



Cold
Spring
Harbor
Laboratory

Introduction to cloud computing

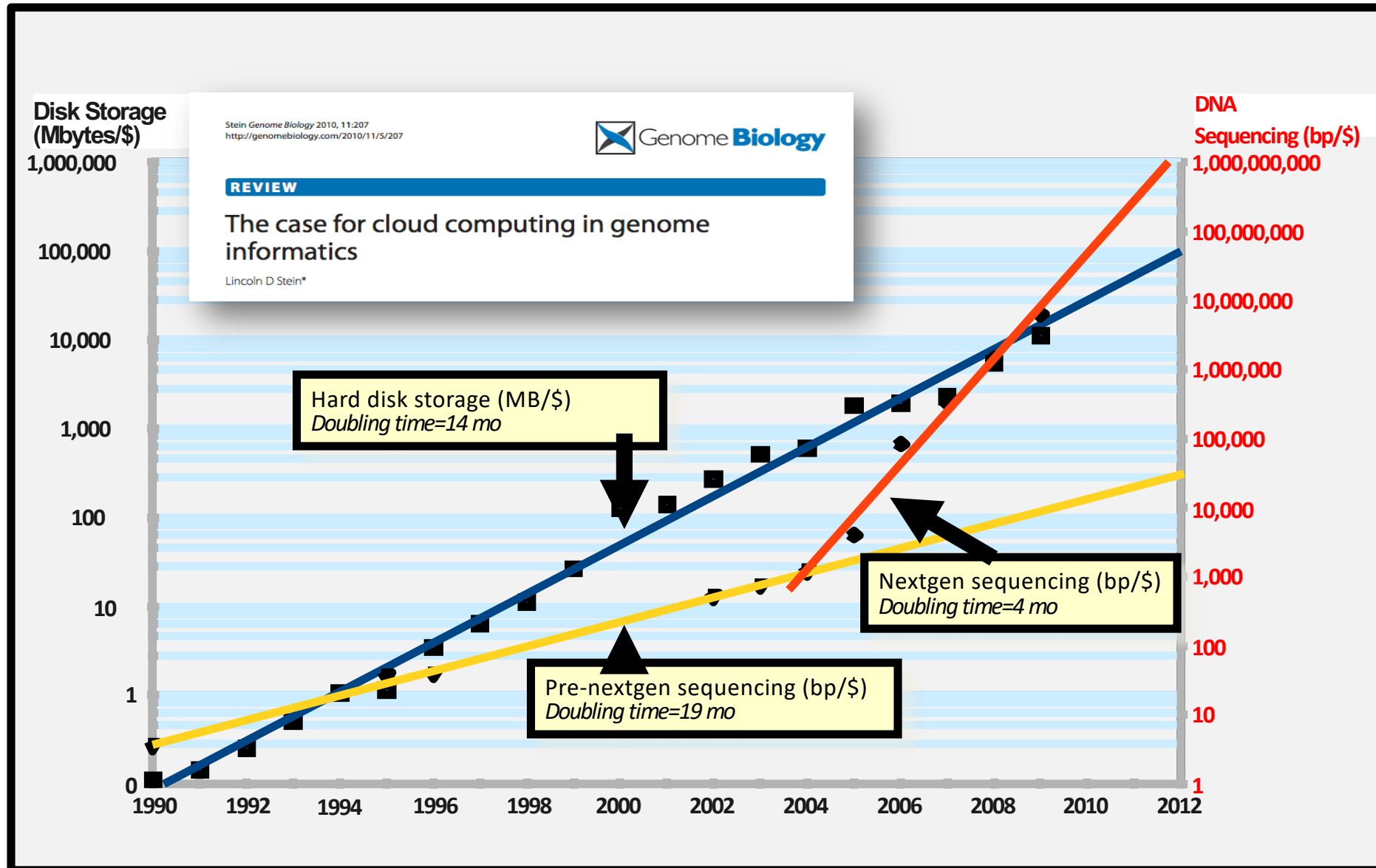
Kelsy Cotto, Arpad Danos, Harriet Dashnow, Felicia Gomez, Sharon Freshour, Obi Griffith,
Malachi Griffith, Jason Kunisaki, Chris Miller, Jonathan Preall, Aaron Quinlan
Advanced Sequencing Technologies & Bioinformatics Analysis November 11-19, 2021



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:

- 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. The 'Linux' tab is selected, and the region is set to 'US West (Oregon)'. The table below lists various EC2 instance types under the 'General Purpose - Current Generation' category. The columns are vCPU, ECU, Memory (GiB), Instance Storage (GB), and Linux/UNIX Usage. The 'Linux/UNIX Usage' column shows the hourly cost for each 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 (genomedata.org)
- 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

IAM user name

Password

Sign In

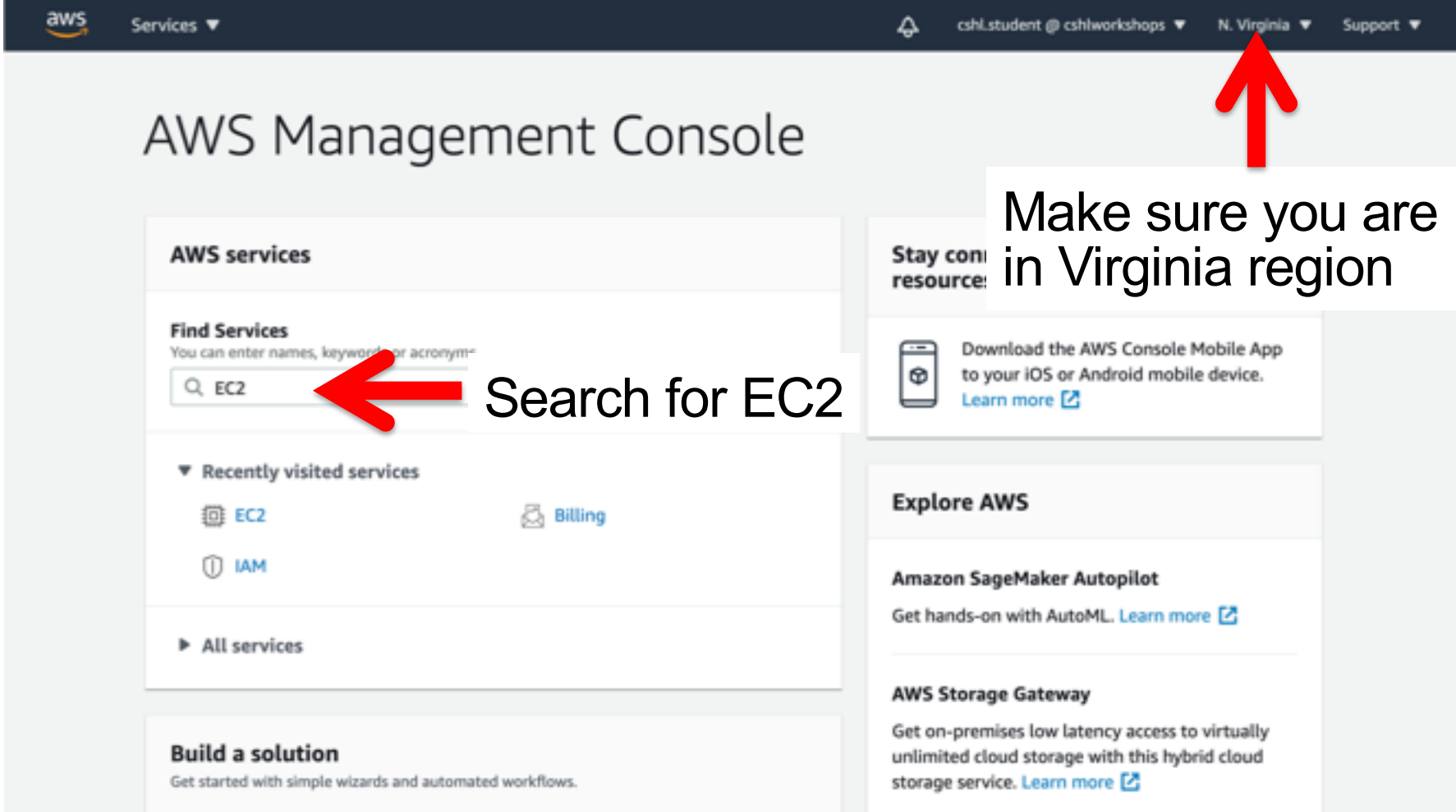


[Sign-in using root account credentials](#)

[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 user profile 'cshl.student @ cshlworkshops', the region 'N. Virginia', and 'Support'. The main heading is 'AWS Management Console'. On the left, under 'AWS services', there is a 'Find Services' search bar containing the text 'EC2'. A red arrow points from the text 'Search for EC2' to this search bar. Another red arrow points from the text 'Make sure you are in Virginia region' to the 'N. Virginia' region dropdown in the top right. Below the search bar, there is a 'Recently visited services' section with icons for 'EC2', 'Billing', and 'IAM'. The right side of the console features a 'Stay connected' section with a mobile app download prompt and an 'Explore AWS' section with links for 'Amazon SageMaker Autopilot' and 'AWS Storage Gateway'.

From EC2 Dashboard, launch a new Instance

The screenshot shows the AWS Management Console interface for the EC2 service. The top navigation bar includes the AWS logo, 'Services', and user information. The left-hand navigation pane lists various EC2-related options, with 'EC2 Dashboard' highlighted and a red arrow pointing to it. The main content area is divided into several sections: 'Resources' (listing metrics for running instances, elastic IPs, key pairs, etc.), 'Account attributes' (showing supported platforms and VPC settings), and 'Additional information' (with links to guides and documentation). At the bottom, the 'Launch instance' section contains a prominent orange 'Launch Instance' button, which is also indicated by a red arrow.

Choose an AMI – Find the CSHL SEQTEC 2021 AMI in the My AMIs

The screenshot shows the AWS console interface for selecting an AMI. The top navigation bar includes the AWS logo, 'Services', a search bar, and user information. Below the navigation bar, a progress indicator shows seven steps: 1. Choose AMI (active), 2. Choose Instance Type, 3. Configure Instance, 4. Add Storage, 5. Add Tags, 6. Configure Security Group, and 7. Review. The main heading is 'Step 1: Choose an Amazon Machine Image (AMI)' with a 'Cancel and Exit' link. A search bar is present with the placeholder text 'Search for an AMI by entering a search term e.g. "Windows"'. On the left, a sidebar contains navigation options: 'Quick Start', 'My AMIs' (highlighted with a red arrow), 'AWS Marketplace', and 'Community AMIs'. Below these are filters for 'Ownership' (Owned by me, Shared with me) and 'Architecture' (32-bit (x86), 64-bit (x86), 64-bit (Arm), 64-bit (Mac)). The main content area displays a table of AMIs. The first row is 'csih-seqtech-2021 - ami-07524de9e52dbd348' with a red arrow pointing to its 'Select' button. The second row is 'csih-seqtech-2020 - ami-09ecbedc3b79937e3' with its own 'Select' button. The table columns include 'Root device type: ebs', 'Virtualization type: hvm', 'Owner: 577255725291', and 'ENA Enabled: Yes'. A pagination control at the bottom right shows '1 to 2 of 2 AMIs'.

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

aws Services cshl.student @ cshworkshops 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".

The screenshot shows the AWS console interface for configuring an instance. The breadcrumb trail at the top indicates the current step is "3. Configure Instance Details". The page title is "Step 3: Configure Instance Details".

The configuration options are as follows:

- 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](#)
Warning: 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, the navigation buttons are: Cancel, Previous, Review and Launch, and Next: Add Storage. The "Next: Add Storage" button is highlighted with a red box.

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 console interface for configuring an EC2 instance. The current step is "Step 4: Add Storage". The interface displays a table of storage volumes with the following columns: Volume Type, Device, Snapshot, Size (GiB), Volume Type, IOPS, Throughput (MB/s), Delete on Termination, and Encryption. Two volumes are listed:

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 Encrypt
EBS	/dev/sdb	snap-07c1c7fa70fa4e	250	General Purpose SSD (gp2)	750 / 3000	N/A	<input checked="" type="checkbox"/>	Not Encrypt

Below the table, there is a button labeled "Add New Volume". At the bottom of the console, there are four buttons: "Cancel", "Previous", "Review and Launch", and "Next: Add Tags". The "Next: Add Tags" button is highlighted with a red box.

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 title 'Step 5: Add Tags' is followed by explanatory text: '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 (128 characters maximum)	Value (256 characters maximum)	Instances (i)	Volumes (i)
Name	KelsyCotto	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Below the table, there is a button 'Add another tag' with the text '(Up to 50 tags maximum)'. At the bottom of the console, there are four buttons: 'Cancel', 'Previous', 'Review and Launch', and 'Next: Configure Security Group'. The 'Next: Configure Security Group' button is highlighted with a red rectangular box. A red arrow points from the text below to the 'Value' field of the tag entry.

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	Copy to new
<input checked="" type="checkbox"/> 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

Step 7: Review Instance Launch

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](#)

ami-07524de9e52dbd348
Root Device Type: ebs Virtualization type: hvm

Instance Type [Edit instance type](#)

Instance Type	ECUs	vCPUs	Memory (GiB)	Instance Storage (GiB)	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-0087dc3af8b6e37a2d	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.

The screenshot shows the AWS Management Console interface for launching an EC2 instance. The main page is at 'Step 7: Review Instance Launch'. A modal dialog is open in the foreground with the title 'Select an existing key pair or create a new key pair'. The dialog contains the following elements:

- A paragraph explaining that a key pair consists of a public key (stored by AWS) and a private key file (stored by the user).
- A note stating that the selected key pair will be added to the set of keys authorized for the instance.
- A dropdown menu labeled 'Choose an existing key pair' with a red arrow pointing to it.
- A second dropdown menu showing the selected key pair: 'cshl_2021_student | RSA', with a red arrow pointing to it.
- A checked checkbox with the text: 'I acknowledge that I have access to the corresponding private key file, and that without this file, I won't be able to log into my instance.', with a red arrow pointing to it.
- A 'Launch Instances' button highlighted with a red box and a red arrow pointing to it.

The background console shows the instance configuration details, including AMI Details (cshl-seqtech-2021 - ami-07524de9e52dbd348), Instance Type (m5.2xlarge), and Security Groups (sg-0087dc3a8b6e37a2d).

View Instances to see your new instance spinning up!

Launch Status

✔ Your instances are now launching
The following instance launches have been initiated: [i-08e73e43f17783273](#) [View launch log](#)

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

How to connect to your instances

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

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

▼ Here are some helpful resources to get you started

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

While your instances are launching you can also

- [Create status check alarms](#) to be notified when these instances fail status checks. (Additional charges may apply)
- [Create and attach additional EBS volumes](#) (Additional charges may apply)
- [Manage security groups](#)

[View Instances](#)

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

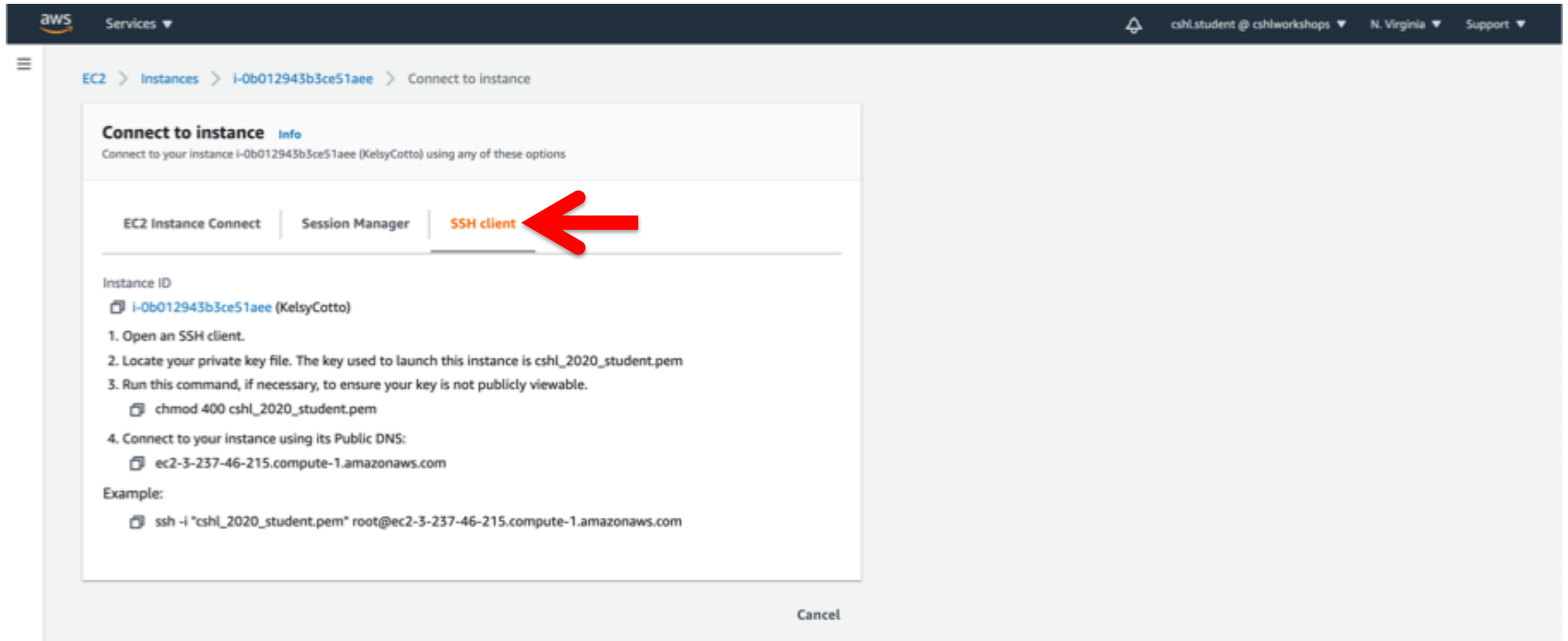
The screenshot displays the AWS Management Console interface for EC2 instances. At the top, the 'Connect' button is highlighted with a red box. Below it, a table lists three instances. The instance named 'KelsyCotto' is selected, indicated by a red arrow. The details for this instance are shown below, with a red arrow pointing to the 'Public IPv4 address' field, which displays '100.24.122.188'.

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)

Field	Value
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

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 Management Console interface for connecting to an EC2 instance. The breadcrumb navigation indicates the path: EC2 > Instances > i-0b012943b3ce51aee > Connect to instance. The main content area is titled "Connect to instance" and provides instructions for connecting to the instance i-0b012943b3ce51aee (KelsyCotto). Three tabs are visible: "EC2 Instance Connect", "Session Manager", and "SSH client", with a red arrow pointing to the "SSH client" tab. The "SSH client" tab contains the following instructions:

Instance ID
i-0b012943b3ce51aee (KelsyCotto)

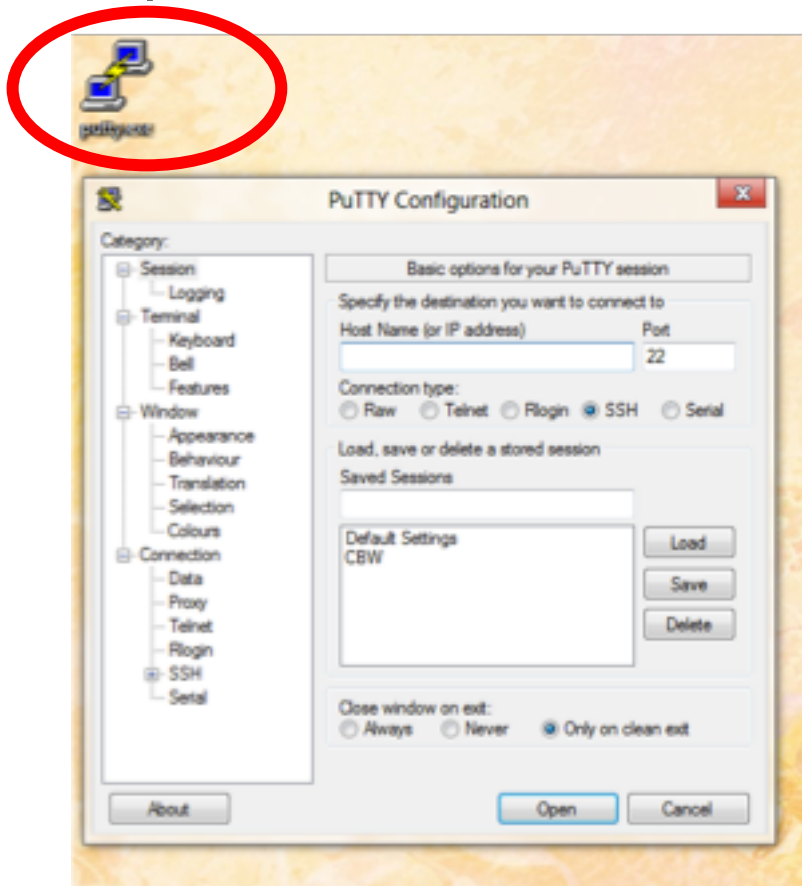
1. Open an SSH client.
2. Locate your private key file. The key used to launch this instance is `csH_2020_student.pem`
3. Run this command, if necessary, to ensure your key is not publicly viewable.
`chmod 400 csH_2020_student.pem`
4. Connect to your instance using its Public DNS:
`ec2-3-237-46-215.compute-1.amazonaws.com`

Example:
`ssh -i "csH_2020_student.pem" root@ec2-3-237-46-215.compute-1.amazonaws.com`

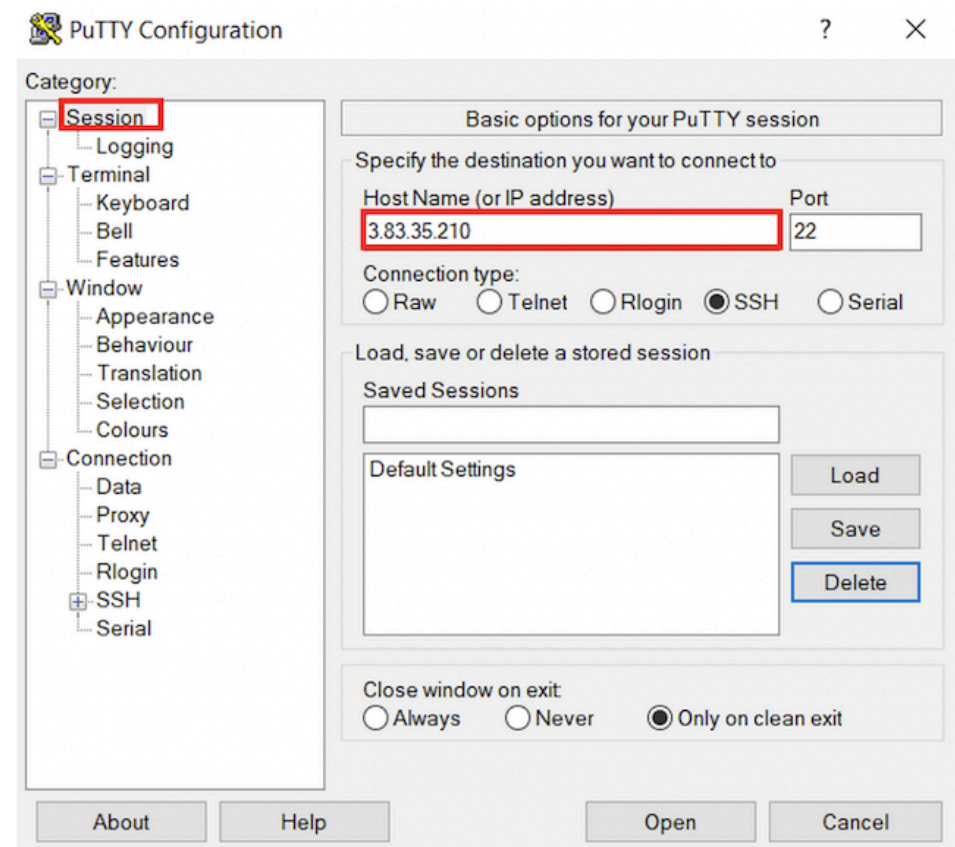
Cancel

Logging into your instance (Windows)

Open PuTTY

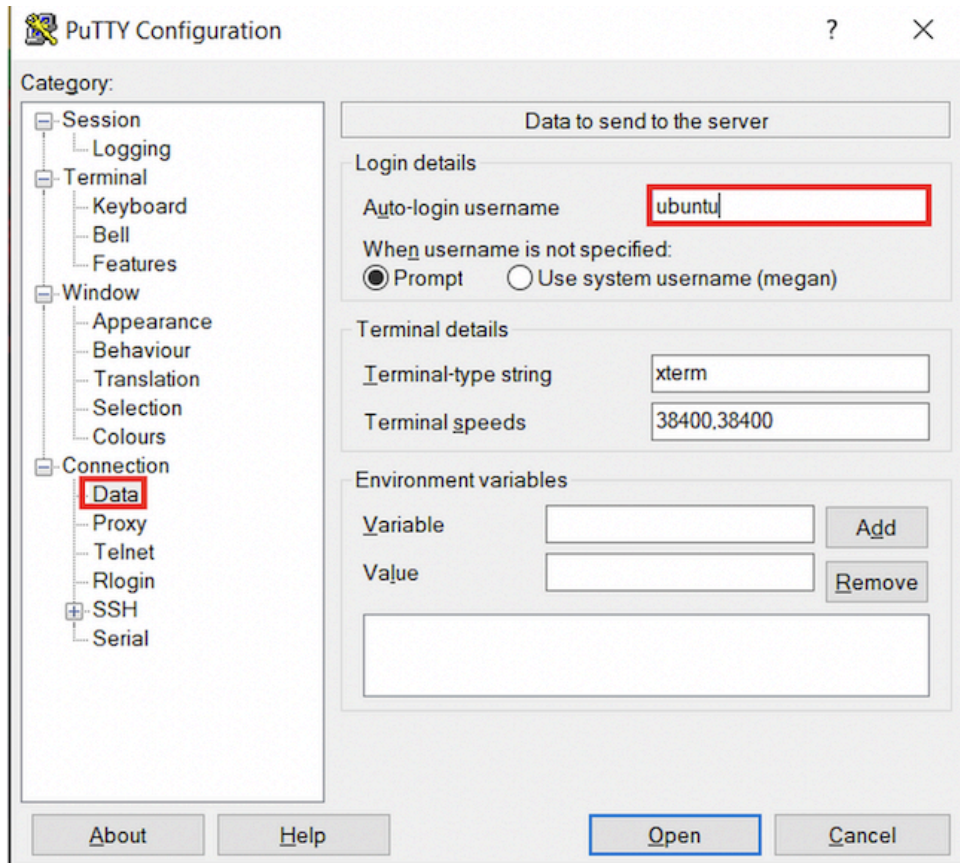


Enter the Host Name (IP address)

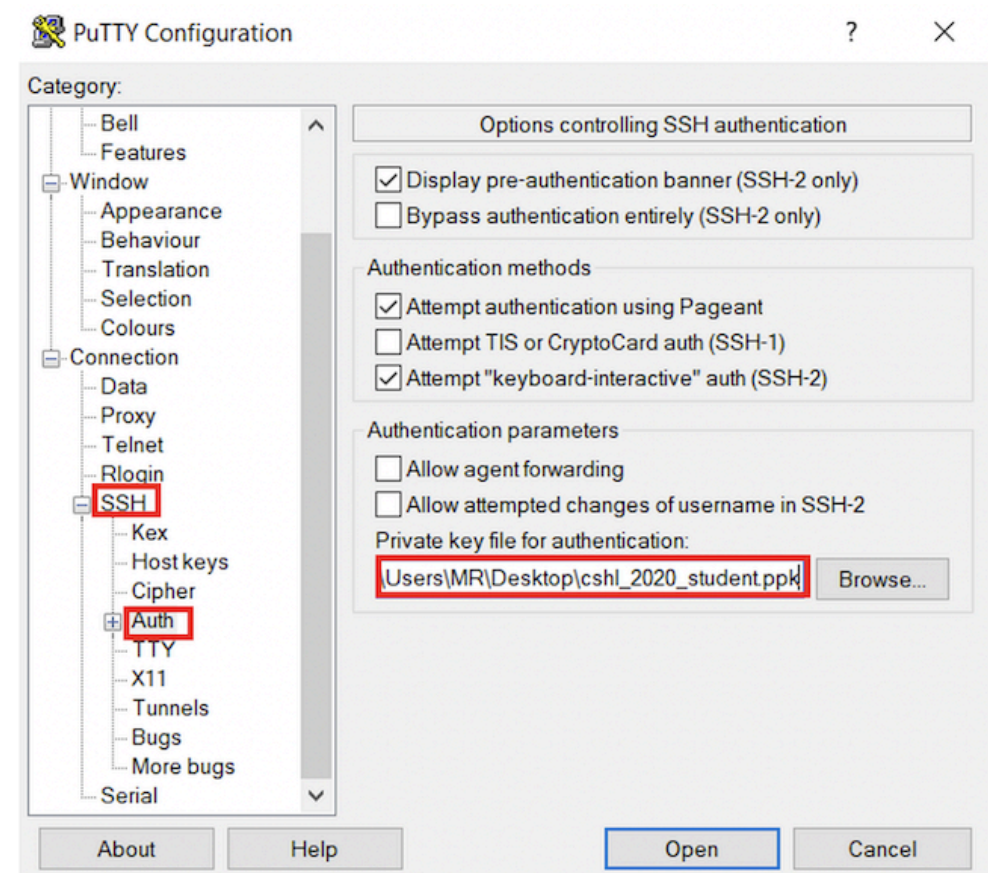


Logging into your instance (Windows)

Choose Connection -> Data
Enter the username 'ubuntu'

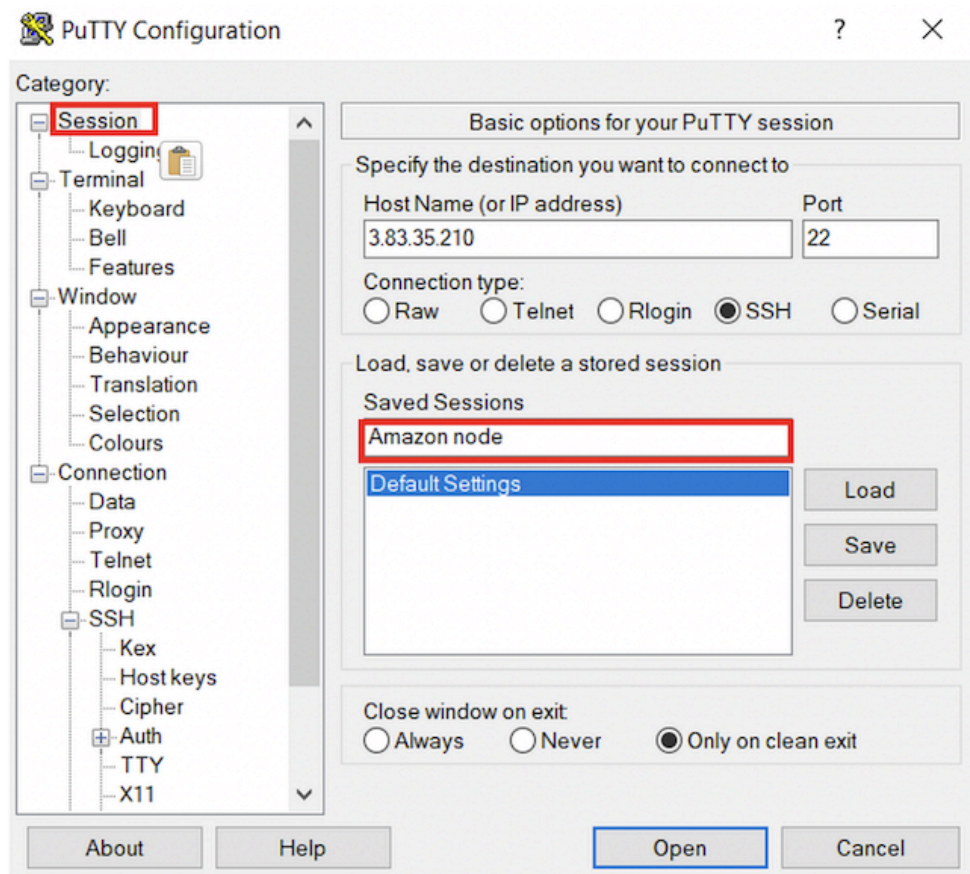


Choose SSH -> Auth
Browse to Private key (ppk) file

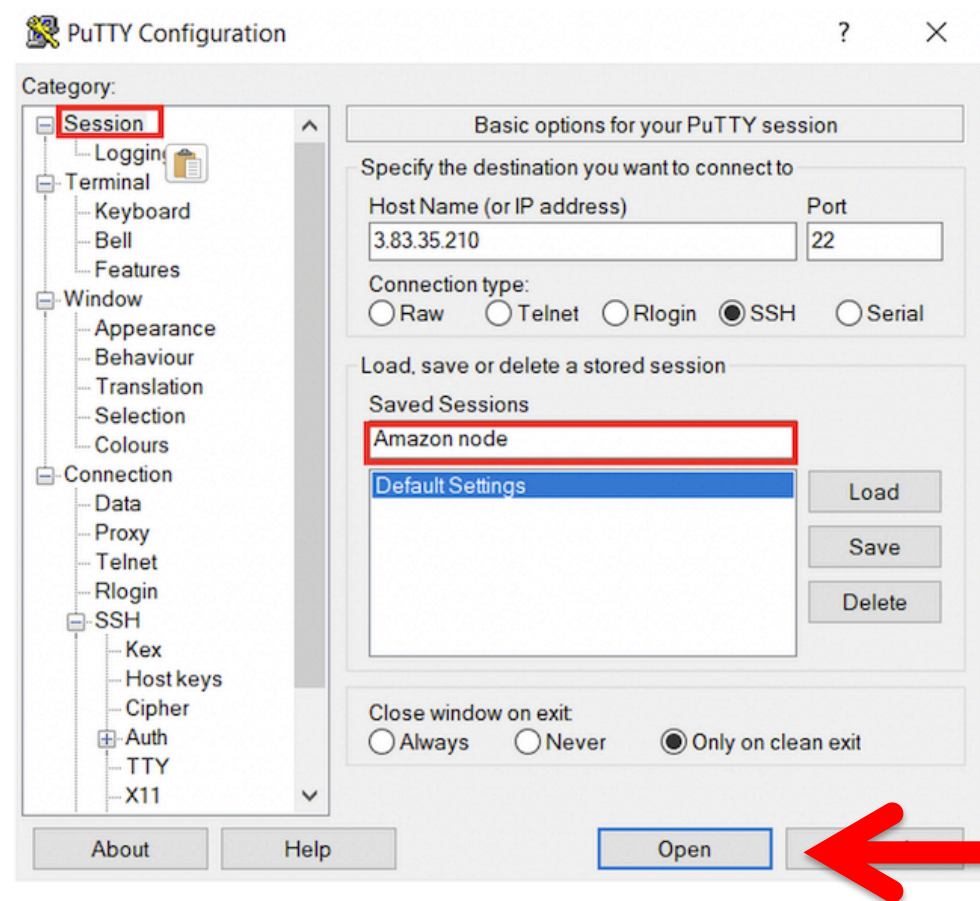


Logging into your instance (Windows)

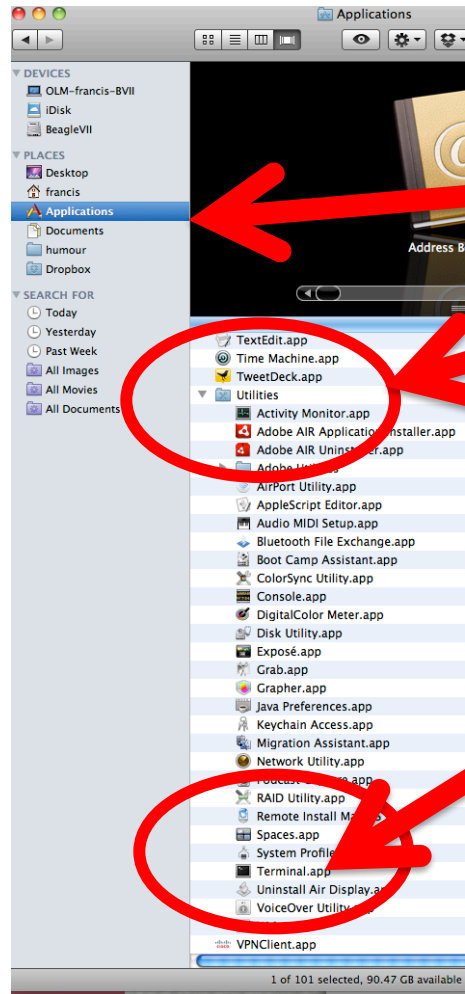
Choose Session
Save your session as "Amazon Node"



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



Logging into your instance (Mac)

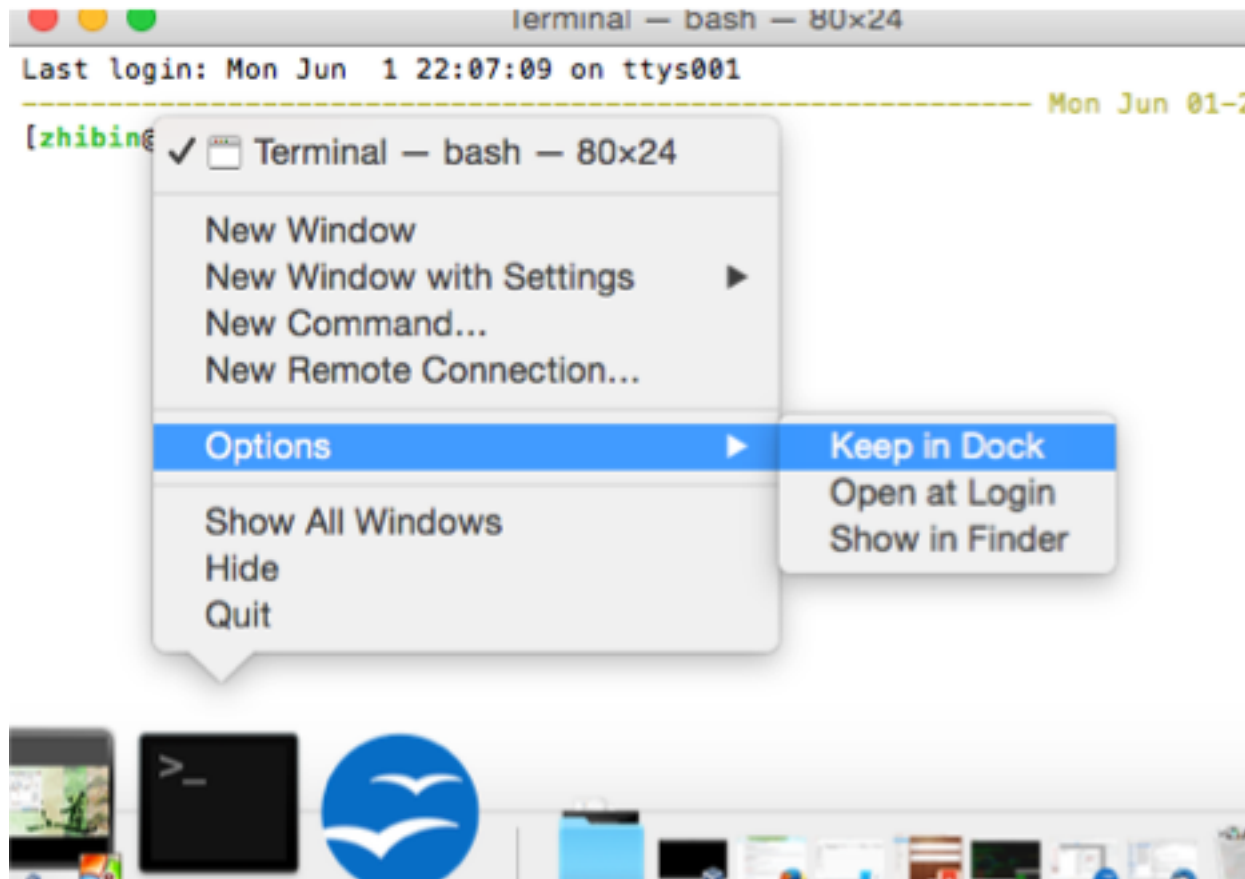


In a Finder window
'Applications' -> 'Utilities' -> 'Terminal'

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_2021_student.pem
```

```
Kelsys-MacBook-Pro:keys kcotto$ cat cshl_2019_student.pem
-----BEGIN RSA PRIVATE KEY-----
MIIEpAIBAACAQEAgtGtpnqERtEu/SCmeF2r1HMESMao0fEJiAwQwk2/SNXK8izr0IH0zTVvmE1D
VUwWq7pkvhhJh05pDb+2U9HiZe3sxLv3S1NrkATYF/Nsrpw8+q1vwqzGW9sQ6uj45RWrPkiZlsaj
TQZmyFRu+tLJTRU3hQDqA0MRWTx1Wxv0gFzuZy/qb+DALuFQsInrEKnijrwdLmd6usaBTvhc0gFS
B9oEelH0bZHJTZfw/wP+Z0uZq0Ujir7Qw0LTM45QH/L0dBdUl3k/mBeez00yvnKMwj8E4Xi0rQ0t
hHtQ7F9iSILK80W1rRH0qwxwt9ycEH1JtNMQmUTi f0vE2XJ6L06chQIDAQABAoIBABG7P/FHu/Qp
WFgg+89myuqR6GvA2X55CFsZzFYg0aQyrj5jDleFtdu2uXiISG8gUBZYvlzxx82a0C0P5j04SBq0
xD/qRlukY/jyXyPn77w/ExmaNoLJjl1W9RUSH0JYLIZvPFPGes3u5zGSGDTS0Nhw3sSdWhq1FX3L
7vY5b6UAQgahXf0dpGFxt6P6qb/BKFQfSthXk7GXMyS/kr4w7ZlhPWHrMSu2UBdS/a/beAVN76qU
E/10HR30oAuqghusZabpbX7MJl00VcviAQgeF8Z+xf2uugDEAKut1PxW0+yvGM4SpZZ0skFZz6YI
pBnX4ELWPhCeMzq4Ml8QY6ptR2UCgYEauc6Yfnlymg24NJu8PANx8navTi50WYuXWXRj0mrLr3SR
/XY+w26cgipM+KSeQfXsr4Yb8BQKjRktMzBzF5nKdaX4pzYIquQH02B0HDhTooHAhkbTMTmKukv4
oEW06wcEE6RQi fw4xbMEnfQfHJBI21am+jwI8Xb7i dwMG4pU/nsCgYEAtD0x4bNC1X3A5by50uY0
WXrtgQszCXYcbkrpbjRET12f9hgZ9MRMHY/xH/XGvMutZSFV2rCZRwd7lm+QWGadk/MQSOkouzW6
gSasyjFq+MKCkqYnS3/JTbx8yrLZmzl0LtX6pwmwmg0Zy8aJjYHo9a2/EI8Tjh2d0SxeadIRVYP8C
gYAToiXww1Vdu+dj/7TDLqYctdH0VAxJX/utI9QQ3yoIryuh+bWmFvEIvAmIGXyyQZRyoZwgIS4A
PNH03+bEa+69wbzlhksiK5g8GKgISVdlC4rZZXB5ehgTmWV7IgJ89y/SF4G/Ityo30K0ohALh597
NcvNEzzqruTja1IIMvTKMwKBgQCkM+QP1Tqc0TbVl fvClviXuJBLsiJLClmYeZL0nZVmIMusbhxX
b8ZQYGSyUz09nuLXaulGLQDvXvf089CzWLLSomxBoHlFJQvGwa9FfYQRIVPhuqut8rs4oPGn0QzC
h7M7QCJcr00oAcrSLLkQmgz+phIw7BzFr039J4HFIRInjQKBgQCtdEvcbtyk8Jh4WH3z0wpkc43f
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>

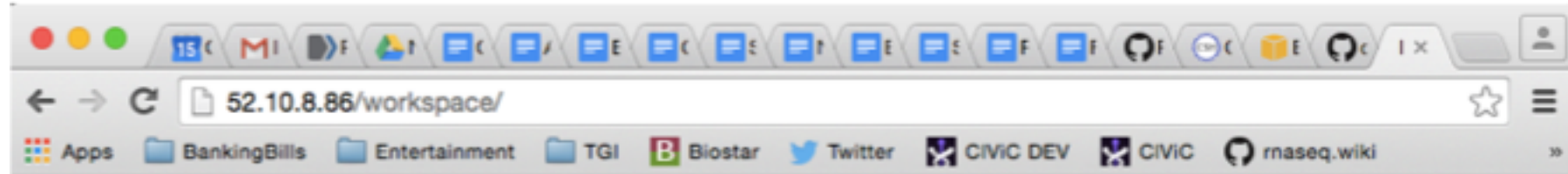
It is "r" for the the file owner **only**

Logging into your instance








Mac/Linux

```
cd ~/cshl  
chmod 400 cshl_2021_student.pem  
ssh -i cshl_2021_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]/

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, named 'KelsyCott...', is selected, and a context menu is open over it. 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-0b01...	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

The screenshot shows the AWS Management Console for EC2. The left sidebar has the 'Instances' tab selected. The main area shows a table of instances. The instance 'KelsyCotto' is selected, and a context menu is open over it, with 'Start instance' highlighted. A green notification bar at the top says 'Successfully stopped i-0b012943b3ce51aee'.

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-3e5ef76ace3f48	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, a navigation sidebar lists various services, with 'Instances' selected. The main content area shows a table of EC2 instances. The instance 'KelsyCotto' is selected, and its details are shown below. 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.