



Cold  
Spring  
Harbor  
Laboratory

# Advanced Sequencing Technologies & Applications

<http://meetings.cshl.edu/courses.html>



Cold  
Spring  
Harbor  
Laboratory

Introduction to cloud computing

Advanced Sequencing Technologies & Applications  
November 6 - 18, 2018

Kelsy Cotto



# 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>



DNA  
Sequencing (bp/\$)

1,000,000

100,000

10,000

1,000

100

10

0

1,000,000,000

100,000,000

10,000,000

1,000,000

100,000

10,000

1,000

100

10

1

REVIEW

The case for cloud computing in genome informatics

Lincoln D Stein\*

Hard disk storage (MB/\$)  
Doubling time=14 mo

Pre-nextgen sequencing (bp/\$)  
Doubling time=19 mo

Nextgen sequencing (bp/\$)  
Doubling time=4 mo0

1990 1992 1994 1996 1998 2000 2002 2003 2004 2005 2006 2007 2008 2010 2012

1 10 100 1,000 10,000 100,000 1,000,000 10,000,000 100,000,000 1,000,000,000

1 10 100 1,000 10,000 100,000 1,000,000 10,000,000 100,000,000 1,000,000,000

# 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**

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

The screenshot shows a web browser window with the following details:

- Title Bar:** C-SEQTEC 2018: Wiki
- Toolbar:** Home, View, Edit, Info, History, Watch, Search: [empty input field], and icons for RSS and Print.
- Page Header:** Accessing The Cloud, Recently visited, and a link to Accessing The Cloud last modified by Malachi Griffith on November 9, 2018 8:55:29 AM EST.
- Section Header:** Instructions for students to access their AWC EC2 cloud instance
- Text:** 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.
- Text:** Visit AWS Console to create instance:
- List:**
  - AWS Console Signin: <https://cshlworkshops.signin.aws.amazon.com/console>
  - User Name: cshl.student
  - Password: seqtec2018
  - Zone: US West (Oregon)
  - Course AMI: cshl-seqtech-2018v2 (ami-0ee7d04a5b96e28c6)
- Text:** Download pem file (right-click and "Save Link As..."; take note of where you download it to):
- List:**
  - [cshl\\_2018\\_student.pem](#)
- Text:** Connect to AWS instance via Terminal:
- Text:**

```
chmod 400 cshl_2018_student.pem
ssh -i cshl_2018_student.pem ubuntu@YOUR_IP_ADDRESS
```

# Login to AWS console

The screenshot shows the AWS sign-in interface. At the top left is the Amazon Web Services logo. Below it is a blue informational box containing the text: "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". The main form fields are: "Account: 364840684323", "User Name: cshl.student" (highlighted with a yellow background), and "Password: .....". There is also a checkbox labeled "I have an MFA Token (more info)". At the bottom is a blue "Sign In" button with a red arrow pointing to it, and a link "Sign-in using root account credentials".

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

# Select "EC2" service

aws Services ▾ Resource Groups ▾

History

Console Home

EC2

Simple Queue Service

Billing

IAM

S3

Find a service by name or feature (for example, EC2, S3 or VM, storage).

Group A-Z

Compute

EC2

Lightsail

ECS

EKS

Lambda

Batch

Elastic Beanstalk

Storage

S3

EFS

S3 Glacier

Storage Gateway

Database

RDS

DynamoDB

ElastiCache

Neptune

Amazon Redshift

Management Tools

CloudWatch

AWS Auto Scaling

CloudFormation

CloudTrail

Config

OpsWorks

Service Catalog

Systems Manager

Trusted Advisor

Managed Services

Media Services

Elastic Transcoder

Kinesis Video Streams

MediaConvert

MediaLive

MediaPackage

MediaStore

MediaTailor

Machine Learning

Amazon SageMaker

Mobile Services

Mobile Hub

AWS AppSync

Device Farm

Security, Identity & Compliance

IAM

Cognito

Secrets Manager

GuardDuty

Inspector

Amazon Macie

AWS Organizations

AWS Single Sign-On

Certificate Manager

Key Management Service

CloudHSM

Directory Service

WAF & Shield

Artifact

Game Development

Amazon GameLift

Desktop & App Streaming

IoT Core

IoT 1-Click

IoT Device Management

IoT Analytics

Greengrass

Amazon FreeRTOS

IoT Device Defender

AR & VR

Make sure you are in Oregon region

# Launch a new Instance

The screenshot shows the AWS EC2 Dashboard. On the left, there's a sidebar with navigation links for EC2 Dashboard, Instances, Images, Elastic Block Store, and Network & Security. The main content area has a title 'Resources' and a summary of resources: 0 Running Instances, 0 Dedicated Hosts, 4 Volumes, 3 Key Pairs, 0 Placement Groups, 1 Elastic IPs, 4 Snapshots, 0 Load Balancers, and 5 Security Groups. Below this is a 'Create Instance' section with a sub-section 'Service Health'. A prominent blue button labeled 'Launch Instance' is highlighted with a large red arrow pointing to it. To the right of the 'Launch Instance' button is a 'Scheduled Events' section showing 'US West (Oregon)' with 'No events'. On the far right, there are sections for 'Account Attributes' (Supported Platforms: VPC, Default VPC), 'Additional Information' (Getting Started Guide, Documentation, All EC2 Resources, Forums, Pricing, Contact Us), and 'AWS Marketplace'.

aws Services ▾ Resource Groups ▾

EC2 Dashboard

- Events
- Tags
- Reports
- Limits

INSTANCES

- Instances
- Launch Templates
- Spot Requests
- Reserved Instances
- Dedicated Hosts
- Scheduled Instances
- Capacity Reservations

IMAGES

- AMIs
- Bundle Tasks

ELASTIC BLOCK STORE

- Volumes
- Snapshots
- Lifecycle Manager

NETWORK & SECURITY

Resources

You are using the following Amazon EC2 resources in the US West (Oregon) region:

0 Running Instances	1 Elastic IPs
0 Dedicated Hosts	4 Snapshots
4 Volumes	0 Load Balancers
3 Key Pairs	5 Security Groups
0 Placement Groups	

Learn more about the latest in AWS Compute from AWS re:Invent 2017 by viewing the [EC2 Videos](#).

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):

Availability Zone Status:

Scheduled Events

US West (Oregon):

No events

Account Attributes

Supported Platforms

- VPC
- Default VPC
- vpc-da25eda2

Resource ID length management

Console experiments

Additional Information

Getting Started Guide

Documentation

All EC2 Resources

Forums

Pricing

Contact Us

AWS Marketplace

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

[Barracuda CloudGen Firewall for AWS -](#)

# Choose an AMI – Find the CSHL SEQTEC 2018 AMI in the My 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.

Search for an AMI by entering a search term e.g. "Windows"

Quick Start

My AMIs ←

AWS Marketplace

Community AMIs

Ownership

Owned by me

Shared with me

cshl-seqtech-2018v1 - ami-05466260c3c44a136

Image for the 2018 Advanced Sequencing Technologies and Applications Course, Cold Spring Harbor Labs

Root device type: ebs Virtualization type: hvm Owner: 577255725291 ENA Enabled: Yes

Select

64-bit (x86)

cshl-seqtech-2018v2 - ami-0ee7d04a5b96e28c6

Advanced Sequencing Technologies and Applications - CSHL 2018

Root device type: ebs Virtualization type: hvm Owner: 577255725291 ENA Enabled: Yes

Select

64-bit (x86)

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

The screenshot shows the AWS EC2 instance creation wizard at Step 2: Choose an Instance Type. The user has selected the m4.2xlarge instance type, which is highlighted in the table below. A red arrow points to the m4.2xlarge row, and a red box highlights the "Next: Configure Instance Details" button at the bottom right.

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	m4.2xlarge	8	32	EBS only	Yes	High
<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'.

The current step is 'Step 3: Configure Instance Details'. The page title is 'Configure Instance Details'. A sub-instruction at the top says: '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.'

The configuration fields include:

- Number of instances:** 1
- Purchasing option:** Request Spot instances (unchecked)
- Network:** vpc-ebcc188e (172.31.0.0/16) (default) - Create new VPC
- Subnet:** No preference (default subnet in any Availability Zone) - Create new subnet
- Auto-assign Public IP:** Use subnet setting (Enable)
- IAM role:** None - Create new IAM role
- Shutdown behavior:** Stop
- Enable termination protection:** Protect against accidental termination (checked)
- Monitoring:** Enable CloudWatch detailed monitoring (unchecked)
- Tenancy:** Shared tenancy (multi-tenant hardware)

A red arrow points to the 'Protect against accidental termination' checkbox. At the bottom right, the 'Next: Add Storage' button is highlighted with a red border.

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

Step 4: Add Storage

Your instance will be launched with the following storage device settings. You can attach additional EBS volumes and instance store volumes to your instance, or edit the settings of the root volume. You can also attach additional EBS volumes after launching an instance, but not instance store volumes. [Learn more](#) about storage options in Amazon EC2.

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

Add New Volume

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.

Cancel Previous **Review and Launch** **Next: Tag Instance**

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

Step 5: Add Tags

A tag consists of a case-sensitive key-value pair. For example, you could define a tag with key = Name and value = Webserver.

A copy of a tag can be applied to volumes, instances or both.

Tags will be applied to all instances and volumes. [Learn more](#) about tagging your Amazon EC2 resources.

Key	(127 characters maximum)	Value	(255 characters maximum)	Instances	Volumes	
Name		KelsyCotto		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<a href="#">X</a>

Add another tag (Up to 50 tags maximum)

**Important: Don't forget to name your instance! (FirstnameLastname)**

Cancel Previous Review and Launch **Next: Configure Security Group**

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

1. Choose AMI    2. Choose Instance Type    3. Configure Instance    4. Add Storage    5. Add Tags    **6. Configure Security Group**    7. Review

### Step 6: Configure Security Group

A security group is a set of firewall rules that control the traffic for your instance. On this page, you can add rules to allow specific traffic to reach your instance. For example, if you want to set up a web server and allow Internet traffic to reach your instance, add rules that allow unrestricted access to the HTTP and HTTPS ports. You can create a new security group or select from an existing one below. [Learn more](#) about Amazon EC2 security groups.

Assign a security group:  Create a new security group  Select an existing security group

Security Group ID	Name	Description	Actions
sg-2ca82259	default	default VPC security group	<a href="#">Copy to new</a>
sg-02a3720173e534fb1	launch-wizard-1	launch-wizard-1 created 2018-11-08T22:48:01.401-05:00	<a href="#">Copy to new</a>
sg-0503d434ee857a234	launch-wizard-2	launch-wizard-2 created 2018-11-08T22:59:54.648-05:00	<a href="#">Copy to new</a>
sg-0eb1322ac3c3d6e54	Mosh	Full access on mosh ports	<a href="#">Copy to new</a>
<input checked="" type="checkbox"/> sg-0e1e3e42604e6f94e	SSH and HTTP	launch-wizard-1 created 2018-11-01T13:09:50.250-05:00	<a href="#">Copy to new</a>

Inbound rules for sg-0e1e3e42604e6f94e (Selected security groups: sg-0e1e3e42604e6f94e)

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

[Cancel](#) [Previous](#) **Review and Launch**

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

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.

**⚠ Improve your instances' security.** Your security group, SSH and HTTP, 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** [Edit AMI](#)

cshl-seqtech-2018v2 - ami-0ee7d04a5b96e28c6

Advanced Sequencing Technologies and Applications - CSHL 2018  
Root Device Type: ebs Virtualization type: hvm

**Instance Type** [Edit instance type](#)

Instance Type	ECUs	vCPUs	Memory (GiB)	Instance Storage (GB)	EBS-Optimized Available	Network Performance
m4.xlarge	13	4	16	EBS only	Yes	High

**Security Groups** [Edit security groups](#)

[Cancel](#) [Previous](#) **Launch**

# Choose an existing key pair: "cshl\_2018\_student" and then Launch.

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  
To launch an instance that's eligible for usage restrictions.

Improve your instances' security  
Your instances may be accessible from the Internet. You can also open additional ports in your security groups.

AMI Details

cshl-seqtech-2018v2 - ami-01234567890123456  
Advanced Sequencing Technology  
Root Device Type: ebs Virtualization Type: hvm

Instance Type

Instance Type	ECUs	vCPUs	Memory (GiB)	Storage	Network	Block Device Mapping
m4.xlarge	13	4	16	EBS only	Yes	

Security Groups

Select an existing key pair or create a new key pair

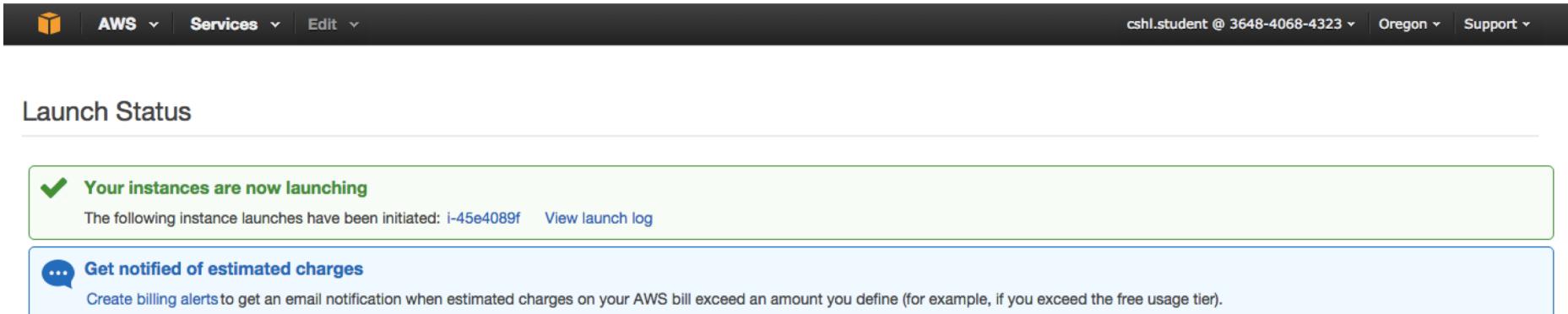
A key pair consists of a **public key** that AWS stores, and a **private key file** that you store. Together, they allow you to connect to your instance securely. For Windows AMIs, the private key file is required to obtain the password used to log into your instance. For Linux AMIs, the private key file allows you to securely SSH into your instance.

Note: The selected key pair will be added to the set of keys authorized for this instance. Learn more about [removing existing key pairs from a public AMI](#).

Choose an existing key pair  
Select a key pair  
cshl\_2018\_student  
 I acknowledge that I have access to the selected private key file (cshl\_2018\_student.pem), and that without this file, I won't be able to log into my instance.

Cancel Launch Instances

# 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, Services, Edit, and user information (cshl.student @ 3648-4068-4323, Oregon, Support). Below the navigation is a section titled "Launch Status". It contains a green box with a checkmark and the text "Your instances are now launching" followed by "The following instance launches have been initiated: i-45e4089f" and a link to "View launch log". There's also a blue box with a speech bubble icon and the text "Get notified of estimated charges" followed by "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. On the left, there's a sidebar with navigation links: EC2 Dashboard, Events, Tags, Reports, Limits, INSTANCES (with 'Instances' selected), Launch Templates, Spot Requests, Reserved Instances, Dedicated Hosts, Scheduled Instances, Capacity Reservations, IMAGES (with 'AMIs' selected), Bundle Tasks, and ELASTIC BLOCK STORE. The main content area has a title bar with 'Launch Instance', 'Connect' (which is highlighted with a red box and arrow), and 'Actions'. Below is a search bar and a table of instances. The table columns are: Name, Instance ID, Instance Type, Availability Zone, Instance State, Status Checks, Alarm Status, and Public DNS (IPv4). There are three instances listed:

Name	Instance ID	Instance Type	Availability Zone	Instance State	Status Checks	Alarm Status	Public DNS (IPv4)
cshl-2018-m...	i-00fe7afc1e68869f5	m5.2xlarge	us-west-2a	stopped	None	None	ec2-34-212-44-228.us...
KelsyCotto	i-00fe7afc1e68869f51	m4.xlarge	us-west-2c	running	Initializing	None	ec2-34-211-56-245.us...
cshl-2018-in...	i-0e3be57bca76f271d	m5.2xlarge	us-west-2a	stopped	None	None	ec2-34-211-56-245.us...

# 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 a modal window titled "Connect To Your Instance". The modal provides instructions for connecting to the instance using a standalone SSH client or a Java SSH Client. It lists four steps for access: 1. Open an SSH client, 2. Locate the private key file, 3. Change permissions on the key file, and 4. Connect using the Public DNS. An example command is provided: `ssh -i "cshl_2018_student.pem" root@ec2-34-211-56-245.us-west-2.compute.amazonaws.com`. A note at the bottom states: "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." A "Close" button is at the bottom right of the modal.

aws Services Resource Groups

EC2 Dashboard Events Tags Reports Limits

INSTANCES Instances Launch Templates Spot Requests Reserved Instances Dedicated Hosts Scheduled Instances Capacity Reservations

IMAGES AMIs Bundle Tasks

ELASTIC BLOCK STORE Volumes Snapshots Lifecycle Manager

NETWORK & SECURITY Security Groups Elastic IPs Placement Groups

Launch

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\_2018\_student.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_2018_student.pem`
4. Connect to your instance using its Public DNS:  
`ec2-34-211-56-245.us-west-2.compute.amazonaws.com`

Example:

```
ssh -i "cshl_2018_student.pem" root@ec2-34-211-56-245.us-west-2.compute.amazonaws.com
```

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](#).

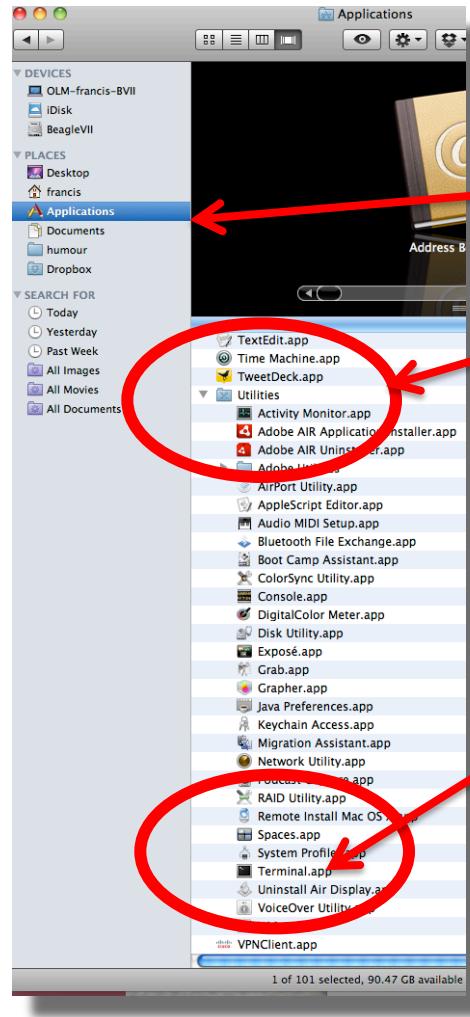
Close

Instance type m4.xlarge IPv6 IPs -

Status Public DNS (IPv4)  
ec2-34-212-44-228.us...  
ec2-34-211-56-245.us...

45.us-west-2.compute.amazonaws.com

# Opening a ‘terminal session’ on a Mac

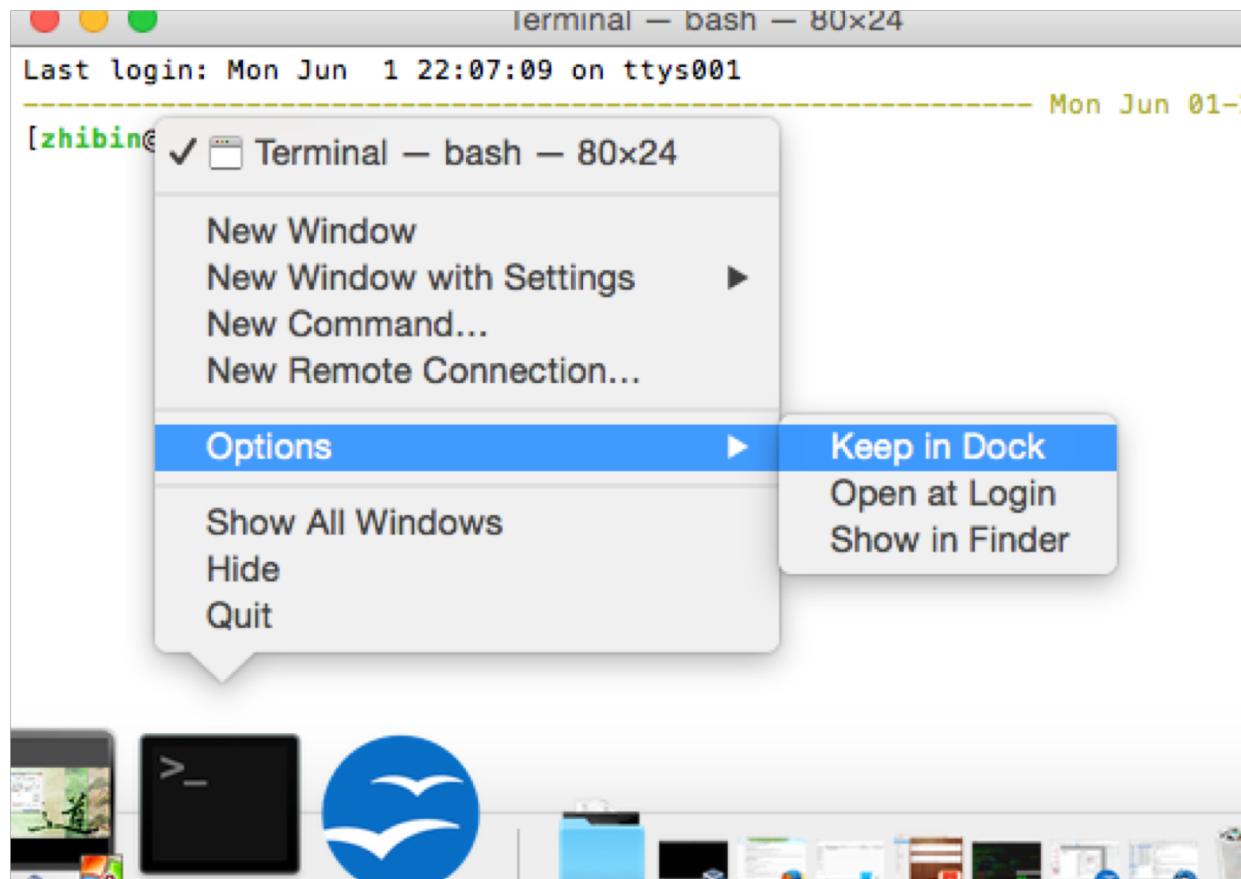


In a Finder window  
'Applications' -> 'Utilities' -> 'Terminal'

Or on your dock



# Add the terminal App to your dock



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

```
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 cshl  
cd cshl

# Download .pem file to cshl directory from course wiki

C-SEQTEC 2018: Wiki

Home View Edit Info History Watch Search:

Accessing The Cloud Recently visited

Accessing The Cloud last modified by Malachi Griffith on November 9, 2018 8:55:29 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://cshlworkshops.signin.aws.amazon.com/console>
- User Name: cshl.student
- Password: seqtec2018
- Zone: US West (Oregon)
- Course AMI: cshl-seqtech-2018v2 (ami-0ee7d04a5b96e28c6)

Download pem file (right-click and "Save Link As..."; take note of where you download it to):

- [cshl\\_2018\\_student.pem](#)

Connect to AWS instance via Terminal:

```
chmod 400 cshl_2018_student.pem  
ssh -i cshl_2018_student.pem ubuntu@YOUR_IP_ADDRESS
```



# Viewing the ‘key’ file once downloaded

cat cshl\_2018\_student.pem

```
GSC-Systemss-MacBook-Pro-3:cshl kcotto$ cat cshl_2018_student.pem
-----BEGIN RSA PRIVATE KEY-----
MIIEpAIBAAKCAQEATgc+HFIf9S1YofCDwMH0C7S/o+t7sr/X660u7P5KhE22EQuiHFVtku4BkPM
oEEUYbtCwBR5dJvoqSoMBfd7/PQnLhPdWuVRVc4XiPZGhsq4E9jZFodzDS9aPAq1aRbHFxbjTCoi
clsKWqcdQQ2g5vLC0P4kqBRnB+KPicuGjPrN5ZrMefZcm57eYTmF9BAWoL67xh+vnNghKH82eXUj
fvk93PuzOKYKAkm5qzkpFTTSjKH5ip264JX44Kr8iNjeycGMpKxD2iZ4CPeiq/w1Zh8PWRfDzlnP
/PzKtrpWW+rkbZ/+hJIe2y8q7IbzUvKCrVzsG1QHfwUzUAZUw5DZ+QIDAQABoIBAQCJqLsLY9d7
U5Uj/T8pfLTl8Ti8JNgotjNDMxfMwND0LRyh0AvhYlncbn4QUlVS1guQoJYb3p3eHZRD6Ka3/VBx
TVWrKOgmYh+6jiJyJQkrETu5MDrbwTlTiMfZtzuifHr0xEPSsCL01Bsh03QNU/cLMrz4avfbRai
c51dyULbAkghWD/FHQzsgQ7KqdM615pLy1u2JvQ0wdlR6emEEB900zfDrSBZzmBGeoJd1EkepEtg
2Lgyjs4CLMxKIiqTMaLY2fCJ3Xm9II5s10ryGoo8lvujKSICJM2zyVIbVgEfwc0lZl020i97IcjU
GSqjDr0/EGu2iCi+8NvQ+u1R0JsVAoGBA070xzZ6k02yLri2Pq7l13D67M9YaG1PeqOrE3M1nwmm
5uAWhmcBAGfGeLH8myJDBmcqtfochsCDyggkI9Ept5UmwxlZ1Ecq8C+Bsj/aDL/jvY+7AxfpK+GIU
L00kXA9FaTICe0VukFDyR4zIlP4sy68iFcKAwIn+tDHywPi+7BzAoGBAoI2bbdvEDRfNB09bi1g
qo5Kbo92oXFewEB55/qHj1hABj+W7hU2V4tavNG1w2RJ9y36Zp0S8QsY1LyiaaZ9i4Zow27rJG1N
2L3B+JeF3ib1JHb94aLiFPsi492k4Wco1QHRSk5dXZykCpwWrkTImgzQErtaQMO35DXG5zWz6gzj
AoGAUqHU7fLvU3PlAmGrRYHhn9E16PHSktJkP4reW/3T9FuCBUXkzmPbRwesHyiSYQ06iByN92LK
SWpzKzqsd7yPCK+XtrgfIfNrS/1ev01YYoYuEZcaH1aj1A8U1RhWOXWwTfnq4rQcNdHCil2moErP
WM25ppwteMrad1c50PtCPB8CgYAkK0Z+PircmPd1Ifof32hi2Tq53YxXGyNuYgQmMi+bXcga1hCH
B30v0ZiLV57Y1Y4vfT24Lhx8xw05jFyzXcvzvhCqCq4MtW66340Nr+bJU0ypmegZ8t2j8tZLura1
2+jPcZiV9/e2rgzbmsDkdIvr59ZrzDhupEjsph5Bk+YVQKBgQC6vNLisnJDX+dbLB6My2D1ZSJ+
paubgw7fi78PsAezNLWX8iHzBBEuDcxVU7TD8fsFIttBEipUERCMAQRxQ0KjrdxRHTc4G7DcDwG
vQ601EwNyPK0Mn8FzEGBb9WT7bk50x9iNox2YYzpRwZdTSCpFaxwErihSaf+VetF2+lijQ==
-----END RSA PRIVATE KEY-----GSC-Systemss-MacBook-Pro-3:cshl kcotto$
```

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

ls -l (long listing)

```
-rw-r--r--@ 1 kcotto staff 1696 Nov 9 09:19 cshl_2018_student.pem
```

    rwx : owner

    rwx : group

    rwx: world

    r read (4)

    w write (2)

    x execute (1)

Which ever 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 400 <file name>**

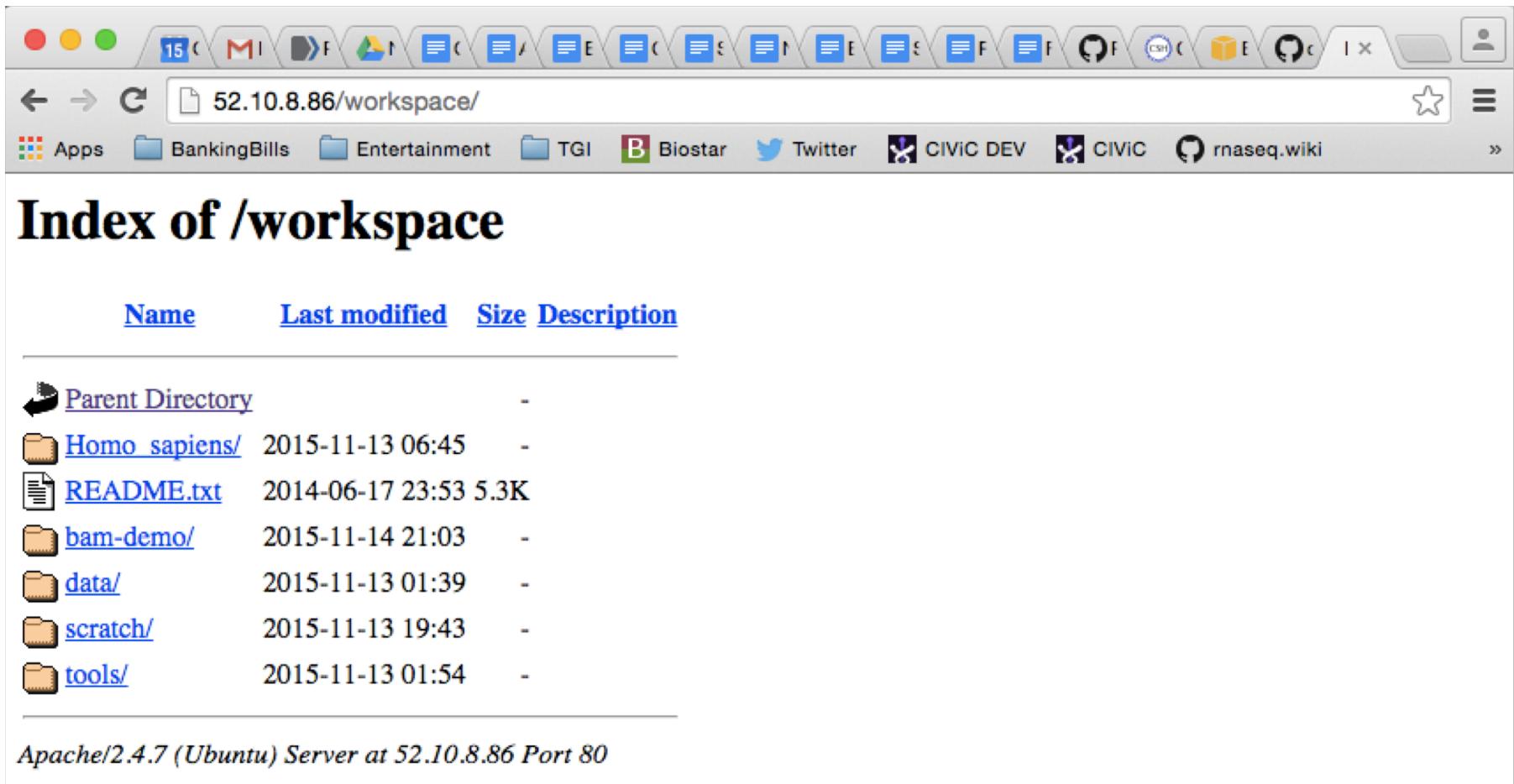
It is “r” for the the file owner **only**

# Logging into your instance

## Mac/Linux

```
cd ~/cshl  
chmod 400 cshl_2018_student.pem  
ssh -i cshl_2018_student.pem ubuntu@[YOUR PUBLIC IP]
```

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



The screenshot shows a Mac OS X desktop environment with a web browser window open. The browser's address bar displays the URL `52.10.8.86/workspace/`. The page content is a file listing titled "Index of /workspace". The table has columns for Name, Last modified, Size, and Description. The "Description" column contains several ellipses ("...") indicating more details are available if expanded.

Name	Last modified	Size	Description
<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. The 'Instances' option is highlighted with a red arrow. In the main content area, there's a table of instances. One instance, 'instructor\_test2', is selected and has a context menu open over it. The menu is titled 'Instance State' and includes options: Connect, Get Windows Password, Launch More Like This, Instance Settings, Image, Networking, CloudWatch Monitoring, Start, Stop, Reboot, and Terminate. The 'Stop' option is highlighted with a red arrow. At the bottom of the dashboard, there's a callout box with 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 sidebar, the 'Instances' tab is selected, indicated by a red arrow. In the main content area, a list of instances is displayed. One instance, 'JasonWalker' (ID i-3246aae8), has a context menu open over it, also indicated by a red arrow. The menu is titled 'Instance State' and includes options: Start, Stop, Reboot, Terminate, and CloudWatch Monitoring. The 'Start' option is highlighted. At the bottom of the dashboard, the instance details are shown: Instance: i-3246aae8 (JasonWalker) and Private IP: 172.31.5.175.

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 Dashboard. The left sidebar lists various services: EC2 Dashboard, Events, Tags, Reports, Limits, Instances, AMIs, Bundle Tasks, Elastic Block Store, Volumes, Snapshots, Network & Security, Security Groups, Elastic IPs, Placement Groups, Key Pairs, Network Interfaces, Load Balancing, Load Balancers, and Auto Scaling. The main content area displays a table of instances. The first instance listed is 'instructor\_test2'. The 'Connect' button in the top navigation bar is highlighted with a red arrow. In the bottom right corner of the instance details, the 'Public IP' field '52.10.8.86' is also highlighted with a red arrow.

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-west-2...
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

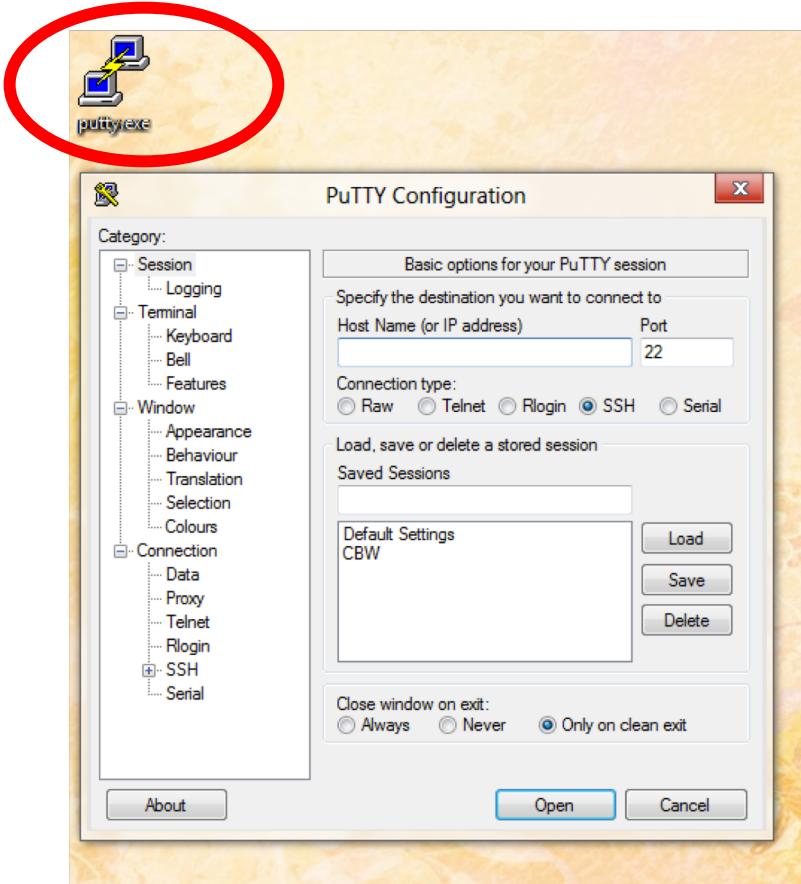
# Break

# Obtain the course SSH key file

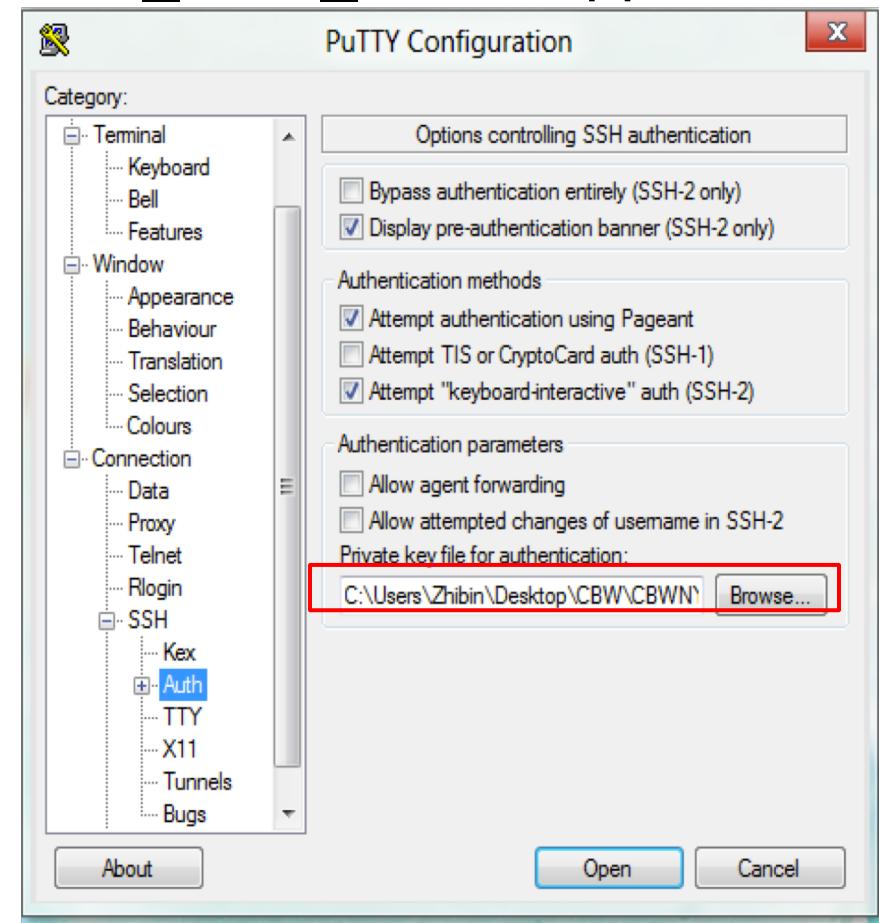
- NOTE for Mac users. You will need to use a “.pem” file
- **NOTE for Windows Users.** You will need to use a “.ppk” file instead.
  - This is created from the “.pem” file.
  - <https://aws.amazon.com/premiumsupport/knowledge-center/convert-pem-file-into-ppk/>
- The SSH key file will be used to securely login to your student instance on the cloud

# Logging into your instance (Windows)

Open PuTTY

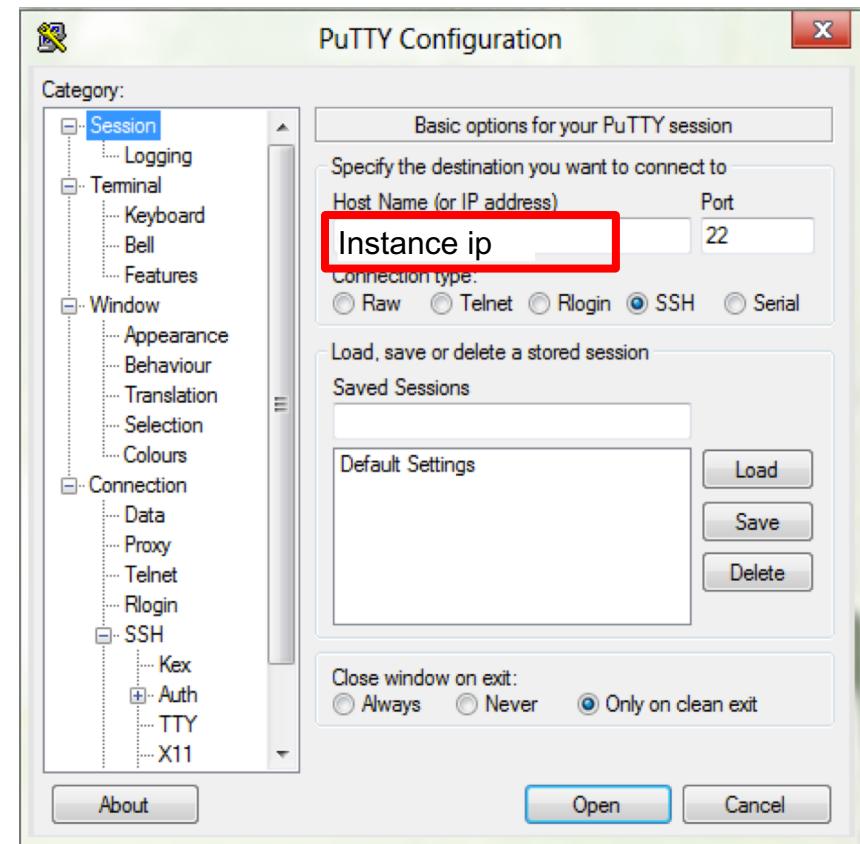
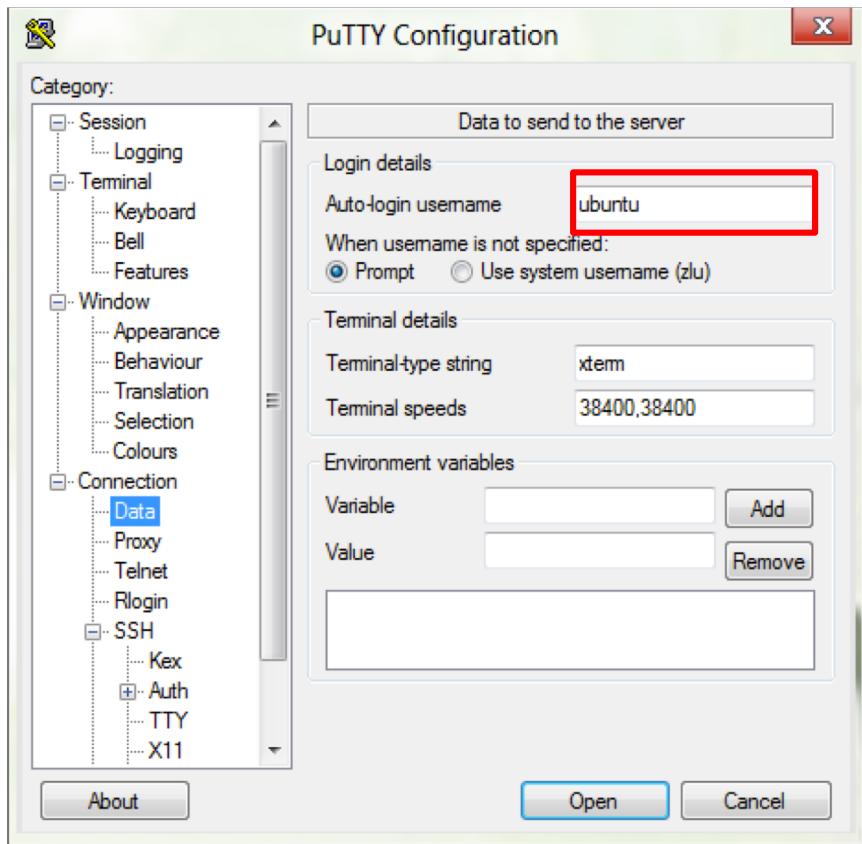


Browse to the  
cshl\_2018\_student.ppk file



# Logging into your instance (Windows)

Enter the user name ‘ubuntu’      Enter the host name



# Logging into your instance (Windows)

Open PuTTY

