CS205: Computing Foundations for Computational Science, Spring 2021

INSTITUTE FOR APPLIED COMPUTATIONAL SCIENCE AT HARVARD UNIVERSITY

HARVARD School of Engineering and Applied Sciences

Guide: First Access to AWS

Ignacio M. Llorente, David Sondak, Dylan Randle, Simon Warchol

3.0 - February 1, 2021

Abstract

This is a screenshot document of how to setup your AWS environment, launch VMs with EC2 and manage storage with S3.

Notes

- First you will have to sign up for an **AWS regular account**, if you don't already have one.
- You will be sent AWS credits to use for this course. However, you will not need to use credits for this lab.
    - See this link for info about redeeming credits: https://aws.amazon.com/awscredits/

Do not apply for an AWS Educate Account

- Set up a **billing alert** to make sure you don't accidentally use up your free credits without noticing.

- **Stop** your instances when you are done for the day to avoid incurring charges. Use your funds wisely. **Terminate** them when you are sure you are done with your instance (disk storage also costs something, and can be significant if you have a large disk footprint). Look into creating custom alarms to automatically stop your instances when they are not doing anything.

- This guide has been prepared considering that you are using Linux or Mac OS to connect to the remote instance. If you are using windows we recommend you read:

    http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/putty.html

- This guide describes a basic use to develop and execute the course hands-on. For further use we recommend you read the AWS guidelines to create IAM users and VPCs:

http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/get-set-up-for-amazon-ec2.html

- We strongly recommend you first read this introductory Amazon EC2 guide that presents the basic concepts: instances, AMIs, security groups, root devices, regions and availability zones.

  http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/concepts.html

- We strongly recommend you first read this introductory Amazon S3 guide that presents the basic concepts: buckets and objects.

  http://docs.aws.amazon.com/AmazonS3/latest/dev/Introduction.html

Acknowledgments

1. Create a key pair to connect to the cloud VMs

You can skip this step if you already have a key pair from other courses.

- Login to AWS, go to the EC2 dashboard, select "key pairs" on the left hand menu, and click "Create Key Pair". We recommend you name it "*CS205*-key".
- **Windows:** Select ppk file format and save it somewhere safe. Skip to #2

EC2 > Key pairs > Create key pair

Create key pair

Key pair
A key pair, consisting of a private key and a public key, is a set of security credentials that you use to prove your identity when connecting to an instance.

Name

```
CS205-key
```

The name can be up to 255 characters long. Valid characters include _, -, a-z, A-Z, and 0-9.

File format

⦿ pem
For use with OpenSSH

◯ ppk
For use with PuTTY

Cancel   **Create key pair**

- Download the private key locally and copy it to the `.ssh` folder. In my case, and for illustration:

```
$ mv ~/Downloads/CS205-key.pem ~/.ssh/CS205-key.pem
```

- Change the permission of the file

```
$ chmod 600 ~/.ssh/CS205-key.pem
```
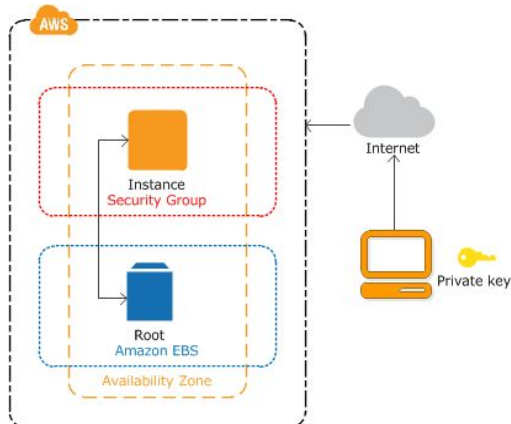
- Remember the key path and name (`~/.ssh/CS205-key.pem`)

2. Launch a VM

We are going to launch an Amazon EBS-backed instance (meaning that the root volume is an EBS[1]

---

[1] EBS stands for "Elastic Block Storage"

volume). We will let Amazon EC2 select an Availability Zone for us .

- Go to the EC2 dashboard and click "Launch Instance".

- Step 1: Select "Ubuntu Server 18.04" as AMI

4

- Step 2: Select "t2.micro" as instance type ("Free Tier Eligible"). It is important to select EBS backed instances for persistency[2].

- Step 3: Click "Next" in the bottom right corner to modify the configurations. Use default configurations for the rest of steps. It is highly recommended to revise and understand all options (especially the "Configure Security Group" option. Changing the "source" from "custom" to "My IP" is safer). At the final step, click "Launch".

- Step 4: Select your key pair and "Launch Instance"

- Step 5: Go to "Running Instances" in EC2 Dashboard and wait for the VM to be "running".

---

[2] EBS will save the data on the drive after the instance is stopped.
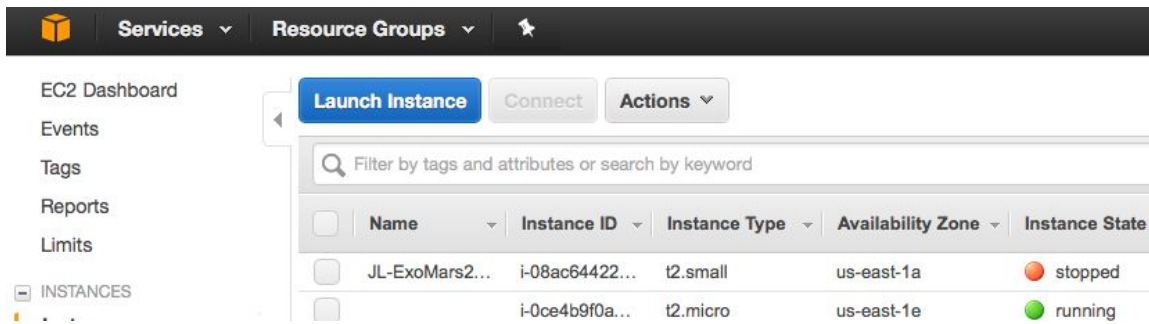
- Step 6: Familiarize yourself with the actions and the status and monitoring information provided by the dashboard.
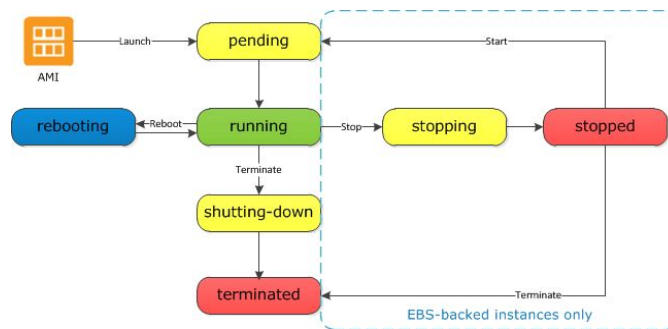
3. Login to the VM

After you launch your instance, you can connect to it and use it the way that you'd use a computer sitting in front of you. It can take a few minutes for the instance to be ready so that you can connect to it. Check that your instance has passed its status checks - you can view this information in the Status Checks column on the Instances page.

To connect to your Linux instance from a computer running Mac or Linux, you'll specify the `.pem` file to your SSH client with the `-i` option and the path to your private key. To connect to your Linux instance from a computer running Windows, you can use either MindTerm or PuTTY. If you plan to use PuTTY, you'll need to install it and use the following procedure to convert the `.pem` file to a `.ppk` file: [Connecting to Your Linux Instance from Windows Using PuTTY](#)

The Linux procedure is as follows:

- Select the instance, and then choose Connect.
- Copy the Public IP address from this pane

Connect to instance   Info

Connect to your instance i-00183dcad65722627 using any of these options

EC2 Instance Connect	Session Manager	SSH client

Instance ID

⧉ i-00183dcad65722627

Public IP address

⧉ 3   ⊘ Public IP address copied

User n

> ubuntu

Connect using a custom user name, or use the default user name ubuntu for the AMI used to launch the instance

- Execute the SSH command to login to your VM, replacing the IP below with the IP address you just copied

```
ssh -i ~/.ssh/CS205-key.pem ubuntu@34.230.37.25
```

```
Welcome to Ubuntu 16.04.2 LTS (GNU/Linux 4.4.0-1022-aws x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:      https://landscape.canonical.com
 * Support:         https://ubuntu.com/advantage

   Get cloud support with Ubuntu Advantage Cloud Guest:
     http://www.ubuntu.com/business/services/cloud

0 packages can be updated.
0 updates are security updates.



The programs included with the Ubuntu system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*/copyright.

Ubuntu comes with ABSOLUTELY NO WARRANTY, to the extent permitted by
applicable law.

To run a command as administrator (user "root"), use "sudo <command>".
See "man sudo_root" for details.

ubuntu@ip-172-30-4-13:~$
```

In most cases, the public IP address is associated with the instance until it's stopped or terminated, after which it's no longer available for you to use. If you require a persistent public IP address that you can associate and disassociate at will, use an Elastic IP address (EIP) instead. You can allocate your own EIP, and associate it to your instance after launch.

- **Windows**

- Download putty
- Copy the public IP address of the instance
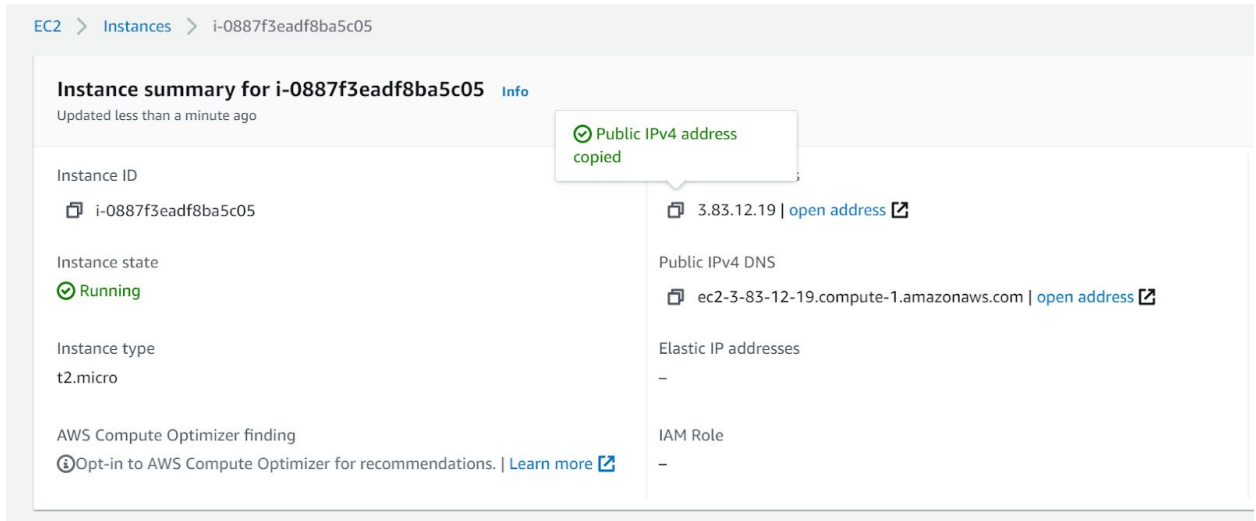
- Open Putty and paste in the IP into Host Name
- On the left side click on SSH -> Auth and browse for your ppk file
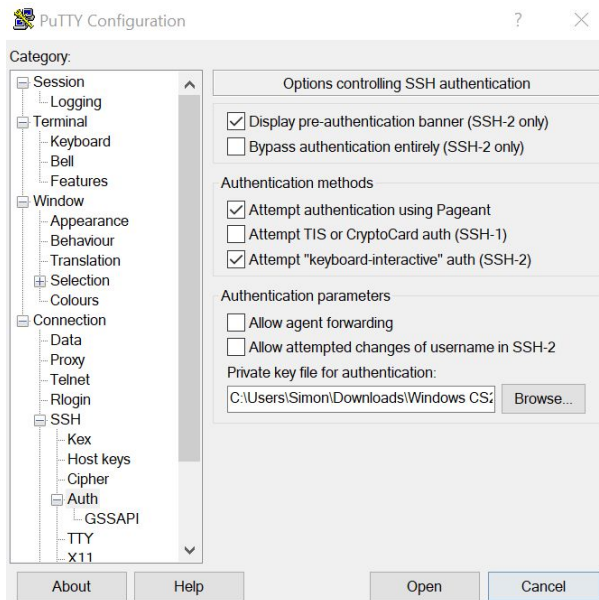
- Click "Open"
- Login as "ubuntu"

4. Not able to login?

If you could not login, then make sure that the security groups (firewalls) of the VM opens the port 22 to the outside world. Under the security access of the cluster console, check the security groups.

- Click on the Security groups of the VM, id:"sg-9d898fed" (your security groups will have different

CS205: Computing Foundations for Computational Science, Spring 2021

IDs but it starts with "sg-"). This can be found on the left pane under "Network and Security".

- Select the Inbound tab and click on Edit

- Click on "Add Rule", a new row will be created and fill out the new row with the following:

- Click "Save"

If you still cannot login, make sure your VPC route table is configured to allow traffic to and from the Internet

- Open the Amazon **VPC** console.

- In the navigation pane, choose **Route Tables** and then select your VPC route table from the list.

- On the Routes tab, ensure that you have a default route pointing to your **Internet gateway** (IGW).

- If you do not see this, choose Internet Gateways from the navigation pane and copy the ID of your Internet gateway. If you do not have an Internet gateway, create one and attach it to your VPC. Be sure to copy the ID of the new IGW

- Go back to Route Tables and select the Routes tab.

- Edit and create a route that points `0.0.0.0/0` to your Internet gateway ID.

- Save the route table.

5. Upload/download files to/from the VM

After you launch your instance, you can use the SCP command to upload and download files to/from your client computer.

Please make a new terminal window/tab, keeping the original window that is connected to the EC2 instance.

- Download the file "shrek.txt" from the lab website

- Copy the file [shrek.txt](#) from the local host to a remote host

```
scp -i ~/.ssh/CS205-key.pem ~/Downloads/shrek.txt
ubuntu@34.230.37.25:/home/ubuntu
```

Confirm that the file was copied into your /home directory
```
less shrek.txt
```

Hit the `q` key to quit less.

Windows
- You have a couple of options, either PSCP or WinSCP, though the WinSCP GUI might be the easiest. Paste your EC2 instance's Public IP address into "Host Name" and input username "ubuntu"
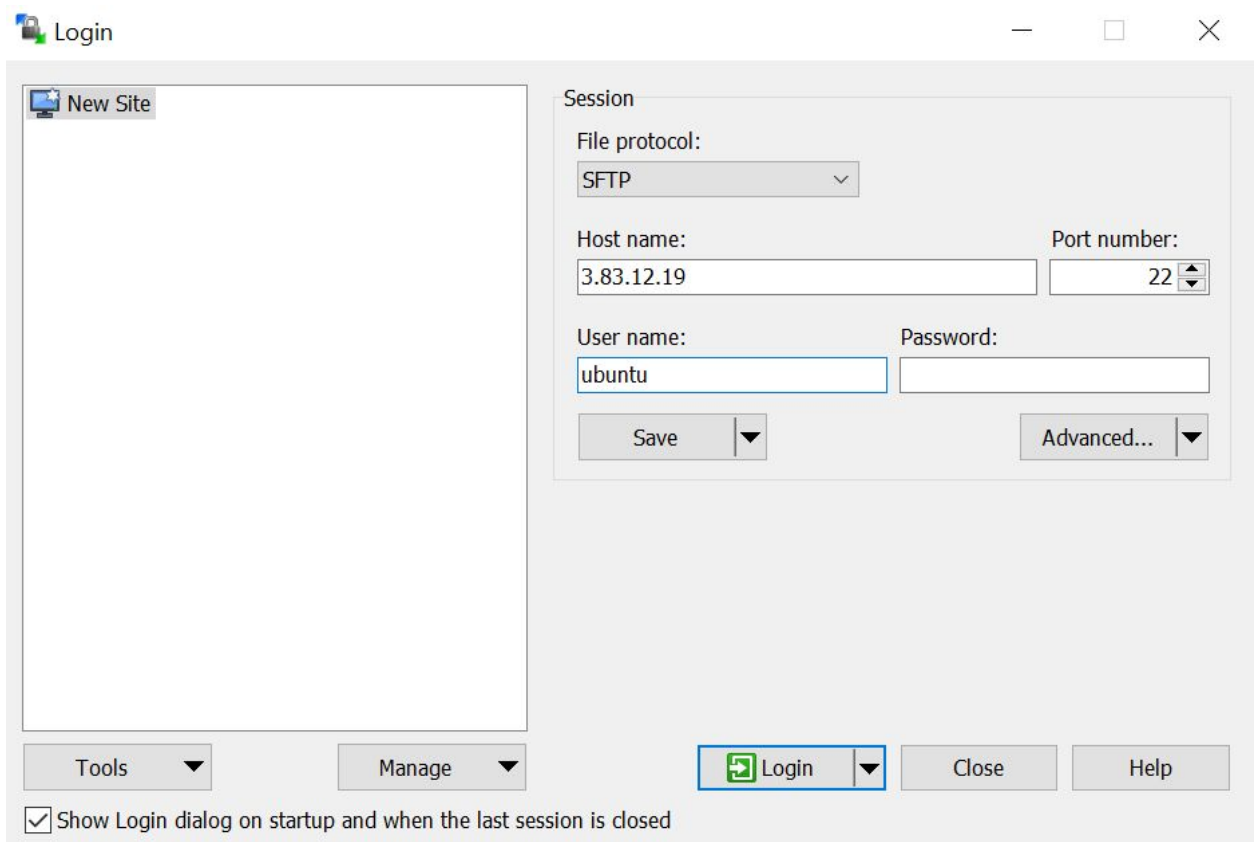
Advanced -> SSH -> Authentication, upload the ppk
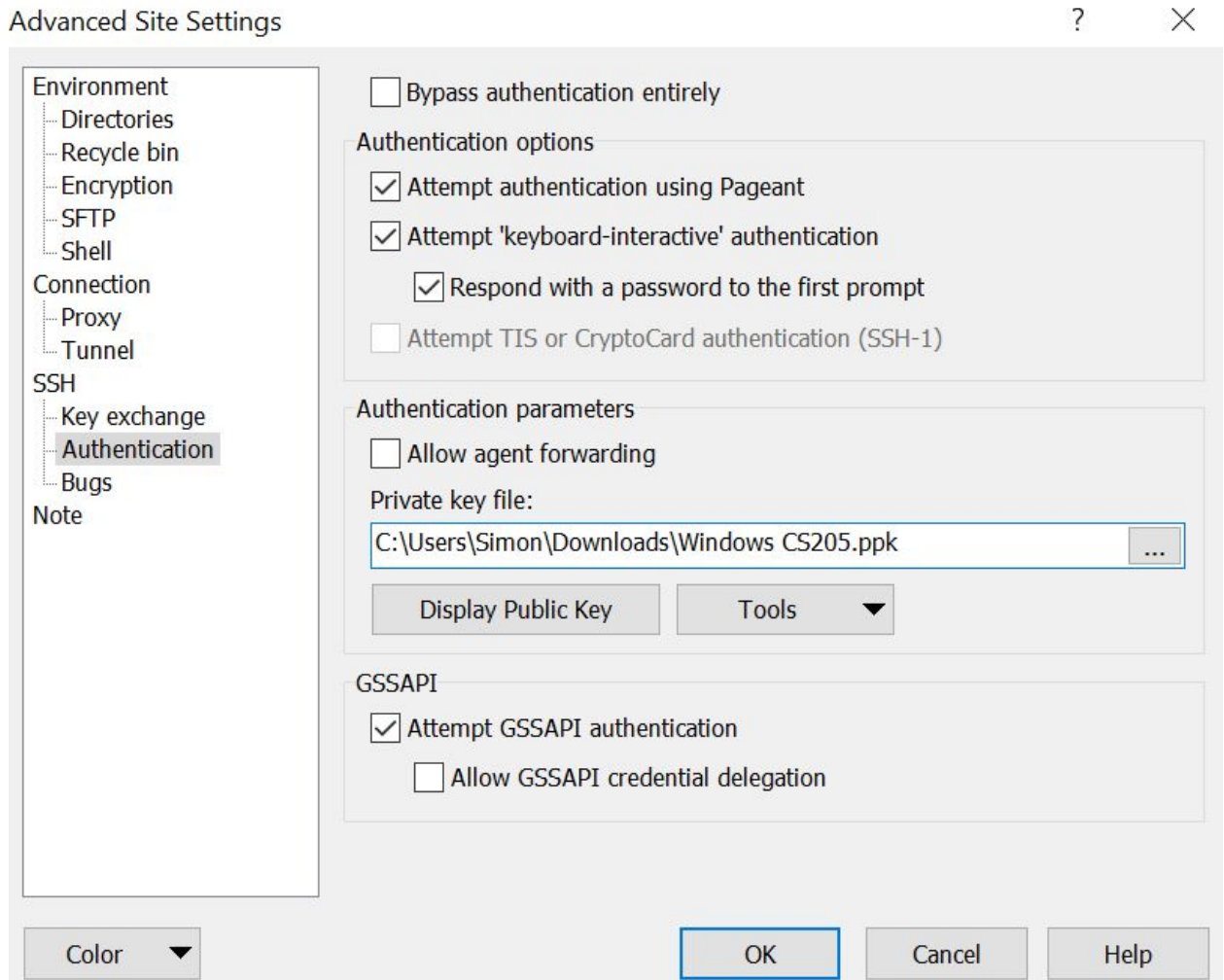
Drag files into the window or from the filesystem on the left.

Optional:

- To copy the file "shrek.txt" from a remote host to the local host

```
scp -i ~/.ssh/CS205-key.pem ubuntu@34.230.37.25:/home/ubuntu/shrek.txt .
```

- You can also use the tool `sftp`, which is integrated into many graphical tools

```
sftp -i ~/.ssh/CS205-key.pem ubuntu@34.230.37.25
```

5. Disconnect and terminate instance

- Type `exit` in your terminal window to disconnect.

- Terminate the instance via the EC2 console

Instances (1/1) Info				C̶	Connect	Instance state ▲	

🔍 Filter instances

Instance state: running ✕	Clear filters

Stop instance

Start instance

Reboot instance

Hibernate instance

Terminate instance

☑	Name ▽	Instance ID	Instance state ▽	Instance type ▽
☑	–	i-0d893d3a786dc34d6	⊘ Running ⊕⊖	t2.micro

6. Bucket and object management in S3

Amazon S3 has a simple web services interface that you can use to store and retrieve any amount of data, at any time, from anywhere on the web.

● Every object in Amazon S3 is stored in a bucket. Before you can store data in Amazon S3, you must create a bucket. Go to the S3 dashboard and click "Create Bucket" in our default region "US East". The name must be unique across all existing bucket names in Amazon S3. After you create the bucket you cannot change the name, so choose wisely. Choose a bucket name that reflects the objects in the bucket because the bucket name is visible in the URL that points to the objects that you're going to put in your bucket.

Proceed through the four steps. You don't need to update the permissions in step 3 at the moment.

Create bucket

Buckets are containers for data stored in S3. Learn more ↗

General configuration

Bucket name

```
some-unique-bucket-name-cs205
```

Bucket name must be unique and must not contain spaces or uppercase letters. **See rules for bucket naming** ↗

Region

```
US East (N. Virginia) us-east-1                                    ▼
```

Copy settings from existing bucket - *optional*
Only the bucket settings in the following configuration are copied.

Choose bucket

Scroll to the bottom and press "Create Bucket". Don't bother changing any settings.

- To upload an object to a bucket, choose the name of the bucket that you want to upload your object to, choose Upload and then "Add Files"

Upload

Add the files and folders you want to upload to S3. To upload a file larger than 160GB, use the AWS CLI, AWS SDK or Amazon S3 REST API. Learn more ↗

Drag and drop files and folders you want to upload here, or choose **Add files**, or **Add folders**.

Files and folders (1 Total, 511.0 B)
All files and folders in this table will be uploaded.

Remove   Add files   Add folder

🔍 Find by name   〈 1 〉

	Name ▲	Folder ▽	Type ▽	Size ▽
☐	shrek.txt	-	text/plain	511.0 B

Scroll to the bottom and press "Upload".

- Click on the destination to bring you back to the bucket. Confirm the file is in the bucket.

some-unique-bucket-name-cs205

Objects   Properties   Permissions   Metrics   Management   Access points

Objects (1)

Objects are the fundamental entities stored in Amazon S3. For others to access your objects, you'll need to explicitly grant them permissions. Learn more

| | Delete | Actions ▼ | Create folder | Upload |

Q Find objects by prefix   ‹   1   ›   ⚙

	Name ▲	Type ▽	Last modified	Size ▽	Storage class ▽
☐	📄 shrek.txt	txt	January 25, 2021, 20:09:36 (UTC-05:00)	511.0 B	Standard

- Delete the file. Then click on "Buckets" on the left to go back to the bucket list. Delete your bucket.

Stop your instances when are done for the day to avoid incurring charges
Terminate them when you are sure you are done with your instance.