



Cold  
Spring  
Harbor  
Laboratory

# Introduction to cloud computing

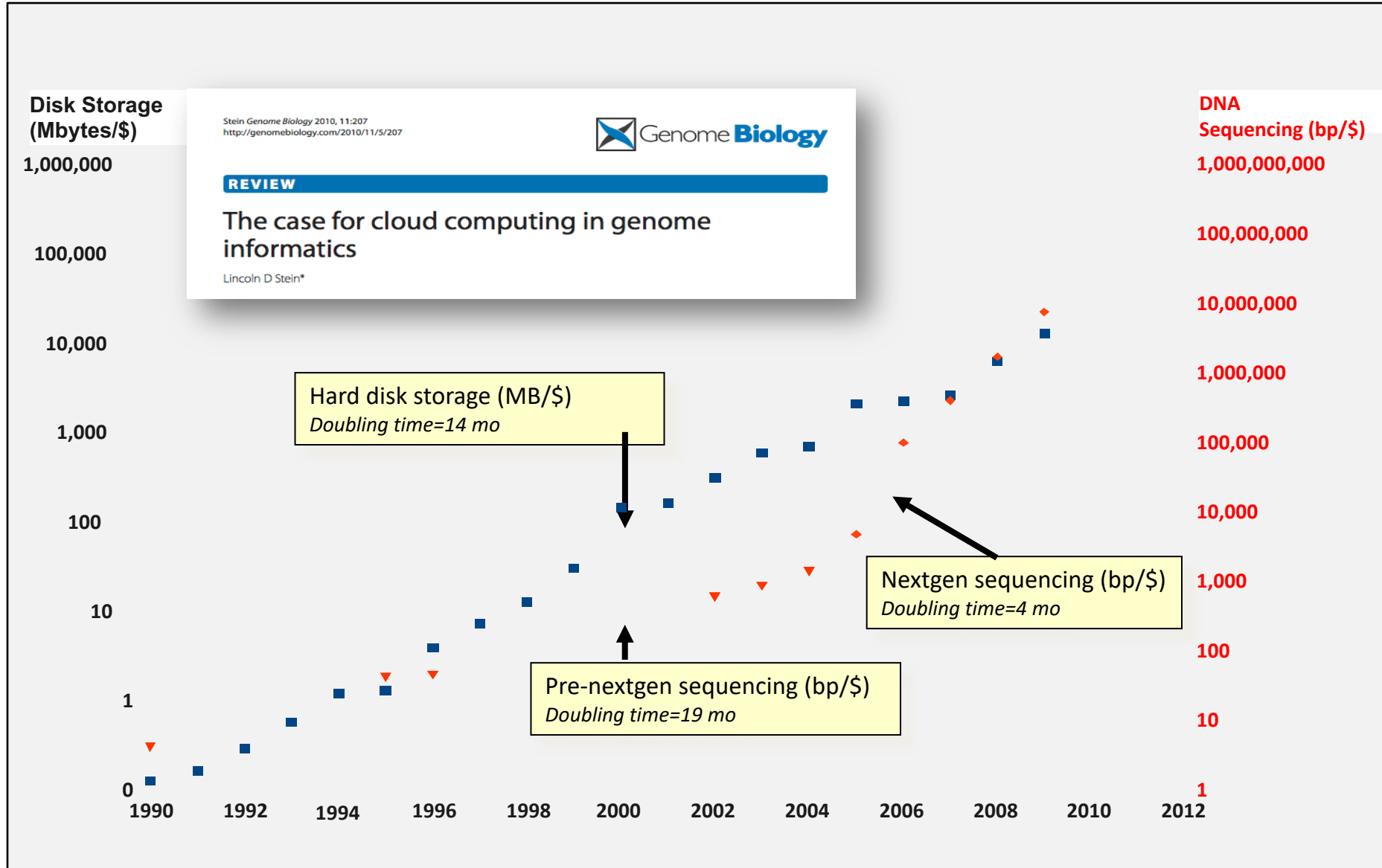
John Chamberlin, Kelsy Cotto, Felicia Gomez, Obi Griffith, Malachi Griffith,  
Simone Longo, Allegra Petti, Aaron Quinlan, Megan Richters, Huiming Xia  
Advanced Sequencing Technologies & Bioinformatics Analysis November 16-20, 2020



# 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



# 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?
- Get bigger hardware?
- Write more grants?

# Cloud computing providers

- Amazon AWS
  - <https://aws.amazon.com/>
- Google cloud
  - <https://cloud.google.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



# Some of the challenges of cloud computing:

- Not cheap
- Getting files to and from there
- Standardization can be a challenge if you don't control hardware
- 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.

# 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 that 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?

The screenshot shows the AWS Pricing Calculator interface. At the top, there are tabs for different operating systems: Linux (selected), RHEL, SLES, Windows, Windows with SQL Standard, and Windows with SQL Web. Below these is a sub-tab for 'Windows with SQL Enterprise'. A 'Region' dropdown menu is set to 'US West (Oregon)'. The main table displays instance types under the heading 'General Purpose - Current Generation'. The columns are vCPU, ECU, Memory (GiB), Instance Storage (GB), and Linux/UNIX Usage. The table lists various instance types with their respective specifications and hourly costs.

	vCPU	ECU	Memory (GiB)	Instance Storage (GB)	Linux/UNIX Usage
<b>General Purpose - Current Generation</b>					
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://rnabio.org/module-00-setup/0000/06/01/Intro\\_to\\_AWS/](https://rnabio.org/module-00-setup/0000/06/01/Intro_to_AWS/)

<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, “Log into AWS” page



## Log into AWS

[« Introduction to AWS](#)

[Course](#)

[Unix »](#)

Using cloud computing to complete this course involves two major components: (1) Launching an instance on the cloud (essentially renting a virtual computer by the hour and turning it on) and (2) logging into that instance).

Covered in this section: logging into AWS EC2 console, starting an instance from the course AMI, configuring it in the console (select instance AMI, instance type, instance details, storage volumes, tags, security group, and key pairs).



[https://rnabio.org/module-00-setup/0000/07/01/Log\\_into\\_AWS/](https://rnabio.org/module-00-setup/0000/07/01/Log_into_AWS/)

# Login to AWS console



Account ID or alias

IAM user name

Password

Sign In



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

[Forgot password?](#)

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

# Select "EC2" service

The screenshot shows the AWS Management Console interface. At the top, the navigation bar includes the AWS logo, 'Services', a notification bell, the user email 'cshl.student @ cshlworkshops', the region 'N. Virginia', and 'Support'. The main heading is 'AWS Management Console'. On the left, the 'AWS services' section has a 'Find Services' search bar containing 'EC2'. A red arrow points to this search bar with the text 'Search for EC2'. Below the search bar, 'Recently visited services' lists 'EC2', 'Billing', and 'IAM'. On the right, a callout box with a red arrow pointing to the 'N. Virginia' region dropdown contains the text 'Make sure you are in Virginia region'. Below this, there are sections for 'Explore AWS' featuring 'Amazon SageMaker Autopilot' and 'AWS Storage Gateway'.

# Launch a new Instance

The screenshot shows the AWS Management Console interface. At the top, the AWS logo and 'Services' dropdown are visible. The user's profile 'cshl.student @ cshlworkshops' and region 'N. Virginia' are shown in the top right. On the left, a navigation menu includes 'New EC2 Experience', 'EC2 Dashboard', 'Events', 'Tags', 'Limits', 'Instances', 'Images', and 'Elastic Block Store'. The main content area is divided into 'Resources' and 'Launch instance' sections. The 'Resources' section displays a table of EC2 resources in the US East (N. Virginia) Region:

Resource	Count
Instances (running)	2
Dedicated Hosts	0
Elastic IPs	0
Instances (all states)	2
Key pairs	5
Load balancers	0
Placement groups	0
Security groups	4
Snapshots	4
Volumes	3

Below the table, a notification box states: 'Easily size, configure, and deploy Microsoft SQL Server Always On availability groups on AWS using the AWS Launch Wizard for SQL Server. [Learn more](#)'.

The 'Launch instance' section contains the text: 'To get started, launch an Amazon EC2 instance, which is a virtual server in the cloud.' Below this text is a prominent orange button labeled 'Launch instance' with a dropdown arrow. A red arrow points to this button. A note at the bottom of the section reads: 'Note: Your instances will launch in the US East (N. Virginia) Region'.

On the right side of the console, there are two panels: 'Account attributes' and 'Additional information'. The 'Account attributes' panel includes a refresh button and links for 'Supported platforms' (listing VPC), 'Default VPC' (vpc-ad2c8fd7), 'Settings', 'EBS encryption', 'Zones', 'Default credit specification', and 'Console experiments'. The 'Additional information' panel includes links for 'Getting started guide', 'Documentation', 'All EC2 resources', and 'Forums'.

# Choose an AMI – Find the CSHL SEQTEC 2020 AMI in the My AMIs

The screenshot shows the AWS console interface for selecting an Amazon Machine Image (AMI). The top navigation bar includes the AWS logo, 'Services', and user information. Below the navigation bar, a progress indicator shows seven steps: 1. Choose AMI (highlighted), 2. Choose Instance Type, 3. Configure Instance, 4. Add Storage, 5. Add Tags, 6. Configure Security Group, and 7. Review. A 'Cancel and Exit' link is visible in the top right.

The main content area is titled 'Step 1: Choose an Amazon Machine Image (AMI)'. Below the title is a search bar with the placeholder text 'Search for an AMI by entering a search term e.g. "Windows"'. Below the search bar is a list of AMIs. The list is filtered to show two AMIs, both owned by the user and 64-bit (x86) architecture. The first AMI is 'cschl-seqtec-2019' and the second is 'cschl-seqtec-2020'. A red arrow points to the 'My AMIs' tab in the left sidebar, and another red arrow points to the 'Select' button for the 'cschl-seqtec-2020' AMI.

AMI ID	Architecture
cschl-seqtec-2019 - ami-018b3bf40f9926ac5	64-bit (x86)
cschl-seqtec-2020 - ami-0cdaba7e6f983f943	64-bit (x86)

# Choose "m5.2xlarge" instance type, then "Next: Configure Instance Details".

aws Services ▼ cshl.student @ cshlworkshops ▼ N. Virginia ▼ Support ▼

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

### Step 2: Choose an Instance Type

<input type="checkbox"/>	m5	m5.large	2	8	EBS only	Yes	Up to 10 Gigabit	Yes
<input type="checkbox"/>	m5	m5.xlarge	4	16	EBS only	Yes	Up to 10 Gigabit	Yes
<input checked="" type="checkbox"/>	m5	m5.2xlarge	8	32	EBS only	Yes	Up to 10 Gigabit	Yes
<input type="checkbox"/>	m5	m5.4xlarge	16	64	EBS only	Yes	Up to 10 Gigabit	Yes
<input type="checkbox"/>	m5	m5.8xlarge	32	128	EBS only	Yes	10 Gigabit	Yes
<input type="checkbox"/>	m5	m5.12xlarge	48	192	EBS only	Yes	10 Gigabit	Yes
<input type="checkbox"/>	m5	m5.16xlarge	64	256	EBS only	Yes	20 Gigabit	Yes
<input type="checkbox"/>	m5	m5.24xlarge	96	384	EBS only	Yes	25 Gigabit	Yes
<input type="checkbox"/>	m5	m5.metal	96	384	EBS only	Yes	25 Gigabit	Yes
<input type="checkbox"/>	m5a	m5a.large	2	8	EBS only	Yes	Up to 10 Gigabit	Yes
<input type="checkbox"/>	m5a	m5a.xlarge	4	16	EBS only	Yes	Up to 10 Gigabit	Yes

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

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

aws Services ▼ 🔔 cshl.student@cshlworkshops ▼ N. Virginia ▼ Support ▼

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

### Step 3: Configure Instance Details

**Auto-assign Public IP** ⓘ

**Placement group** ⓘ  Add instance to placement group

**Capacity Reservation** ⓘ


**Domain join directory** ⓘ  [🔄 Create new directory](#)

**IAM role** ⓘ  [🔄 Create new IAM role](#)  
**⚠️** You do not have permissions to list instance profiles. Contact your administrator, or check your IAM permissions.

**CPU options** ⓘ  Specify CPU options

**Shutdown behavior** ⓘ

**Stop - Hibernate behavior** ⓘ  Enable hibernation as an additional stop behavior

**Enable termination protection** ⓘ  Protect against accidental termination 

**Monitoring** ⓘ  Enable CloudWatch detailed monitoring  
[Additional charges apply.](#)

**EBS-optimized instance** ⓘ  Launch as EBS-optimized instance

[Cancel](#) [Previous](#) [Review and Launch](#) [Next: Add Storage](#)

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

aws Services cshl.student@cshlworkshops N. Virginia Support

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

### 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.

Volume Type	Device	Snapshot	Size (GiB)	Volume Type	IOPS	Throughput (MB/s)	Delete on Termination	Encryption
Root	/dev/sda1	snap-00fec8d273403984a	32	General Purpose SSD (gp2)	100 / 3000	N/A	<input checked="" type="checkbox"/>	Not Encrypted
EBS	/dev/sdb	snap-07c1c7fa70fa4e	250	General Purpose SSD (gp2)	750 / 3000	N/A	<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: Add Tags](#)



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

The screenshot shows the AWS console interface for the 'Add Tags' step. The breadcrumb trail at the top indicates the current step is '5. Add Tags'. Below the breadcrumb, the 'Step 5: Add Tags' section provides instructions on how to create a tag. The main area contains a table for adding tags. The first row has 'Name' in the 'Key' column and 'KelsyCotto' in the 'Value' column. There are checkboxes for 'Instances' and 'Volumes', both of which are checked. A red arrow points to the 'KelsyCotto' value. At the bottom right, the 'Next: Configure Security Group' button is highlighted with a red box.

Key	Value	Instances	Volumes
Name	KelsyCotto	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Buttons: Cancel, Previous, Review and Launch, Next: Configure Security Group

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

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

**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
<input type="checkbox"/> sg-384f5b79	default	default VPC security group	<a href="#">Copy to new</a>
<input checked="" type="checkbox"/> sg-0087dc3a8b6e37a2d	SSH and HTTP	created 2019-11-08T09:43:29.293-05:00	<a href="#">Copy to new</a>

Inbound rules for sg-0087dc3a8b6e37a2d (Selected security groups: sg-0087dc3a8b6e37a2d)

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

aws Services

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

### 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.

**Warning:** 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](#)

**Warning:** 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-2020 - ami-0cdaba7e6f983f943**  
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
m5.2xlarge	-	8	32	EBS only	Yes	Up to 10 Gigabit

Security Groups [Edit security groups](#)

Security Group ID	Name	Description
sg-0087dc3a8b6e37a2d	SSH and HTTP	created 2019-11-08T09:43:29.293-05:00

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

# Choose an existing key pair: "cshl\_2020\_student" and then Launch.

aws Services

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

### 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**  
Your instances may be accessible from any IP address. We recommend you restrict access to your instances. You can also open additional ports in your security group to facilitate access.

**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\_2020\_student

I acknowledge that I have access to the selected private key file (cshl\_2020\_student.pem), and that without this file, I won't be able to log into my instance.

Cancel **Launch Instances**

**AMI Details**  
cshl-seqtech-2020 - ami-0cdaba7e6f983f943  
Root Device Type: ebs Virtualization type: hvm

**Instance Type**

Instance Type	ECUs	vCPUs	Memory (GB)
m5.2xlarge	-	8	32

**Security Groups**

Cancel Previous **Launch**

# View Instances to see your new instance spinning up!

The screenshot shows the AWS Management Console interface. At the top, there is a navigation bar with the AWS logo, 'Services', 'Resource Groups', and user information for 'cshl.student @ cshlworkshops' in 'N. Virginia'. The main content area is titled 'Launch Status'. It features a green notification box stating 'Your instances are now launching' with a checkmark icon and a link to 'View launch log'. Below this is a blue information box about estimated charges. The section 'How to connect to your instances' explains that instances will be in a 'running' state and provides a link to 'View Instances'. A list of helpful resources is provided, including links to 'How to connect to your Linux instance', 'Learn about AWS Free Usage Tier', 'Amazon EC2: User Guide', and 'Amazon EC2: Discussion Forum'. At the bottom, there are links for 'Create status check alarms', 'Create and attach additional EBS volumes', and 'Manage security groups'. A red box highlights the 'View Instances' button in the bottom right corner.

# 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)

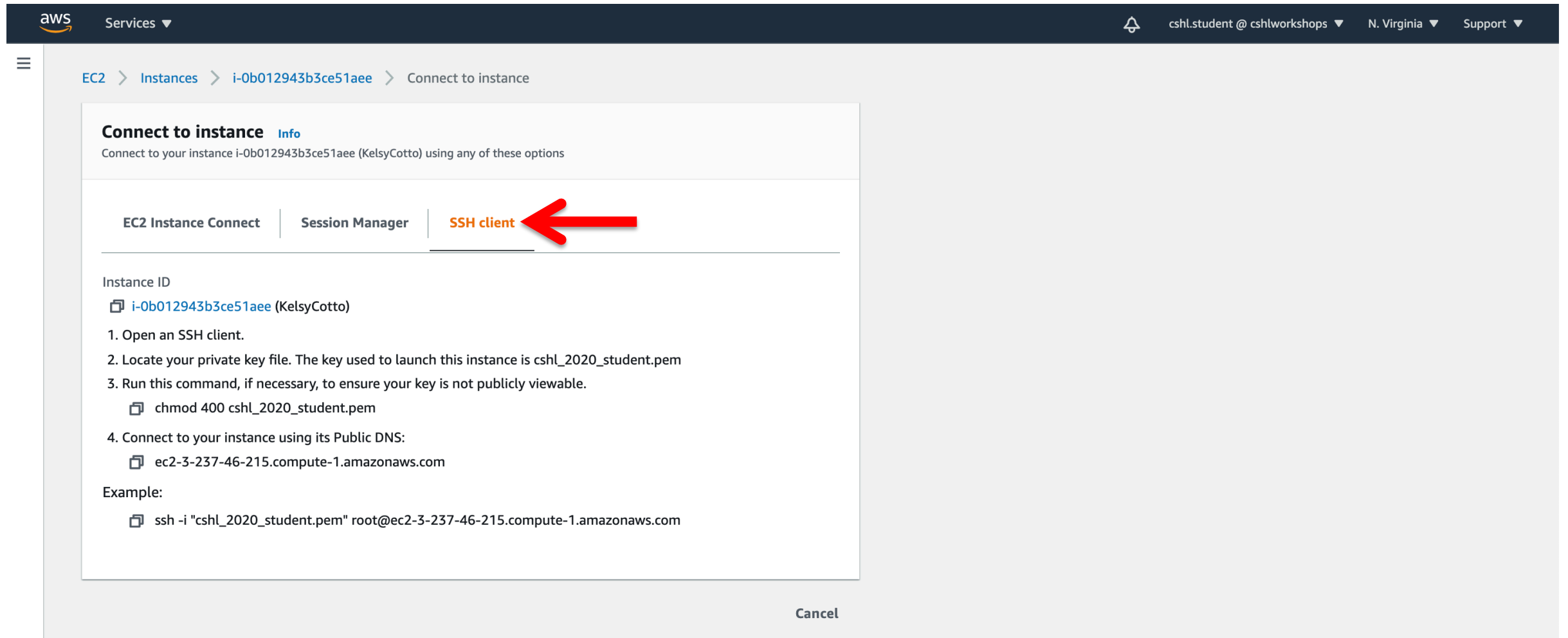
The screenshot shows the AWS Management Console interface. At the top, there's a navigation bar with the AWS logo, 'Services', and user information. Below that, a sidebar on the left contains navigation options like 'New EC2 Experience', 'EC2 Dashboard', 'Events', 'Tags', 'Limits', 'Instances', 'Instance Types', 'Launch Templates', 'Spot Requests', 'Savings Plans', 'Reserved Instances', 'Dedicated Hosts', 'Scheduled Instances', 'Capacity Reservations', 'Images', and 'Elastic Block Store'. The main content area is titled 'Instances (1/3) Info'. It features a search bar, a 'Connect' button (highlighted with a red box), and an 'Instance state' dropdown. Below this is a table of instances:

Name	Instance ID	Instance state	Instance type	Status check	Alarm status	Availability Zone	Public IPv4 DNS
test_login_in...	i-061dc4cf116f108b5	Running	t2.micro	2/2 checks ...	No alarms +	us-east-1e	ec2-54-208-183-137.c...
instructor_in...	i-01f3e5ef76ace3f48	Running	m5.2xlarge	2/2 checks ...	No alarms +	us-east-1f	ec2-3-83-35-210.com...
<input checked="" type="checkbox"/> KelsyCotto	i-0b012943b3ce51aee	Running	m5.2xlarge	2/2 checks ...	No alarms +	us-east-1f	ec2-100-24-122-188.c...

Below the table, the details for the selected instance 'i-0b012943b3ce51aee (KelsyCotto)' are shown. The 'Details' tab is active, displaying an 'Instance summary' with the following information:

- Instance ID: i-0b012943b3ce51aee (KelsyCotto)
- Instance state: Running
- Instance type: m5.2xlarge
- Public IPv4 address: 100.24.122.188 | open address
- Public IPv4 DNS: ec2-100-24-122-188.compute-1.amazonaws.com | open address
- Elastic IP addresses: -
- Private IPv4 addresses: 172.31.72.162
- Private IPv4 DNS: ip-172-31-72-162.ec2.internal
- VPC ID: vpc-ad2c8fd7

# 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)



aws Services ▾ cshl.student@cshlworkshops ▾ N. Virginia ▾ Support ▾

EC2 > Instances > i-0b012943b3ce51aee > Connect to instance

### Connect to instance [Info](#)

Connect to your instance i-0b012943b3ce51aee (KelsyCotto) using any of these options

EC2 Instance Connect | Session Manager | **SSH client** ←

Instance ID  
i-0b012943b3ce51aee (KelsyCotto)

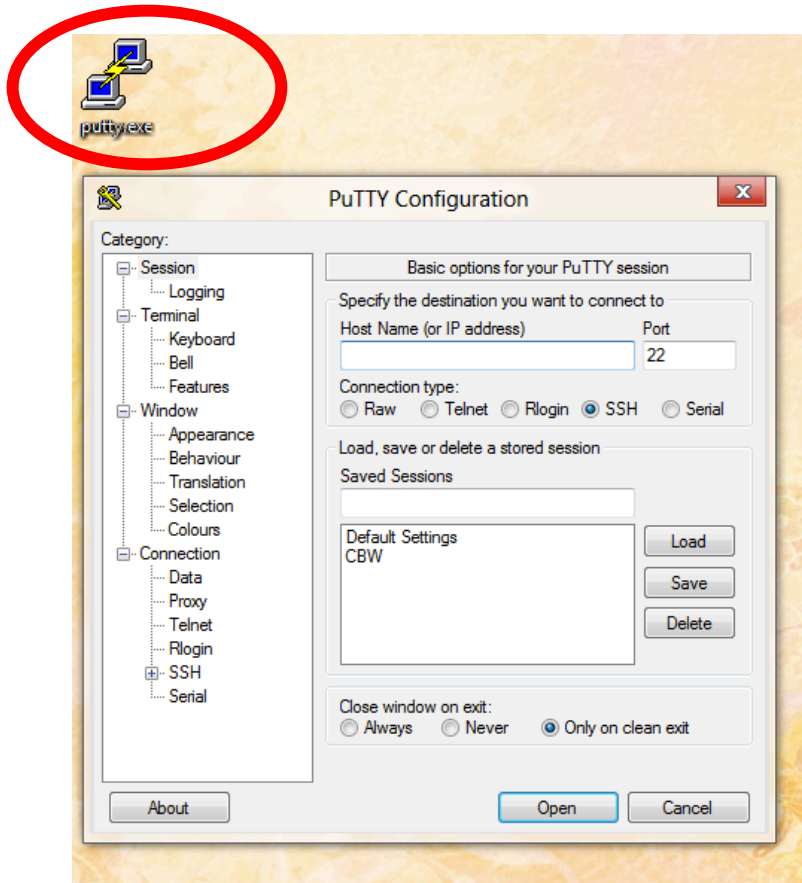
1. Open an SSH client.
2. Locate your private key file. The key used to launch this instance is cshl\_2020\_student.pem
3. Run this command, if necessary, to ensure your key is not publicly viewable.  
`chmod 400 cshl_2020_student.pem`
4. Connect to your instance using its Public DNS:  
`ec2-3-237-46-215.compute-1.amazonaws.com`

Example:  
`ssh -i "cshl_2020_student.pem" root@ec2-3-237-46-215.compute-1.amazonaws.com`

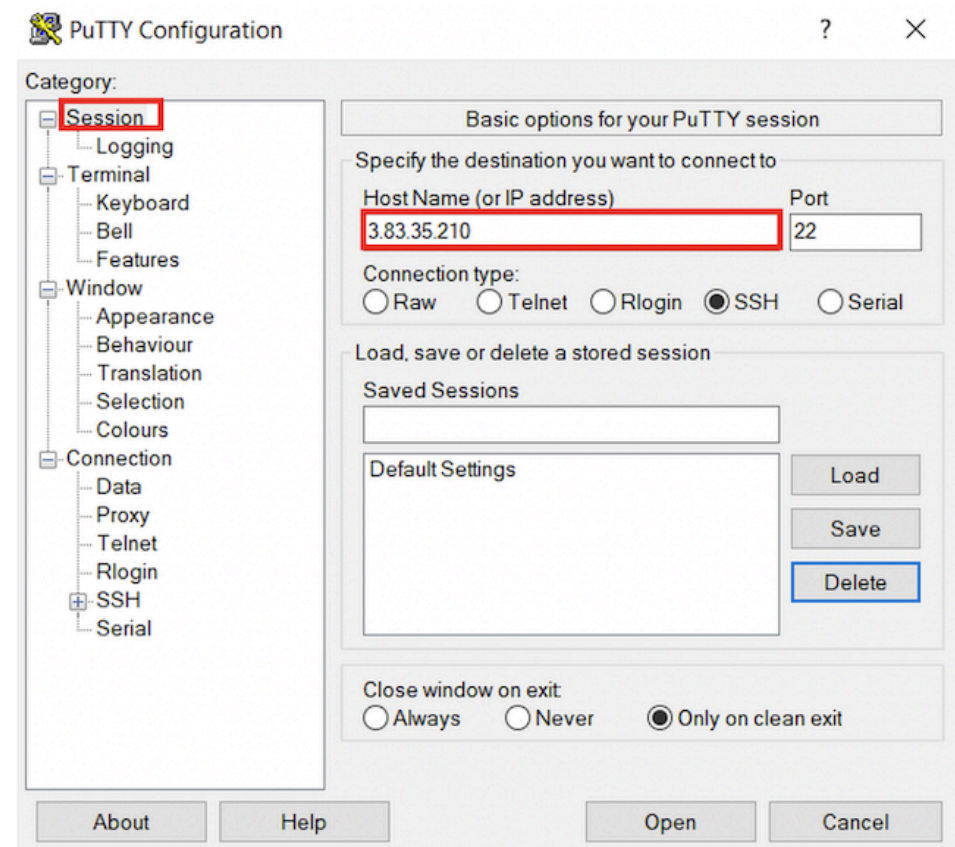
Cancel

# Logging into your instance (Windows)

Open PuTTY



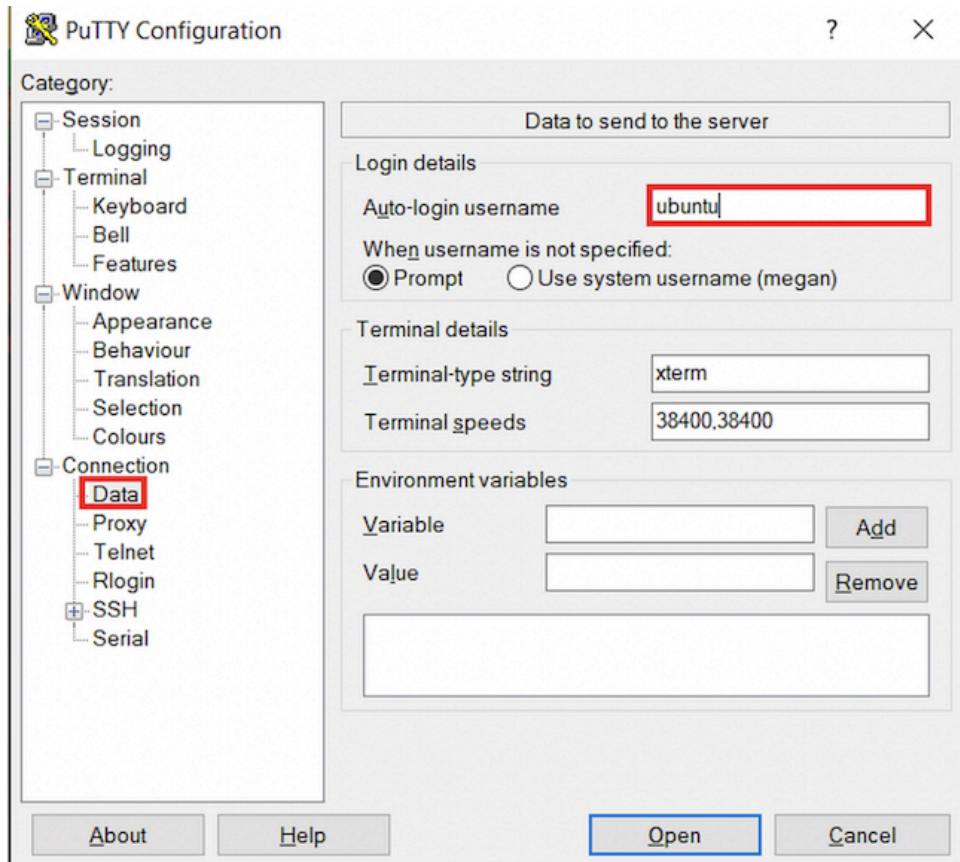
Enter the Host Name (IP address)



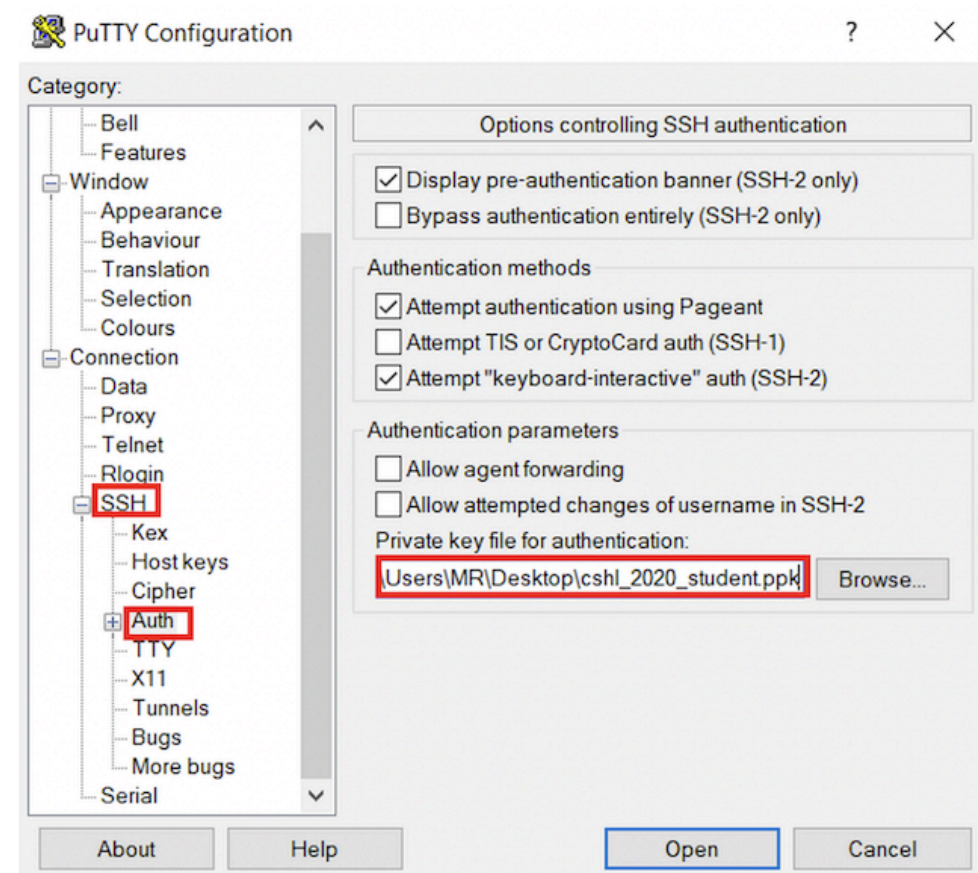


# Logging into your instance (Windows)

Choose Connection -> Data  
Enter the username 'ubuntu'

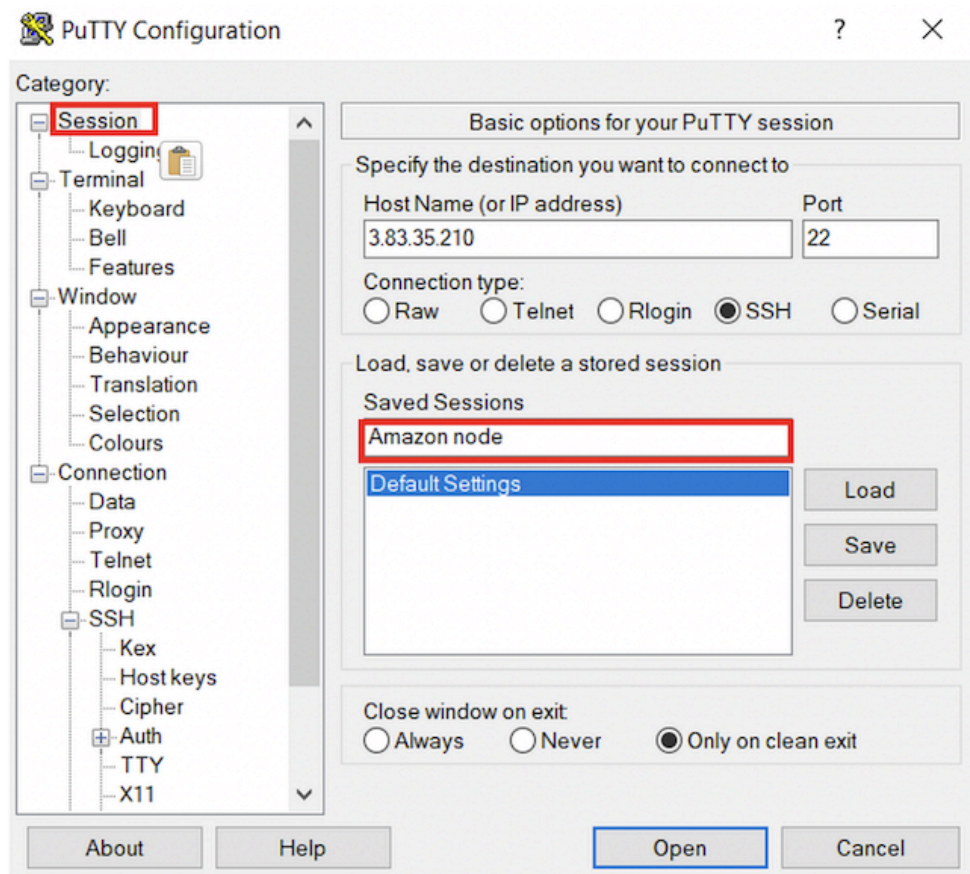


Choose SSH -> Auth  
Browse to Private key (ppk) file

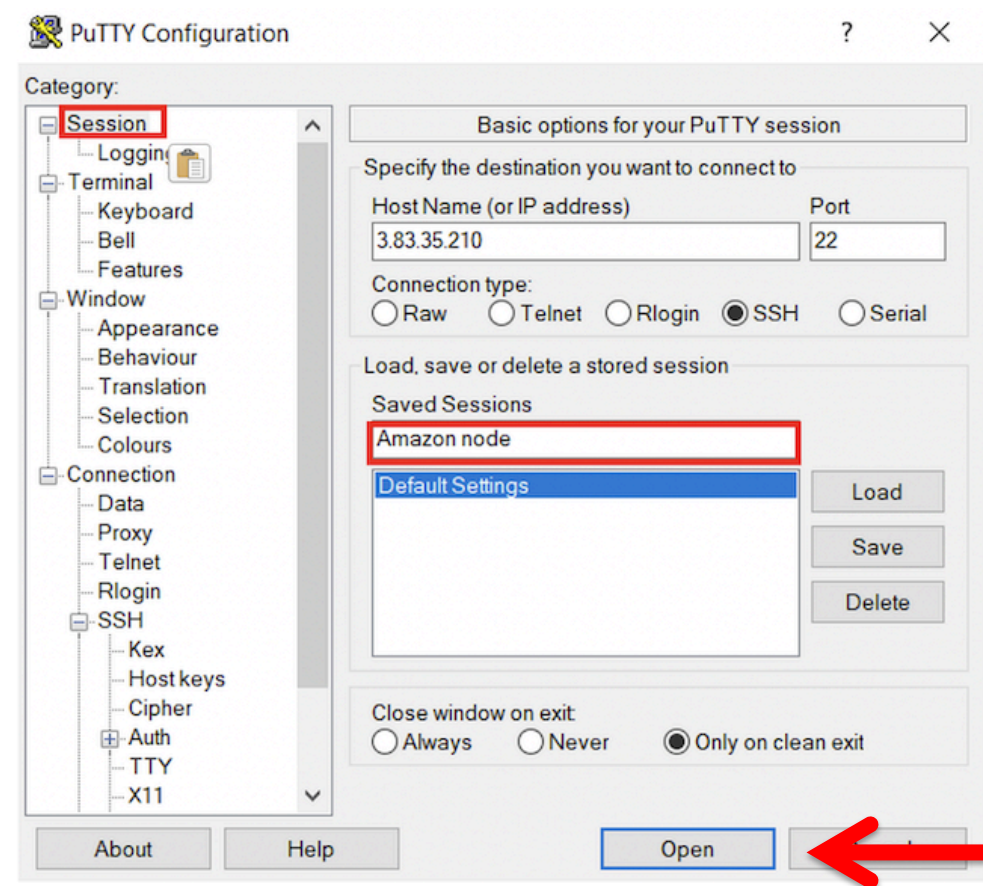


# Logging into your instance (Windows)

Choose Session  
Save your session as "Amazon Node"

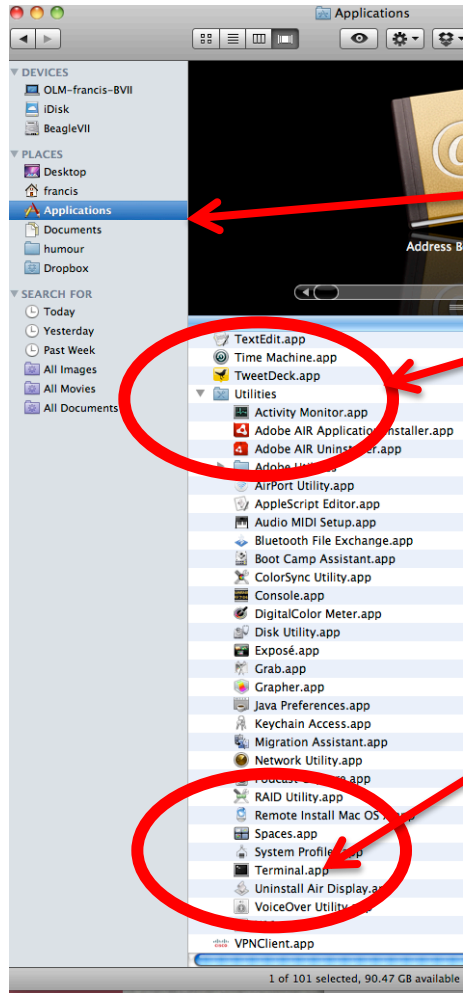


Double-click saved "Amazon Node" session OR  
Select "Amazon Node" session and click Open



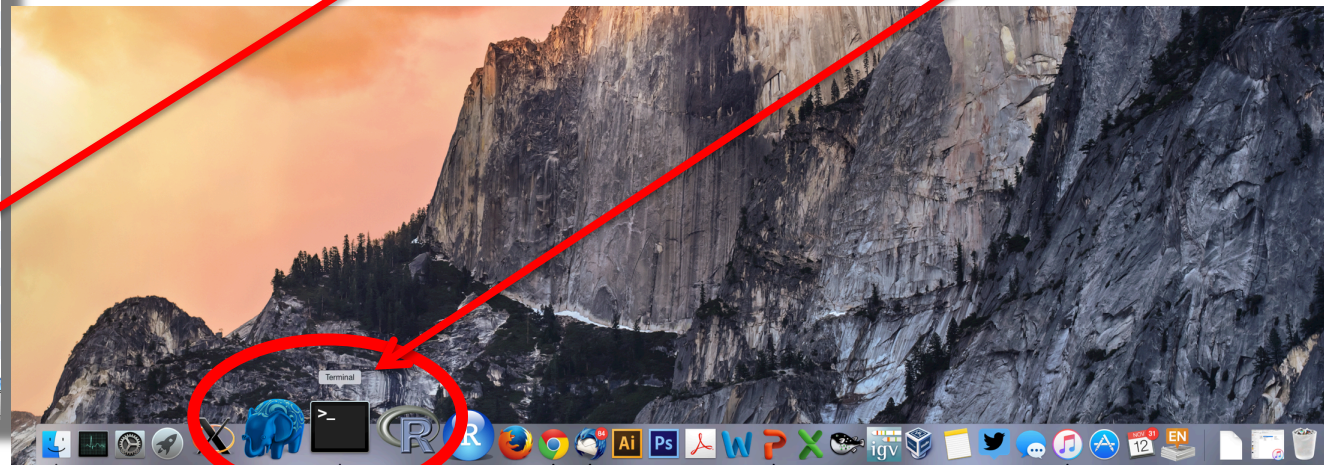


# Logging into your instance (Mac)

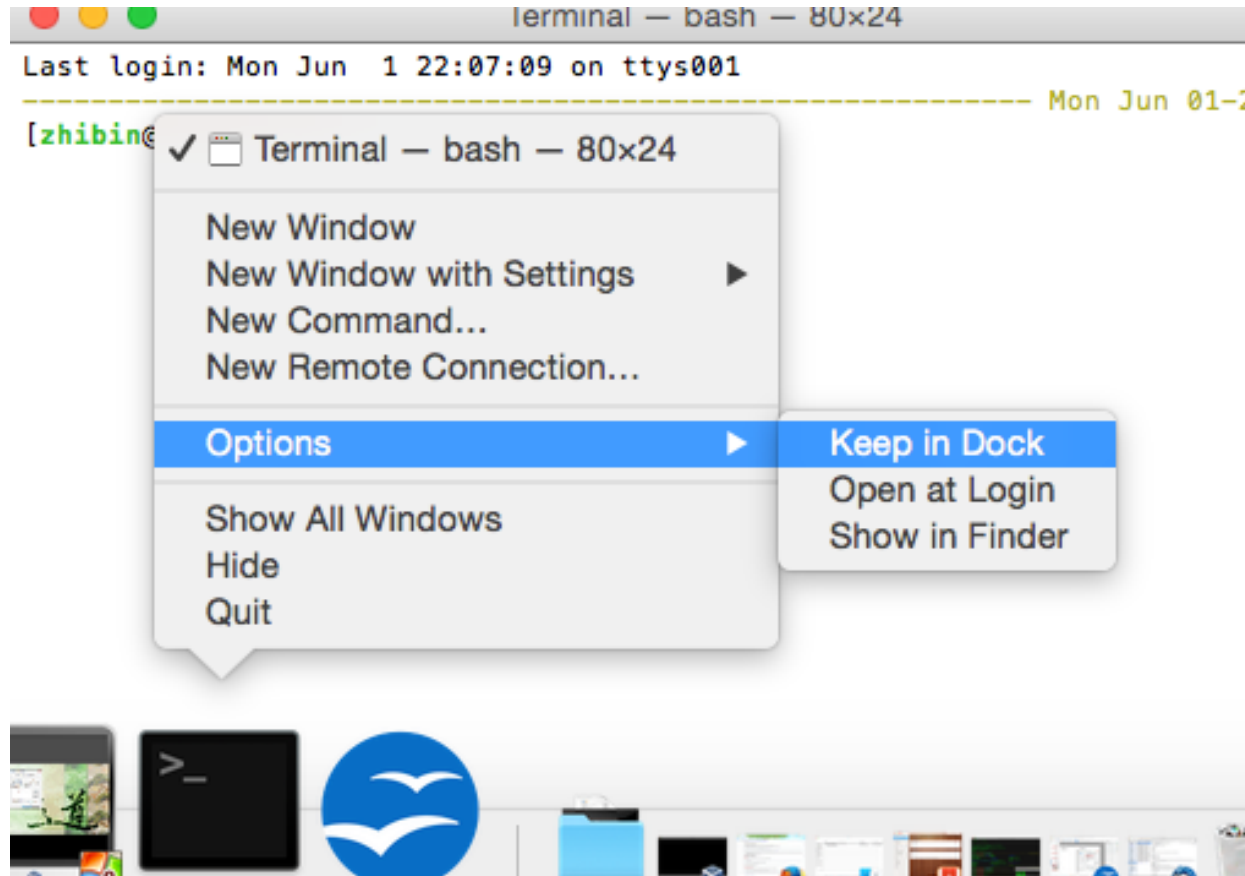


In a Finder window  
'Applications' -> 'Utilities' -> 'Terminal'

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
```

# 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

**Save the pem/ppk file you received via email/slack to your new cshl folder**

# Viewing the 'key' file once downloaded

```
cat cshl_2020_student.pem
```

```
Kelsys-MacBook-Pro:keys kcotto$ cat cshl_2019_student.pem
-----BEGIN RSA PRIVATE KEY-----
MIIEpAIBAACAQEAgtGtpnqERtEu/SCmeF2r1HMESMao0fEJiAwQwk2/SNXK8izr0IH0zTVvmE1D
VUwWq7pkvhhJh05pDb+2U9HiZe3sXlv3S1NrKATYF/NsrpwB+q1vwqzGW9sQ6uj45RWrpkiZlsaj
TQZmyFRu+tLJTRU3hQDqA0MRWTx1Wxv0GfZuZy/qb+DALuFQsInrEKnijrwdLmd6usaBTvhc0gFS
B9oEelH0bZHJTzFw/wP+Z0uZq0Ujir7Qw0LTM45QH/L0dBdU13k/mBeez00yvnKMwj8E4Xi0rQ0t
hHtQ7F9iSILK80W1rRH0qwxwt9ycEH1JtNMQmUtif0vE2XJ6L06chQIDAQABAoIBABG7P/FHu/Qp
WFgg+89myuqR6GvA2X55CFSzZfYg0aQyrj5jDleFtdu2uXiISG8gUBZYvlzxx82a0C0P5j04SBq0
xD/qRlukY/jyXyPn77w/ExmaNoLJj1L1W9RUSH0JYLIZvPFPGes3u5zGSGDTSdNhw3sSdWhq1FX3L
7vY5b6UAQgahXf0dpGFxt6P6qb/BKFQfSthXk7GXMyS/kr4w7Z1hPWHrMSu2UBdS/a/beAVN76qU
E/10HR30oAuqghusZabpbX7MJL00VcviAQgeF8Z+xf2uugDEAKut1PxW0+yvGM4SpZZ0skFZz6YI
pBnX4ELWPhCeMzq4Ml8QY6ptR2UCgYEAuc6Yfnlymg24NJu8PANx8navTi50WYuXWxrj0mrLr3SR
/XY+w26cgipM+K5eQfXsr4Yb8BQKjRktMzBzf5nKdaX4pzYIquQH02B0HDhTooHAhkbTMTmKukv4
oEW06wcEE6RQi fw4xbMEnfQfHJBI21am+jwI8Xb7idwMG4pU/nsCgYEAtD0x4bNC1X3A5by50uY0
WXrtgQszCXyCbkrpbjRET12f9hgZ9MRMHY/xH/XGvMutZSFV2rCZRwd7lm+QWGadk/MQS0kouzW6
gSasyjFq+MKCKqYnS3/JTbx8yrLZmz10LtX6pwwmg0Zy8aJjYHo9a2/EI8Tjh2d0SxeadIRVYP8C
gYAToiXww1Vdu+dj/7TDLqYctdH0VaxJX/utI9QQ3yoIryuh+bWmFvEIvAmIGXyyQZRyoZwgIS4A
PNH03+bEa+69wbzlhksiK5g8GKgISVdlC4rZZXB5ehgTmWV7Igj89y/SF4G/Ityo30K0ohALh597
NcvNEzzqruTja1IIMvTKMwKBgQCkM+QP1Tqc0TbVlFvClviXuJBLsiJLClmYeZL0nZVmIMusbhxX
b8ZQYGSyUz09nuLXaulGLQDvXvf089CzWLLSomxBoHlFJQvGwa9FfYQRIVPHuqut8rs4oPGn0QzC
h7M7QCJcr00oAcrSLLkQmgz+phIw7BzFr039J4HFIRinjQKBgQCtdEvcbytk8Jh4WH3z0wpkc43f
U8DZhZwjRQpGwLD8CPj9RgRnE4+1PCH6s/RLQf7SiE1ZjX/0Ud0WPEvr0j5sVjy0IujohRbty0CM
oqWeSeUb1sLogRvMrTfCEpl/rz3GpoQLSC/5s6XvjnnKK8RN8s7MseLuuJ63T/CRBpIs8A==
-----END RSA PRIVATE KEY-----Kelsys-MacBook-Pro:keys 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_2020_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**

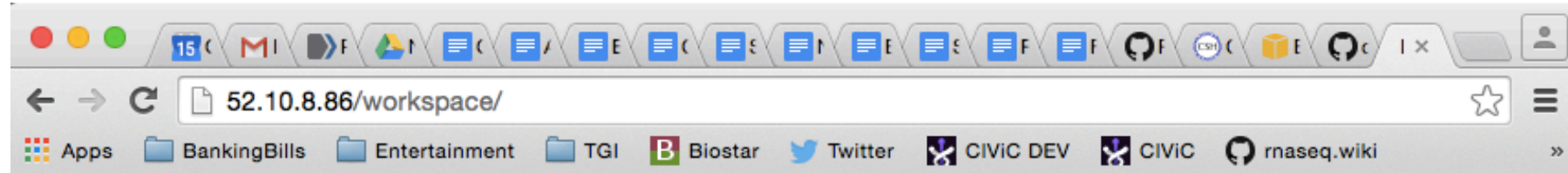


# Logging into your instance







## Mac/Linux

```
cd ~/cshl  
chmod 400 cshl_2020_student.pem  
ssh -i cshl_2020_student.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 Management Console interface for the EC2 service. The left-hand navigation pane is visible, with the 'Instances' tab selected. The main content area displays a table of EC2 instances. One instance, 'KelsyCott...', is selected, and a context menu is open over it, showing various actions. The 'Stop instance' option is highlighted in the menu. A red box is drawn around the text 'Go to AWS EC2 Dashboard, select “Instances” tab, then find your instance. Right-click and chose ‘Stop instance’', with red arrows pointing to the 'Instances' tab, the instance row, and the 'Stop instance' menu item.

Name	Instance ID	Instance state	Instance type	Status check	Alarm status	Availability Zone	Public IPv4 DNS
test_login_in...	i-061dc4cf116f108b5	Running	t2.micro	2/2 checks ...	No alarms +	us-east-1e	ec2-54-208-183-137.c...
instructor_in...	i-01f3e5ef76ace3f48	Running	m5.2xlarge	2/2 checks ...	No alarms +	us-east-1f	ec2-3-83-35-210.com...
KelsyCott...	i-05012942b7c0e51ae	Running	m5.2xlarge	2/2 checks ...	No alarms +	us-east-1f	ec2-3-237-46-215.co...

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

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

Successfully stopped i-0b012943b3ce51aee

Instances (1/3) Info

Name	Instance ID	Instance state	Instance type	Status check	Alarm status	Availability Zone	Public IPv4 DNS
test_login_in...	i-061dc4cf116f108b5	Running	t2.micro	2/2 checks ...	No alarms +	us-east-1e	ec2-54-208-183-137.c...
instructor_in...	i-01f3e5ef76ace3f48	Running	m5.2xlarge	2/2 checks ...	No alarms +	us-east-1f	ec2-3-83-35-210.com...
KelsyCotto	i-0b012943b3ce51aee	Stopped	m5.2xlarge	2/2 checks ...	No alarms +	us-east-1f	-

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

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 displays the AWS Management Console interface. On the left, there is a navigation sidebar with categories like 'New EC2 Experience', 'EC2 Dashboard', 'Events', 'Tags', 'Limits', 'Instances', 'Instance Types', 'Launch Templates', 'Spot Requests', 'Savings Plans', 'Reserved Instances', 'Dedicated Hosts', 'Scheduled Instances', 'Capacity Reservations', 'Images', and 'Elastic Block Store'. The main content area shows a table of EC2 instances. The instance 'KelsyCotto' is selected, and its details are shown below. The 'Instance summary' tab is active, displaying fields such as Instance ID, Instance state (Running), Instance type (m5.2xlarge), Public IPv4 address (100.24.122.188), Public IPv4 DNS (ec2-100-24-122-188.compute-1.amazonaws.com), Private IPv4 addresses (172.31.72.162), Private IPv4 DNS (ip-172-31-72-162.ec2.internal), and VPC ID (vpc-ad2c8fd7). A red arrow points to the 'Connect' button in the instance list, and another red arrow points to the 'Public IPv4 address' field in the instance details.

Name	Instance ID	Instance state	Instance type	Status check	Alarm status	Availability Zone	Public IPv4 DNS
test_login_in...	i-061dc4cf116f108b5	Running	t2.micro	2/2 checks ...	No alarms +	us-east-1e	ec2-54-208-183-137.c...
instructor_in...	i-01f3e5ef76ace3f48	Running	m5.2xlarge	2/2 checks ...	No alarms +	us-east-1f	ec2-3-83-35-210.com...
KelsyCotto	i-0b012943b3ce51aee	Running	m5.2xlarge	2/2 checks ...	No alarms +	us-east-1f	ec2-100-24-122-188.c...

Instance: i-0b012943b3ce51aee (KelsyCotto)

Instance summary

Instance ID	Public IPv4 address	Private IPv4 addresses
i-0b012943b3ce51aee (KelsyCotto)	100.24.122.188   open address	172.31.72.162
Instance state	Public IPv4 DNS	Private IPv4 DNS
Running	ec2-100-24-122-188.compute-1.amazonaws.com   open address	ip-172-31-72-162.ec2.internal
Instance type	Elastic IP addresses	VPC ID
m5.2xlarge	-	vpc-ad2c8fd7

# So, at this point:

- Your laptop/pc 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



# 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 22 regions: (US East (N.Virginia), US East (Ohio), US West (Oregon), US West (N. California), GovCloud (US-West), GovCloud (US-East), Canada (Montreal), EU (Ireland), EU (Frankfurt), EU (London), EU (Paris), EU (Milan), EU (Stockholm), Middle East (Bahrain), Asia Pacific (Singapore), Asia Pacific (Sydney), Asia Pacific (Seoul), Asia Pacific (Tokyo), Asia Pacific (Mumbai), Asia Pacific (Hong Kong), Asia Pacific (Beijing), 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.