda1	snap-6f450833	32	General Purpose (SSD)	96 / 3000	<input checked="" type="checkbox"/>	Not Encrypted
EBS	/dev/sdb	snap-11e6954e	500	General Purpose (SSD)	1500 / 3000	<input checked="" type="checkbox"/>	Not Encrypted

A red box highlights the two selected volumes. Below the table, there is a note about free tier usage:

 Free tier eligible customers can get up to 30 GB of EBS General Purpose (SSD) or Magnetic storage. [Learn more](#) about free usage tier eligibility and usage restrictions.

At the bottom right, the "Next: Tag Instance" button is highlighted with a red box.

**Create a tag like “Name=ObiGriffith” [use your own name]. Then hit "Next: Configure Security Group".**

The screenshot shows the AWS EC2 instance creation wizard at Step 5: Tag Instance. The top navigation bar includes icons for Home, AWS Services, Edit, and account information (cshl.student @ 3648-4068-4323, Oregon, Support). Below the navigation is a progress bar with steps 1 through 7. Step 5, "Tag Instance", is highlighted with an orange underline. The main area is titled "Step 5: Tag Instance" with the sub-instruction: "A tag consists of a case-sensitive key-value pair. For example, you could define a tag with key = Name and value = Webserver. [Learn more](#) about tagging your Amazon EC2 resources." It features two input fields: "Key" (127 characters maximum) and "Value" (255 characters maximum). The "Value" field contains the text "ObiGriffith". A red arrow points upwards from the bottom of the page towards this value field. At the bottom are buttons for "Create Tag" (disabled), "Cancel", "Previous", "Review and Launch" (highlighted in blue), and "Next: Configure Security Group" (surrounded by a red box).

# Select an Existing Security Group, choose "SSH\_HTTP". Then hit "Review and Launch".

Screenshot of the AWS EC2 instance creation wizard Step 6: Configure Security Group.

The screenshot shows the configuration of an existing security group named "SSH\_HTTP". A red arrow points to the "Select an existing security group" radio button, and another red arrow points to the selected security group "SSH\_HTTP".

**Inbound rules for sg-4e1b6128 (Selected security groups: sg-4e1b6128)**

Type	Protocol	Port Range	Source
HTTP	TCP	80	0.0.0.0/0
SSH	TCP	22	0.0.0.0/0

At the bottom right, the "Review and Launch" button is highlighted with a red box.

# Review the details of your instance, note the warnings, then hit Launch

The screenshot shows the AWS Step 7: Review Instance Launch interface. At the top, there's a navigation bar with tabs: 1. Choose AMI, 2. Choose Instance Type, 3. Configure Instance, 4. Add Storage, 5. Tag Instance, 6. Configure Security Group, and 7. Review. The 7. Review tab is highlighted.

**Step 7: Review Instance Launch**

Please review your instance launch details. You can go back to edit changes for each section. Click **Launch** to assign a key pair to your instance and complete the launch process.

**Your instance configuration is not eligible for the free usage tier**

To launch an instance that's eligible for the free usage tier, check your AMI selection, instance type, configuration options, or storage devices. Learn more about [free usage tier eligibility and usage restrictions](#).

[Don't show me this again](#)

**Improve your instances' security. Your security group, SSH\_HTTP\_8081\_IN\_ALL\_OUT, is open to the world.**

Your instances may be accessible from any IP address. We recommend that you update your security group rules to allow access from known IP addresses only. You can also open additional ports in your security group to facilitate access to the application or service you're running, e.g., HTTP (80) for web servers. [Edit security groups](#)

**AMI Details**

cshl\_seqtec\_2015\_v2 - ami-28130249

Root Device Type: ebs Virtualization type: hvm

**Instance Type**

Instance Type	ECUs	vCPUs	Memory (GiB)	Instance Storage (GB)	EBS-Optimized Available	Network Performance
m3.2xlarge	26	8	30	2 x 80		High

**Security Groups**

[Edit security groups](#)

Cancel Previous **Launch**

# Choose an existing key pair: "cshl\_2017" and then Launch.

The screenshot shows the AWS Step 7: Review Instance Launch interface. The main page displays instance launch details, including an AMI selection (cshl\_seqtec\_2015\_v4), an instance type (t2.micro), and a security group (sg-4e1b6128). A modal dialog box titled "Select an existing key pair or create a new key pair" is overlaid on the page. The dialog contains instructions about key pairs and a note about selecting a key pair for the instance. It features a dropdown menu for selecting an existing key pair, which has "CSHL\_2016" selected. A checkbox for acknowledging access to the private key file is checked. Red arrows point from the text "Choose an existing key pair" to the dropdown menu, from "Select a key pair" to the dropdown menu, and from the acknowledgement note to the checkbox. A red box highlights the "Launch Instances" button at the bottom right of the dialog. The overall background shows the AWS navigation bar and other launch steps.

# View Instances to see your new instance spinning up!

The screenshot shows the AWS CloudWatch Launch Status page. At the top, there's a navigation bar with icons for CloudWatch Metrics, AWS Lambda, Services, Edit, and Support, along with user information (cshl.student @ 3648-4068-4323, Oregon). Below the navigation is a section titled "Launch Status" with a green checkmark icon and the message "Your instances are now launching". It indicates that instance launch i-45e4089f has been initiated and provides a link to view the launch log. There's also a section for "Get notified of estimated charges" with a speech bubble icon, explaining how to create billing alerts for email notifications when charges exceed a defined amount. The main content area is titled "How to connect to your instances" and includes instructions for monitoring instance status and connecting to them. A sidebar lists helpful resources like connecting to a Linux instance, learning about the AWS Free Usage Tier, and using the Amazon EC2 User Guide and Discussion Forum. At the bottom, there are links for creating status check alarms, attaching EBS volumes, and managing security groups, followed by a prominent blue "View Instances" button.

AWS Services Edit Support cshl.student @ 3648-4068-4323 Oregon

## Launch Status

**Your instances are now launching**

The following instance launches have been initiated: i-45e4089f [View launch log](#)

**Get notified of estimated charges**

Create billing alerts to get an email notification when estimated charges on your AWS bill exceed an amount you define (for example, if you exceed the free usage tier).

### How to connect to your instances

Your instances are launching, and it may take a few minutes until they are in the **running** state, when they will be ready for you to use. Usage hours on your new instances will start immediately and continue to accrue until you stop or terminate your instances.

Click [View Instances](#) to monitor your instances' status. Once your instances are in the **running** state, you can **connect** to them from the Instances screen. [Find out](#) how to connect to your instances.

▼ Here are some helpful resources to get you started

- [How to connect to your Linux instance](#)
- [Amazon EC2: User Guide](#)
- [Learn about AWS Free Usage Tier](#)
- [Amazon EC2: Discussion Forum](#)

While your instances are launching you can also

[Create status check alarms](#) to be notified when these instances fail status checks. (Additional charges may apply)

[Create and attach additional EBS volumes](#) (Additional charges may apply)

[Manage security groups](#)

[View Instances](#)

# Find YOUR instance, select it, and then hit connect for instructions on how to connect (It may take some time for your instance to be ready)

The screenshot shows the AWS EC2 Dashboard. The 'Instances' section is selected, displaying a list of running instances. A red arrow points to the 'ObiGriffith' instance, which is highlighted with a blue selection bar. Above the instance list, there is a navigation bar with tabs: 'Launch Instance', 'Connect' (which is highlighted with a red box), and 'Actions'. Below the instance list, a detailed view for the selected instance (i-45e4089f) is shown. The instance details include:

Description	Value
Instance ID	i-45e4089f
Instance state	running
Instance type	m3.2xlarge
Private DNS	ip-172-31-4-176.us-west-2.compute.internal
Public DNS	ec2-52-33-240-196.us-west-2.compute.amazonaws.com
Public IP	52.33.240.196
Elastic IP	-
Availability zone	us-west-2c

# Take note of your Public DNS/IP and the instructions on changing permissions for the key file (Note, we will login as ubuntu NOT root)

The screenshot shows the AWS EC2 Dashboard with the 'Instances' section selected. A modal window titled 'Connect To Your Instance' is open, providing instructions for connecting to the instance. The modal includes options for connecting with a standalone SSH client or a Java SSH Client, steps for accessing the instance using an SSH client, and a specific command to change file permissions. It also provides an example SSH command and a note about the default AMI username. The background shows a list of instances with their Public DNS and IP addresses.

Public DNS | Public IP

Public DNS	Public IP	State
ec2-52-33-240-196.us...	52.33.240.196	C
ec2-52-34-44-168.us-w...	52.34.44.168	in
ec2-52-10-59-49.us-w...	52.10.59.49	C
ec2-52-34-43-79.us-w...	52.34.43.79	C
ec2-52-11-219-138.us...	52.11.219.138	in

**Connect To Your Instance**

I would like to connect with  A standalone SSH client  A Java SSH Client directly from my browser (Java required)

To access your instance:

1. Open an SSH client. (find out how to [connect using PuTTY](#))
2. Locate your private key file (CSHL.pem). The wizard automatically detects the key you used to launch the instance.
3. Your key must not be publicly viewable for SSH to work. Use this command if needed:  
`chmod 400 CSHL.pem`
4. Connect to your instance using its Public IP:  
**52.33.240.196**

Example:

```
ssh -i "CSHL.pem" root@52.33.240.196
```

Please note that in most cases the username above will be correct, however please ensure that you read your AMI usage instructions to ensure that the AMI owner has not changed the default AMI username.

If you need any assistance connecting to your instance, please see our [connection documentation](#).

**Close**

Instance: i-45e40891

Description

Instance type: m3.xlarge

Elastic IP

Private DNS: ip-172-31-4-176.us-west-2.compute.internal

Availability zone: us-west-2c

Feedback English

© 2008 - 2015, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use

# Use ‘wget’ at command line to download pem file

```
Obis-MacBook-Air:~ ogriffit$ cd cshl/
Obis-MacBook-Air:cshl ogriffit$ wget http://genomedata.org/seq-tec-workshop/cshl_2017.pem
--2017-11-10 11:29:03--  http://genomedata.org/seq-tec-workshop/cshl_2017.pem
Resolving genomedata.org... 34.210.251.2
Connecting to genomedata.org|34.210.251.2|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 1696 (1.7K)
Saving to: 'cshl_2017.pem'

cshl_2017.pem          100%[=====] 1.66K --.-KB/s in 0s

2017-11-10 11:29:03 (116 MB/s) - 'cshl_2017.pem' saved [1696/1696]
```

```
cd ~/cshl
wget genomedata.org/seq-tec-workshop/cshl_2017.pem
```

# Viewing the ‘key’ file once downloaded

```
cat CBWNY.pem
```

```
obis-air:cshl ogriffit$ cd ~/cshl/
obis-air:cshl ogriffit$ ls -la
total 8
drwxr-xr-x  3 ogriffit  staff   102 Nov 13 22:21 .
drwxr-xr-x+ 58 ogriffit  staff  1972 Nov 13 22:18 ..
-rw-r-----@ 1 ogriffit  staff  1696 Nov 13 22:21 CSHL.pem
obis-air:cshl ogriffit$ cat CSHL.pem
-----BEGIN RSA PRIVATE KEY-----
MIIEpgIBAAKCAQEAvJ5gwmtby9QZ2Idz+ugiEQQHW6Ps0ZAZFvr+mWDnM4pKpccaVmDh7XjcE0LF
0kJzaP9+jj0kSF0yNinitoB32DgrmVhgNhyheEqH5XMn28szxUj1Eu0NXAogNuY7mWMo6MoWssSW
Rqy+rj19vMGQn5rsnMLjCM1smebPoqY0L8EPa1ccRbdGXG1dMTlCC1ho/Hk9bZweamGiZLaAWVm
z0K/L0zxgY3K4cwaL48HV6oGuMh5lTDpnobxXghQ4oC5Mej+DpCRF8C+EG2uNDuyuLzRJfQmFBV2
GKDWDwhdgGmKmX9IpMT9ubvNoQPy0vYLvM80eG3cMbZ2IzpaNryihwIDAQABoIBAQCYT0TvF04
a3DdCEE/C/rN9HMaS+bjFkm0kp9RTi15XJhTPvBmptjzibA6gWJfDaXgKIQGbzxJrEkxwCR2IB03v
0LV7jEcomZ2ggRMDPeJitFoUCuDnkZZtivppSk2az0zeaD+0/ZeqPx0L+Yr+7HSbpVLVoxEV/l5a
xDuCawBMSY2cnGWkfEB1SPnB6fGZj8luGzv0aP/CETx/K78TIS56m4yrTIQIeEPfFt/PQr/EUqoL
7co5oy9K3sD1noPLDhk3vJa1VNrMjHkMZLkbZua0HPzgSQHninm80Ca25WWTGsSZ8vQsBIUTlGI1
W7lzXH3wD1jJNd+03QK4bnKaZ+DZAoGBAPVpisa49JY/6K2f9B8naqtX/ljzVWTl3Q7r6t6uh21Y
oexmC8ej2wQwd0qNjZWVxSMVksIwdM6xcsBIJRMmltWTVdmD0fkDv0fjd8CM4nctH76tvSvZz02e
qI9wSshHY1fh+09CoLZeefSURxqWbkJfREjoZ4UGUWMi3k1rxC9AoGBAMTB1BB0WQ+5ojzQYu0L
Q4YrsIPg1/ni0WmJ+05vcTCJ2aeI88VhK5c2PoXPWWiJ9CdD2VFZDiCm2XuJA5iwJmnhuwGGHHe
BFBqEF/ueJrW+r43pRcYRuRIXjih4mQQLK4Zemecym5fAHvxZxq4fs2kWfMPySfaVufcP0VC7X6T
AoGBAMhro0xbxFQwaU0yh9oRhMneGPhn8WtvVjNjc/LcMfmZEtRPGnuhF965/hJCvEhXgiH+8lXo
4NwUixBVtXnA/P0WX5Ea2ykIth2Kkx0Qlb14SEGh7RZ0saRiLqmcZ9gXFpkm6rimByrDMezVr
nU7CcwNWSB0jaOgluZoJv6k5AoGBAJJuFsmD5ZhkaS+lTpnlZtXDIk5XsMkYQGQpS0clzqufQPI
UtPEm3Jv9lwTktDQSpqmTifShUcbpaPgtoJ+JjiKvGh7QbxKK7II00kULG760SD+S0U972Rdj3Q
M1aRWHWx1h1kH0vDXFLhuAAU6poVBLR2PRPLbf4k1hmvt05xtAoGBAJVQy1GF8uVNwkOCNzLIqmkY
uk9M24hfqn3N2GY3Zgqf43bD4kdYgL4rvsgp08QzotPf+19kVlCv0ciolSjEHLyUdlyPGzj4CTTH
1f1RoGHmYzVn9VuFTu4hJ17J+uwgXgIr9Sx/UTjwkmCjPf7CEyIuGxaThG/ZoR9stufZB5db
-----END RSA PRIVATE KEY-----obis-air:cshl ogriffit$
```

# Changing file permissions of your ‘key’ file (Mac/Linux)

ls -l (long listing)

```
drwx-----+ 67 ogriffit staff 2278 22 May 21:25 ../  
-rw-r--r--@ 1 ogriffit staff 1696 22 May 21:31 CBWNY.pem  
rwx : owner  
rwx : group  
rwx: world  
r read (4)  
w write (2)  
x execute (1)
```

Whichever way you add these 3 numbers, you know which integers were used (6 is always 4+2, 5 is 4+1, 4 is by itself, 0 is none of them etc ...)

So, when you have:

**chmod 600 <file name>**

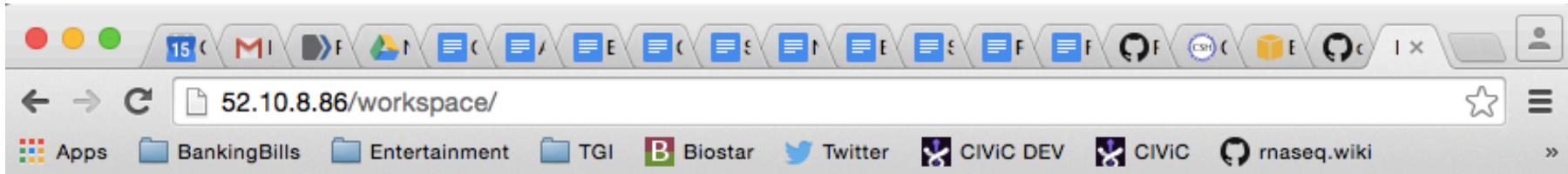
It is “r” for the file owner **only**

# Logging into your instance

## Mac/Linux

```
cd ~/cbw  
chmod 600 CBWNY.pem  
ssh -i CBWNY.pem ubuntu@[YOUR PUBLIC IP]
```

# Copying files from AWS to your computer (using a web browser)



## Index of /workspace

	<a href="#">Name</a>	<a href="#">Last modified</a>	<a href="#">Size</a>	<a href="#">Description</a>
	<a href="#">Parent Directory</a>	-		
	<a href="#">Homo_sapiens/</a>	2015-11-13 06:45	-	
	<a href="#">README.txt</a>	2014-06-17 23:53	5.3K	
	<a href="#">bam-demo/</a>	2015-11-14 21:03	-	
	<a href="#">data/</a>	2015-11-13 01:39	-	
	<a href="#">scratch/</a>	2015-11-13 19:43	-	
	<a href="#">tools/</a>	2015-11-13 01:54	-	

*Apache/2.4.7 (Ubuntu) Server at 52.10.8.86 Port 80*

**http://[YOUR PUBLIC DNS OR IP]/**

# Logging out of your instance

Mac/Linux – simply type exit

```
exit
```

Note, this disconnects the terminal session (ssh connection) to your cloud instance. But, your cloud instance is still running! See next slide for how to stop your instance.

# When you are done for the day you can “Stop” your instance – Don’t Terminate!

The screenshot shows the AWS EC2 Dashboard. On the left, there's a sidebar with various navigation options like EC2 Dashboard, Events, Tags, Reports, Limits, Instances, Images, AMIs, and more. A red arrow points to the 'Instances' tab. In the main area, a table lists several instances. One instance, 'instructor\_test2', is selected. A context menu is open over this instance, with 'Instance State' expanded. Under 'Instance State', another red arrow points to the 'Stop' option, which is highlighted. Below the table, it says 'Instance: i-068e6cdc (instructor\_test2) Public DNS: ec2-52-10-8-86.us-west-2.compute.amazonaws.com'. A large red box encloses the following text:

Go to AWS EC2 Dashboard, select “Instances” tab, then find your instance. Right-click and chose ‘Instance State’ -> ‘Stop’

# Next morning, you can “Start” your instance again

The screenshot shows the AWS EC2 Dashboard. On the left, there's a sidebar with various navigation options like EC2 Dashboard, Events, Tags, Reports, Limits, Instances, Images, AMIs, and more. A red arrow points to the 'Instances' link in the sidebar. In the main content area, there's a table of instances. One instance named 'JasonWalker' is highlighted with a red box and has a context menu open over it. The context menu has several options: Connect, Get Windows Password, Launch More Like This, Instance State, Instance Settings, Image, Networking, CloudWatch Monitoring, and Terminate. The 'Instance State' option is expanded, and its sub-option 'Start' is also highlighted with a red arrow. Below the table, there's a status bar showing 'Instance: i-3246aae8 (JasonWalker)' and 'Private IP: 172.31.5.175'. A large red box encloses the entire central area of the dashboard.

Go to AWS EC2 Dashboard, select “Instances” tab, then find your instance. Right-click and chose ‘Instance State’ -> ‘Start’

**When you restart your instance you will need to find your new Public DNS or IP address. Select your instance and “Connect” or look in Description tab. Then go back to instructions for “Logging into your instance”**

The screenshot shows the AWS EC2 Instances page. On the left, there's a sidebar with navigation links like EC2 Dashboard, Events, Tags, Reports, Limits, Instances, AMIs, and more. The main area displays a table of instances with columns for Name, Instance ID, Instance Type, Availability Zone, Instance State, Status Checks, Alarm Status, and Public DNS. A red arrow points to the 'Connect' button at the top of the table. Below the table, a specific instance is selected: 'instructor\_test2' (Instance ID: i-068e6cdc). The 'Description' tab is active, showing details such as Instance ID, Public DNS (ec2-52-10-8-86.us-west-2.compute.amazonaws.com), Instance state (running), and Public IP (52.10.8.86). Another red arrow points to the Public IP address.

Name	Instance ID	Instance Type	Availability Zone	Instance State	Status Checks	Alarm Status	Public DNS
instructor_test2	i-068e6cdc	m3.2xlarge	us-west-2c	running	2/2 checks ...	None	ec2-52-10-8-86.us-wes...
JasonWalker	i-3246aae8	m3.2xlarge	us-west-2c	stopped		None	
pengpeng	i-6740acbd	m3.2xlarge	us-west-2c	stopped		None	
ALesiak	i-0d42aed7	m3.2xlarge	us-west-2c	stopped		None	
djcoughlin	i-3540acef	m3.2xlarge	us-west-2c	stopped		None	
jakesaunders	i-a747ab7d	m3.2xlarge	us-west-2c	stopped		None	
YunjuSung	i-6540acb	m3.2xlarge	us-west-2c	stopped		None	
Jonathan.Wan	i-6640acbc	m3.2xlarge	us-west-2c	stopped		None	
KateD	i-a241ad78	m3.2xlarge	us-west-2c	stopped		None	
JenTudor	i-0e42aed4	m3.2xlarge	us-west-2c	stopped		None	
YanZhang	i-0342aed9	m3.2xlarge	us-west-2c	stopped		None	
ArenMarshall	i-0242aed8	m3.2xlarge	us-west-2c	stopped		None	

Instance: i-068e6cdc (instructor\_test2)    Public DNS: ec2-52-10-8-86.us-west-2.compute.amazonaws.com

Description    Status Checks    Monitoring    Tags

Instance ID: i-068e6cdc    Public DNS: ec2-52-10-8-86.us-west-2.compute.amazonaws.com  
Instance state: running    Public IP: 52.10.8.86

# So, at this point:

- Your Mac desktop is ready for the workshop
- If it is not, you know where to get the information you need
- You know how to login to AWS
- The next step is to login to your linux machine on AWS and learn the basics of a linux command line

We are on a Coffee Break &  
Networking Session