

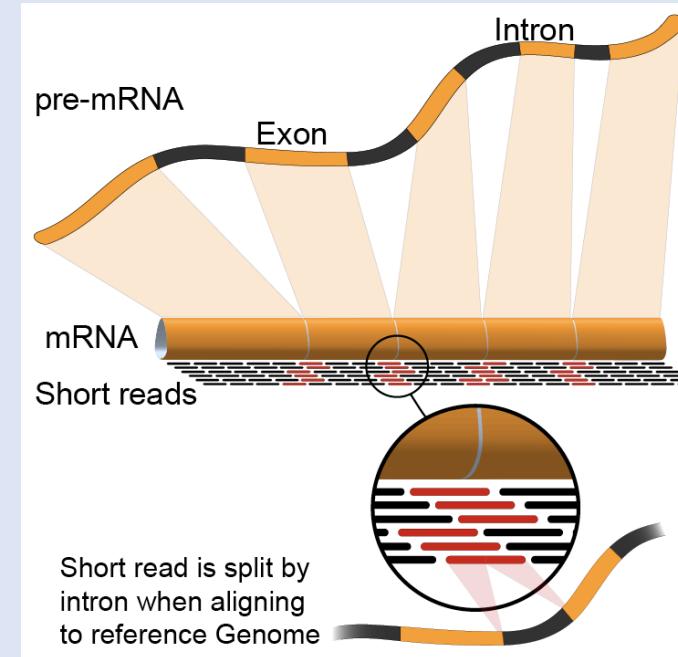
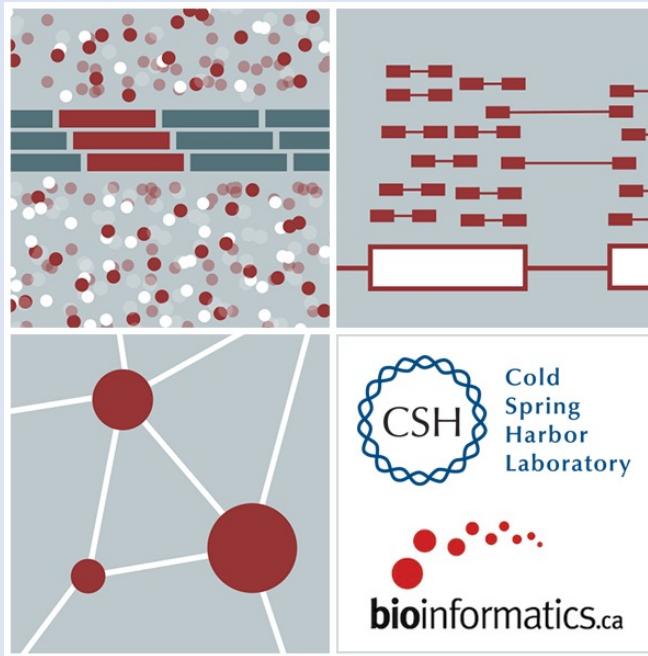


Cold
Spring
Harbor
Laboratory

Introduction to cloud computing

Arpad Danos, Felicia Gomez, Obi Griffith, Malachi Griffith,
My Hoang, Mariam Khanfar, Chris Miller, Kartik Singhal

Advanced Sequencing Technologies & Bioinformatics Analysis November 5-19, 2023

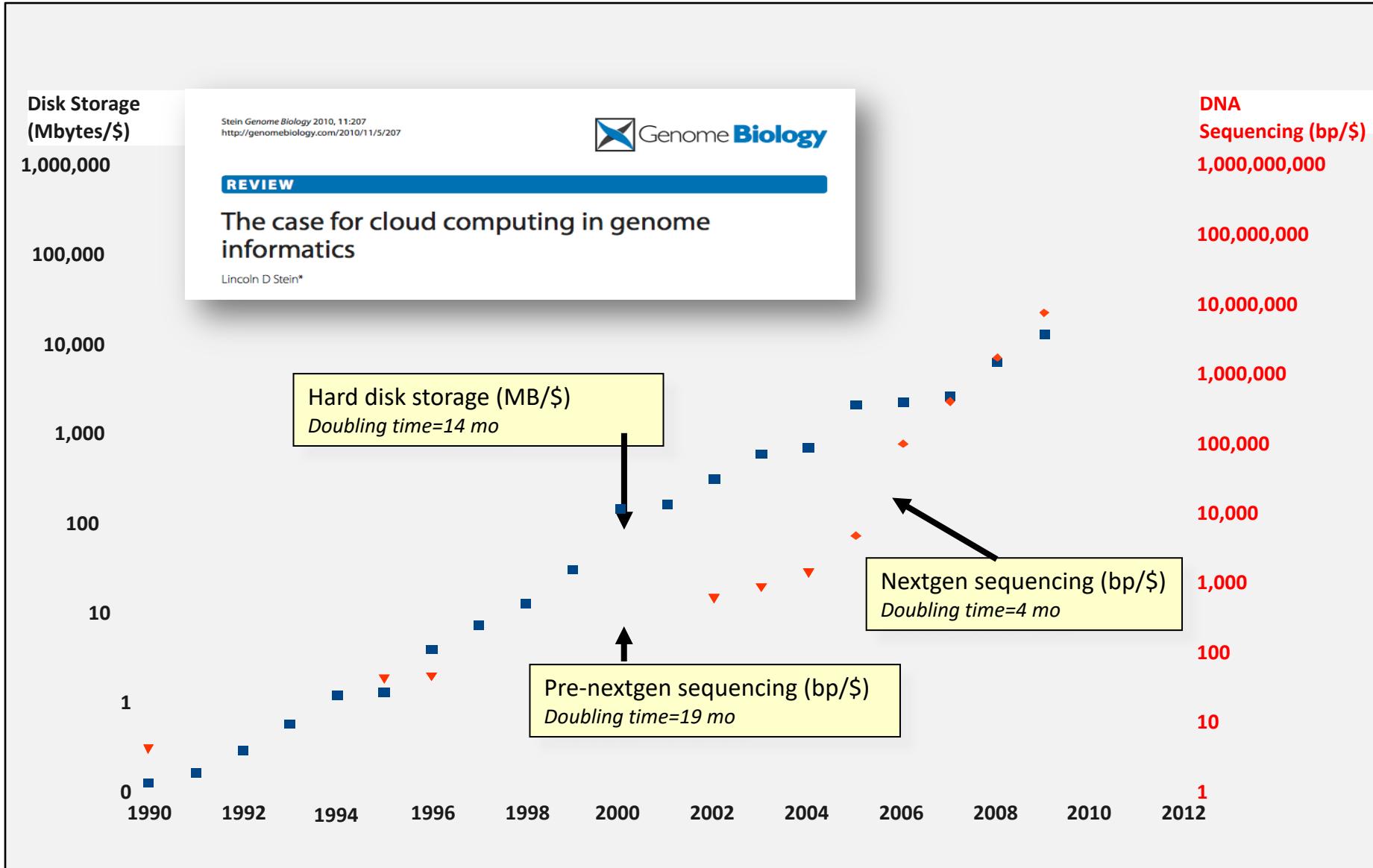


Washington University in St. Louis
SCHOOL OF MEDICINE

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)



All services

Services by category

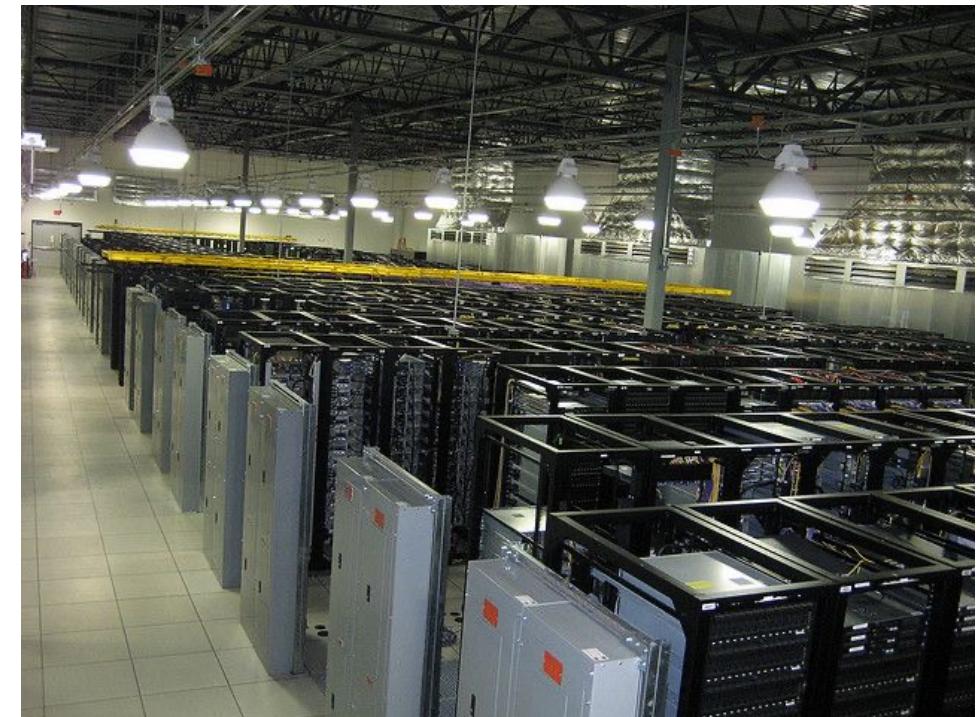
Compute	Quantum Technologies	Security, Identity, & Compliance
EC2 Lightsail Lambda Batch Elastic Beanstalk Serverless Application Repository AWS Outposts EC2 Image Builder AWS App Runner AWS SimSpace Weaver	Amazon Braket	Resource Access Manager Cognito Secrets Manager GuardDuty Amazon Inspector Amazon Macie IAM Identity Center Certificate Manager Key Management Service CloudHSM Directory Service WAF & Shield AWS Firewall Manager AWS Artifact Security Hub Detective AWS Signer AWS Private Certificate Authority AWS Audit Manager Security Lake Amazon Verified Permissions IAM AWS Payment Cryptography
Containers	Management & Governance	AWS Cost Management
Elastic Container Registry Elastic Container Service Elastic Kubernetes Service Red Hat OpenShift Service on AWS	AWS Organizations CloudWatch AWS Auto Scaling CloudFormation Config OpsWorks Service Catalog Systems Manager Trusted Advisor Control Tower AWS License Manager AWS Well-Architected Tool AWS Health Dashboard AWS Chatbot Launch Wizard AWS Compute Optimizer Resource Groups & Tag Editor Amazon Grafana Amazon Prometheus AWS Resilience Hub Incident Manager Service Quotas AWS Proton CloudTrail AWS Resource Explorer AWS User Notifications AWS Telco Network Builder	AWS Cost Explorer AWS Budgets AWS Marketplace Subscriptions AWS Application Cost Profiler AWS Billing Conductor Billing
Storage	Media Services	Front-end Web & Mobile
S3 EFS FSx S3 Glacier Storage Gateway AWS Backup AWS Elastic Disaster Recovery	Kinesis Video Streams MediaConvert MediaLive MediaPackage MediaStore MediaTailor Elemental Appliances & Software Elastic Transcoder Nimble Studio	AWS Amplify AWS AppSync Device Farm Amazon Location Service
Database	Migration & Transfer	Application Integration
RDS ElastiCache Neptune Amazon QLDB Amazon DocumentDB Amazon Keyspaces Amazon Timestream DynamoDB Amazon MemoryDB for Redis	AWS Migration Hub AWS Application Migration Service	Step Functions

EC2 (ECC)

- Elastic: can increase or decrease capacity dynamically
 - Cloud: internet, or intranet
 - Compute: program
- > An AWS service that allows you to configure and **rent computers** to meet your compute needs on an **as needed basis**.

Amazon Web Services (AWS)

- Compute per hour: EC2 (elastic cloud computing)
- Infinite storage (scalable): S3 (simple storage service)
- 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, TCGA).
- 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

Quick summary

EC2 = where we rent virtual computers (instances)

That we can start, stop, expand as needed to save computing cost

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?

Viewing 489 of 489 available instances						
Instance name	On-Demand hourly rate	vCPU	Memory	Storage	Network performance	
a1.medium	\$0.0255	1	2 GiB	EBS Only	Up to 10 Gigabit	
a1.large	\$0.051	2	4 GiB	EBS Only	Up to 10 Gigabit	
a1.xlarge	\$0.102	4	8 GiB	EBS Only	Up to 10 Gigabit	
a1.2xlarge	\$0.204	8	16 GiB	EBS Only	Up to 10 Gigabit	
a1.4xlarge	\$0.408	16	32 GiB	EBS Only	Up to 10 Gigabit	
a1.metal	\$0.408	16	32 GiB	EBS Only	Up to 10 Gigabit	
t4g.nano	\$0.0042	2	0.5 GiB	EBS Only	Up to 5 Gigabit	
t4g.micro	\$0.0084	2	1 GiB	EBS Only	Up to 5 Gigabit	
t4g.small	\$0.0168	2	2 GiB	EBS Only	Up to 5 Gigabit	
t4g.medium	\$0.0336	2	4 GiB	EBS Only	Up to 5 Gigabit	
t4g.large	\$0.0672	2	8 GiB	EBS Only	Up to 5 Gigabit	
t4g.xlarge	\$0.1344	4	16 GiB	EBS Only	Up to 5 Gigabit	

Data transfer (GB): In: free; Out: first 100 GB free, \$0.05-0.09 per GB depending on amount

EBS storage (GB/Month): \$0.08

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-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 started with a base Ubuntu 22.04 (Linux) instance and loaded a whole bunch of software for NGS analysis.
- We will copy 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 to protect your data.

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_Compute/

Login to AWS console



Sign in as IAM user

Account ID (12 digits) or account alias

cshlworkshops

IAM user name

cshl_student_2023

Password

.....

Remember this account

Sign in



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

Select "EC2" service

The screenshot shows the AWS search interface with the search bar set to 'EC2'. A red arrow points to the search bar. To the right of the search results, there is a 'Search for EC2' input field and a 'Reset to default layout' button. A red arrow also points to the '+ Add widget' button. On the right side of the screen, there is a 'AWS Health' info card with the message 'Make sure you are in N. Virginia region' and 'No health data'. The user is in the N. Virginia region.

aws Services Q EC2 X

Search results for 'EC2'
Try searching with longer queries for more relevant results

Services (13)

- Features (53)
- Resources New
- Documentation (33,750)
- Knowledge Articles (20)
- Marketplace (3,181)
- Blogs (2,067)
- Events (30)
- Tutorials (21)

Services

- EC2** ☆
Virtual Servers in the Cloud
- EC2 Image Builder ☆
A managed service to automate build, customize and deploy OS images
- Recycle Bin
Protect resources from accidental deletion
- Amazon Inspector ☆
Continual vulnerability management at scale

See all 13 results ▶

Features

- Dashboard
EC2 feature

See all 53 results ▶

Reset to default layout + Add widget

AWS Health Info

Make sure you are in N. Virginia region

No health data

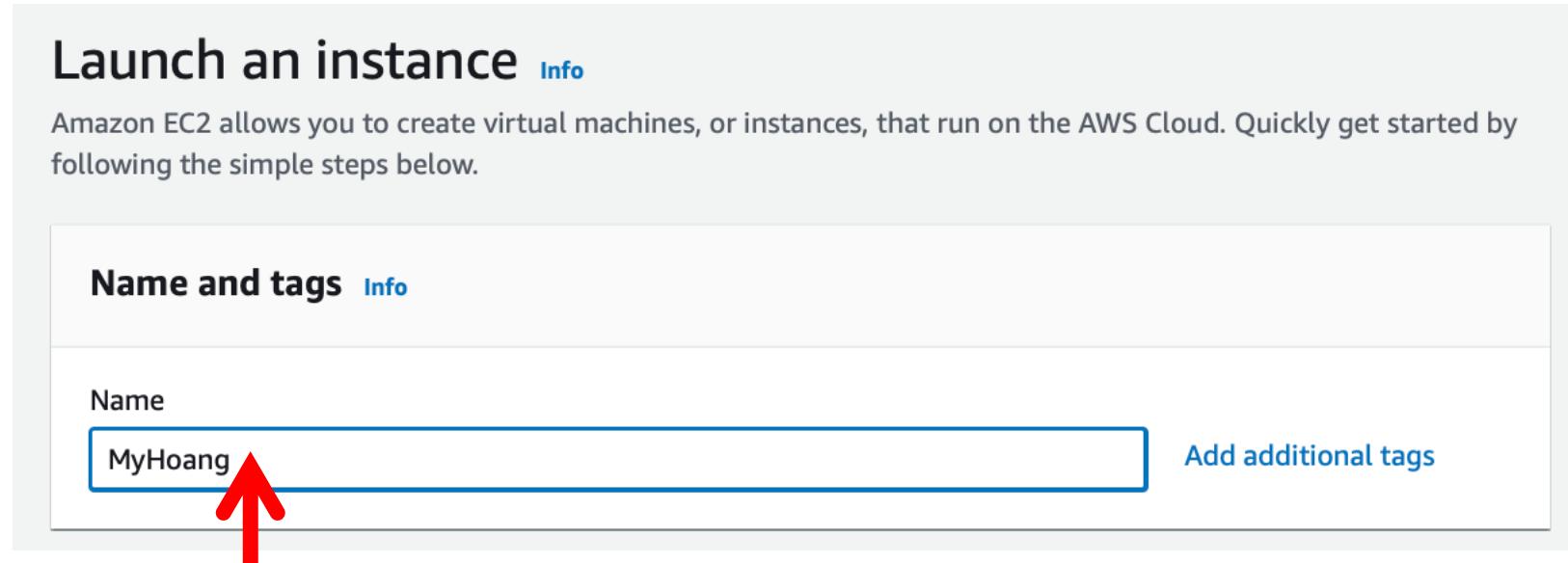
You don't have permissions to access AWS Health.

Go to AWS Health

From EC2 Dashboard, launch a new Instance

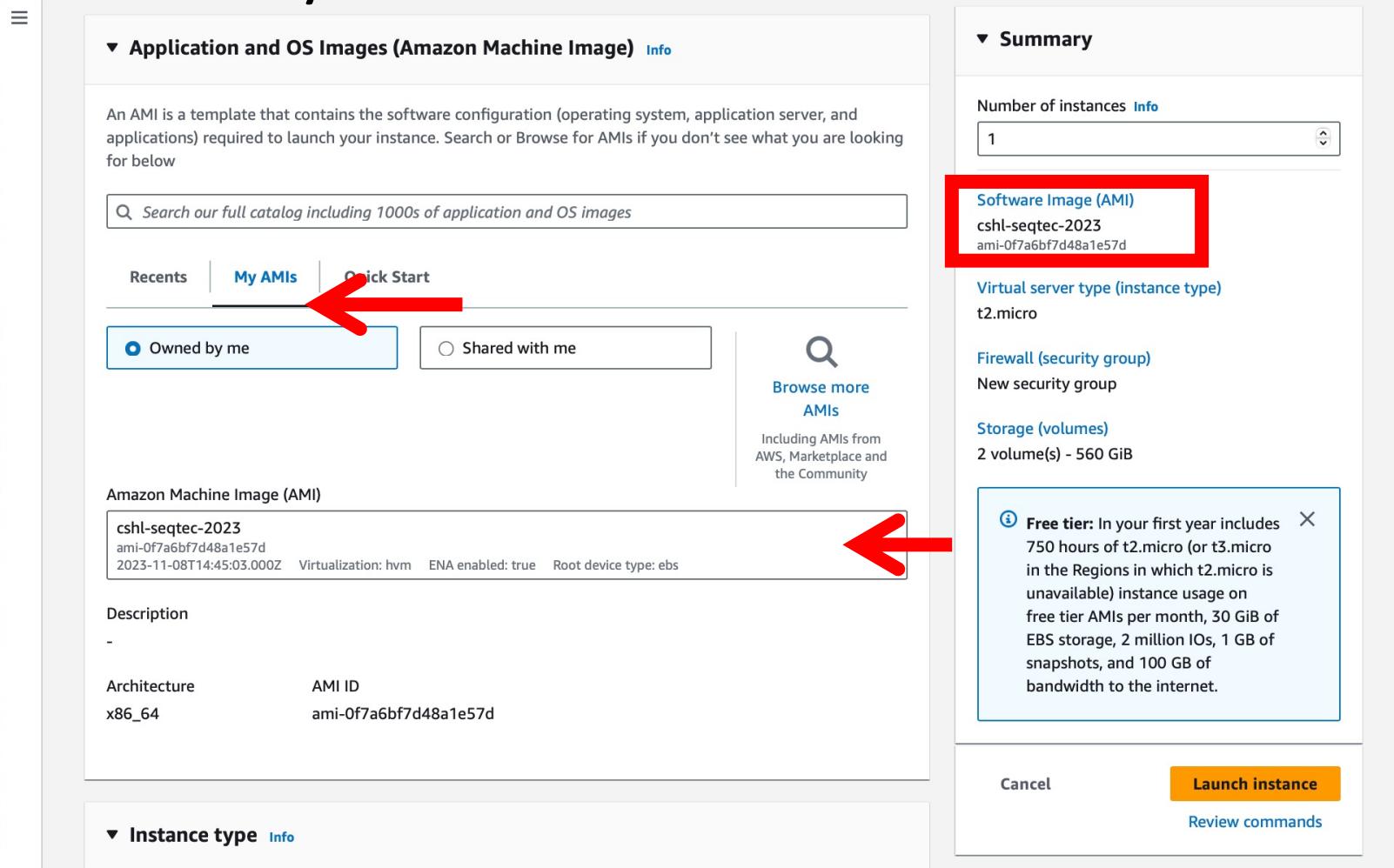
The screenshot shows the AWS EC2 Dashboard. On the left, a navigation sidebar lists various services like EC2 Dashboard, EC2 Global View, Events, Instances, Images, Elastic Block Store, and Network & Security. A red arrow points to the 'EC2 Dashboard' link. In the center, under the 'Resources' section, there's a summary of resources: Instances (running) 2, Auto Scaling Groups 0, Dedicated Hosts 0, Elastic IPs 0, Instances 6, Key pairs 2, Load balancers 0, Placement groups 0, Security groups 6, Snapshots 8, and Volumes 12. Below this, a 'Launch instance' button is highlighted with a red arrow. To the right, there are sections for Account attributes (Default VPC vpc-ad2c8fd7, Settings, Data protection and security, Zones, EC2 Serial Console, Default credit specification, Console experiments), Service health (AWS Health Dashboard), Zones (listing us-east-1a, us-east-1b, us-east-1c with Zone IDs use1-az4, use1-az6, use1-az1), and Explore AWS (sections for 10 Things You Can Do Today to Reduce AWS Costs, Amazon GuardDuty Malware Protection, and Get Up to 40% Better Price Performance).

Name your instance “FirstnameLastname” (e.g. MyHoang)



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

Choose an AMI – Find the cshl-seqtec-2023 AMI in the My AMIs



Choose “m6a.xlarge” instance type, then "Next: Configure Instance Details".

The screenshot shows the AWS Lambda console interface. On the left, under the 'My AMIs' tab, an 'Amazon Machine Image (AMI)' section displays 'cshl-seqtec-2023' with details: ami-0f7a6bf7d48a1e57d, created on 2023-11-08T14:45:03.000Z, using hvm virtualization, ENA enabled, and ebs as the root device type. Below this, the 'Instance type' section is expanded, showing 'm6a.xlarge' selected. A red arrow points from the text 'Choose “m6a.xlarge” instance type,' located in the main text above, to the 'm6a.xlarge' dropdown in the screenshot. To the right, the 'Summary' section of the 'Configure Instance Details' step shows the selected instance type 'm6a.xlarge' highlighted with a red box. Other summary details include 1 instance, software image 'cshl-seqtec-2023', and storage '2 volume(s) - 560 GiB'. A callout box provides information about the free tier.

Recents | My AMIs | Quick Start

Owned by me | Shared with me

Amazon Machine Image (AMI)

cshl-seqtec-2023
ami-0f7a6bf7d48a1e57d
2023-11-08T14:45:03.000Z Virtualization: hvm ENA enabled: true Root device type: ebs

Description

-

Architecture x86_64 AMI ID ami-0f7a6bf7d48a1e57d

▼ Instance type Info

Instance type

m6a.xlarge

Family: m6a 4 vCPU 16 GiB Memory Current generation: true
On-Demand Windows base pricing: 0.3568 USD per Hour
On-Demand Linux base pricing: 0.1728 USD per Hour
On-Demand SUSE base pricing: 0.2291 USD per Hour
On-Demand RHEL base pricing: 0.2328 USD per Hour

Additional costs apply for AMIs with pre-installed software

All generations

Compare instance types

▼ Summary

Number of instances Info

1

Software Image (AMI)

cshl-seqtec-2023
ami-0f7a6bf7d48a1e57d

Virtual server type (instance type)
m6a.xlarge

New security group

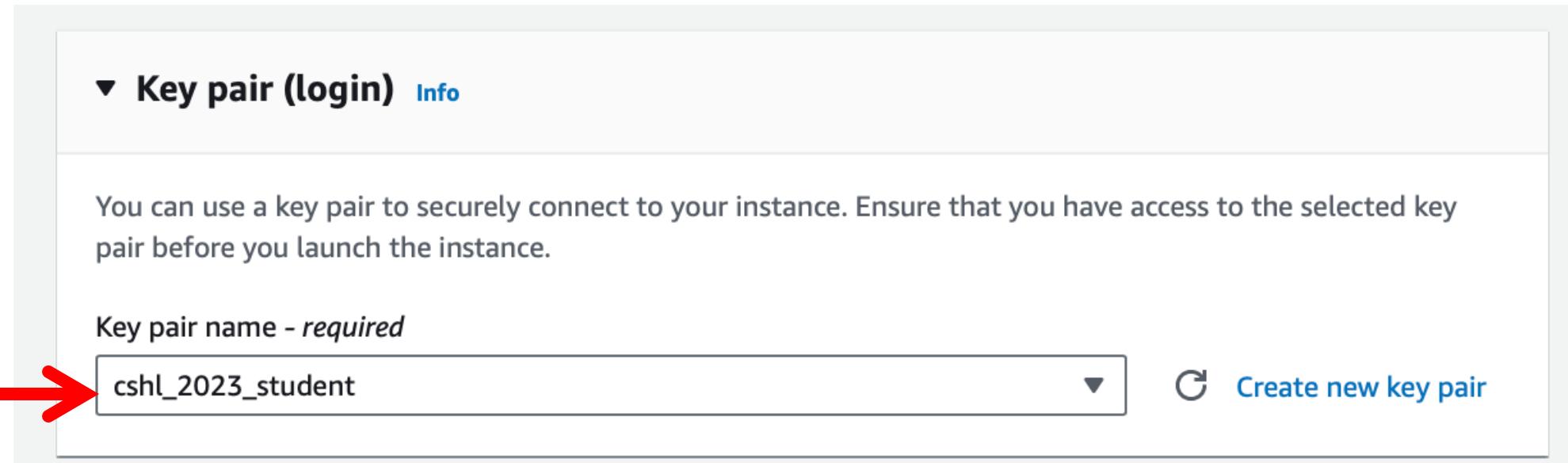
Storage (volumes)

2 volume(s) - 560 GiB

Free tier: In your first year includes 750 hours of t2.micro (or t3.micro in the Regions in which t2.micro is unavailable) instance usage on free tier AMIs per month, 30 GiB of EBS storage, 2 million IOs, 1 GB of snapshots, and 100 GB of bandwidth to the internet.

Cancel | Launch instance | Review commands

Choose an existing key pair:
"cshl_2023_student"



Select an Existing Security Group, choose "SSH/HTTP/Jupyter". Then hit "Review and Launch".

The screenshot shows the AWS Lambda 'Create Function' wizard at the 'Review and Launch' step. On the left, under 'Network settings', the 'Select existing security group' option is selected, and a security group named 'SSH/HTTP/Jupyter' is listed. On the right, the 'Summary' section shows the chosen instance type as 'm6a.xlarge' and the AMI as 'cshl-seqtec-2023'. A red box highlights the 'Firewall (security group)' field, which contains 'SSH/HTTP/Jupyter'. Another red arrow points to the 'Select existing security group' button, and a third red arrow points to the listed security group.

Network settings

Network [Info](#)
vpc-ad2c8fd7

Subnet [Info](#)
No preference (Default subnet in any availability zone)

Auto-assign public IP [Info](#)
Enable

Firewall (security groups) [Info](#)
A security group is a set of firewall rules that control the traffic for your instance. Add rules to allow specific traffic to reach your instance.

Create security group Select existing security group

Common security groups [Info](#)
Select security groups ▾

SSH/HTTP/Jupyter sg-0fd7356e013ca81a3 X
VPC: vpc-ad2c8fd7

Compare security group rules

Summary

Number of instances [Info](#)
1

Software Image (AMI)
cshl-seqtec-2023
ami-0f7a6bf7d48a1e57d

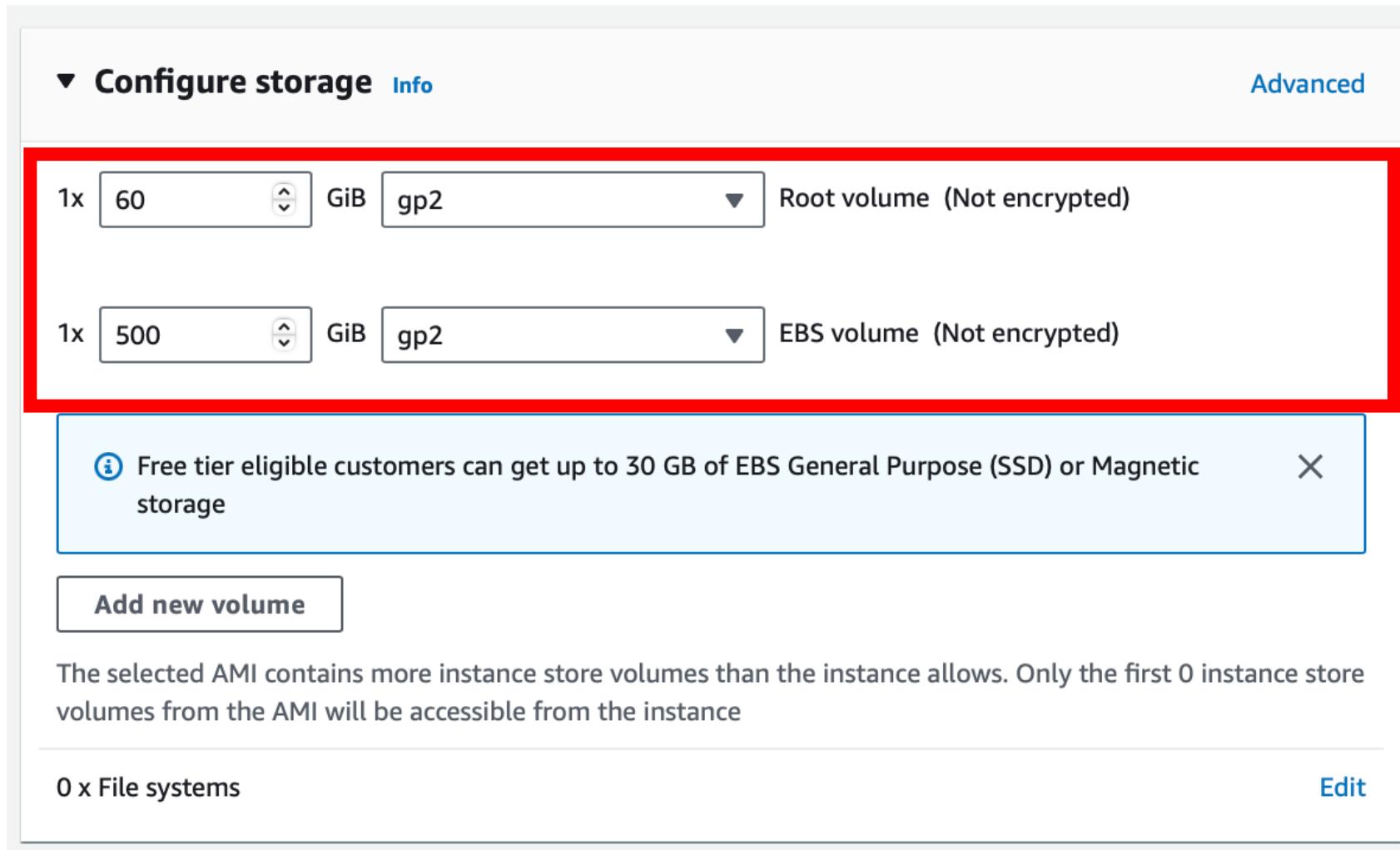
Virtual server type (instance type)
m6a.xlarge

Firewall (security group)
SSH/HTTP/Jupyter

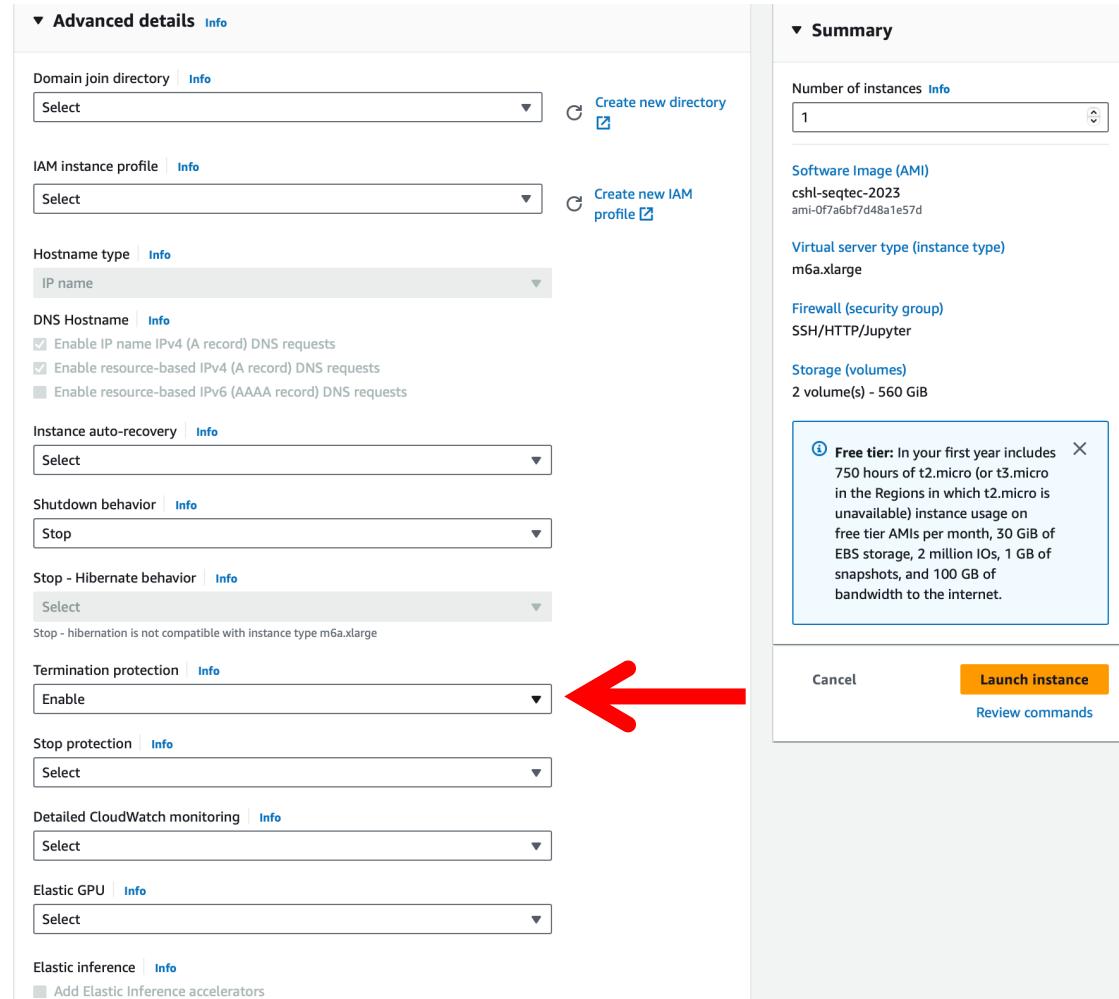
Storage (volumes)
2 volume(s) - 560 GiB

Free tier: In your first year includes 750 hours of t2.micro (or t3.micro in the Regions in which t2.micro is unavailable) instance usage on

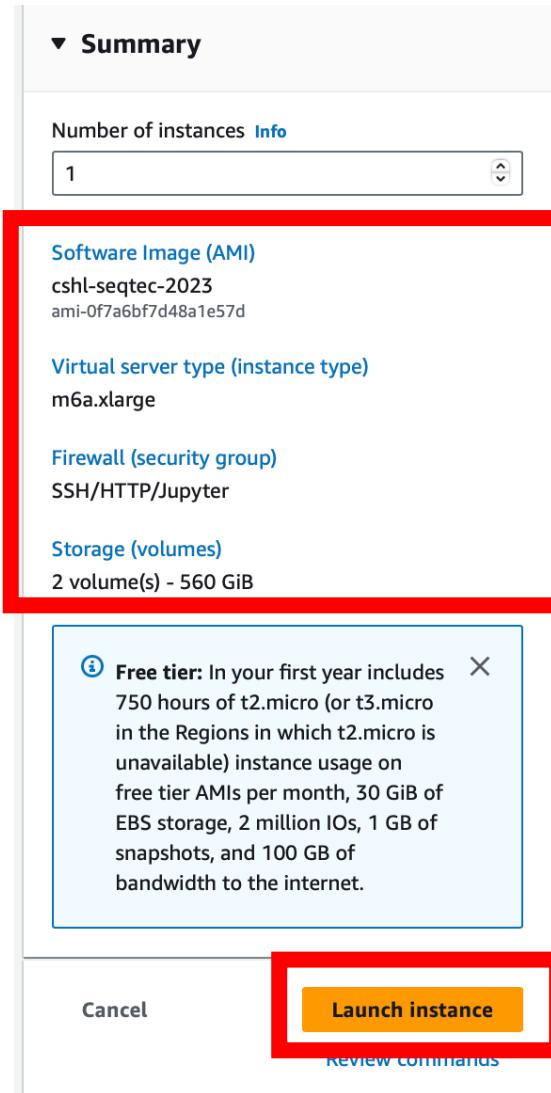
You should see 1 60 GiB root volume and 1 500 GiB EBS volume as the two storage volumes.



Select “Enable” for termination protection.



Review the details of your instance and hit Launch



View Instances to see your new instance spinning up!

EC2 > Instances > Launch an instance



Success

Successfully initiated launch of instance (i-00e1dc9cd6843478d)

▶ Launch log

Next Steps

Create billing and free tier usage alerts

To manage costs and avoid surprise bills, set up email notifications for billing and free tier usage thresholds.

[Create billing alerts](#)

Connect to your instance

Once your instance is running, log into it from your local computer.

[Connect to instance](#)

[Learn more](#)

Connect an RDS database

New

Configure the connection between an EC2 instance and a database to allow traffic flow between them.

[Connect an RDS database](#)

[Create a new RDS database](#) [Learn more](#)

[View all 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 navigation sidebar with links like EC2 Dashboard, EC2 Global View, Events, Tags, Limits, Instances (selected), Images, Elastic Block Store, Network & Security, and more. The main area displays a table of instances:

Name	Instance ID	Instance state	Instance type	Status check	Alarm status	Availability Zone	Public IPv4 DNS	Public IPv4 IP	Elasti
Kcotto	i-0302a728fbd4095f	Running	m5.2xlarge	2/2 checks passed	No alarms	us-east-1a	ec2-18-234-50-3.compute...	18.234.50.3	-
Instructor_test	i-088f1b7da8b8d0656	Running	m5.2xlarge	2/2 checks passed	No alarms	us-east-1a	ec2-18-212-77-248.compute...	18.212.77.248	-
KelsyCotto	i-00e1dc9cd6843478d	Running	m5.2xlarge	Initializing	No alarms	us-east-1a	ec2-100-26-22-19.compute...	100.26.22.19	-

At the top right of the table, there's a 'Connect' button with a red box around it. Below the table, a specific instance is selected: 'KelsyCotto' (i-00e1dc9cd6843478d). The detailed view shows:

- Details tab (selected)
- Security, Networking, Storage, Status checks, Monitoring, Tags tabs
- Instance summary section:
 - Instance ID: i-00e1dc9cd6843478d (KelsyCotto)
 - Public IPv4 address: 100.26.22.19 | open address (with a red arrow pointing here)
 - Private IPv4 addresses: 172.31.28.82
 - IPv6 address: -
 - Instance state: Running
 - Private IP DNS name (IPv4 only): ip-172-31-28-82.ec2.internal
 - Instance type: m5.2xlarge

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, the navigation path is EC2 > Instances > i-00e1dc9cd6843478d > Connect to instance. Below this, there's a section titled 'Connect to instance' with a 'Info' link. It says 'Connect to your instance i-00e1dc9cd6843478d (KelsyCotto) using any of these options'. There are four tabs: 'EC2 Instance Connect', 'Session Manager', 'SSH client' (which is highlighted with a red arrow), and 'Terminal console'. Under each tab, there are instructions and links. For 'SSH client': 1. Open an SSH client. 2. Locate your private key file. The key used to launch this instance is cshl_2022_student.pem. 3. Run this command, if necessary, to ensure your key is not publicly viewable. 4. Connect to your instance using its Public DNS. For 'Terminal console': A single link to ssh -i "cshl_2022_student.pem" root@ec2-100-26-22-19.compute-1.amazonaws.com. A note at the bottom says: 'Note: In most cases, the guessed user name is correct. However, read your AMI usage instructions to check if the AMI owner has changed the default AMI user name.'

Congratulations, you just fire
your own instance!!

Let's log into it

Steps

- Launch instance : Done !
- Step 0: Download key file (.pem file for Mac, .ppk file for Windows)
& Change permission of downloaded key
- Step 1: Log into your instance :
`ssh -i [student-key].pem ubuntu@[public.ip.address]`
- When done for the day, ‘stop’ instance (don’t ‘terminate’). Next morning, ‘start’ instance again (don’t launch new instance).

Logging into your instance

Mac

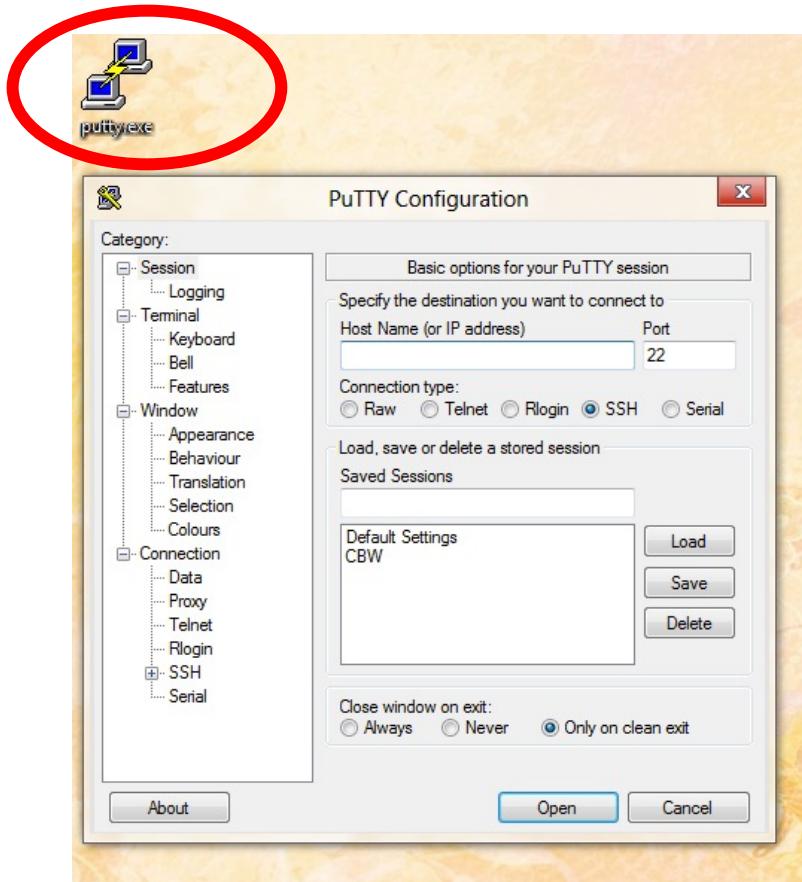
- Terminal

Window

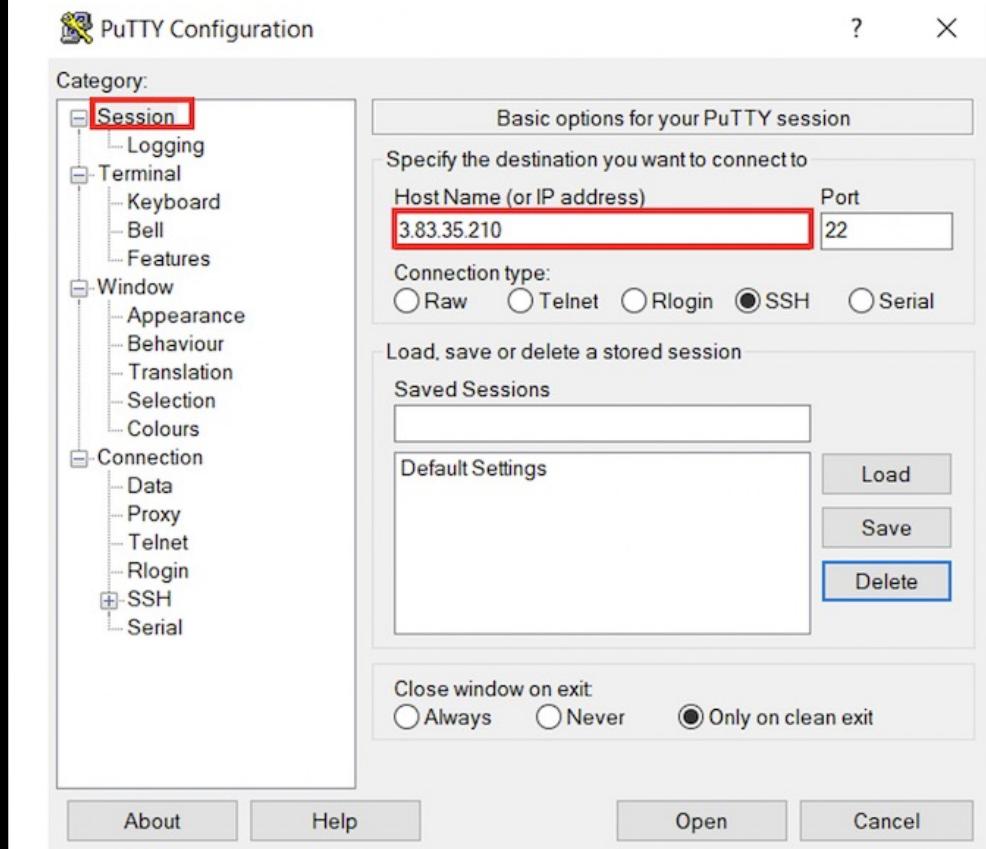
- PowerShell (preferred, similar to Terminal)
- PuTTy (more complicated)

Logging into your instance (Windows)

Open PuTTY

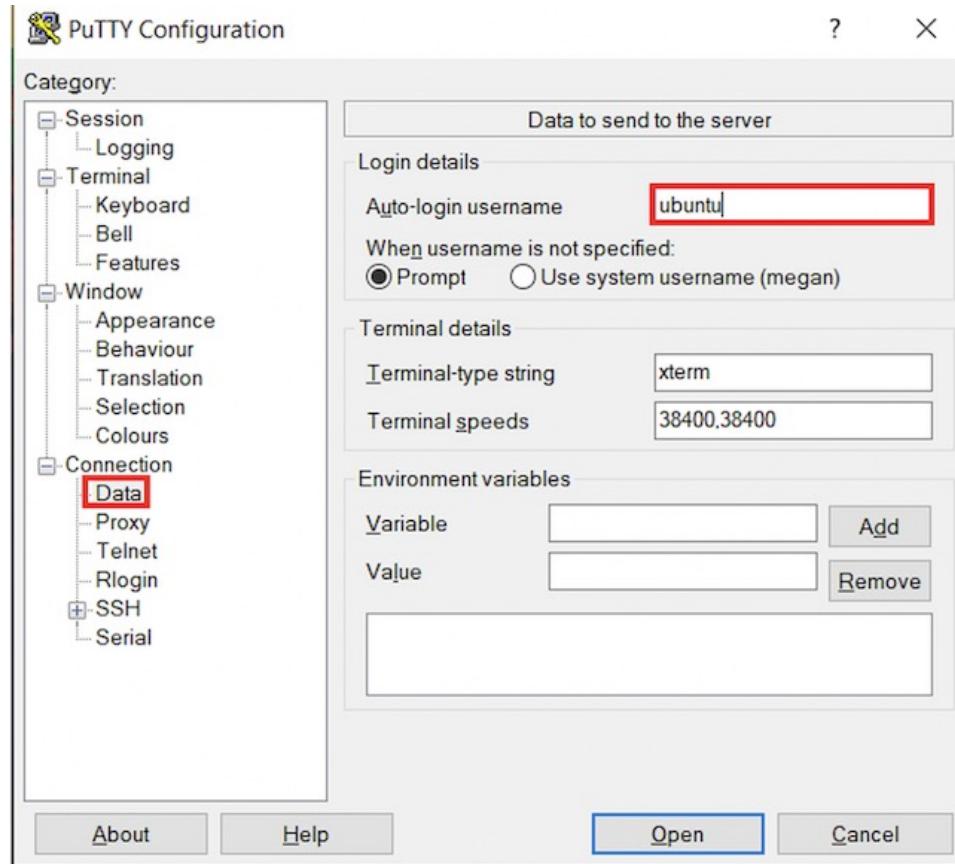


Enter the Host Name (IP address)

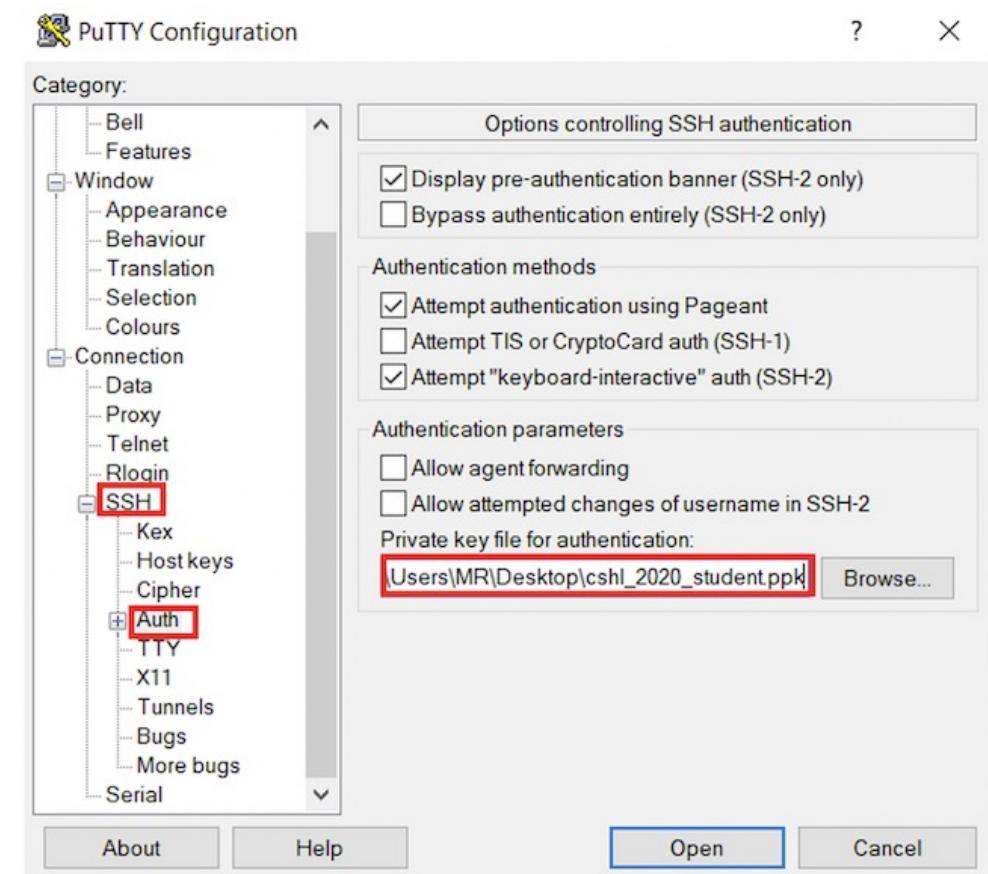


Logging into your instance (Windows)

Choose Connection -> Data
Enter the username 'ubuntu'



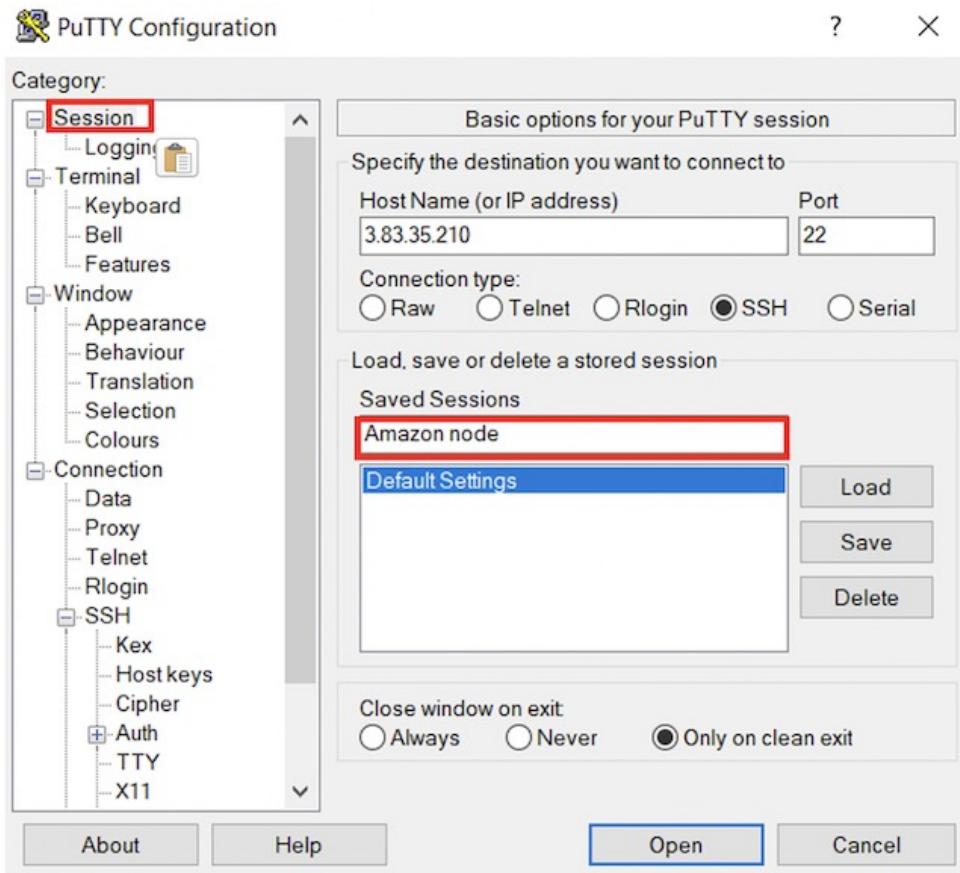
Choose SSH -> Auth
Browse to Private key (ppk) file



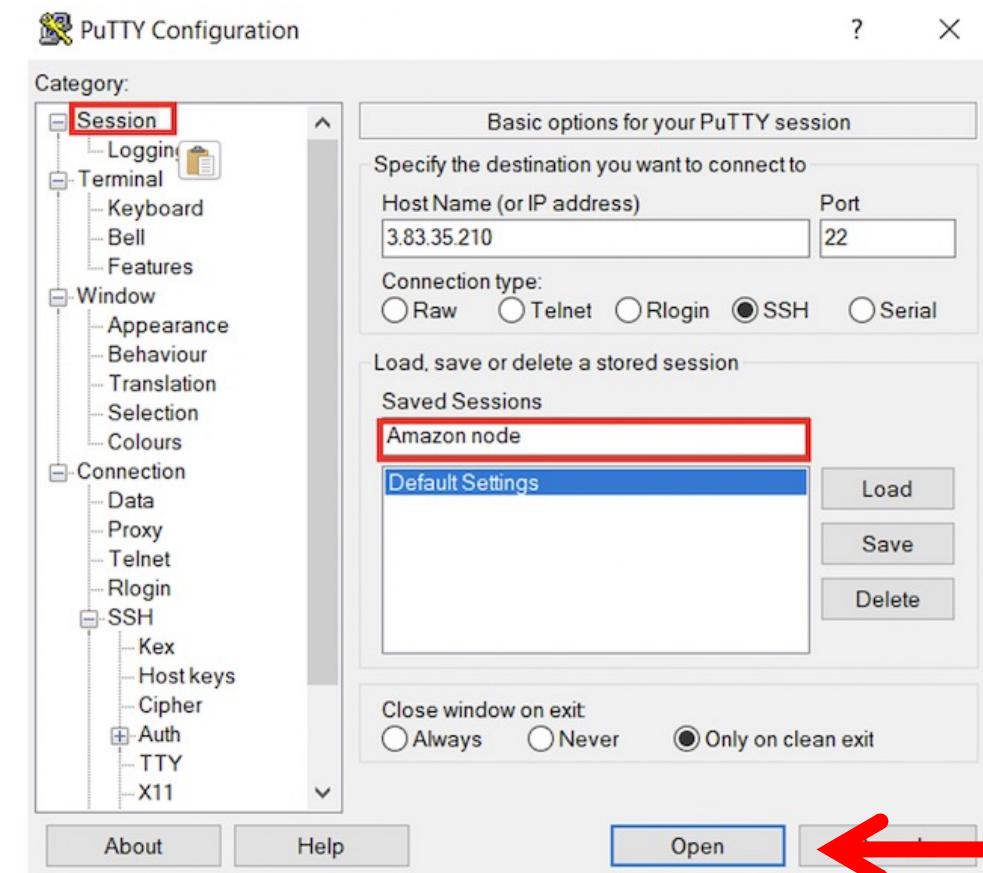
Logging into your instance (Windows)

Choose Session

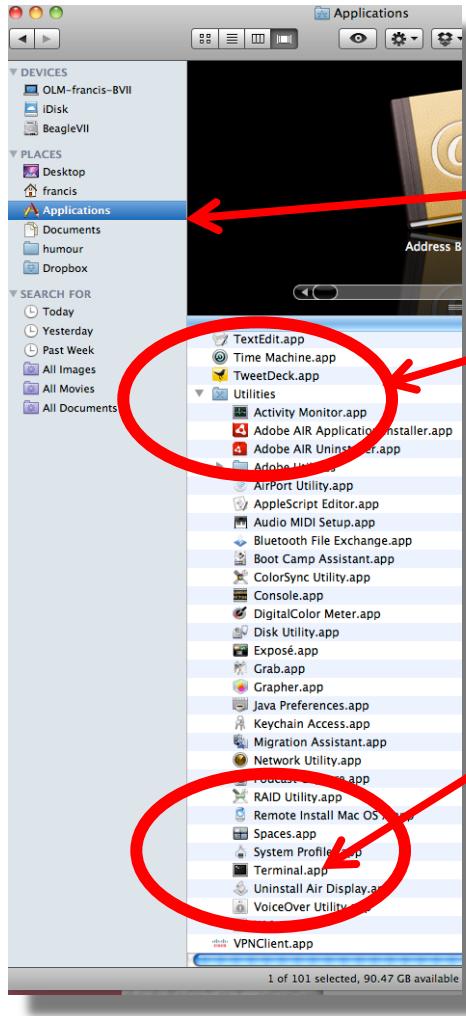
Save your session as "Amazon Node"



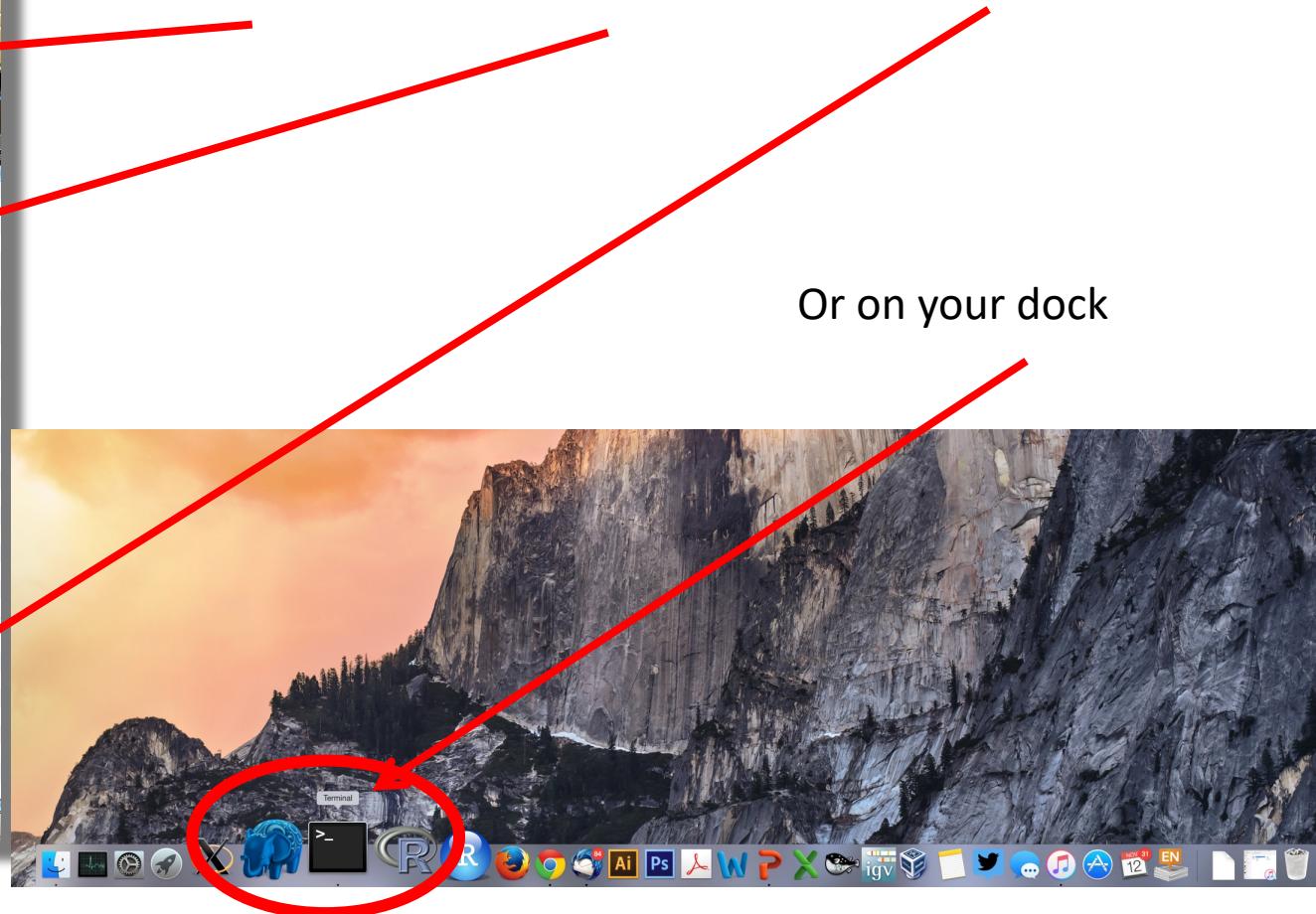
Double-click saved “Amazon Node” session OR
Select “Amazon Node” session and click Open



Logging into your instance (Mac)

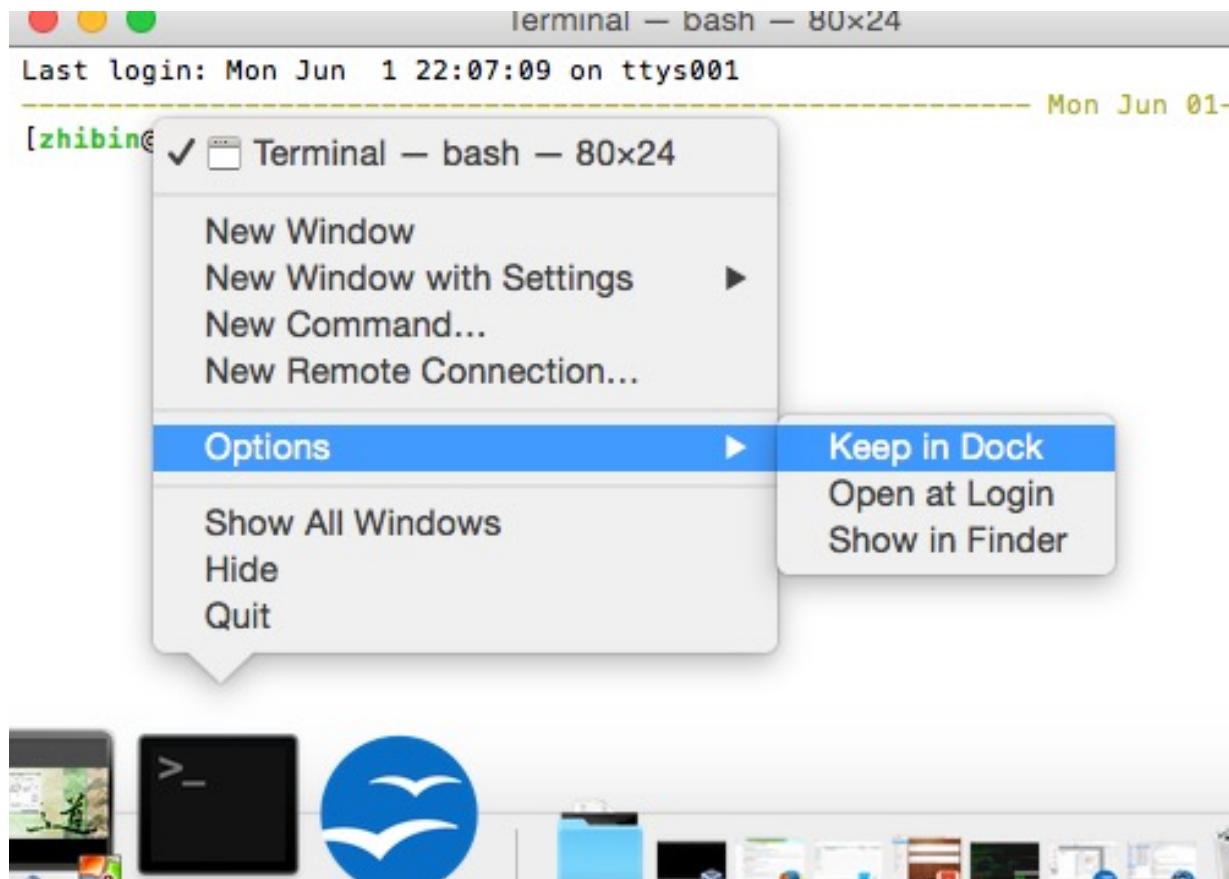


In a Finder window
'Applications' -> 'Utilities' -> 'Terminal'



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 can use the same ‘.pem’ file if use PowerShell
 - You will need to use a ‘.ppk’ file instead. This can be 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  
less cshl_2021_student.pem
```

```
Kelsys-MacBook-Pro:keys kcotto$ cat cshl_2019_student.pem  
-----BEGIN RSA PRIVATE KEY-----  
MIIEpAIBAAKCAQEAgGtpnqERtEu/SCmeF2r1HMESMao0fEJiAwQwk2/SNXK8izr0IH0zTVvmE1D  
VUwWq7pkvhjh05pDb+2U9HiZe3sxLv3S1NrkATYF/NsrpwB+q1vwqzGW9sQ6uj45RWrPkiZlsaj  
TQZmyFRu+t1JTRU3hQDqA0MRWTx1Wxv0gFzuZy/qb+DALuFQsInrEKnijrwdLmd6usaBTvhc0gFS  
B9oEelH0bZHJTZFW/wP+Z0uZq0Ujir7Qw0LTM45QH/L0dBdUl3k/mBeez00yvnKMwj8E4Xi0rQ0t  
hHtQ7F9iSILK80W1rRH0qwxwt9ycEH1JtNMQmUTif0vE2XJ6106chQIDAQABAoIBABG7P/FHu/Qp  
WFgg+89myuqR6GvA2X55CFSzFyG0aQyrj5jDleFtdu2uXiISG8gUBZYvlzxx82aOC0P5j04SBq0  
xD/qRlukY/jyXyPn77w/ExmaNoLJj1W9RUSH0JYLIzVpFPGes3u5zGSGDTSDNh3sSdWhq1FX3L  
7vY5b6UAQgahXf0dpGFxt6P6qb/BKFQFsThXk7GXMyS/kr4w7ZlhPWHRMSu2UBdS/a/beAVN76qU  
E/10HR30oAuqghusZabpbX7MJ100VcviaQgeF8Z+xf2uugDEAKut1PxW0+yvGM4SpZZ0skFZz6YI  
pBnX4ELWPhCeMzq4M18QY6ptR2UCgYEAc6YfNlymg24NJu8PANx8navTi50WYuXWXrj0mrLr3SR  
/XY+w26cgipM+K5eQfxSr4Yb8BQKjRktMzBzf5nKdaX4pzYIquQH02B0HDhTooHAhkbTMTmKukv4  
oEW06wcEE6RQi fw4xbMEnfQfHJBI21am+jwI8Xb7idwMG4pU/nsCgYEAtD0x4bNC1X3A5by50uY0  
WXrtgQszCXYcbkrpbjRET12f9hgz9MRMHY/xH/XGvMutZSFV2rCZRwd7lm+QWGadk/MQS0kouzW6  
gSasyjFq+MKCkqYnS3/JTbx8yrLZmz10LtX6pwmwg0Zy8aJjYHo9a2/EI8Tjh2d0SxeadIRVYP8C  
gYAToiXww1Vdu+dj/7TDLqYCtdHOVAxJX/utI9Q3yoIryuh+bWmFvEIvAmIGXyyQZRyoZwgIS4A  
PNH03+bEa+69wbzlhksiK5g8GKgISVdLC4rZZXB5ehgTmWV7IgJ89y/SF4G/Ityo30K0ohALh597  
NcvNEzzqrutja1IIMvTKMwKBgQCkM+QP1Tqc0TbV1fvClviXuJBLsiJLCImYeZL0nZVmIMusbhxX  
b8ZQYGSyUz09nulXau1G1QDvXvf089CzWL1SomxBoH1FJQvGwa9FfYQRIVPHuqut8rs4oPGn0QzC  
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_2023_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:

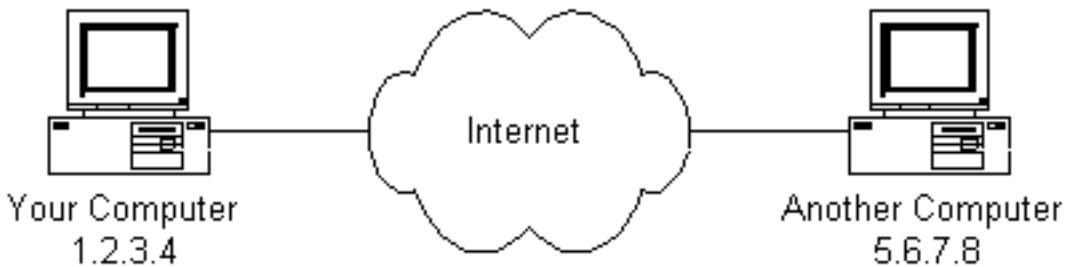
chmod 400 <file name>

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

Logging into your instance

Mac/Linux

```
cd ~/cshl  
chmod 400 cshl_2023_student.pem  
ssh -i cshl_2023_student.pem ubuntu@[YOUR PUBLIC IP]
```



Screenshot of the AWS CloudWatch Instances console showing the details of an EC2 instance.

The instance ID is i-00e1dc9cd6843478d (KelsyCotto).

The Public IPv4 address is 100.26.22.19 (highlighted with a red arrow).

The instance state is Running.

The Private IP DNS name (IPv4 only) is ip-172-31-28-82.ec2.internal.

The instance type is m5.2xlarge.

Category	Value
Instance ID	i-00e1dc9cd6843478d (KelsyCotto)
Public IPv4 address	100.26.22.19 open address
Private IPv4 addresses	172.31.28.82
IPv6 address	-
Instance state	Running
Public IPv4 DNS	ec2-100-26-22-19.compute-1.amazonaws.com open address
Hostname type	IP name: ip-172-31-28-82.ec2.internal
Private IP DNS name (IPv4 only)	ip-172-31-28-82.ec2.internal
Answer private resource DNS name	IPv4 (A)
Instance type	m5.2xlarge
Elastic IP addresses	-

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

Index of /

<u>Name</u>	<u>Last modified</u>	<u>Size</u>	<u>Description</u>
 Notebooks/	2023-10-27 21:36	-	
 ensembl-vep/	2023-10-27 19:35	-	

Apache/2.4.52 (Ubuntu) Server at 52.205.115.252 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!

The screenshot shows the AWS EC2 Instances page. On the left, there's a sidebar with navigation links like EC2 Dashboard, EC2 Global View, Events, Tags, Limits, Instances (with sub-links like Instances, Instance Types, Launch Templates, Spot Requests, Savings Plans, Reserved Instances, Dedicated Hosts, Scheduled Instances, Capacity Reservations), and Images (AMIs, AMI Catalog). The main area has a table titled "Instances (1/3) Info" with columns: Name, Instance ID, Instance state, Instance type, Status check, Alarm status, Availability Zone, Public IPv4 DNS, and Public IPv6 DNS. There are three instances listed: "Kcotto" (Instance ID i-0302a728fbdc4095f, Running, m5.2xlarge, 2/2 checks passed, No alarms, us-east-1a, ec2-18-234-50-3.comp..., 18.234.50.3), "Instructor_test" (Instance ID i-088f1b7da8b8d0656, Running, m5.2xlarge, 2/2 checks passed, No alarms, us-east-1a, ec2-18-212-77-248.co..., 18.212.77.248), and "KelsyCotto" (Instance ID i-00e1dc9cd6843478d, Running, m5.2xlarge, 2/2 checks passed, No alarms, us-east-1a, ec2-54-221-110-174.co..., 54.221.110.174). A red arrow points from the top-left to the "Instances" link in the sidebar. Another red arrow points from the middle-left to the "KelsyCotto" instance in the table. A third red arrow points from the bottom-right to the "Stop instance" option in the context menu that appears when right-clicking on the instance row.

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 EC2 Instances page. On the left, there's a navigation sidebar with links like EC2 Dashboard, EC2 Global View, Events, Tags, Limits, and Instances (with sub-links for Instances, Instance Types, Launch Templates, Spot Requests, Savings Plans, Reserved Instances, Dedicated Hosts, Scheduled Instances, and Capacity Reservations). A red arrow points from the 'Instances' link in the sidebar to the 'Instances' tab in the main header. Another red arrow points from the 'Instances' tab to the table where an instance named 'KelsyCotto' is listed. This instance has a status of 'Stopped'. A third red arrow points from the 'Actions' dropdown menu to the 'Start instance' option, which is highlighted. The main content area shows a table with columns: Name, Instance ID, Instance state, Instance type, Status check, Alarm status, Availability Zone, Public IPv4 DNS, Public IPv4 ..., and Elast. There are three instances listed: 'Kcotto' (Running), 'Instructor_test' (Running), and 'KelsyCotto' (Stopped). Below the table, a modal window titled 'Instance: i-00e1dc9cd6843478d (KelsyCotto)' is open, showing tabs for Details, Security, and Network. A large red box highlights the text: 'Go to AWS EC2 Dashboard, select “Instances” tab, then find your instance. Right-click and chose ‘Start Instance’'.

Instances (1/3) Info

Find instance by attribute or tag (case-sensitive)

Name	Instance ID	Instance state	Instance type	Status check	Alarm status	Availability Zone	Public IPv4 DNS	Public IPv4 ...	Elast
Kcotto	i-0302a728fbdc4095f	Running	m5.2xlarge	2/2 checks passed	No alarms	us-east-1a	ec2-18-234-50-3.comp...	18.234.50.3	-
Instructor_test	i-088f1b7da8b8d0656	Running	m5.2xlarge	2/2 checks passed	No alarms	us-east-1a	ec2-18-212-77-248.co...	18.212.77.248	-
<input checked="" type="checkbox"/> KelsyCotto	i-00e1dc9cd6843478d	Stopped	m5.2xlarge	-	-	t-1a	-	-	-

Launch instances

Launch instance from template

Migrate a server

Connect

Stop instance

Start instance

Reboot instance

Hibernate instance

Terminate instance

Instance settings

Networking

Instance: i-00e1dc9cd6843478d (KelsyCotto)

Details Security Network

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 shows the AWS EC2 Instances page. On the left, there's a sidebar with various navigation links like EC2 Dashboard, Events, Tags, Limits, Instances, Images, Elastic Block Store, and Network & Security. The main area displays a table of instances with columns for Name, Instance ID, Instance state, Instance type, Status check, Alarm status, Availability Zone, Public IPv4 DNS, and Public IPv4 IP. Three instances are listed: Kcotto, Instructor_test, and KelsyCotto. The KelsyCotto instance is selected, highlighted with a blue border. A red arrow points to the 'Connect' button in the top right corner of the table header. Below the table, a detailed view for the KelsyCotto instance is shown. It includes tabs for Details, Security, Networking, Storage, Status checks, Monitoring, and Tags. Under the Details tab, there's an 'Instance summary' section with fields for Instance ID (i-00e1dc9cd6843478d), Public IPv4 address (100.26.22.19), Instance state (Running), and Private IP DNS name (ip-172-31-28-82.ec2.internal). A second red arrow points to the Public IPv4 address '100.26.22.19'.

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

We are on a Coffee Break & Networking Session

Key AWS concepts and terminology

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

Key AWS concepts and terminology

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

Key AWS concepts and terminology

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

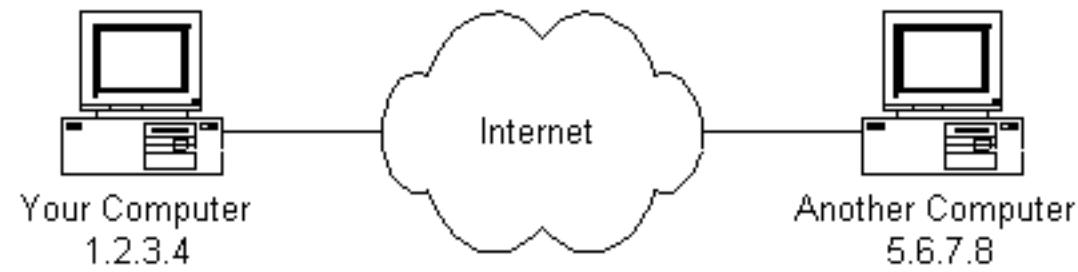
What is a Region?

- An AWS Region is set of compute resources that Amazon maintains (like the Data Center image shown before)
- Each Region corresponds to a physical warehouse of compute hardware (computers, storage, networking, etc.).
- At the time of writing there are 27 regions with more planned to come online soon: (US East (N.Virginia), US East (Ohio), US West (Oregon), US West (N. California), GovCloud (US-West), GovCloud (US-East), Canada (Central), EU (Ireland), EU (Frankfurt), EU (London), EU (Paris), EU (Milan), EU (Stockholm), Middle East (Bahrain), Middle East (UAE), Africa (Cape Town), Asia Pacific (Singapore), Asia Pacific (Sydney), Asia Pacific (Seoul), Asia Pacific (Tokyo), Asia Pacific (Mumbai), Asia Pacific (Hong Kong), Asia Pacific (Beijing), Asia Pacific (Osaka), Asia Pacific (Jakarta), Asia Pacific (Ningxia), 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.

We are on a Coffee Break & Networking Session

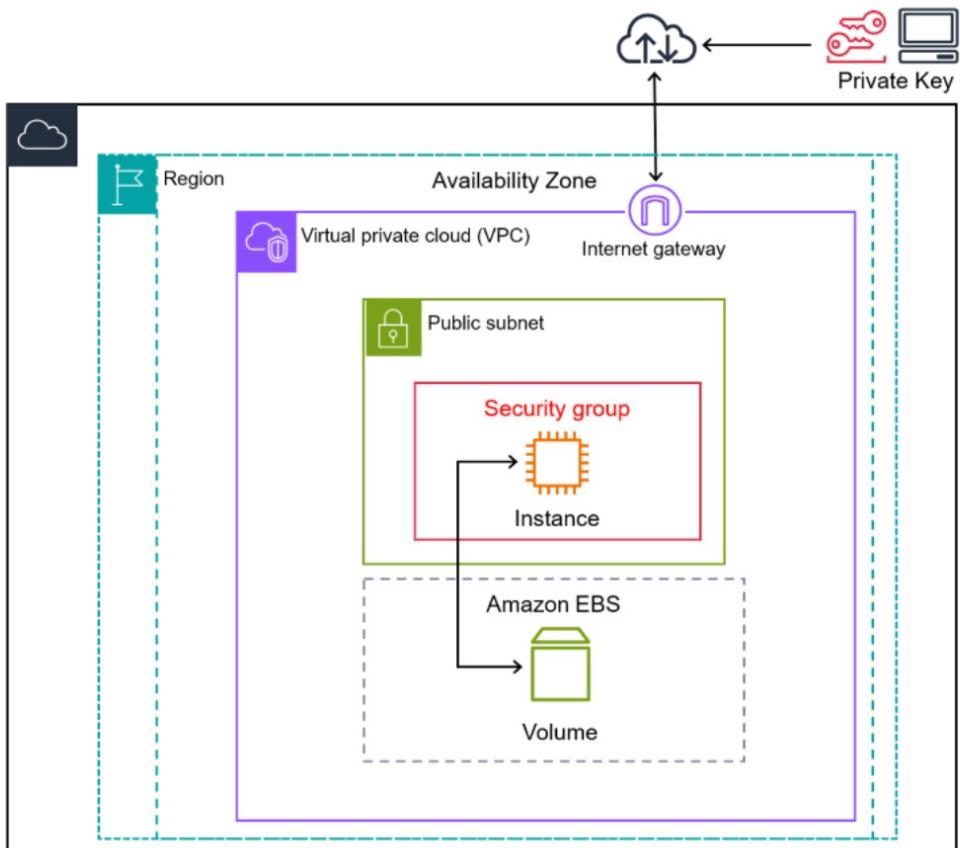


Steps

- Launch instance
- Download key file (.pem file for Mac, .ppk file for Windows)
& Change permission of downloaded key
- Log into your instance :
`ssh -i [student-key].pem ubuntu@[public.ip.address]`
- When done for the day, ‘stop’ instance (don’t ‘terminate’). Next morning, ‘start’ instance again (don’t launch new instance).

Explain to me like im 4 yo

- AWS = Amazon Web Service
- EC2 : ECC = Elastic Compute Cloud -> where I rent computers
- Instance: a virtual server in EC2
- Server: In computing, a server is a piece of computer hardware or software (computer program) that provides functionality for other programs or devices, called "clients".
- EBS: Elastic Block Storage -> I rent block storage that can be attached to my instance (the computer I rented)
[S3: Simple Storage Service. Cheaper than EBS, allows storage of larger amount, but not attached to the instance.]
- AMI: Amazon Machine Image
- IAM: Identity and Access Management
- Key pair: consisting of a public key and a private key, is a set of security credentials that you use to prove your identity when connecting to an instance. Amazon stores the public key on your instance. You store the private key.
- IP address: Internet Protocol address, a unique address that identifies a device on the internet or a local network.



Features of Amazon EC2

Amazon EC2 provides the following high-level features:

Instances

Virtual servers.

Amazon Machine Images (AMIs)

Preconfigured templates for your instances that package the components you need for your server (including the operating system and additional software).

Instance types

Various configurations of CPU, memory, storage, networking capacity, and graphics hardware for your instances.

Key pairs

Secure login information for your instances. AWS stores the public key and you store the private key in a secure place.

Instance store volumes

Storage volumes for temporary data that is deleted when you stop, hibernate, or terminate your instance.

Amazon EBS volumes

Persistent storage volumes for your data using Amazon Elastic Block Store (Amazon EBS).

Regions, Availability Zones, Local Zones, AWS Outposts, and Wavelength Zones

Multiple physical locations for your resources, such as instances and Amazon EBS volumes.

Security groups

A virtual firewall that allows you to specify the protocols, ports, and source IP ranges that can reach your instances, and the destination IP ranges to which your instances can connect.

Elastic IP addresses

Static IPv4 addresses for dynamic cloud computing.

Tags

Metadata that you can create and assign to your Amazon EC2 resources.

Virtual private clouds (VPCs)

Virtual networks you can create that are logically isolated from the rest of the AWS Cloud. You can optionally connect these virtual networks to your own network.

For details about all of the features of Amazon EC2, see [Amazon EC2 features](#).

For options to run your website on AWS, see [Web Hosting](#).