AWS EC2/EFS Guide

Lijun Yu [lijun@cmu.edu](mailto:lijun@cmu.edu)

# Objectives

* Launch an EC2 virtual machine for computing.
* Create an EFS file system for data storage.
* Attach the EFS file system to the EC2 virtual machine.

# Prerequisites

A logged-in AWS account. The account ID should be in the format of *<andrewid>+11775@andrew.cmu.edu*.

If applicable, redeem your credit code following <https://aws.amazon.com/premiumsupport/knowledge-center/add-aws-promotional-code/>.

# Notes

* This guide operates in the AWS us-east-1 region which is near Pittsburgh. Feel free to change to another region near your location.
* This guide creates a t2.micro instance for minimum cost. You may later resize it or create a new one with more CPUs and memory.

# Procedure

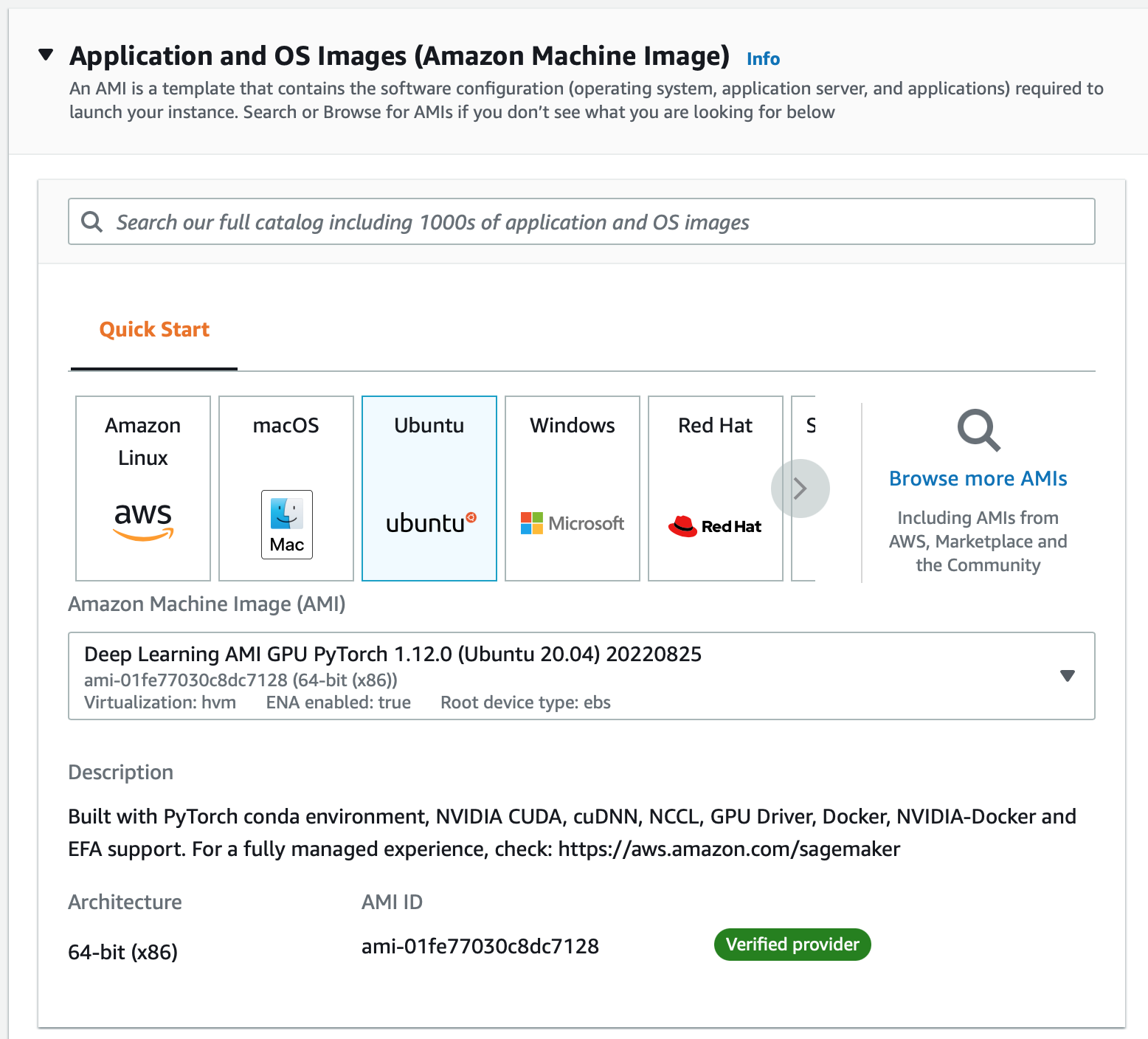
## Step 1 Launch a virtual machine

Go to <https://console.aws.amazon.com/ec2/v2/home?region=us-east-1#Instances>



Click *Launch instances*.

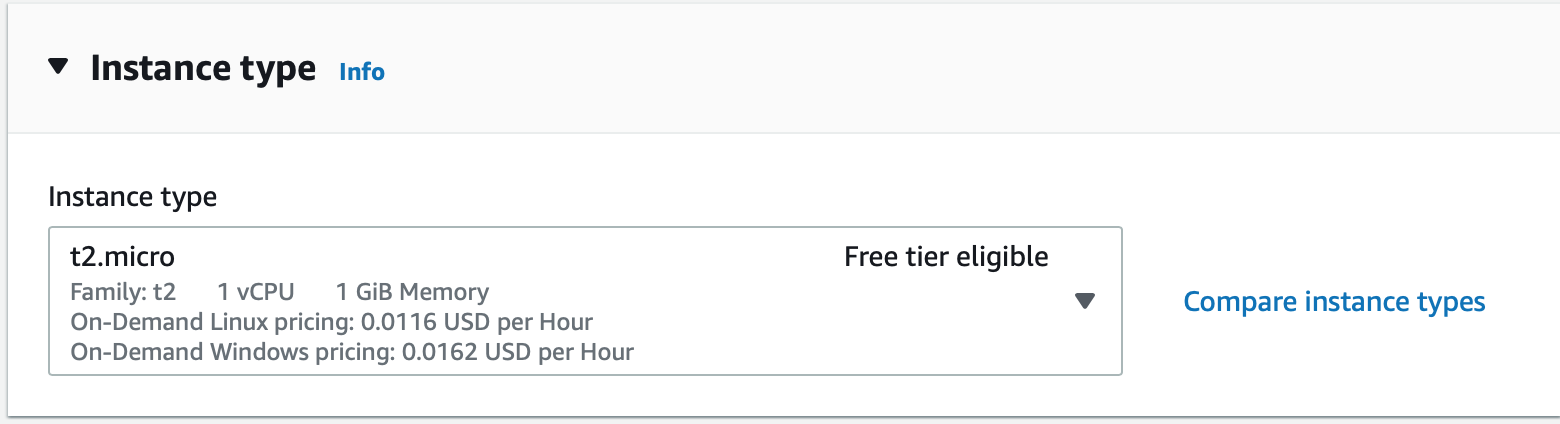
In *Application and OS Images (Amazon Machine Image)*, click *Ubuntu* and select ***Deep Learning AMI GPU PyTorch 1.12.0 (Ubuntu 20.04)*** *20220825* from the dropdown.



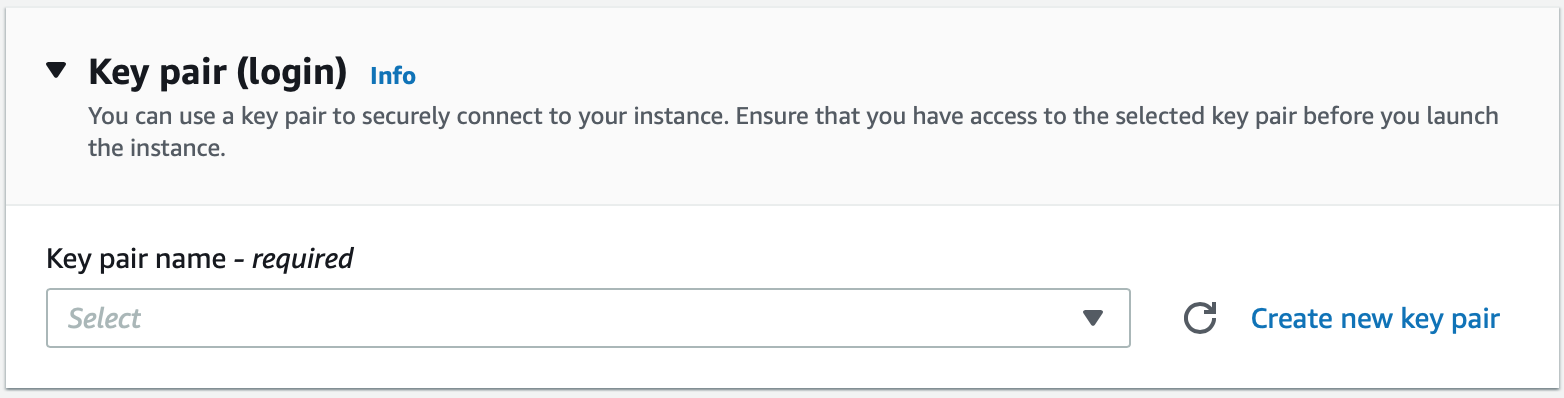
In *Instance type*, choose *t2.micro* for now. Later we’ll choose larger *t2* series for more CPUs or *g4dn* series for GPU access.

For a newly opened AWS account, the default quota is 5 vCPUs ([Quota details](https://us-east-1.console.aws.amazon.com/servicequotas/home/services/ec2/quotas/L-1216C47A)), you can try to request for an increase but the aws customer service won’t respond very fast.

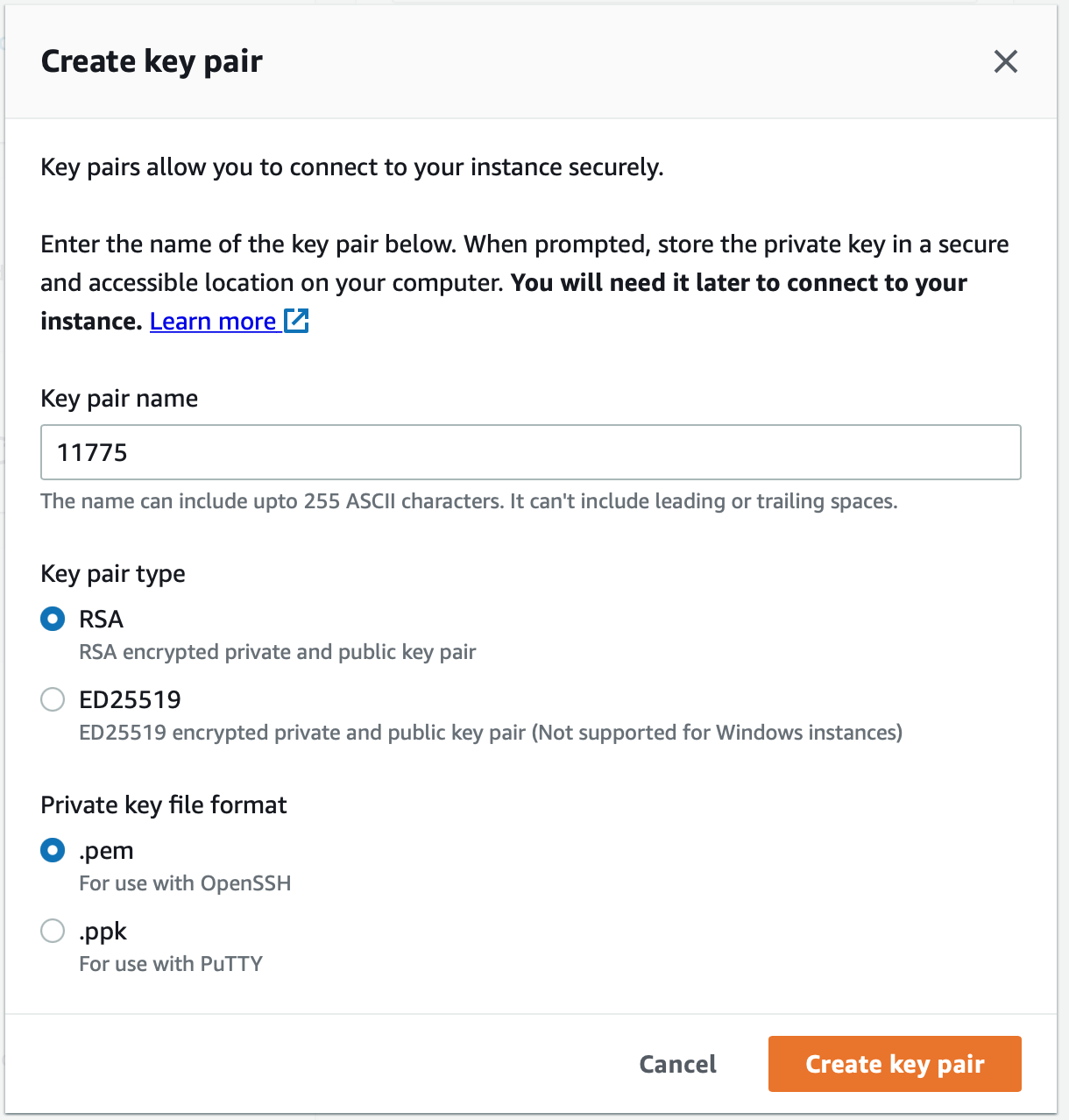
When your AWS account is fully activated (likely within 7 days), you will have a quota of 64 vCPUs for all instances including GPUs, which is sufficient for all homework.



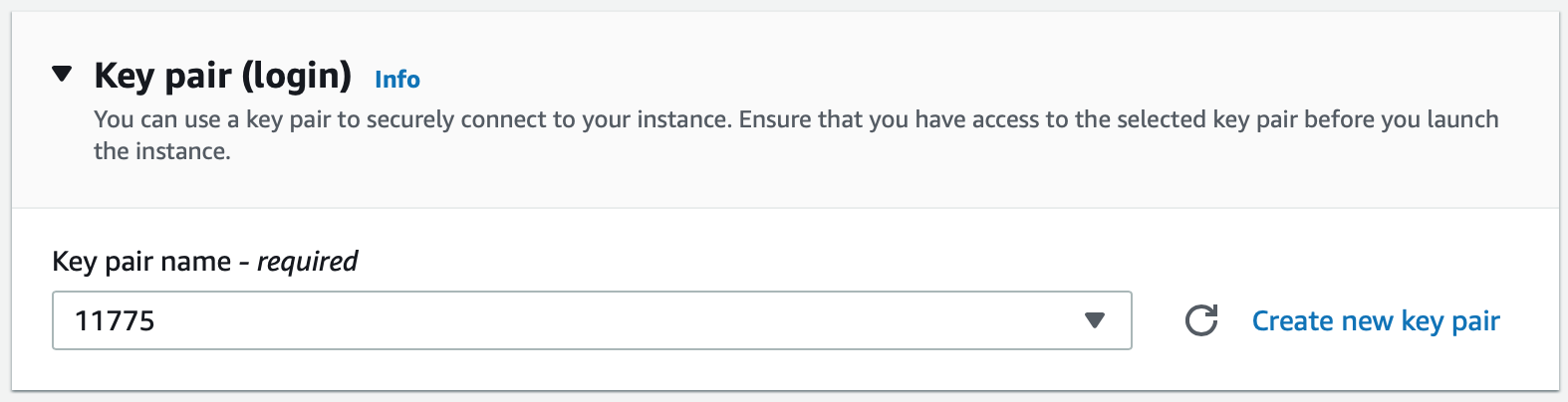
In *Key pair (login)*, click *Create new key pair* for the first time. (Later we’ll just select *11775* in *Key pafir name*.)



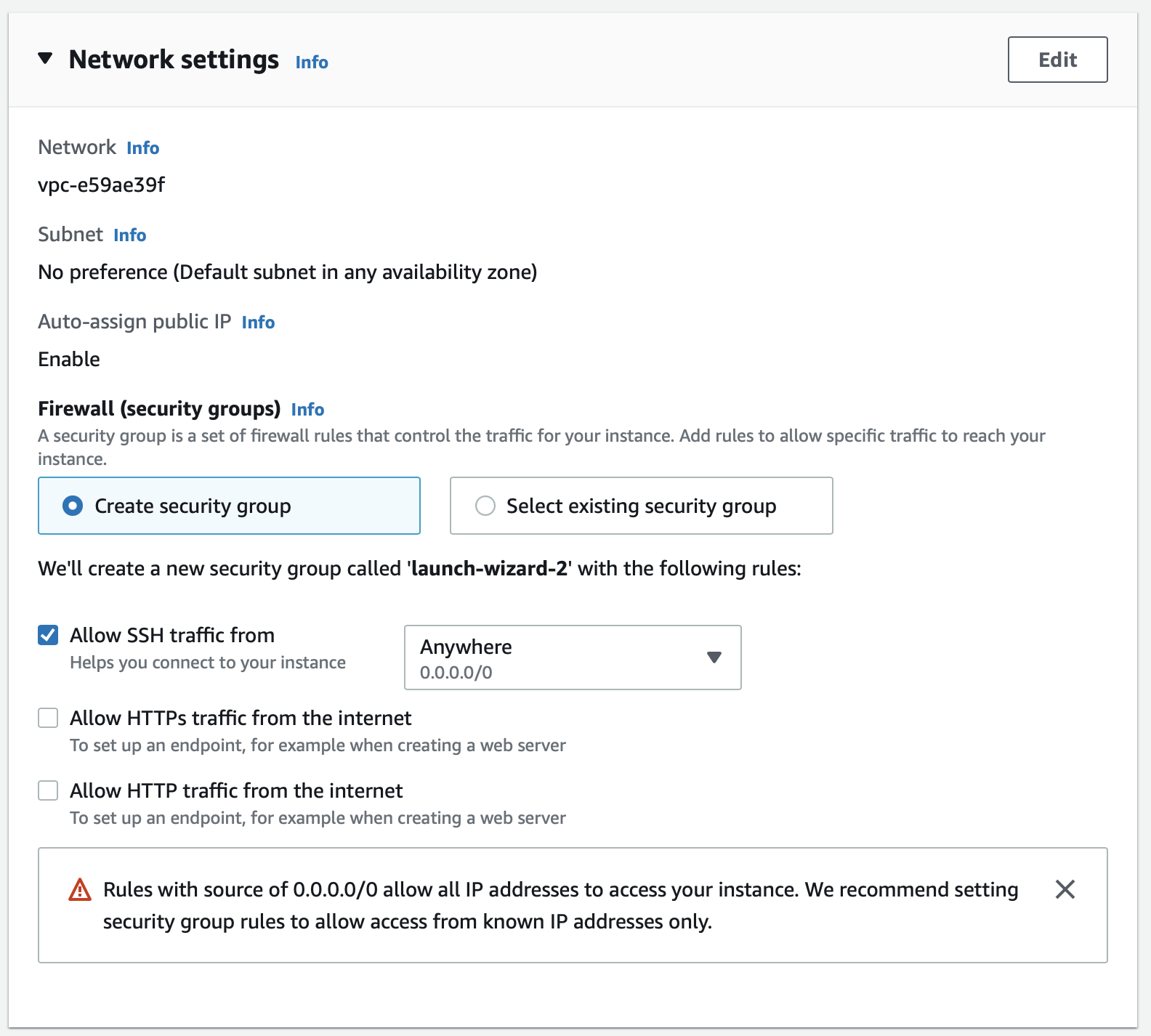
Enter ***11775*** in *Key pair name*, and press *Create key pair*. Keep the downloaded key file.



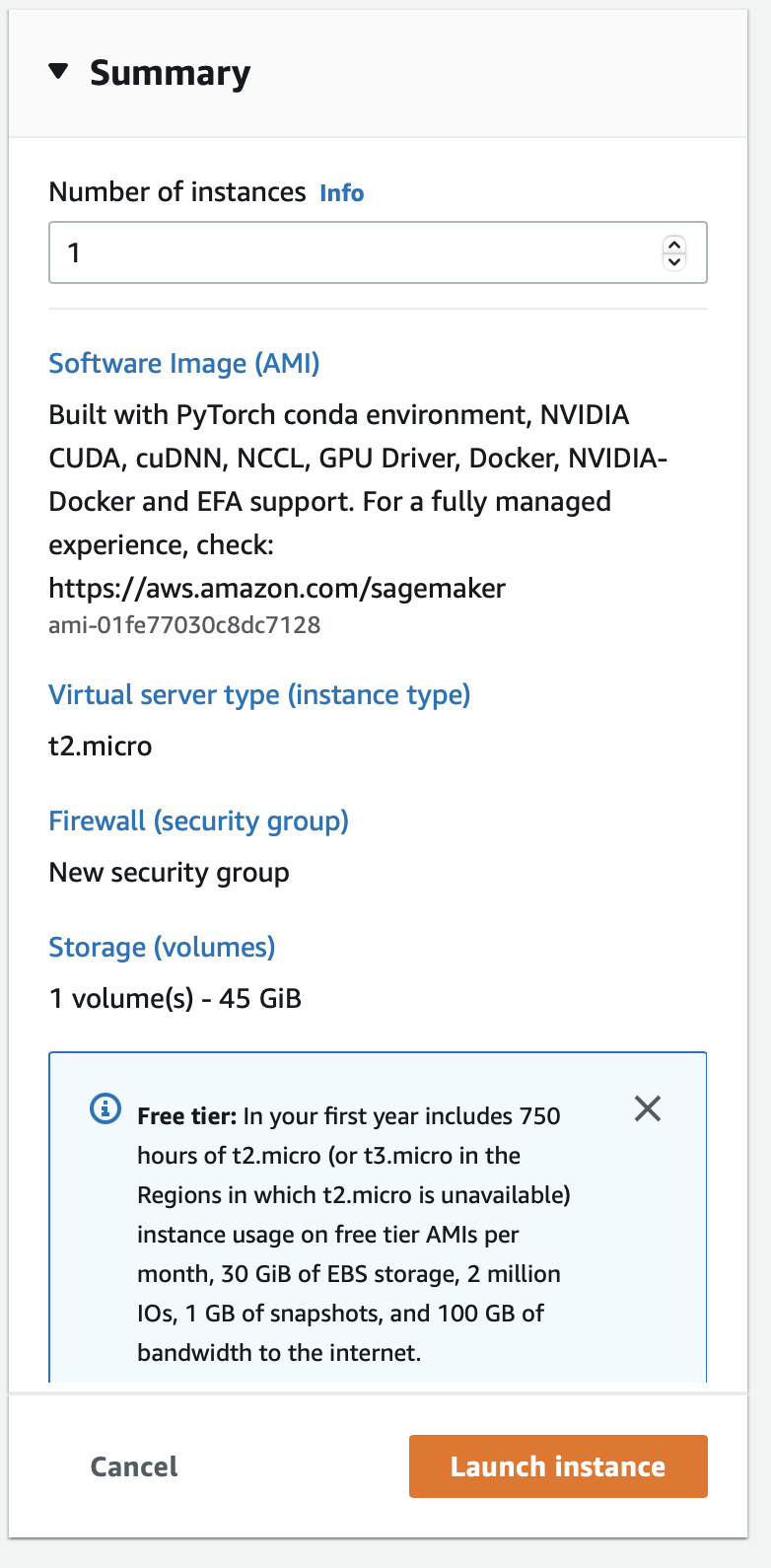
And select *11775* in *Key pair name*.



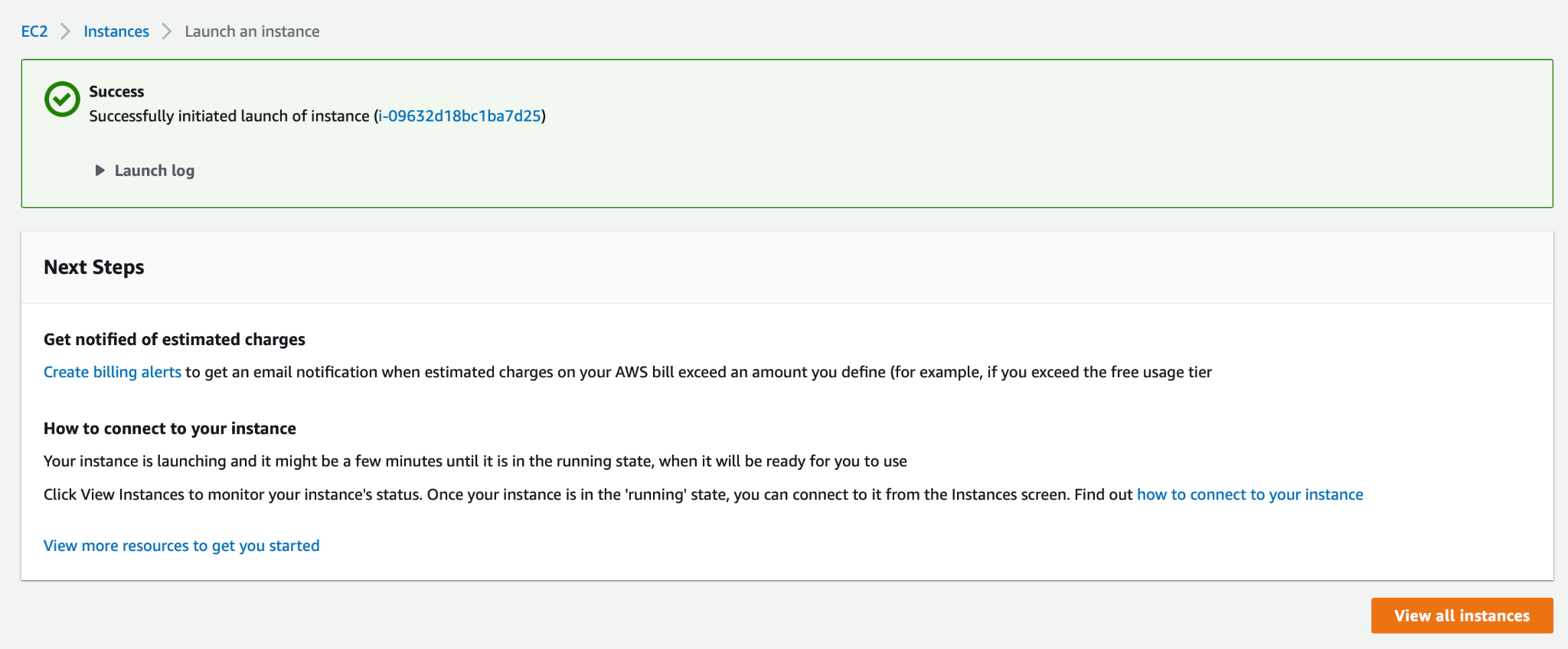
In *Network settings*, select *Create security group* and *Allow SSH traffic from Anywhere*.



Now check the *Summary*, and Then click *Launch instance*.



Finally, click *Launch Instances*.

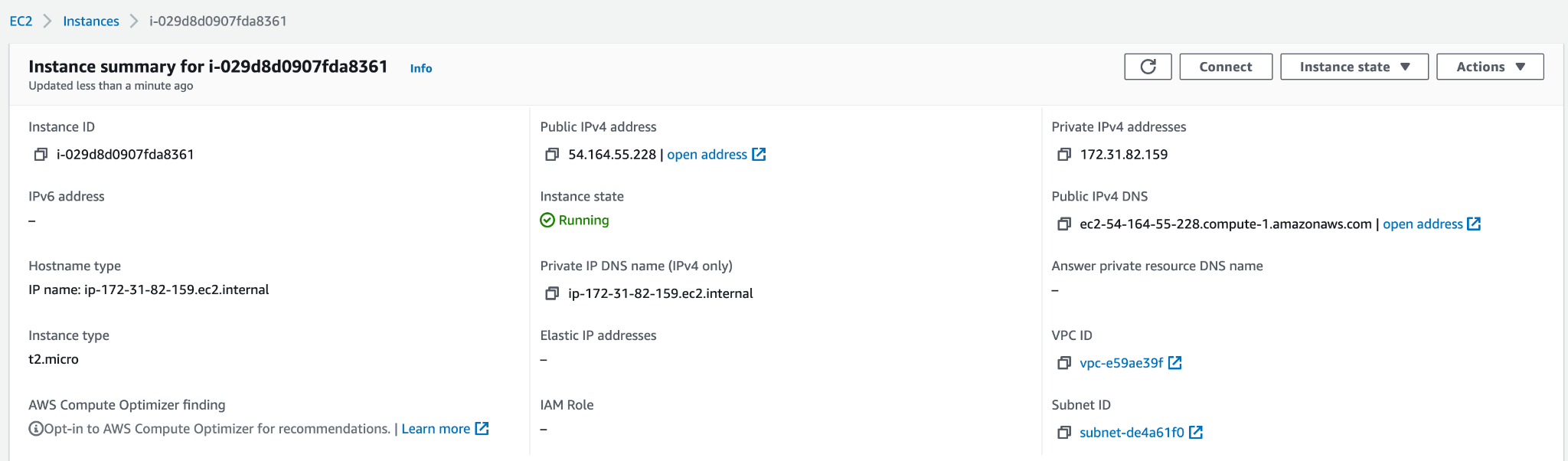


## Step 2 Connect to the virtual machine

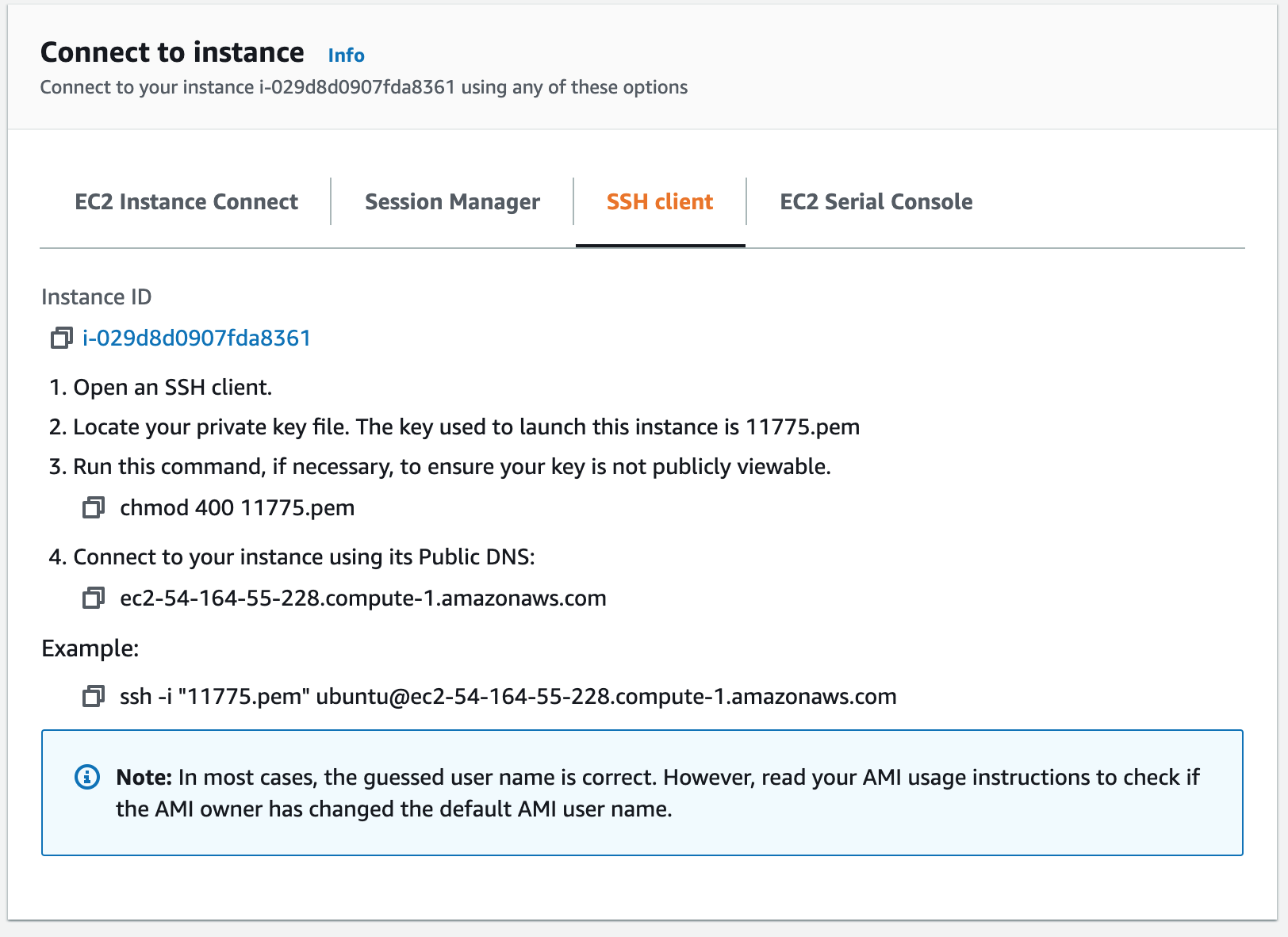
Click *View all instances*.

## 

Click the link under *Instance ID* to enter the details page.

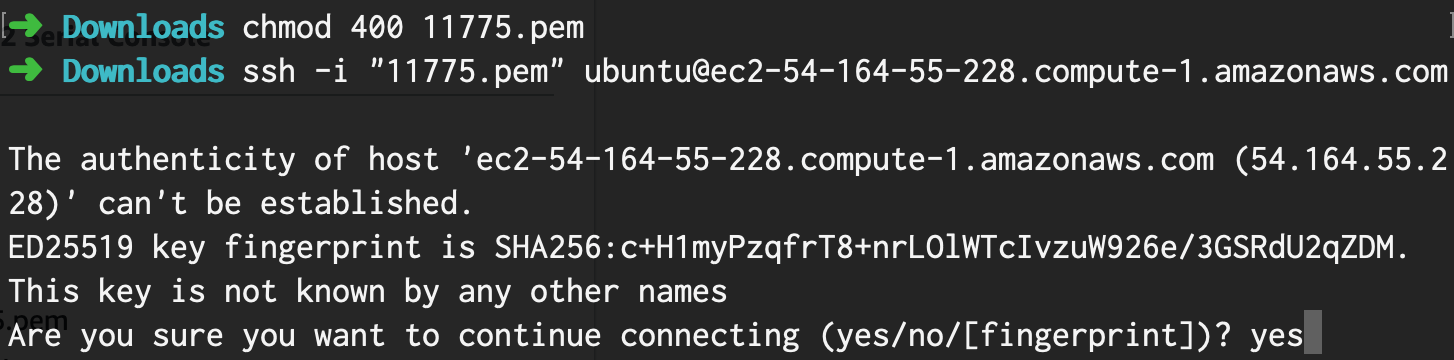


Click *Connect*, then select *SSH client* and follow the instructions.

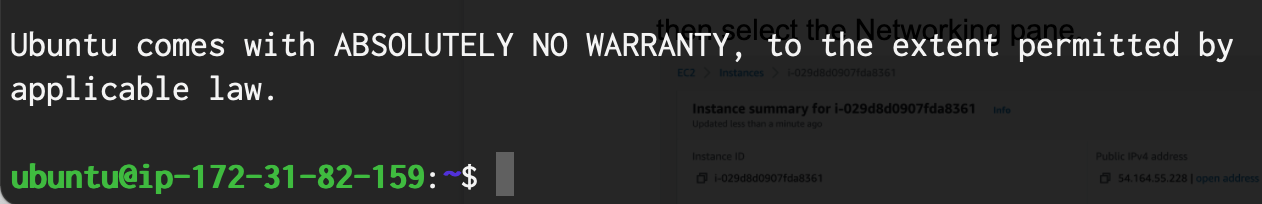


For Mac, open the Terminal.app. For Windows, see ​​<https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/AccessingInstances.html>

In the terminal, first **chmod**, then **ssh**. When prompted, type *yes*.



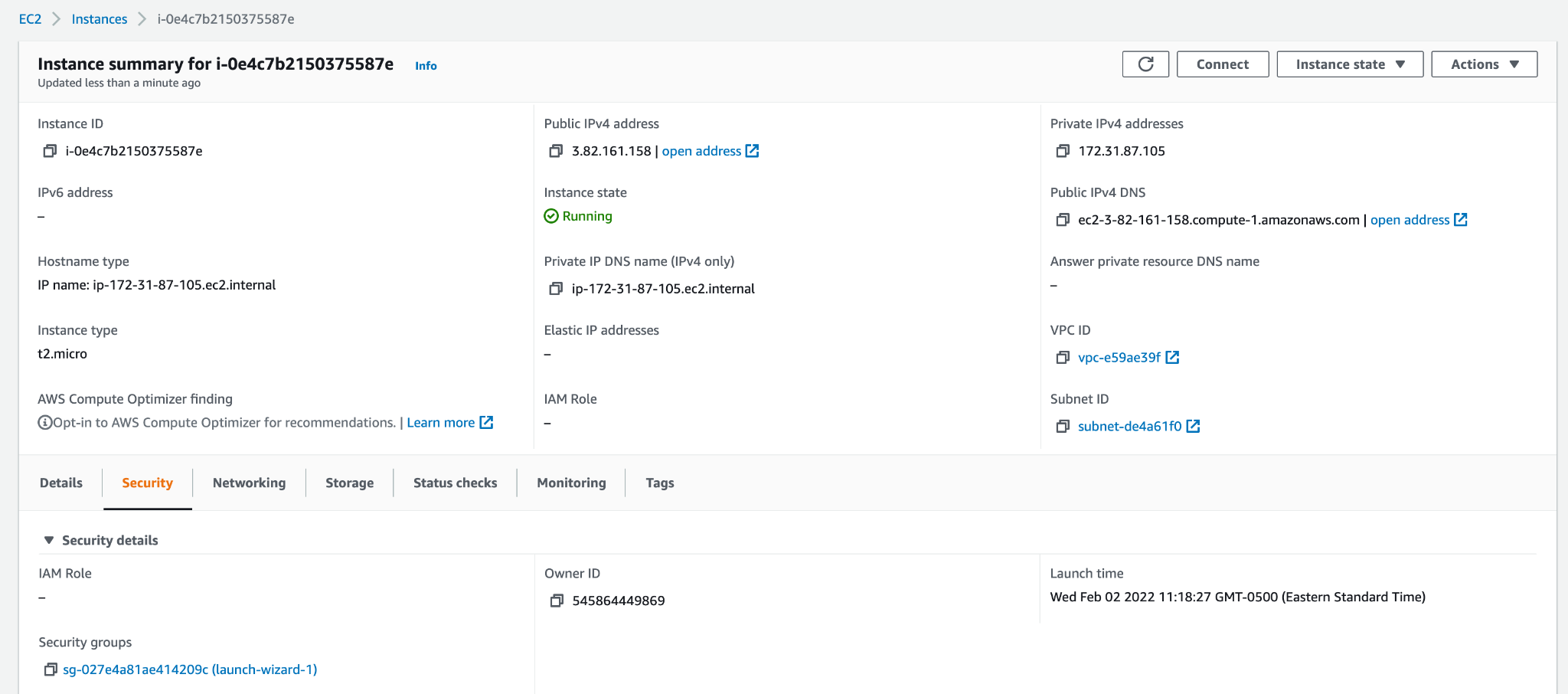
Now the virtual machine shell appears.



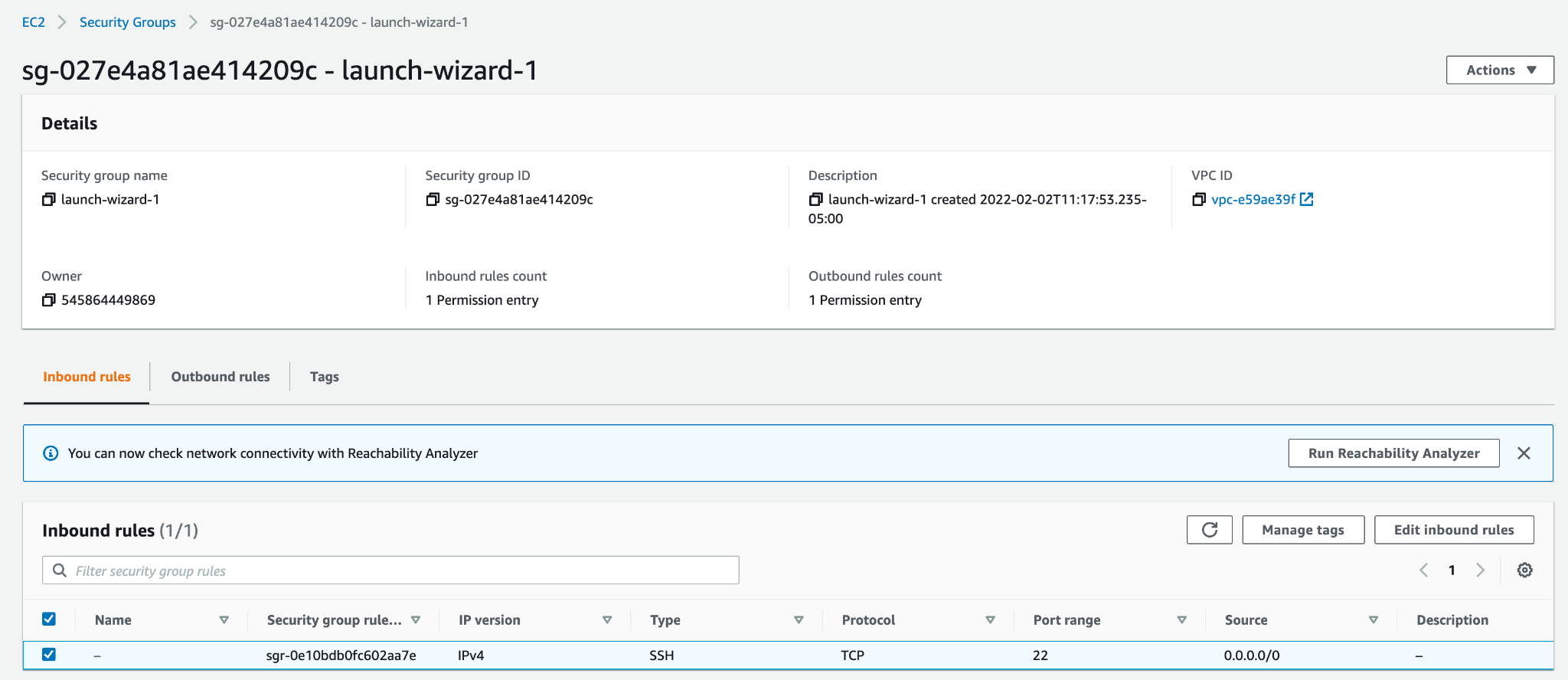
## Step 3 Configure the security group

Click *Cancel* on the *Connect* page.

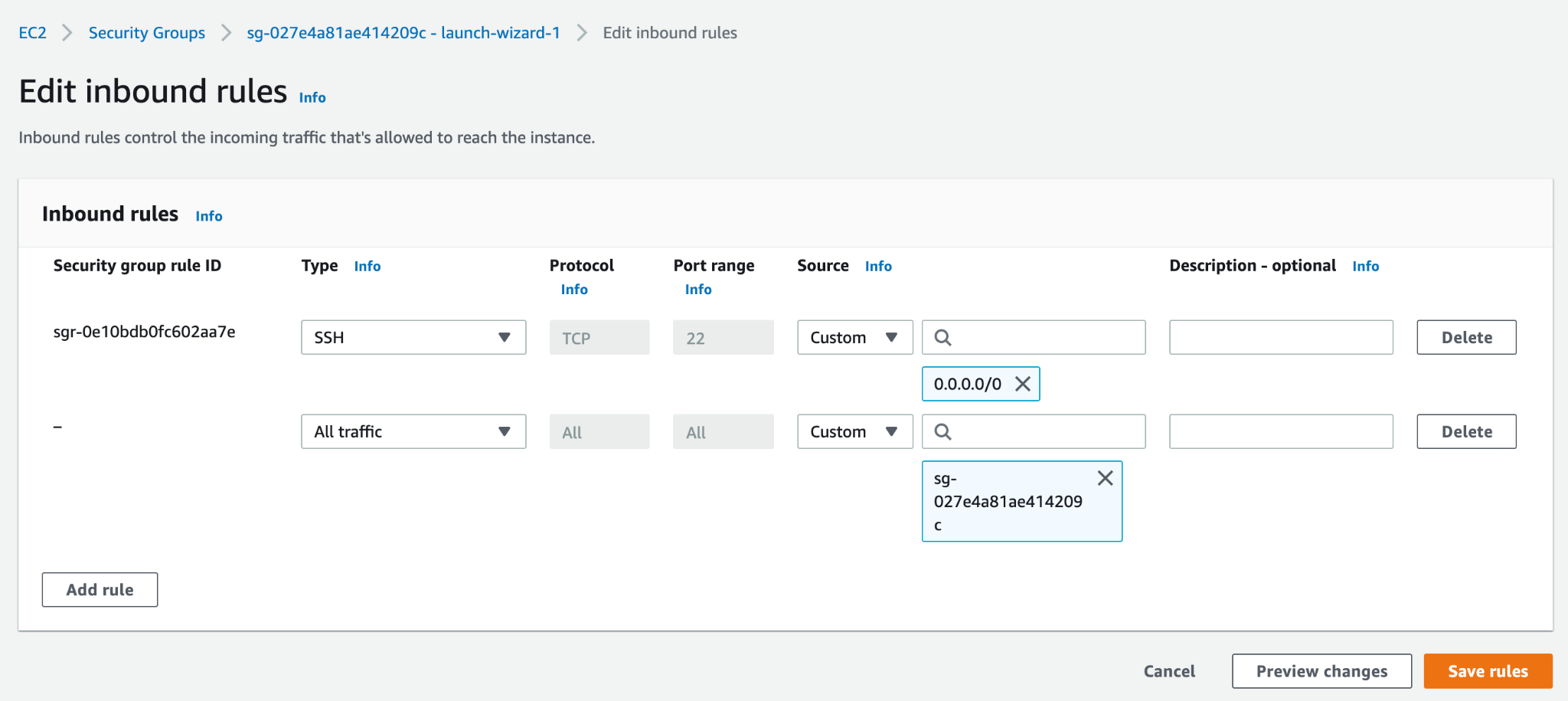
Select the *Security* pane at the bottom.



Click the link under *Security groups* to enter the security group page



Click *Edit inbound rules*, then *Add rule*. In the newly added rule, select *All traffic* for *Type* and the **current security group** as *Source*. In this way, all internal traffic is allowed.

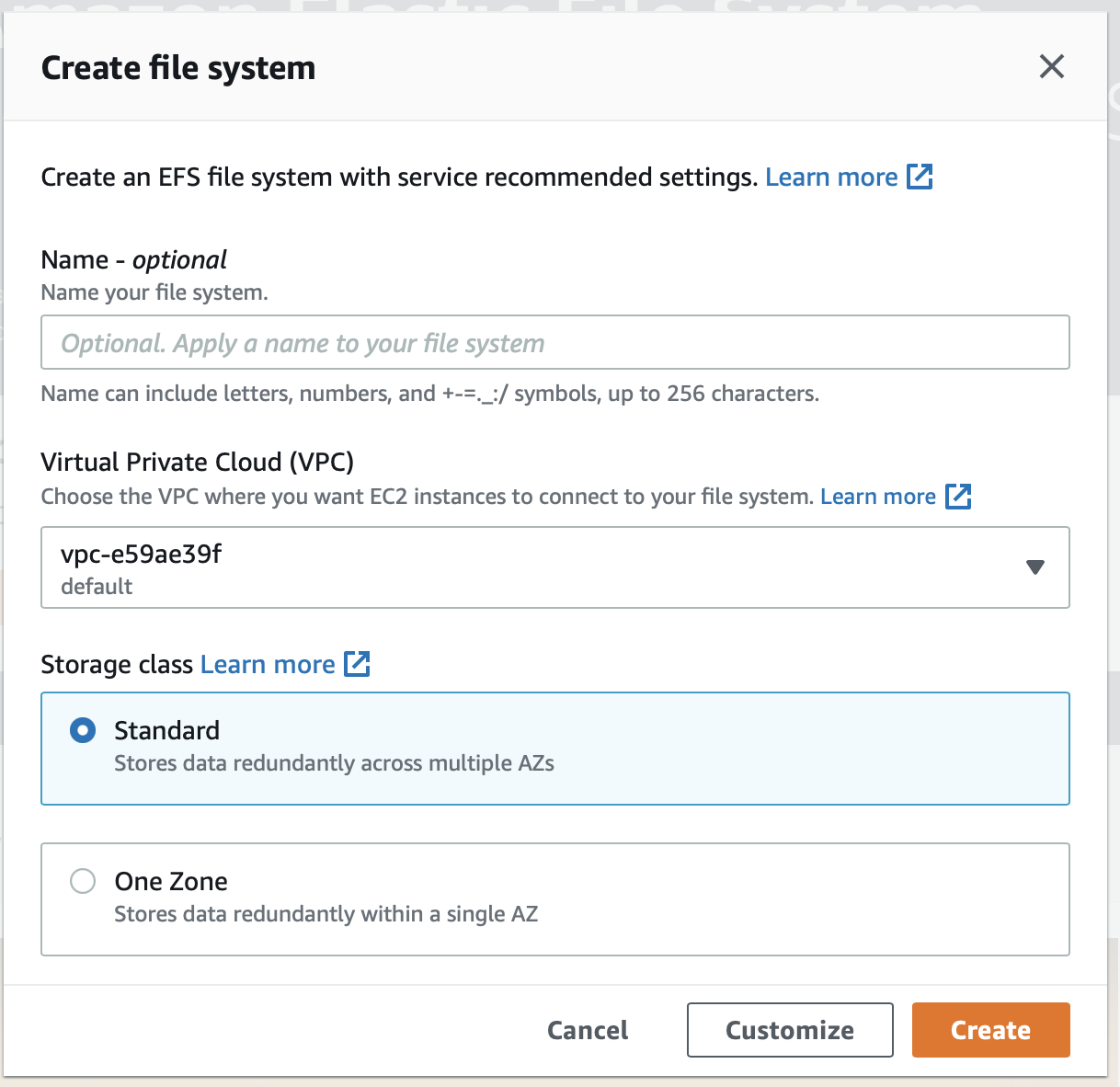


Then click *Save rules*.

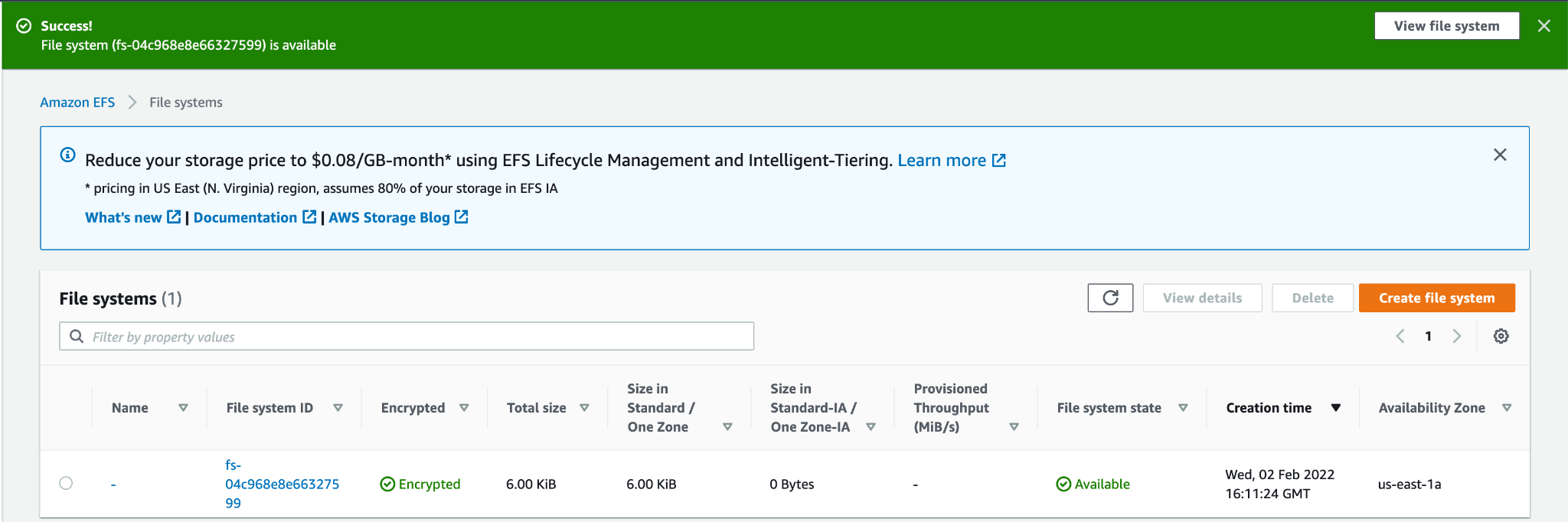
## Step 4 Create a file system

Go to <https://console.aws.amazon.com/efs/home?region=us-east-1#/get-started>

Click *Create file system*, and then *Create*.

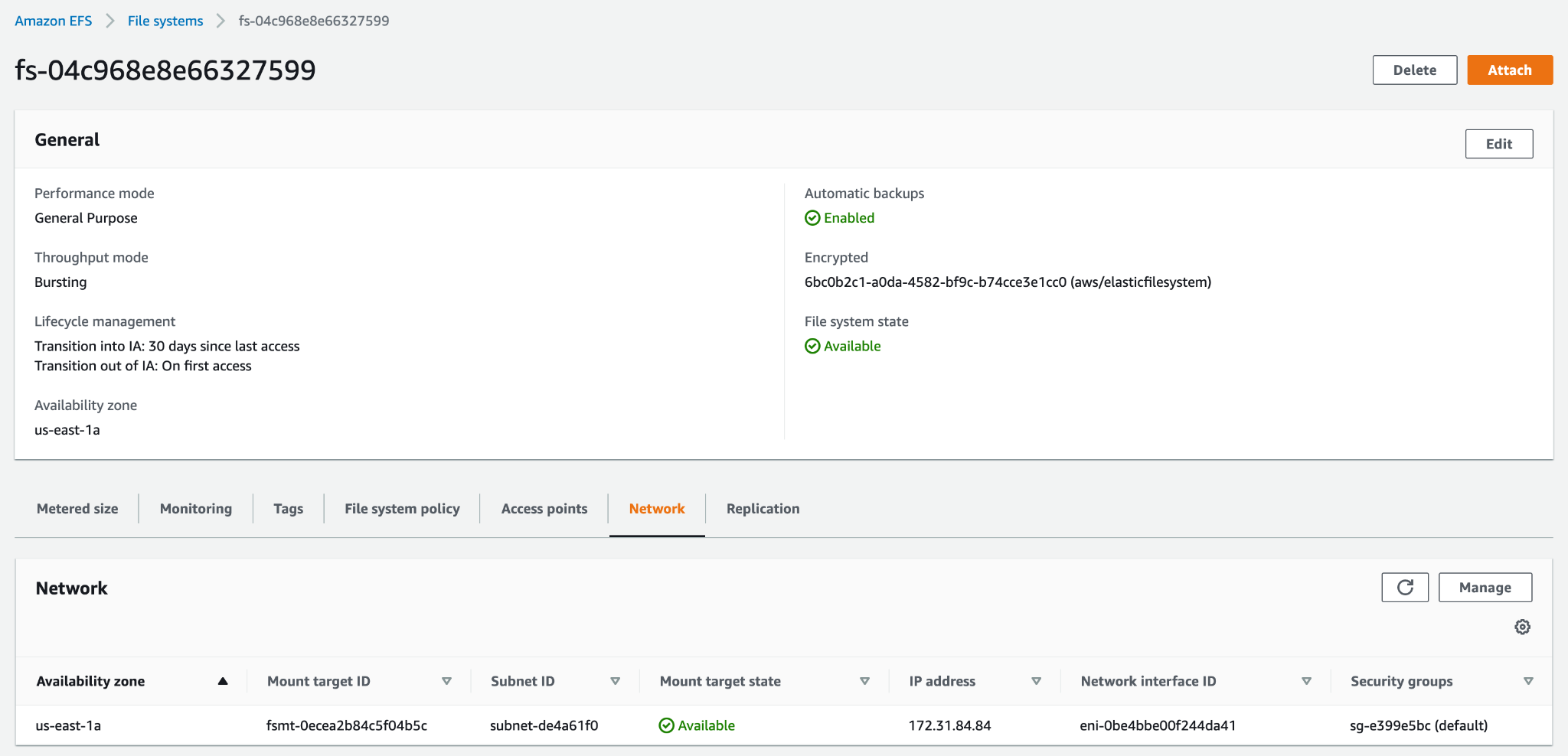


It will jump to the file system list page as below.

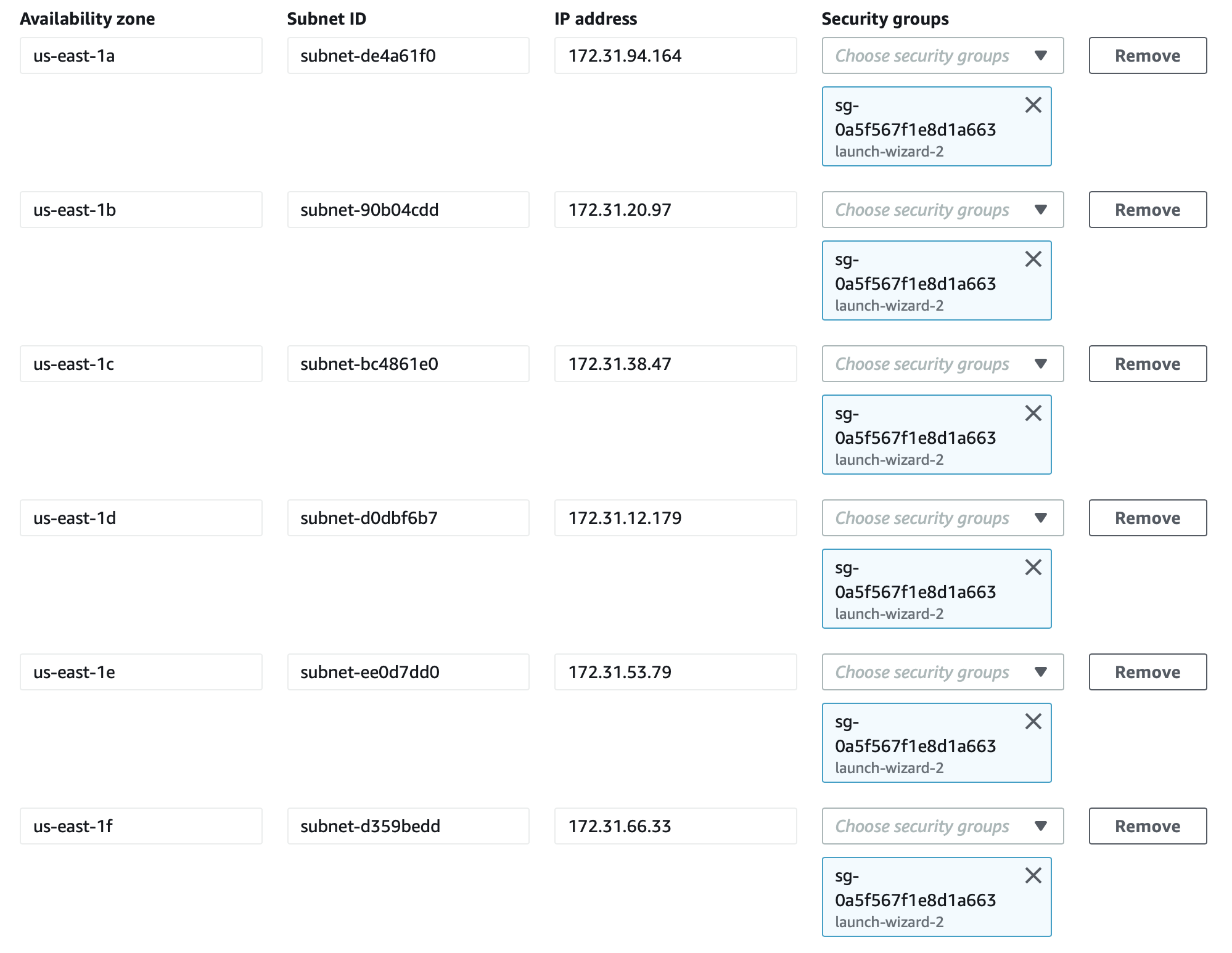


## Step 5 Configure access to EFS

Click the link under *File system ID* to enter the details page, then select *Network* pane. Wait for all *Mount target state* values to become *available*.



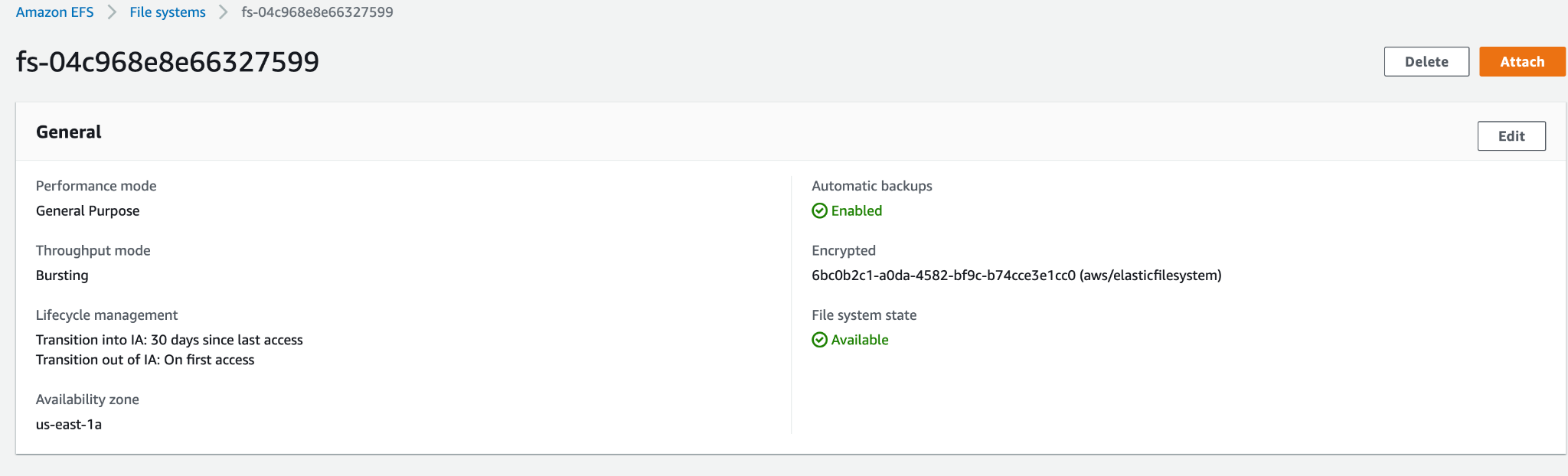
Click *Manage*, and change the *Security groups* for all *Mount targets* to only the same security group as your EC2 instance.



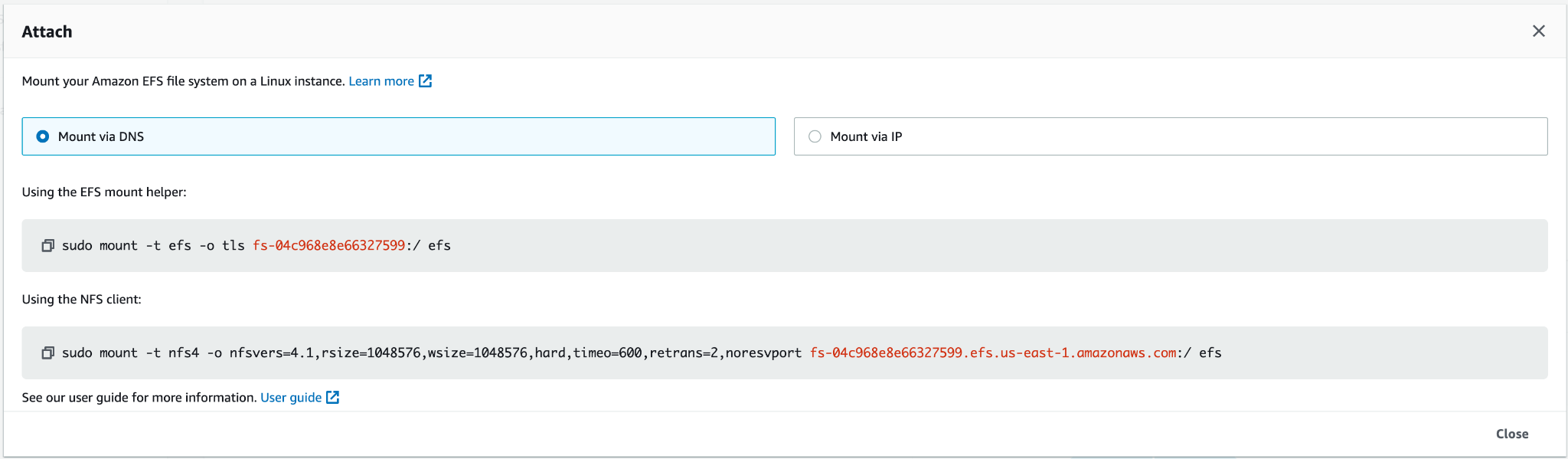
Then click *Save*.

## Step 6 Mount EFS to the virtual machine

You should now be back on the file system details page.



Click *Attach* at the top right corner.



In your EC2 instance, create a directory as the mount point, e.g.

*mkdir efs*

Copy the command under *Using the* ***NFS*** *client* and execute it in your EC2 (optionally change the last parameter to your preferred mount point). The command should finish shortly, otherwise, you should check Steps 3 and 5 to make sure your security group config is correct.

## Step 7 Check and access the file system

In your EC2, run the following command

*df -h*

In the output, there should be one line of the file system you just mounted.



Then you can access its content under the mounted path.

You will have to repeat Step 7 every time you reboot your EC2 instance. Alternatively, you can try to add the mount point to */etc/fstab* with instructions at <https://docs.aws.amazon.com/efs/latest/ug/efs-mount-helper.html#mount-fs-auto-mount-onreboot>.

When doing your homework, try to put everything on the *efs* filesystem you created following the above guide. It is elastic and persistent so you don’t need to worry about disk size and instance termination. It is also easy to scale to different machine type and multiple machines. Things to put on *efs* include:

* The downloaded dataset, which you will be using for all 3 homeworks
* The extracted features in homework 1 and 2, which you will need in homework 3
* The source code
* The conda environment (optional, see below)

You can create your conda environment inside the file system. For example, suppose your environment name is *11775-hw2*, instead of doing

*conda env create -f environment.yml -n 11775-hw2*

*conda activate 11775-hw2*

do

*conda env create -f environment.yml -p /home/ubuntu/efs/11775-hw2*

and then you can activate it on all machines via

*conda activate /home/ubuntu/efs/11775-hw2*