



Cold
Spring
Harbor
Laboratory

Introduction to cloud computing

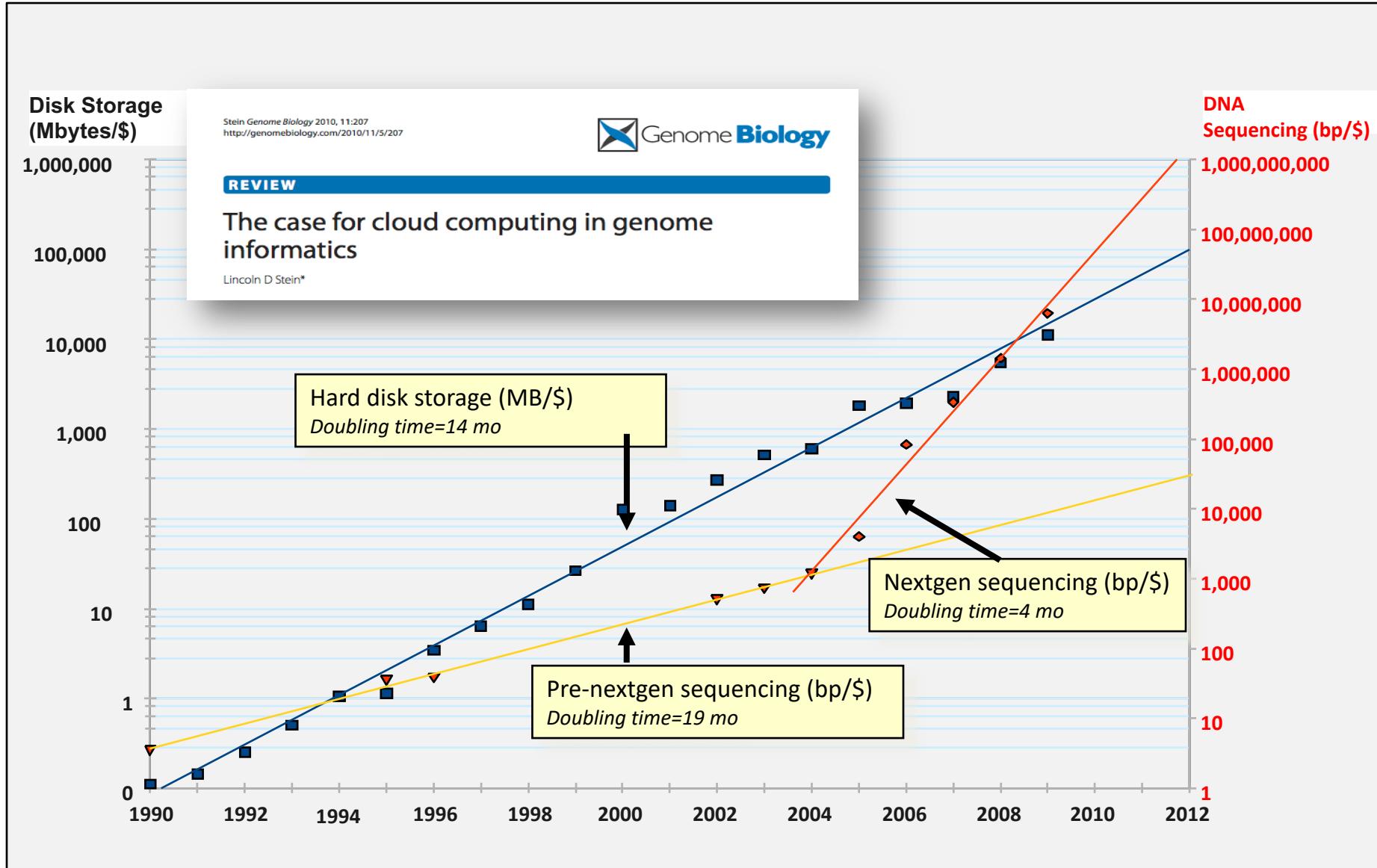
Kelsy Cotto, Arpad Danos, Harriet Dashnow, Felicia Gomez, Sharon Freshour, Obi Griffith,
Malachi Griffith, Jason Kunisaki, Chris Miller, Jonathan Preall, Aaron Quinlan
Advanced Sequencing Technologies & Bioinformatics Analysis November 11-19, 2021



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

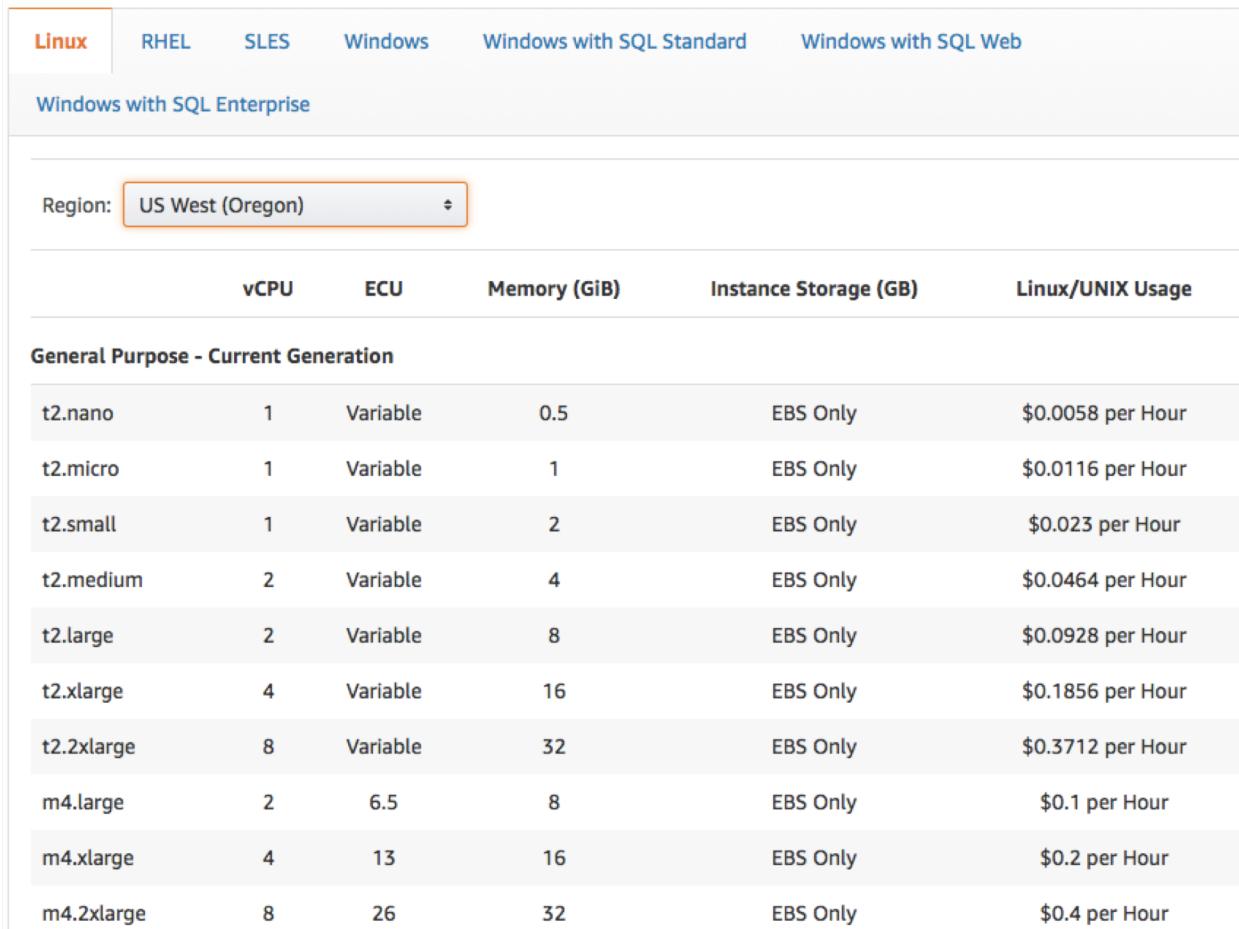
What is an AMI/snapshot?

- 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 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 CloudWatch Metrics Insights search interface. At the top, there are tabs for Linux, RHEL, SLES, Windows, Windows with SQL Standard, and Windows with SQL Web. Below these tabs, the text "Windows with SQL Enterprise" is displayed. A dropdown menu labeled "Region:" shows "US West (Oregon)". The main table has columns for vCPU, ECU, Memory (GiB), Instance Storage (GB), and Linux/UNIX Usage. The table is titled "General Purpose - Current Generation" and lists ten instance types:

Instance Type	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://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/

Login to AWS console



Account ID or alias

cshlworkshops

IAM user name

cshl.student

Password

.....

Sign In



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

[Forgot password?](#)

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

Select "EC2" service

AWS Management Console

AWS services

Find Services
You can enter names, keywords or acronyms

EC2 ← Search for EC2

Recently visited services

- EC2
- Billing
- IAM

All services

Build a solution
Get started with simple wizards and automated workflows.

Stay connected to your resources
Download the AWS Console Mobile App to your iOS or Android mobile device. [Learn more](#)

Explore AWS

Amazon SageMaker Autopilot
Get hands-on with AutoML. [Learn more](#)

AWS Storage Gateway
Get on-premises low latency access to virtually unlimited cloud storage with this hybrid cloud storage service. [Learn more](#)

N. Virginia ↑ Make sure you are in Virginia region

Launch a new Instance

The screenshot shows the AWS EC2 Dashboard. On the left, there's a sidebar with navigation links like 'EC2 Dashboard', 'Instances', 'Images', and 'Elastic Block Store'. The main area is titled 'Resources' and shows various EC2 metrics: 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, and Volumes 3. Below these metrics is a callout box with text about launching Microsoft SQL Server Always On availability groups. At the bottom of the dashboard, there's a section titled 'Launch instance' with a large orange 'Launch instance' button. A red arrow points to this button. A note at the bottom says 'Note: Your instances will launch in the US East (N. Virginia) Region'.

aws Services ▾

New EC2 Experience Tell us what you think X

EC2 Dashboard New

Events New

Tags

Limits

Instances

- Instances New
- Instance Types
- Launch Templates
- Spot Requests
- Savings Plans
- Reserved Instances
- Dedicated Hosts New
- Scheduled Instances
- Capacity Reservations

Images

- AMIs

Elastic Block Store

Resources

You are using the following Amazon EC2 resources in the US East (N. Virginia) Region:

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

i Easily size, configure, and deploy Microsoft SQL Server Always On availability groups on AWS using the AWS Launch Wizard for SQL Server. [Learn more](#) X

Launch instance

To get started, launch an Amazon EC2 instance, which is a virtual server in the cloud.

Launch instance ▾

Note: Your instances will launch in the US East (N. Virginia) Region

Account attributes

Supported platforms

- VPC

Default VPC vpc-ad2c8fd7

Settings

EBS encryption

Zones

Default credit specification

Console experiments

Additional information

Getting started guide

Documentation

All EC2 resources

Forums

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

The screenshot shows the AWS Step 1: Choose an Amazon Machine Image (AMI) interface. The top navigation bar includes the AWS logo, Services dropdown, user information (cshl.student @ cshlworkshops), location (N. Virginia), and Support. Below the navigation, a progress bar shows steps 1 through 7. Step 1 is highlighted with an orange underline. The main content area is titled "Step 1: Choose an Amazon Machine Image (AMI)". It explains that an AMI is a template containing software configuration required to launch an instance. A search bar at the top allows searching by term like "Windows". On the left, a sidebar has buttons for Quick Start, My AMIs (which is highlighted with a red arrow), AWS Marketplace, and Community AMIs. Under "Ownership", there are checkboxes for "Owned by me" (checked) and "Shared with me". Under "Architecture", there are checkboxes for "32-bit (x86)" and "64-bit (x86)". The main list displays two AMIs: "cshl-seqtec-2019 - ami-018b3bf40f9926ac5" and "cshl-seqtech-2020 - ami-0cdaba7e6f983f943". Each entry includes details like Root device type: ebs, Virtualization type: hvm, Owner: 577255725291, ENA Enabled: Yes, and Architecture: 64-bit (x86). To the right of each entry is a blue "Select" button, with another red arrow pointing to the one next to the 2020 AMI.

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"

Cancel and Exit

Quick Start

My AMIs

AWS Marketplace

Community AMIs

Ownership

Owned by me

Shared with me

Architecture

32-bit (x86)

64-bit (x86)

cshl-seqtec-2019 - ami-018b3bf40f9926ac5

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

64-bit (x86)

Select

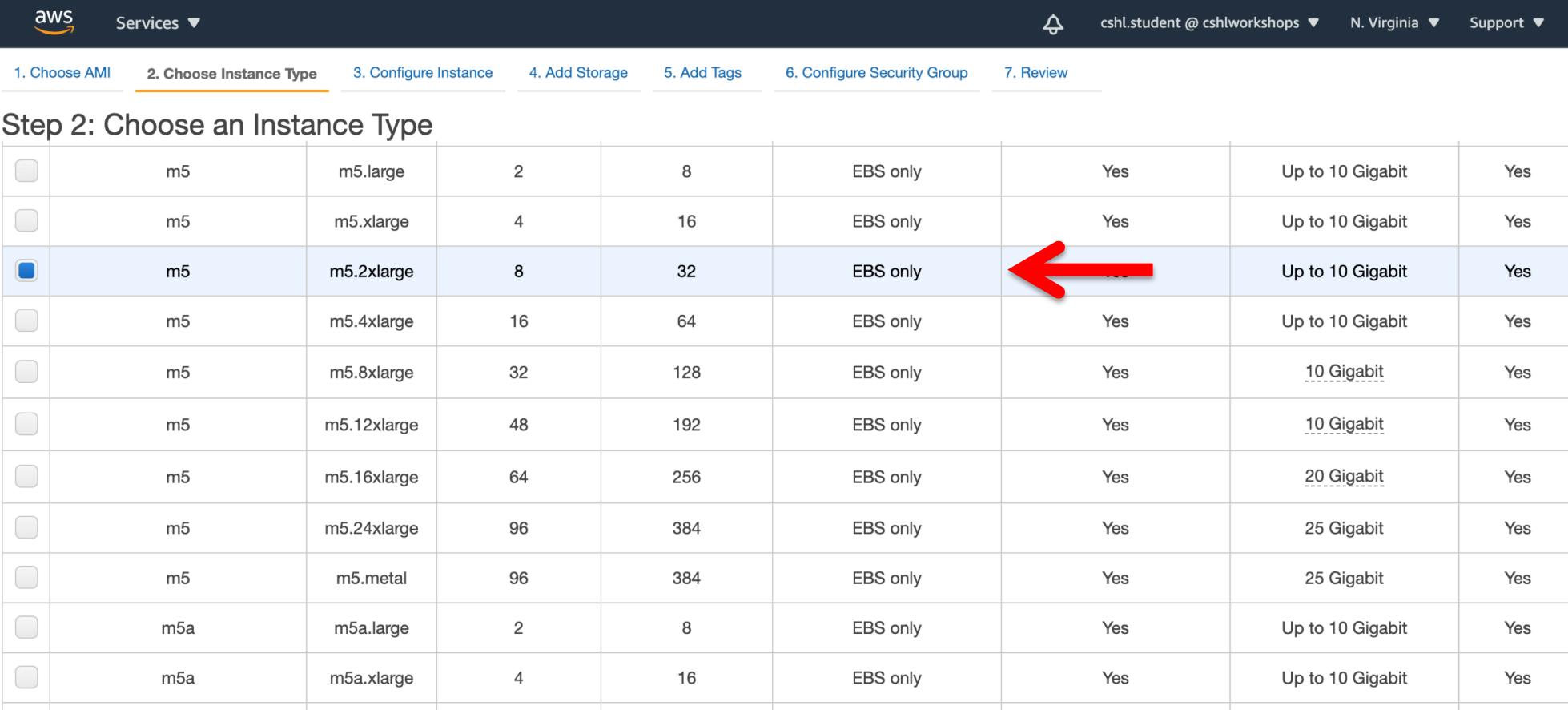
cshl-seqtech-2020 - ami-0cdaba7e6f983f943

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

64-bit (x86)

Select

Choose “m5.2xlarge” instance type, then “Next: Configure Instance Details”.



The screenshot shows the AWS CloudFormation console interface for creating a new stack. The top navigation bar includes the AWS logo, Services dropdown, and user information (cshl.student @ cshlworkshops, N. Virginia, Support). Below the navigation is a progress bar with seven steps: 1. Choose AMI, 2. Choose Instance Type (which is underlined in orange), 3. Configure Instance, 4. Add Storage, 5. Add Tags, 6. Configure Security Group, and 7. Review. The main content area is titled "Step 2: Choose an Instance Type". A table lists various Amazon M5 instance types with their details. The "m5.2xlarge" row is highlighted with a blue background and has a checked checkbox in its first column. A red arrow points to the "Yes" entry in the "Network" column for the "m5.2xlarge" row. At the bottom right are four buttons: "Cancel", "Previous", "Review and Launch" (which is highlighted in blue), and "Next: Configure Instance Details" (which is enclosed in a red box).

	Instance Type	Memory (GiB)	Cores	Storage (GiB)	Volume Type	Network	Up to 10 Gigabit	Yes
<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

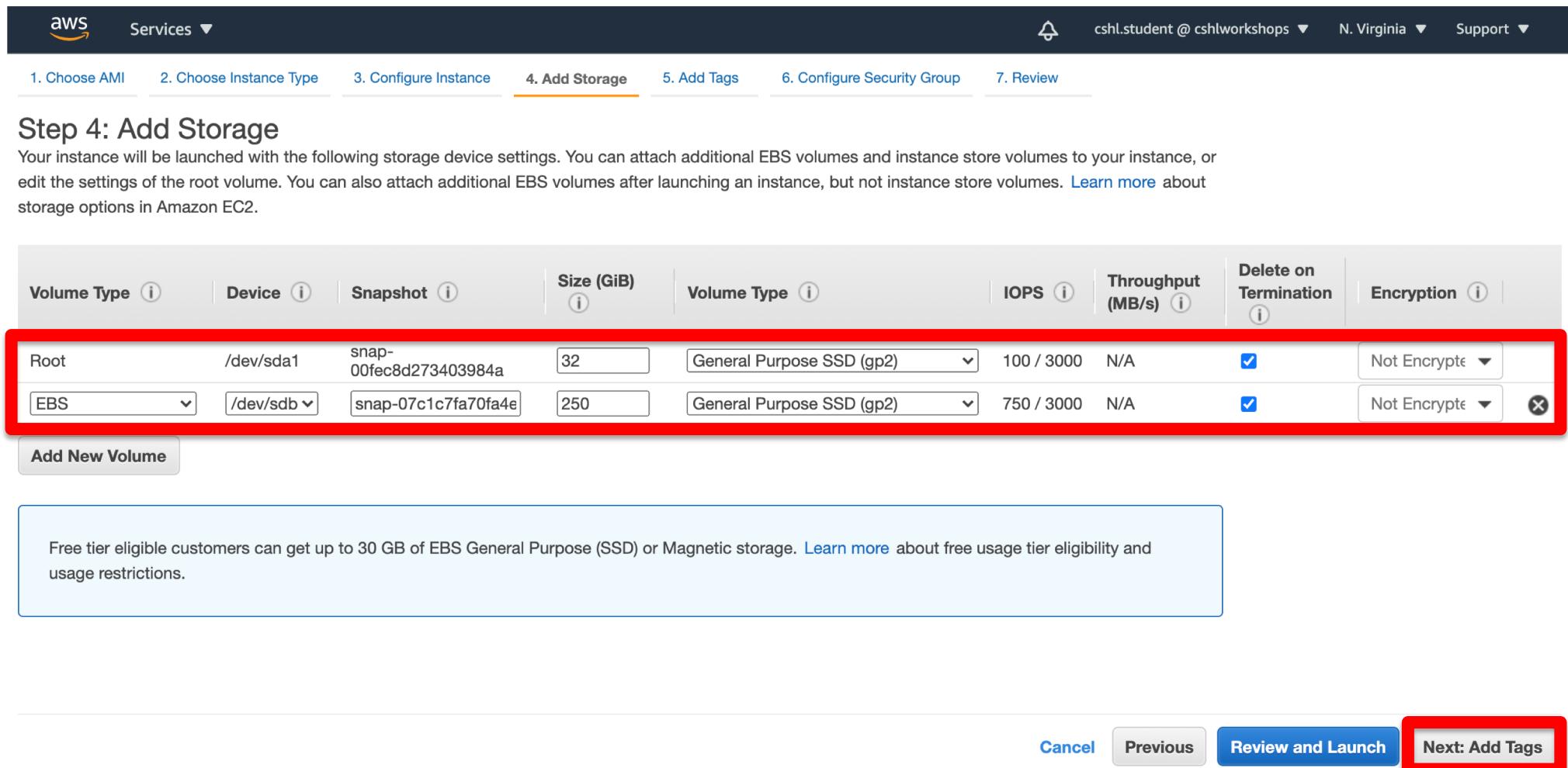
Select "Protect against accidental termination", then "Next: Add Storage".

The screenshot shows the AWS Step 3: Configure Instance Details wizard. The top navigation bar includes the AWS logo, Services dropdown, a bell icon, user information (cshl.student @ cshlworkshops), region (N. Virginia), and support links. Below the navigation, a progress bar shows steps 1 through 7. Step 3, 'Configure Instance', is selected. The main form contains the following fields:

- Auto-assign Public IP:** Use subnet setting (Enable)
- Placement group:** Add instance to placement group
- Capacity Reservation:** Open
- Domain join directory:** No directory [Create new directory](#)
- IAM role:** None [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
- 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

At the bottom right are buttons: Cancel, Previous, Review and Launch (highlighted with a red border), and Next: Add Storage.

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



The screenshot shows the AWS EC2 instance creation wizard at Step 4: Add Storage. The top navigation bar includes the AWS logo, Services dropdown, user cshl.student @ cshlworkshops, region N. Virginia, and Support. Below the navigation is a progress bar with steps 1-7: Choose AMI, Choose Instance Type, Configure Instance, Add Storage (highlighted in orange), Add Tags, Configure Security Group, and 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.

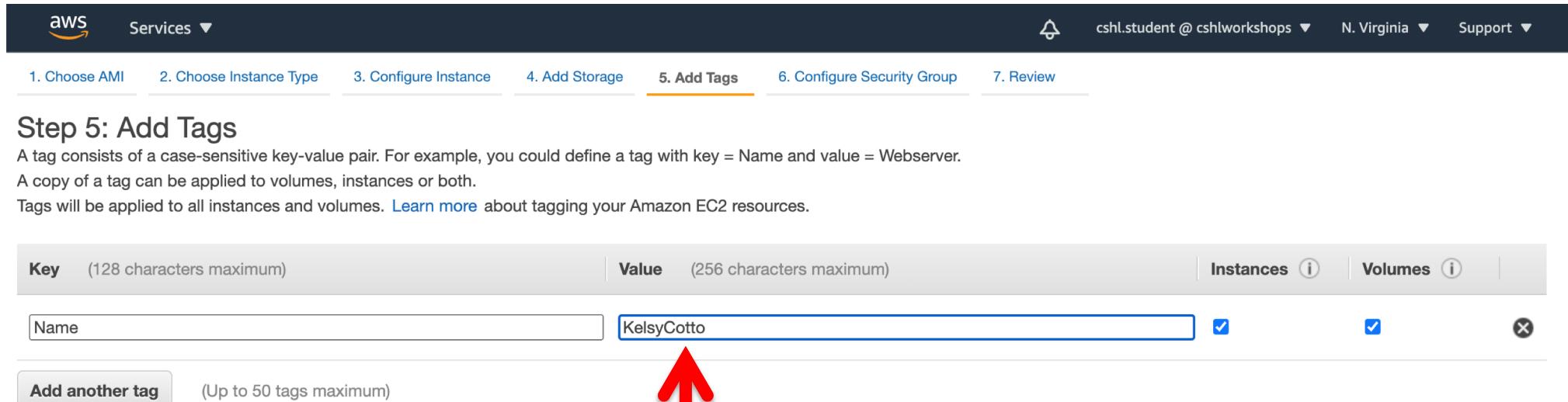
Volume Type	Device	Snapshot	Size (GiB)	Volume Type	IOPS	Throughput (MB/s)	Delete on Termination	Encryption
Root	/dev/sda1	snap-00fec8d273403984a	32	General Purpose SSD (gp2)	100 / 3000	N/A	<input checked="" type="checkbox"/>	Not Encrypted
EBS	/dev/sdb	snap-07c1c7fa70fa4e	250	General Purpose SSD (gp2)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

At the bottom are buttons: Cancel, Previous, Review and Launch (blue), and Next: Add Tags (red box).

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



The screenshot shows the AWS EC2 instance creation wizard at Step 5: Add Tags. The navigation bar includes links for Choose AMI, Choose Instance Type, Configure Instance, Add Storage, Add Tags (which is highlighted), Configure Security Group, and Review. The main area is titled "Step 5: Add Tags" with a 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." It also notes that tags can be applied to volumes, instances, or both, and that tags will be applied to all instances and volumes. A link to learn more about tagging is provided. The "Value" field is highlighted with a red arrow and contains the text "KelsyCotto". Below the fields are buttons for "Add another tag" and "Instances" and "Volumes" checkboxes. At the bottom, there are buttons for Cancel, Previous, Review and Launch (which is highlighted with a red box), and Next: Configure Security Group.

Key (128 characters maximum) | Value (256 characters maximum) | Instances | Volumes |

Name

Add another tag (Up to 50 tags maximum)

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

Select an Existing Security Group, choose "SSH_HTTP". Then hit "Review and Launch".

The screenshot shows the AWS EC2 instance creation process at Step 6: Configure Security Group. The user has selected the 'Select an existing security group' option and chosen the 'SSH and HTTP' group. The 'Review and Launch' button is highlighted with a red box.

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-384f5b79	default	default VPC security group	Copy to new
sg-0087dc3a8b6e37a2d	SSH and HTTP	created 2019-11-08T09:43:29.293-05:00	Copy to new

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

The screenshot shows the AWS Step 7: Review Instance Launch page. At the top, there are tabs: 1. Choose AMI, 2. Choose Instance Type, 3. Configure Instance, 4. Add Storage, 5. Add Tags, 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.

Warning 1: 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 2: 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 (highlighted by a red arrow)

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

Instance Type (highlighted by a red arrow)

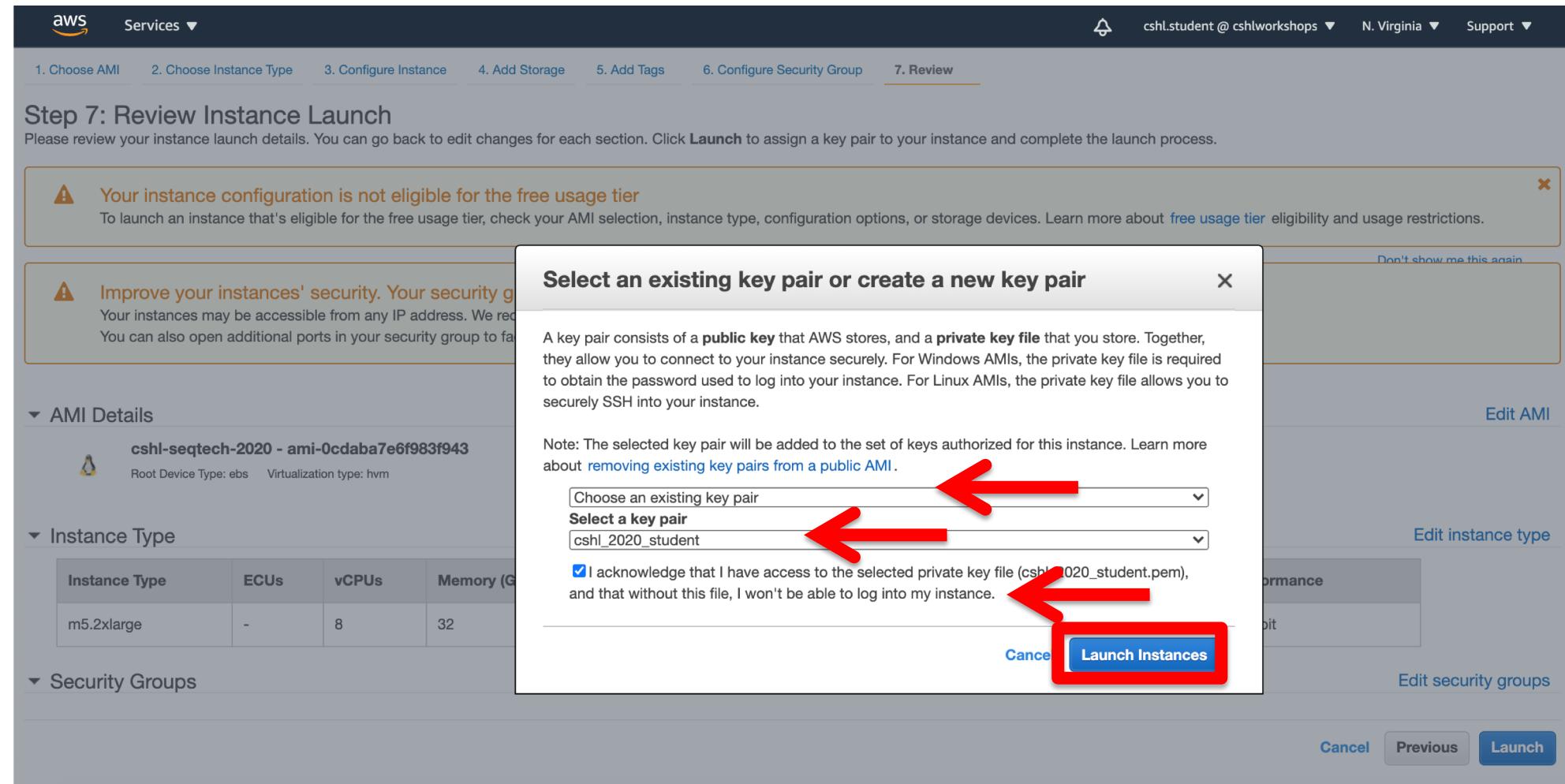
Instance Type	ECUs	vCPUs	Memory (GiB)	Instance Storage (GB)	EBS-Optimized Available	Network Performance
m5.2xlarge	-	8	32	EBS only	Yes	Up to 10 Gigabit

Security Groups

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

Buttons at the bottom: Cancel, Previous, **Launch** (highlighted with a red box).

Choose an existing key pair: "cshl_2020_student" and then Launch.



View Instances to see your new instance spinning up!

The screenshot shows the AWS Launch Status page. At the top, there's a navigation bar with the AWS logo, 'Services' dropdown, 'Resource Groups' dropdown, a star icon, a bell icon, user information 'cshl.student @ cshlworkshops', location 'N. Virginia', and a 'Support' dropdown.

The main content area is titled 'Launch Status'. It contains a green box with a checkmark icon and the text 'Your instances are now launching'. Below this, it says 'The following instance launches have been initiated: i-08e73e43f17783273' and a 'View launch log' link.

Below this is a blue box with an info icon and the text 'Get notified of estimated charges'. It includes a link to 'Create billing alerts' and a descriptive text about receiving email notifications for estimated charges.

A section titled 'How to connect to your instances' follows. It explains that instances are launching and may take a few minutes to reach the 'running' state. It suggests clicking 'View Instances' to monitor status and provides a link to find out how to connect.

Under 'Helpful resources', there are two columns of links:

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

At the bottom, there are links for creating status check alarms, attaching EBS volumes, and managing security groups. A red button at the bottom right is labeled '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 Instances page. On the left, there's a sidebar with navigation links like EC2 Dashboard, Events, Tags, Limits, Instances (selected), and more. The main area displays a table of instances with columns for Name, Instance ID, Instance state, Instance type, Status check, Alarm status, Availability Zone, and Public IPv4 DNS. A red arrow points to the 'KelsyCotto' instance, which is selected (indicated by a blue border). Another red arrow points to the 'Public IPv4 address' field, which contains '100.24.122.188 | open address'. The 'Connect' button in the top right is also highlighted with a red box.

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...

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 'Connect to instance' page for an EC2 instance. At the top, there are tabs for 'EC2 Instance Connect', 'Session Manager', and 'SSH client'. A red arrow points to the 'SSH client' tab, which is highlighted in orange. Below the tabs, the instance ID 'i-0b012943b3ce51aee (KelsyCotto)' is listed. Step-by-step instructions for connecting via SSH are provided, along with examples of command-line inputs. A 'Cancel' button is at the bottom right.

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`

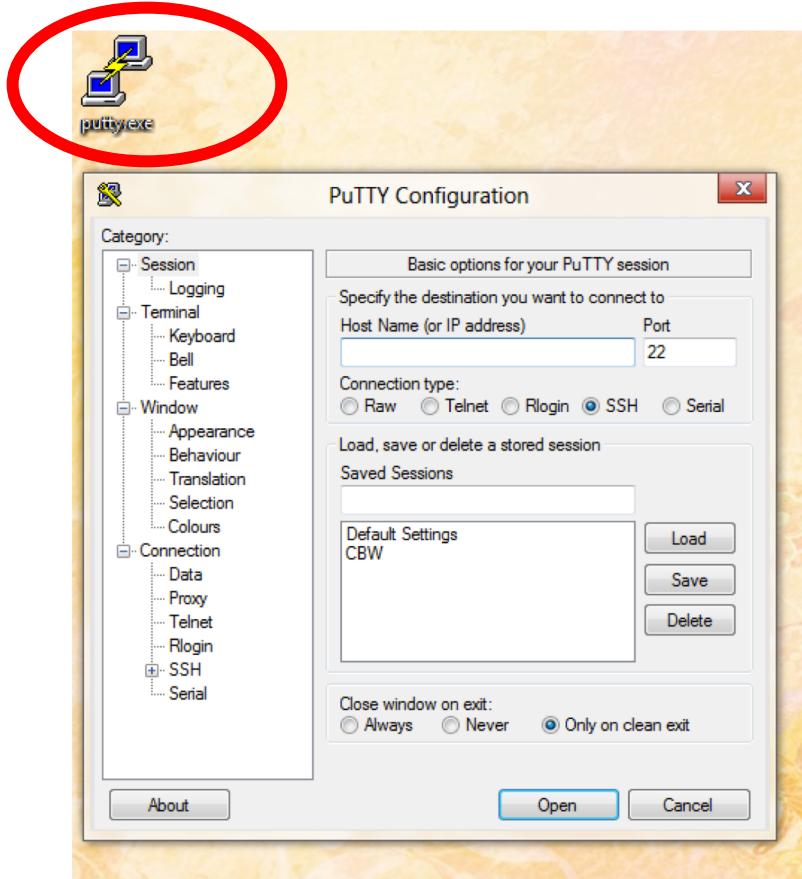
Example:

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

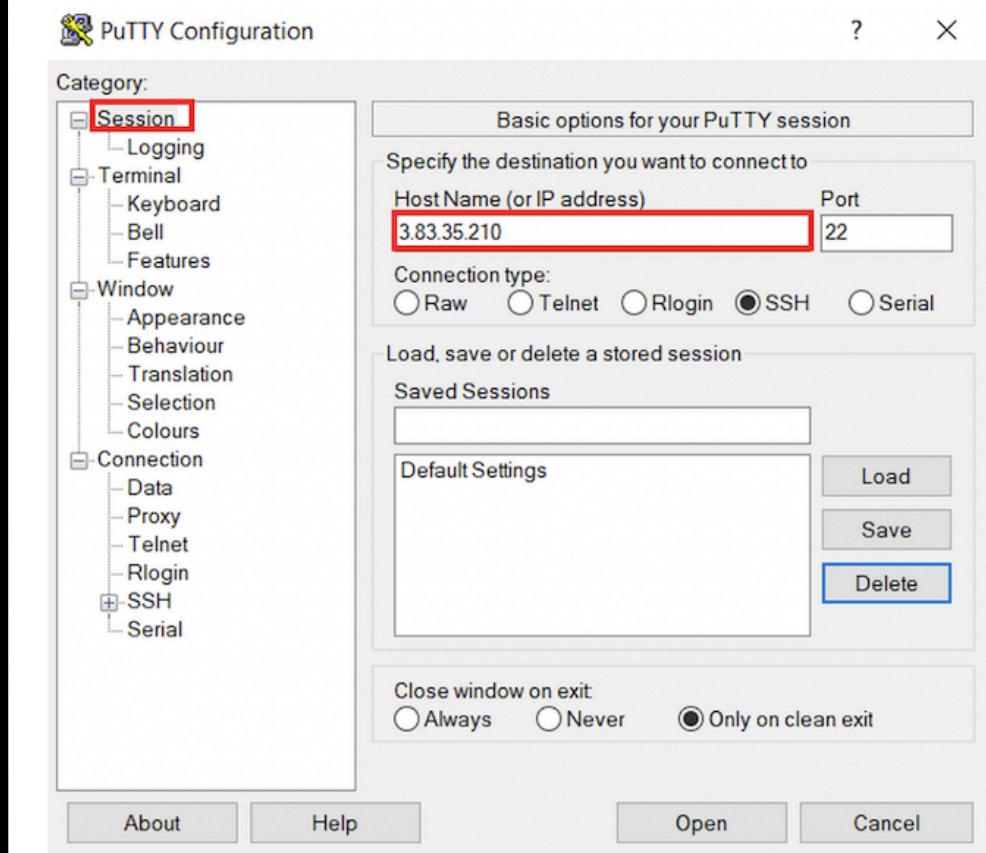
Cancel

Logging into your instance (Windows)

Open PuTTY

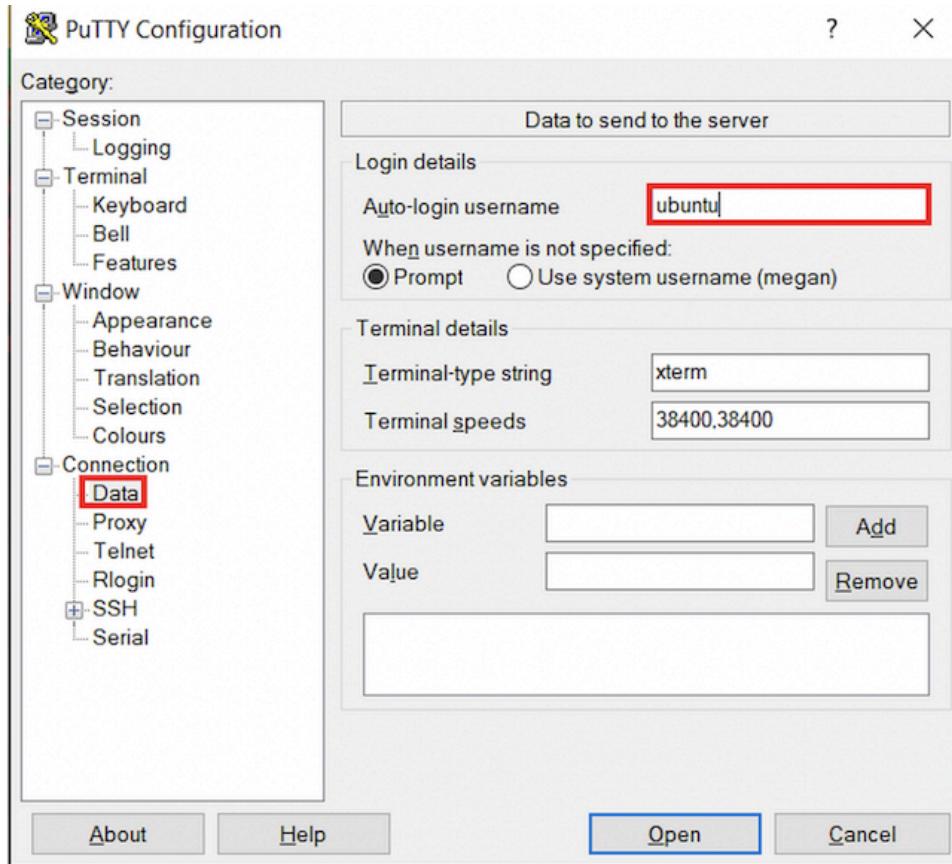


Enter the Host Name (IP address)

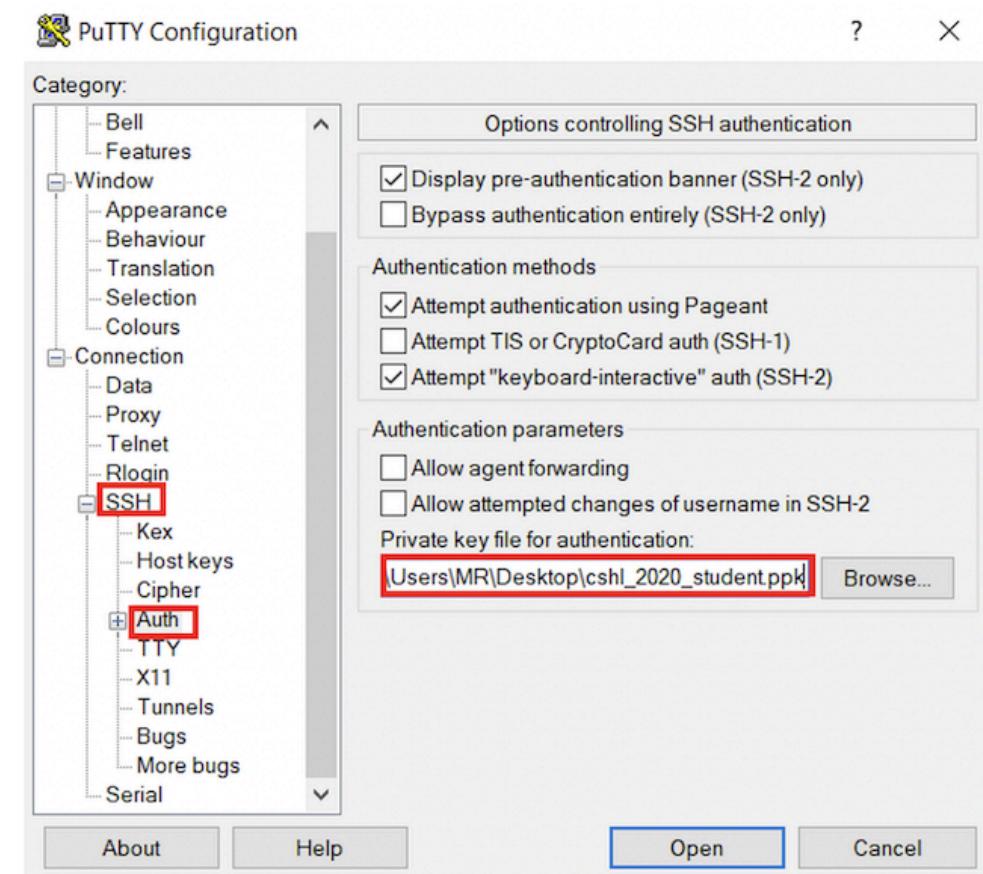


Logging into your instance (Windows)

Choose Connection -> Data
Enter the username 'ubuntu'



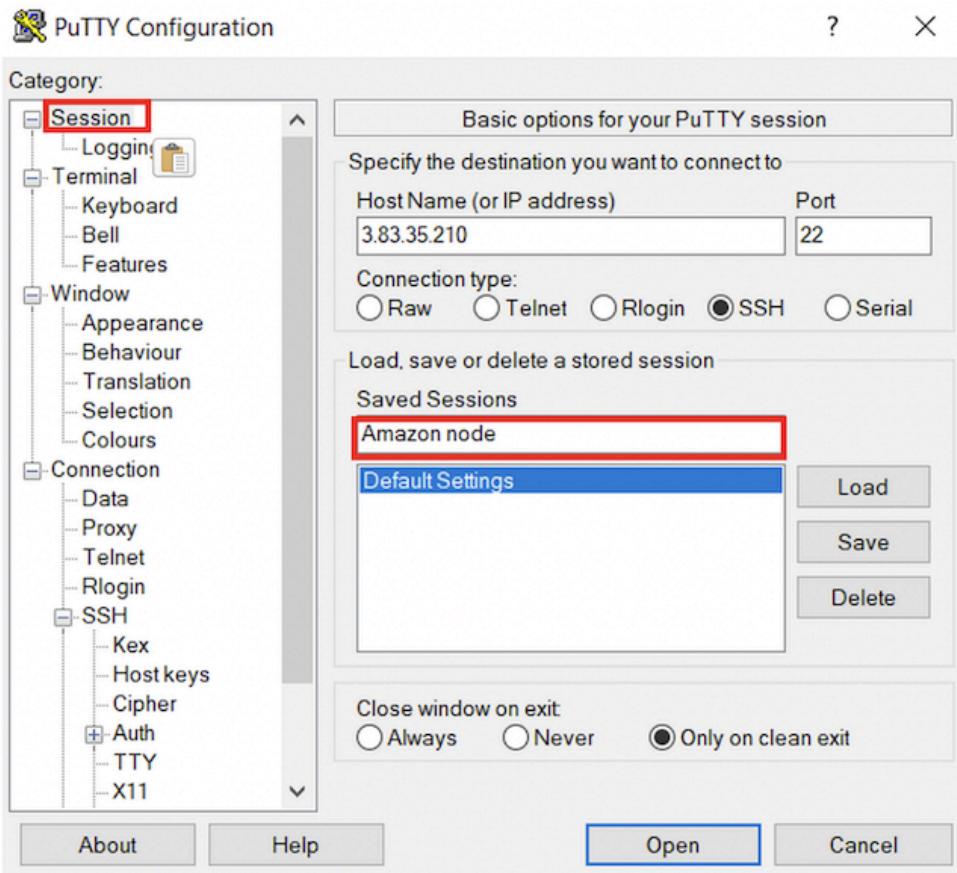
Choose SSH -> Auth
Browse to Private key (ppk) file



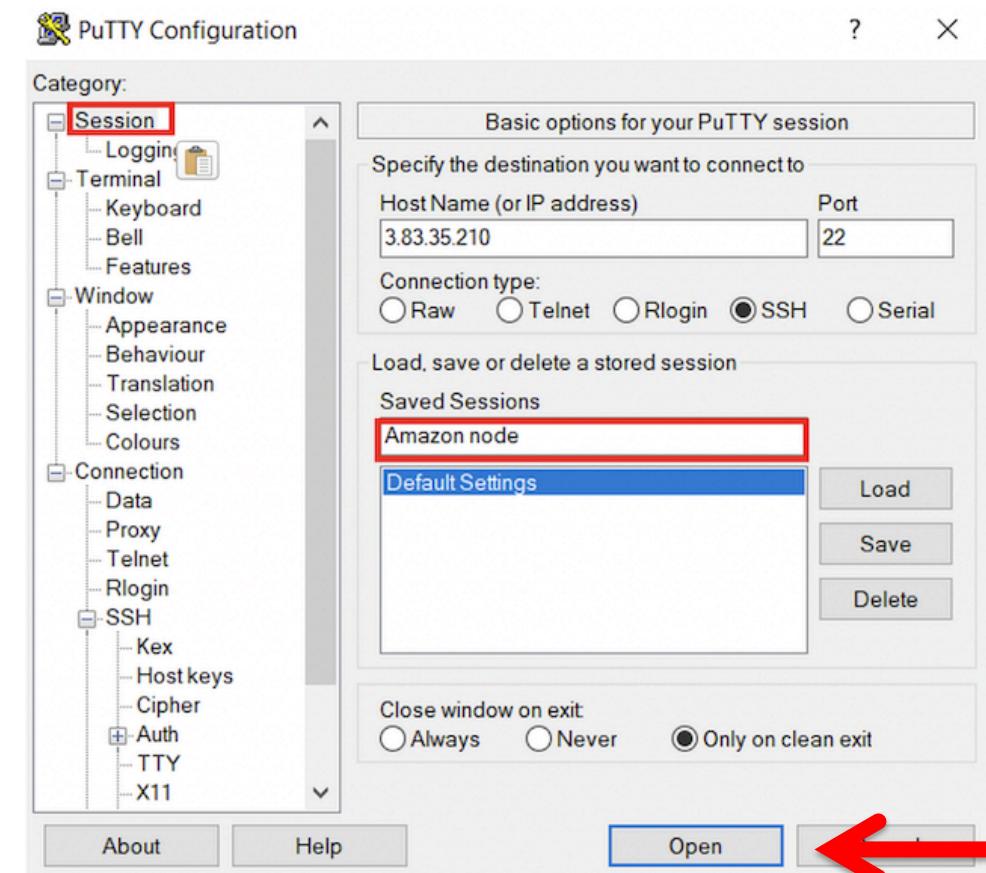
Logging into your instance (Windows)

Choose Session

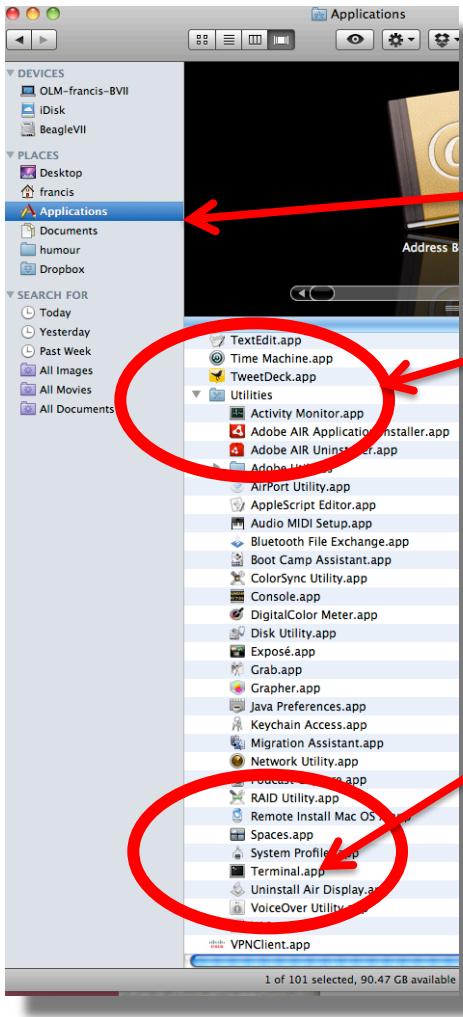
Save your session as "Amazon Node"



Double-click saved “Amazon Node” session OR
Select “Amazon Node” session and click Open



Logging into your instance (Mac)

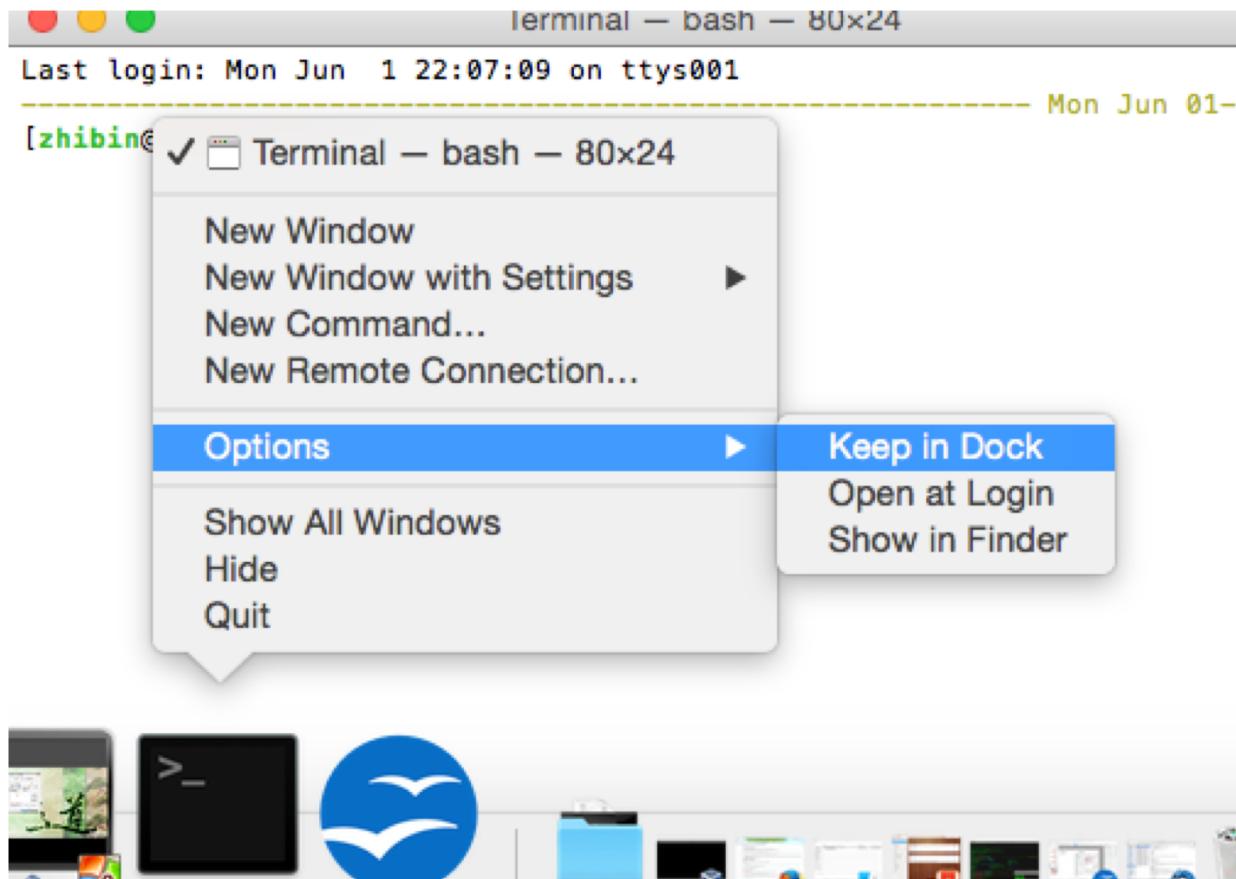


In a Finder window
'Applications' -> 'Utilities' -> 'Terminal'



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-----
MIIEpAIBAAKCAQEAgGtpnqERtEu/SCmeF2r1HMESMao0fEJiAwQwk2/SNXK8izr0IH0zTVvmE1D
VUwWq7pkvhjh05pDb+2U9HiZe3sxLv3S1NrkATYF/NsrpwB+q1vwqzGW9sQ6uj45RWrPkiZlsaj
TQZmyFRu+tLJTRU3hQDqA0MRWTx1Wxv0gFzuZy/qb+DALuFQsInrEKnijrwdLmd6usaBTvhc0gFS
B9oEelH0bZHJTZFW/wP+Z0uZq0Ujir7Qw0LTM45QH/L0dBdUl3k/mBeez00yvnKMwj8E4Xi0rQ0t
hHtQ7F9iSILK80W1rRH0qwxwt9ycEH1JtNMQmUTif0vE2XJ6l06chQIDAQABAoIBABG7P/FHu/Qp
WFgg+89myuqR6GvA2X55CFSzFyG0aQyrj5jDleFtdu2uXiISG8gUBZYvlzxx82aOC0P5j04SBq0
xD/qRlukY/jyXyPn77w/ExmaNoLJjl1W9RUSH0JYLIzVpFPGes3u5zGSGDTSDNh3sSdWhq1FX3l
7vY5b6UAQgahXf0dpGFxt6P6qb/BKFQFsThXk7GXMyS/kr4w7ZlhPWHRMSu2UBdS/a/beAVN76qU
E/10HR30oAuqghusZabpbX7MJl00VcviaQgeF8Z+xf2uugDEAKut1PxW0+yvGM4SpZZ0skFZz6YI
pBnX4ELWPhCeMzq4M18QY6ptR2UCgYEAc6YfNlymg24NJu8PANx8navTi50WYuXWXrj0mrLr3SR
/XY+w26cgipM+K5eQfxSr4Yb8BQKjRktMzBzf5nKdaX4pzYIquQH02B0HDhTooHAhkbTMTmKukv4
oEW06wcEE6RQi fw4xbMEnfQfHJB121am+jwI8Xb7idwMG4pU/nsCgYEAtD0x4bNC1X3A5by50uY0
WXrtgQszCXYcbkrpbjRET12f9hgz9MRMHY/xH/XGvMutZSFV2rCZRwd7lm+QWGadk/MQS0kouzW6
gSasyjFq+MKCkqYnS3/JTbx8yrLZmzl0LtX6pwmwg0Zy8aJjYHo9a2/EI8Tjh2d0SxeadIRVYP8C
gYAToiXww1Vdu+dj/7TDLqYCtdHOVAxJX/utI9Q3yoIryuh+bWmFvEIvAmIGXyyQZRyoZwgIS4A
PNH03+bEa+69wbzlhksiK5g8GKgISVdLC4rZZXB5ehgTmWV7IgJ89y/SF4G/Ityo30K0ohALh597
NcvNEzzqrutja1IIMvTKMwKBgQCkM+QP1Tqc0TbVlfvClviXuJBLsiJLCImYeZL0nZVmIMusbhxX
b8ZQYGSyUz09nulXau1G1QDvXvf089CzWL1SomxBoHlFJQvGwa9FfYQRIVPHuqut8rs4oPGn0QzC
h7M7QCJcr00oAcrSLLkQmgz+phIw7BzFr039J4HFiRInjQKBgQCtdEvcbtyk8Jh4WH3z0wpkc43f
U8DZhZwjRQpGWlD8CPj9RgRnE4+1PCH6s/RLQf7SiE1ZjX/0Ud0WPEvr0j5sVjy0IujohRbty0CM
oqWeSeUb1sLogRvMrTfCEpl/rz3GpoQ1SC/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>

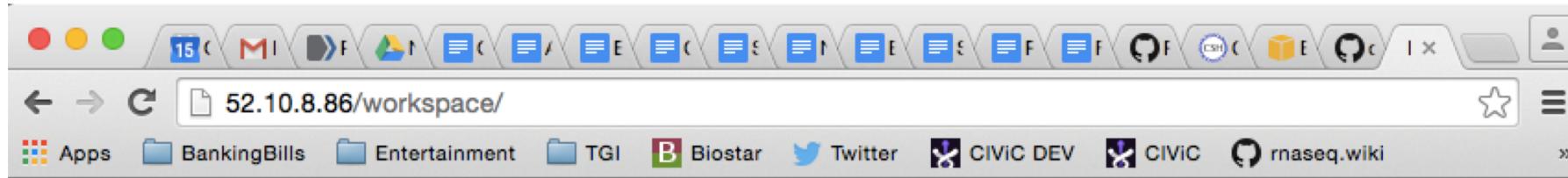
It is “r” for the the file owner **only**

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

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

Apache/2.4.7 (Ubuntu) Server at 52.10.8.86 Port 80

[http://\[YOUR PUBLIC DNS OR IP\]/](http://[YOUR PUBLIC DNS OR IP]/)

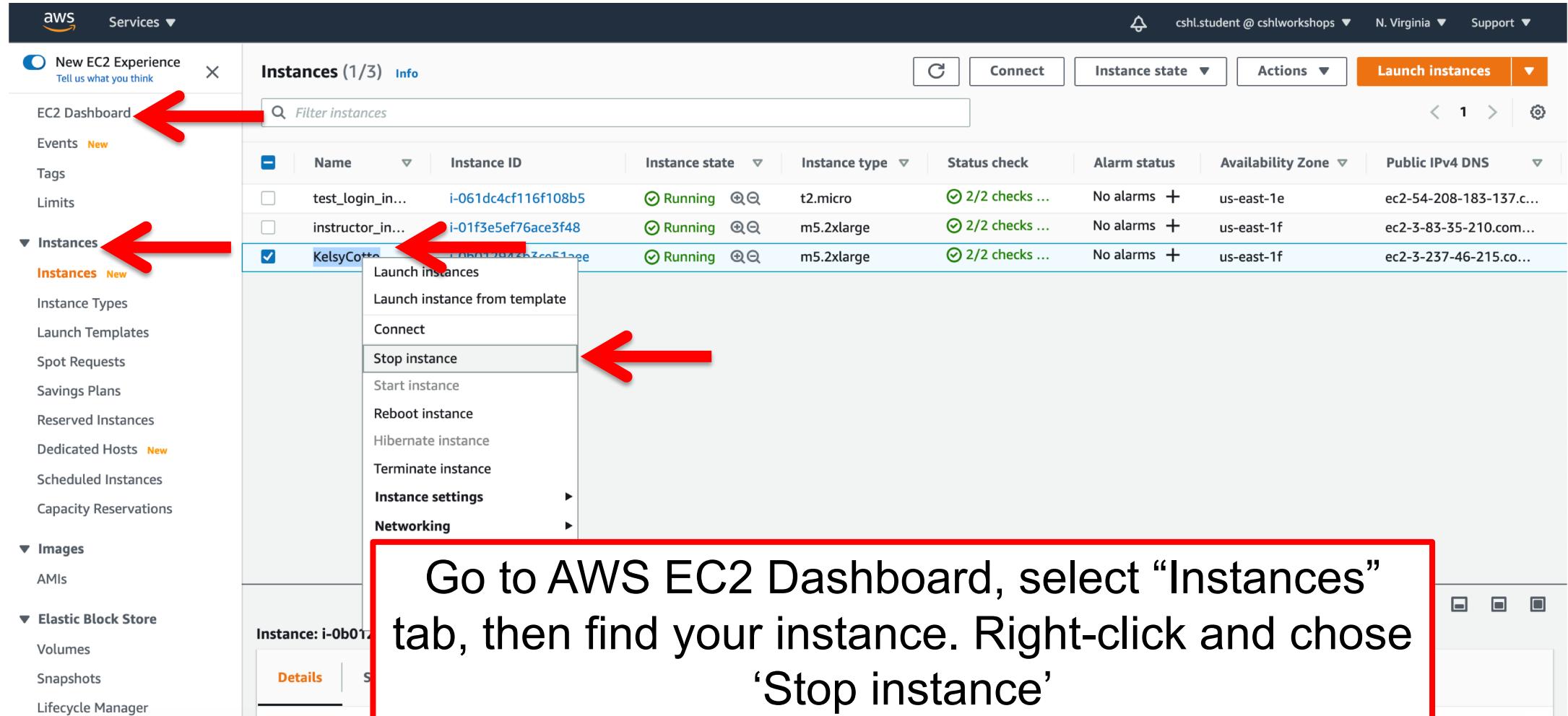
Logging out of your instance

Mac/Linux – simply type exit

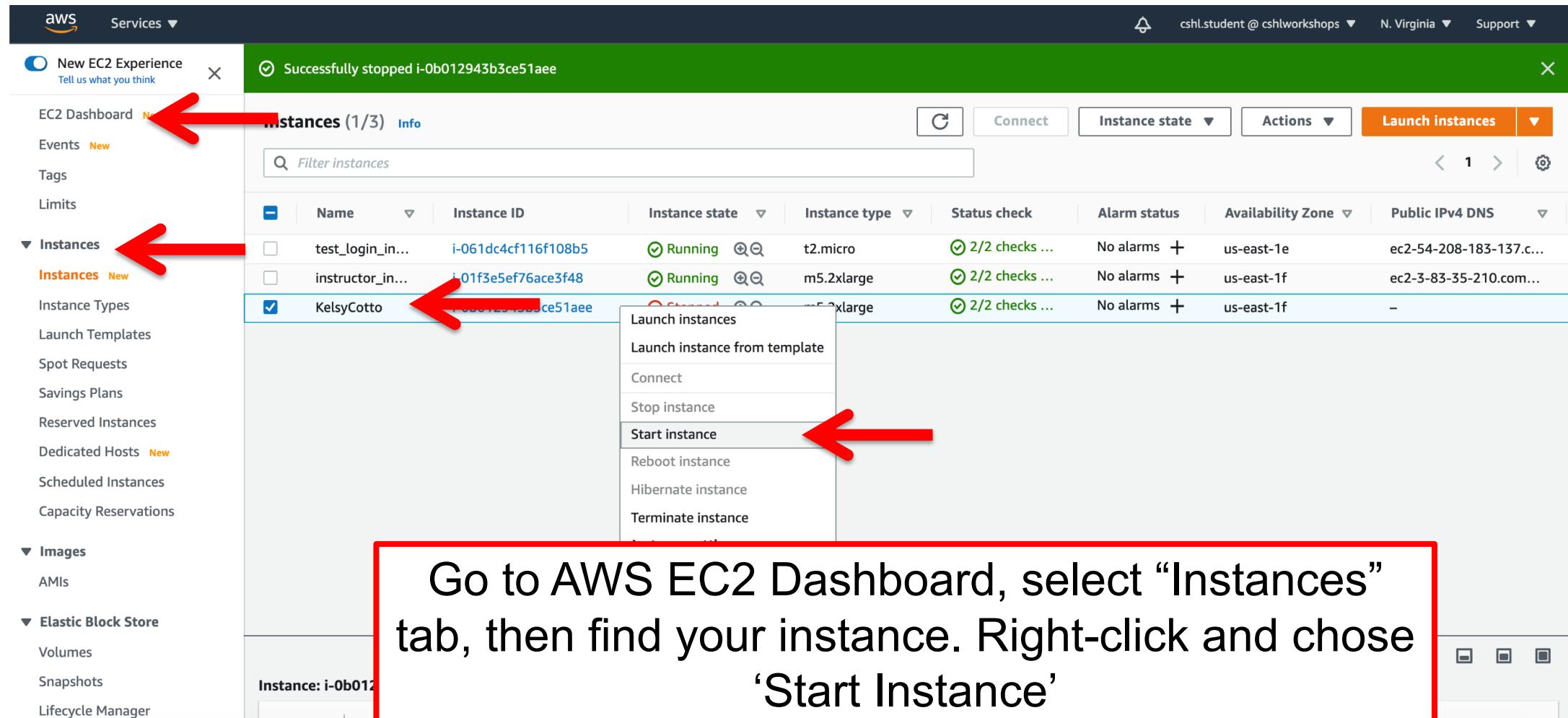
exit

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!



Next morning, you can “Start” your instance again



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. A red arrow points to the 'Connect' button in the top navigation bar. Another red arrow points to the Public IPv4 address '100.24.122.188' in the Instance Details section.

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	Running	m5.2xlarge	2/2 checks ...	No alarms	us-east-1f	ec2-100-24-122-188.c...

Instance: i-0b012943b3ce51aee (KelsyCotto)

Details | Security | Networking | Storage | Status Checks | Monitoring | Tags

Instance summary

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