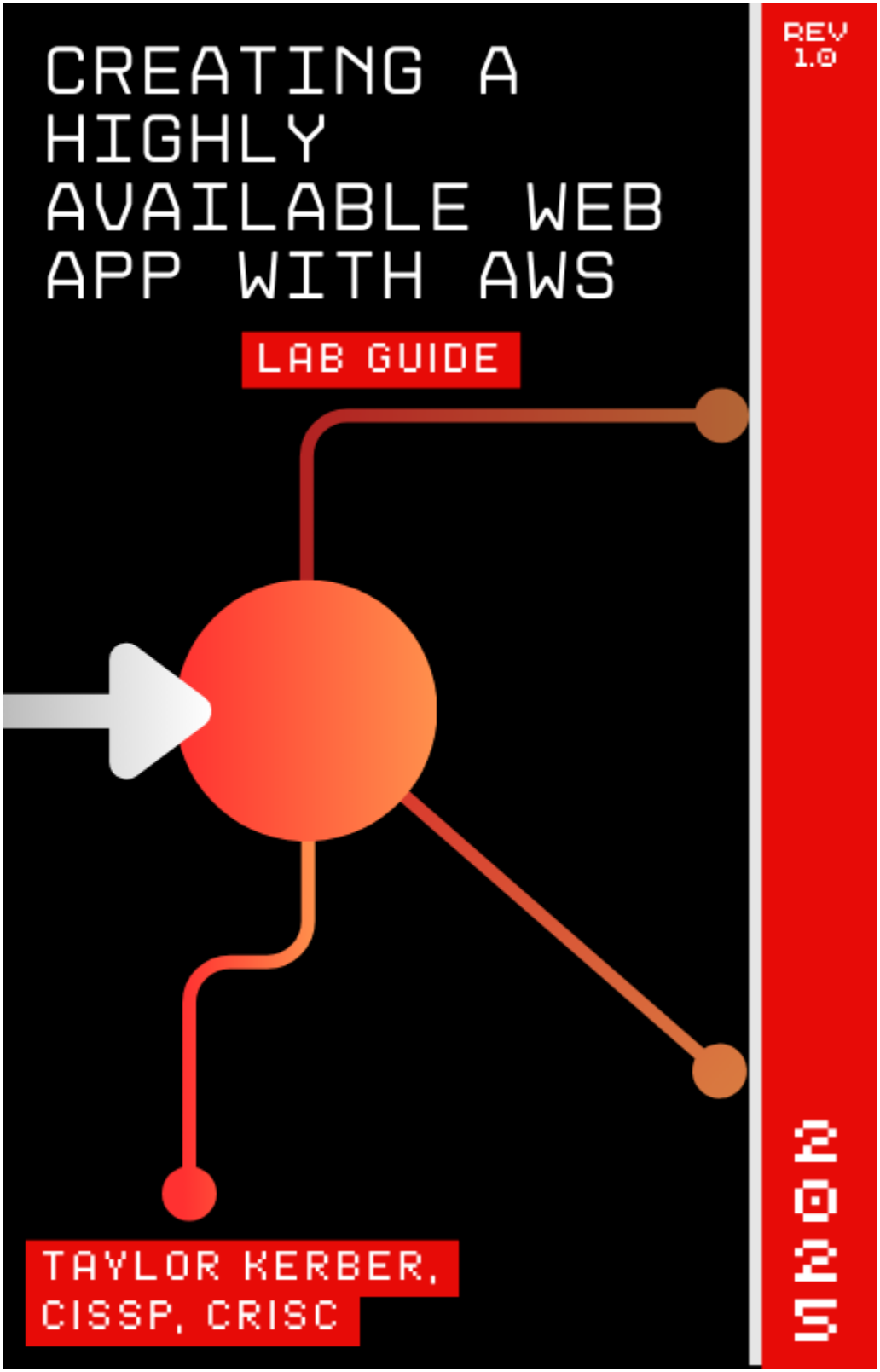

CREATING A HIGHLY AVAILABLE WEB APP WITH AWS
LAB GUIDE
REV 1.0
TAYLOR KERBER, CISSP, CRISC
2024

AWS Web App with Load Balancing Lab

Objective

This lab will demonstrate how to create a basic web application with two EC2 instances in different availability zones, behind a load balancer. Upon completion of this lab, you should have a better understanding of how to navigate the AWS console,

Pre-requisites

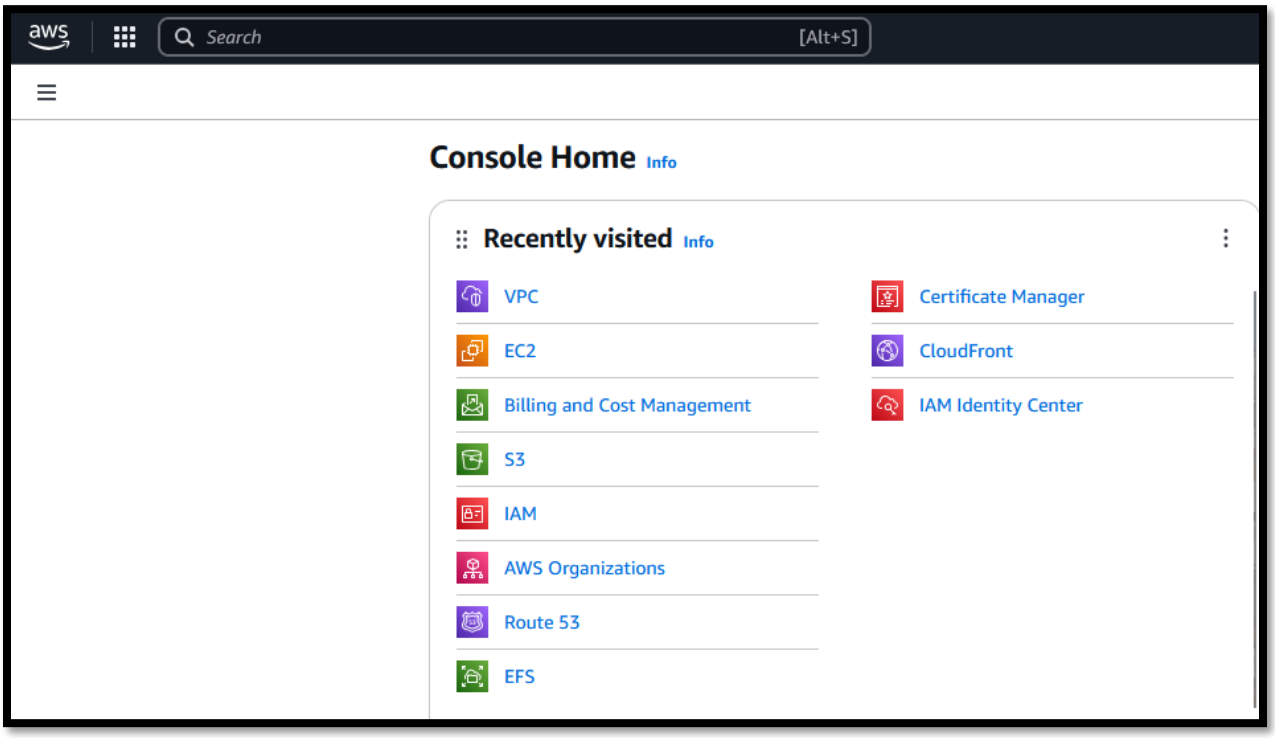
- A personal or lab AWS account (This lab does not walk-through how-to set up an AWS account.)
- Access to an internet / WAN connection to use the AWS console.
- A modern web browser like: Chrome, Firefox, Edge, etc.
- Windows: Install the application PuTTY and PuTTYgen.
- Linux: Able to use SSH over the web.

Table of Content

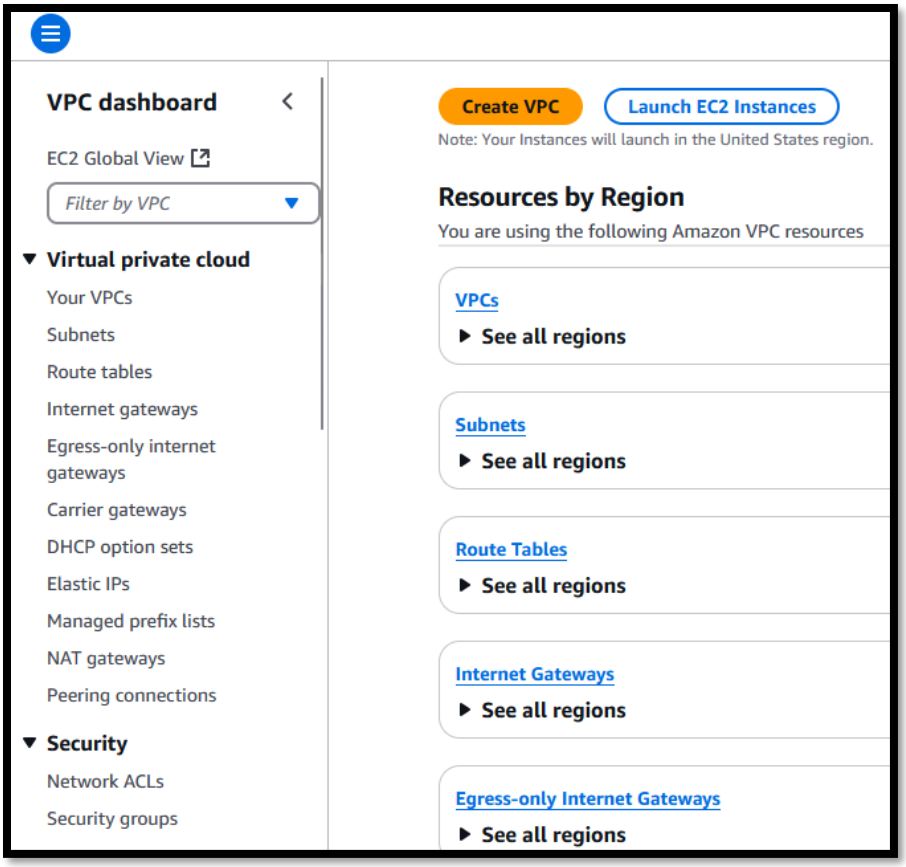
VPC Creation

In this section we will create our VPC, subnets, internet gateway, and routing.

1) Log into the AWS console and on the main page search for VPC or click on it (if you recently used it).

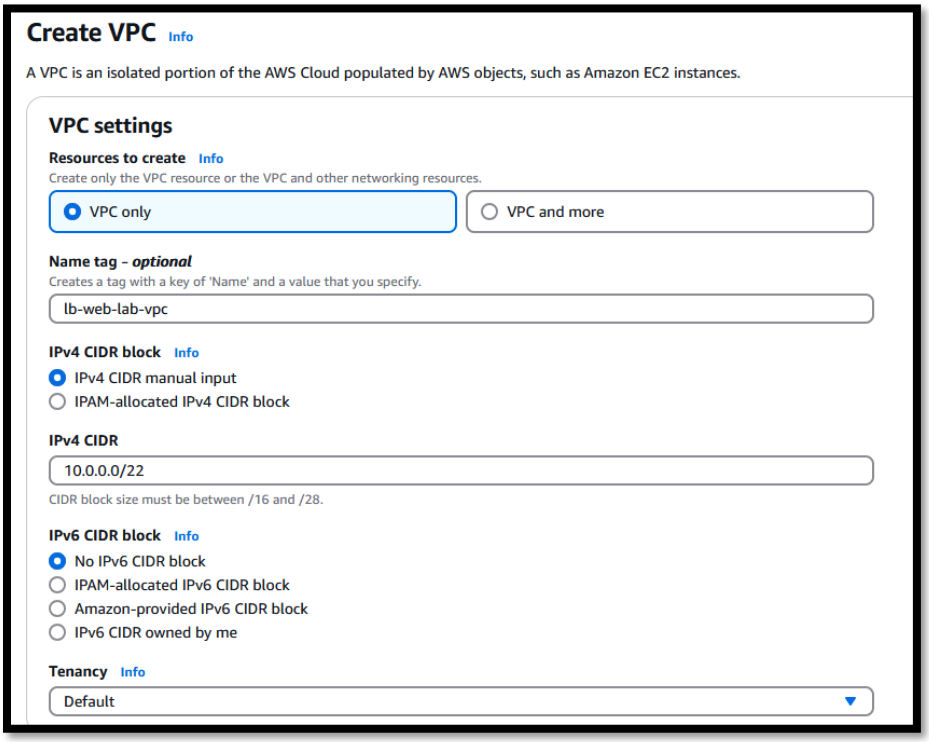


2) Once the VPC dashboard is open you will see a menu on the left-hand side. Click on Your VPCs and in the top right-hand side click the create VPC button. 



3) Fill out a name and choose a CIDR block. In my example I'm choosing a /22.

4) When you're done click the Create VPC button. 

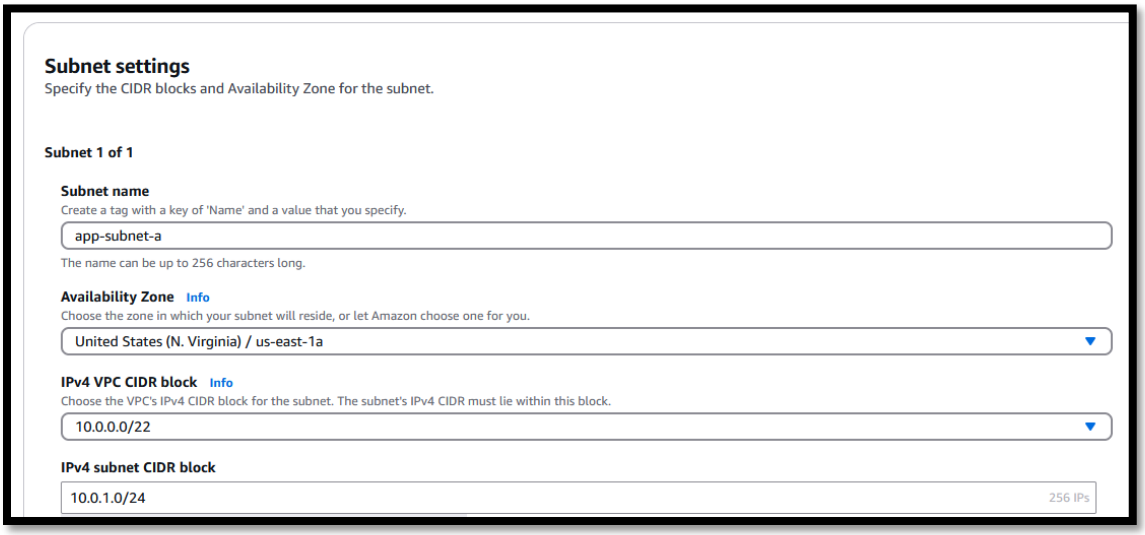


Subnets

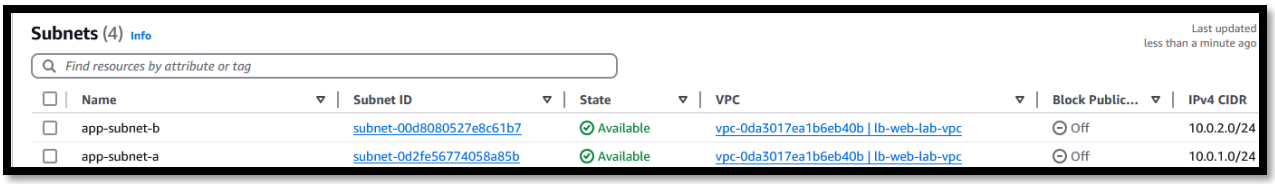
1) On the VPC dashboard you will look on the left-hand side again and click on Subnets. The Subnets dashboard will display, and you will click the Create Subnet button.

2) I've chosen to carve out (2) /24 networks of my /22 VPC. This will give me subnets with 256 Ips each. Remember AWS reserves (5) IP addresses from each subnet CIDR for it's own use.

3) I also chose the 'A' availability zone and have named my subnet accordingly.



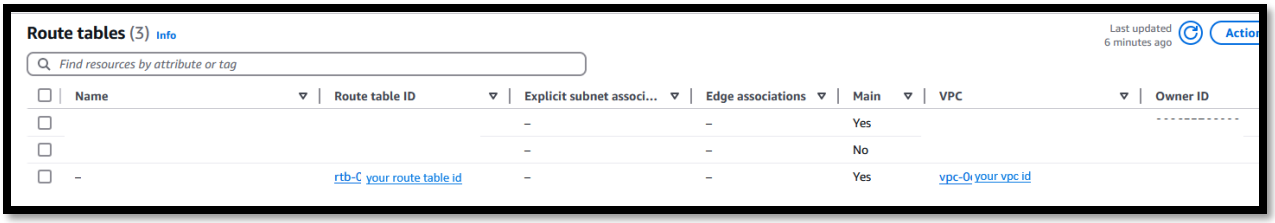
Repeat Steps 1-3 for the other subnet. When you're finished it should appear similar to the screenshot below:



Routes

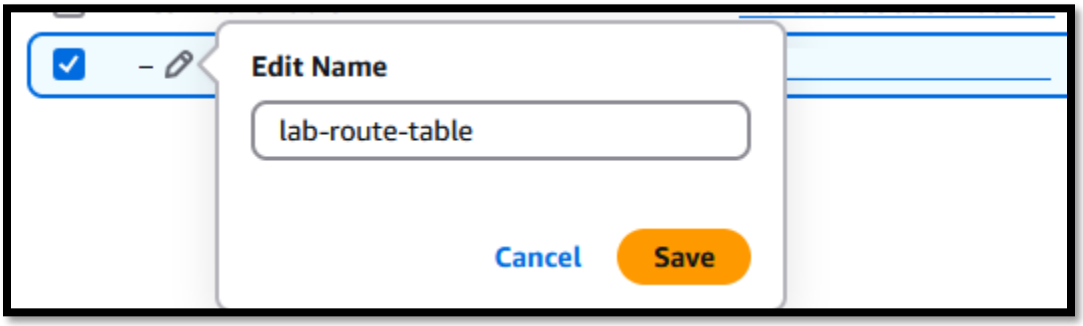
In AWS when you create a VPC, a route table is created by default. When we created our VPC earlier, unbeknownst to us, it created a route table automatically. We are going to name the route table, create an internet gateway, and create a default route so our EC2 instances can get to the internet.

1) On the VPC dashboard click on Route tables. **Route tables**



*For brevity I've removed information on some of my other routes. You should see just (1) route without a name. This route will have the VPC ID that matches your VPC and it should be a **Main** route table.*

2) Click on the route and there should be an edit icon under the blank name. Rename it and click Save.



3) Next, click on routes at the bottom and you'll notice it automatically has a route for your entire /22 VPC. Your route table does not however, have a default route to the internet.

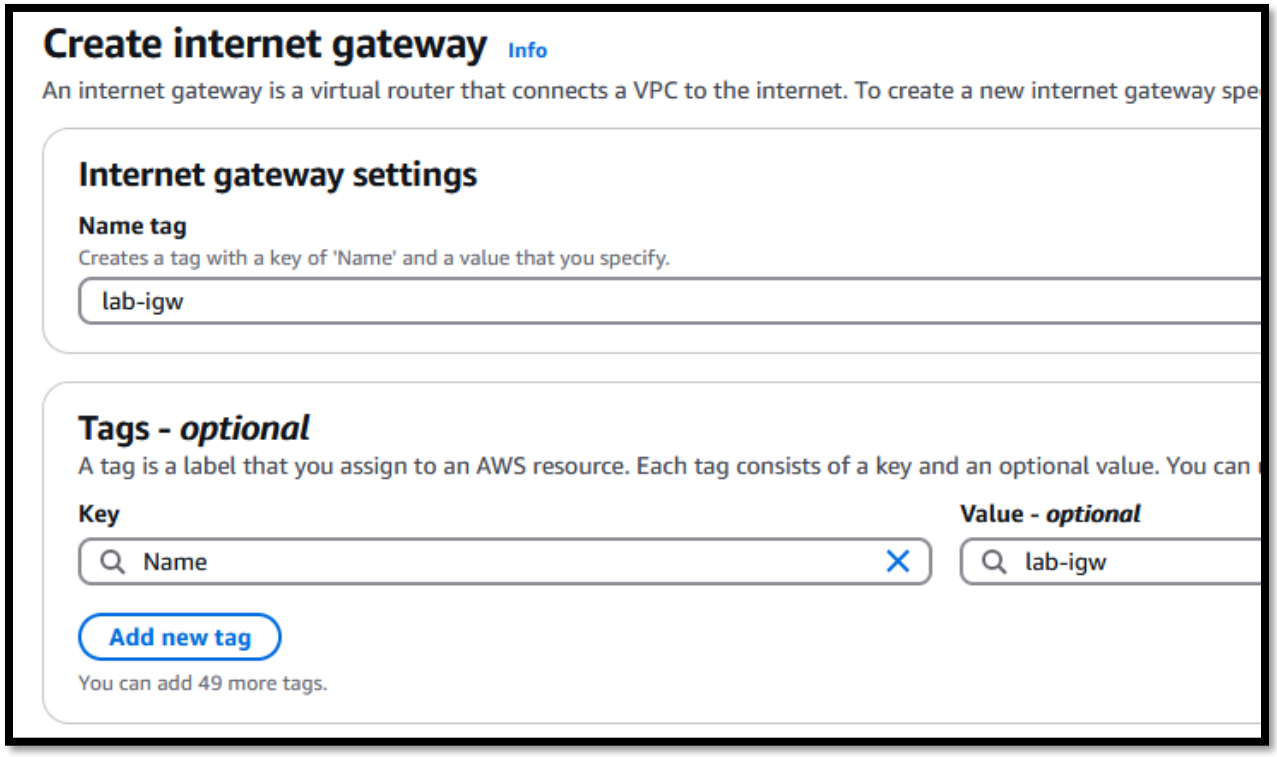
Internet Gateway

Now we've named our routing table and verified intra-VPC routing, we need to create an internet gateway. An internet gateway allows inbound and outbound connectivity from our VPC to the internet.

1) On the VPC dashboard we will click on Internet gateways. Internet gateways

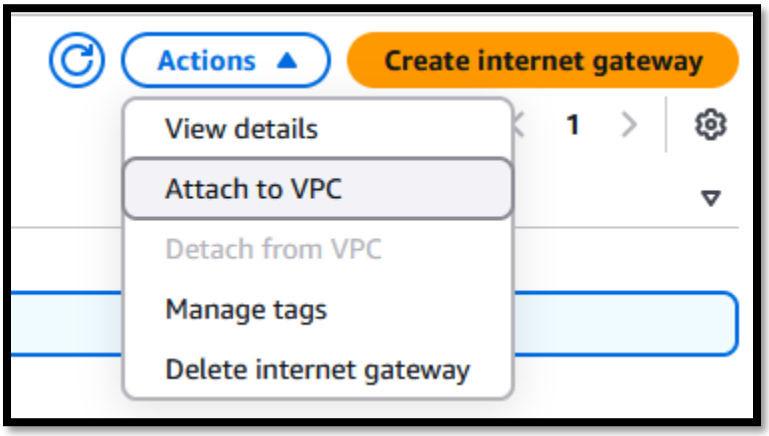
2) Next you will click Create internet gateway. Create internet gateway

3) All you must do is provide a name. For me, I am going to abbreviate internet gateway with igw. It will keep my name shorter, and it is an industry recognized abbreviation.



4) When you are finished click Create internet gateway.

5) Lastly, the igw may take a few minutes to complete and you will want to check the state. It will most likely be detached and that is expected behavior. All you have left to do is attach it to your VPC.



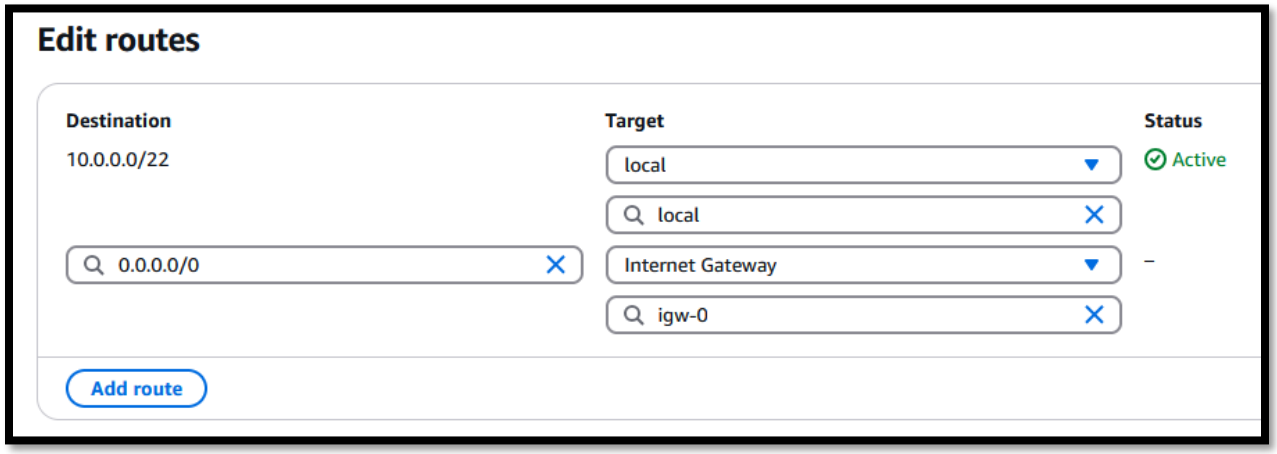
Default Route

Now you've created a VPC, two subnets, and an internet gateway attached to your VPC. All that is left is to create default route pointing to your internet gateway.

1) Go back to Route tables on the VPC dashboard and open your route table you named earlier.

2) In the route table settings click on the Routes button **Routes** and next click the Edit routes button. **Edit routes**

3) Click on Add route and create your route so it looks like the screenshot below:



Recap

Before moving on to the next section ensure you've created and configured the following:

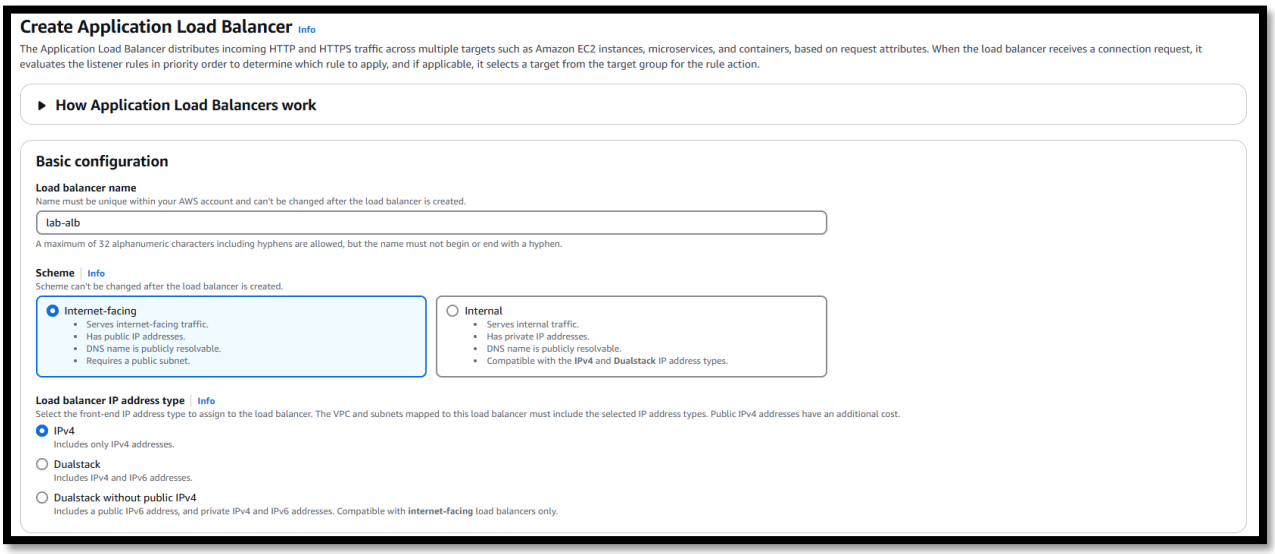
- ✓ A VPC.
- ✓ (2) Subnets, one in each availability zone.
- ✓ A default route table.
- ✓ An internet gateway attached to your VPC.
- ✓ A default 0.0.0.0/0 route pointing to the internet gateway.

Server and Load Balancer Creation

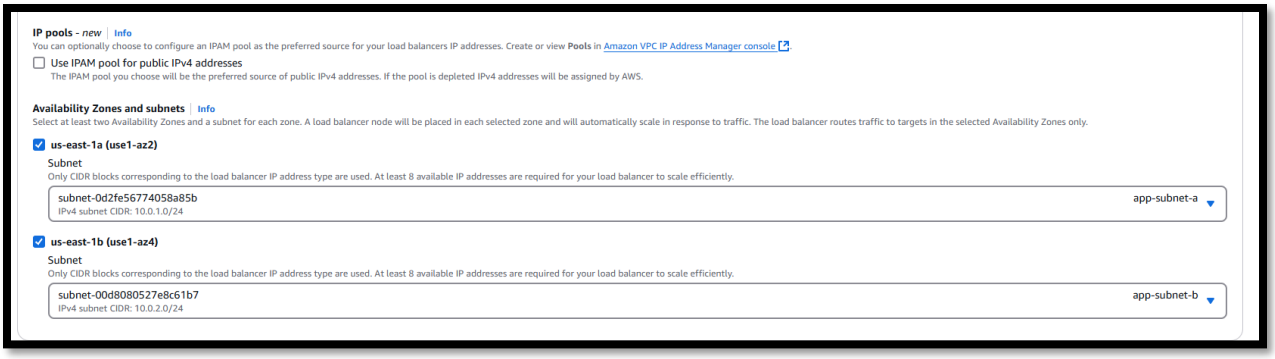
In this section of the lab, we will create our Application Load Balancer (ALB), then create two instances serving a simple HTTP website.

1) Under EC2 you will go to Load balancers and click create an Application Load Balancer.

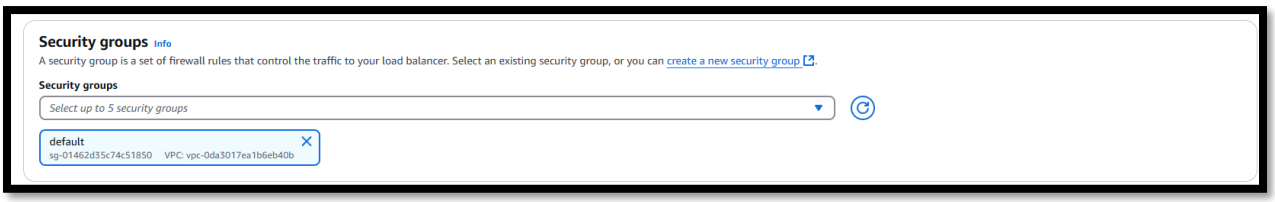
2) You will want to configure this as IPv4 only and it will be an internet-facing scheme.



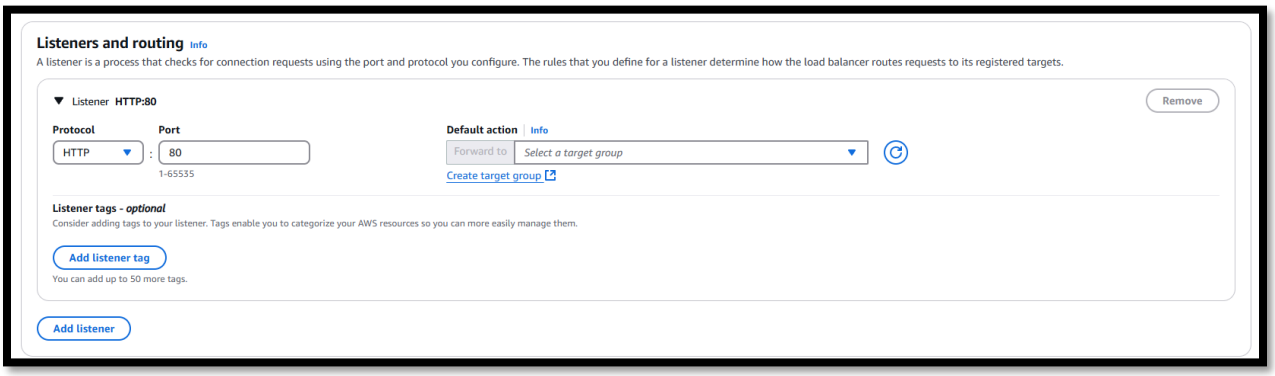
3) Next, ensure you select both availability zones you created earlier.



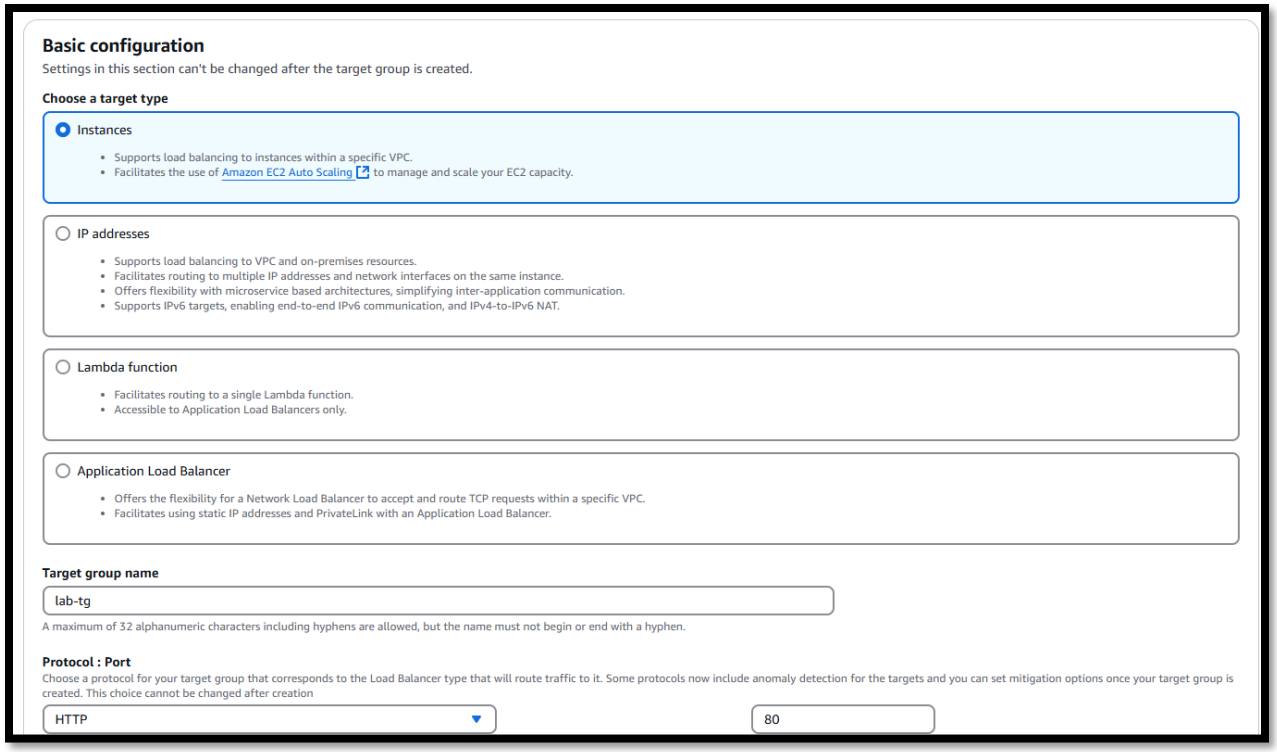
4) Leave the Security group settings default for now and continue.

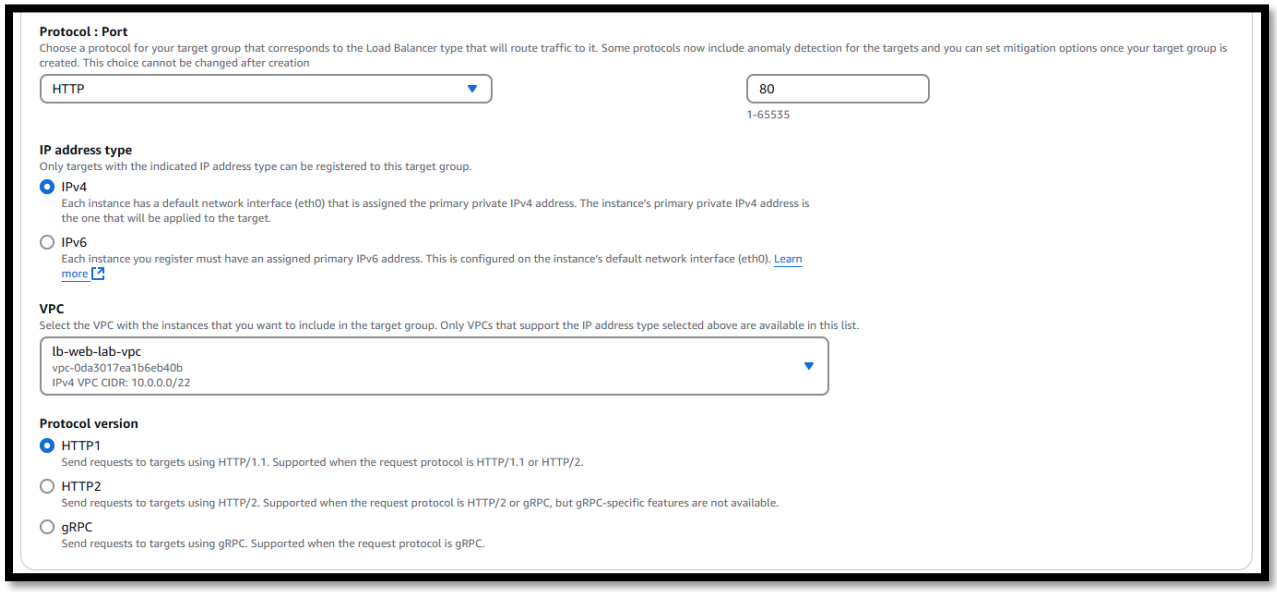


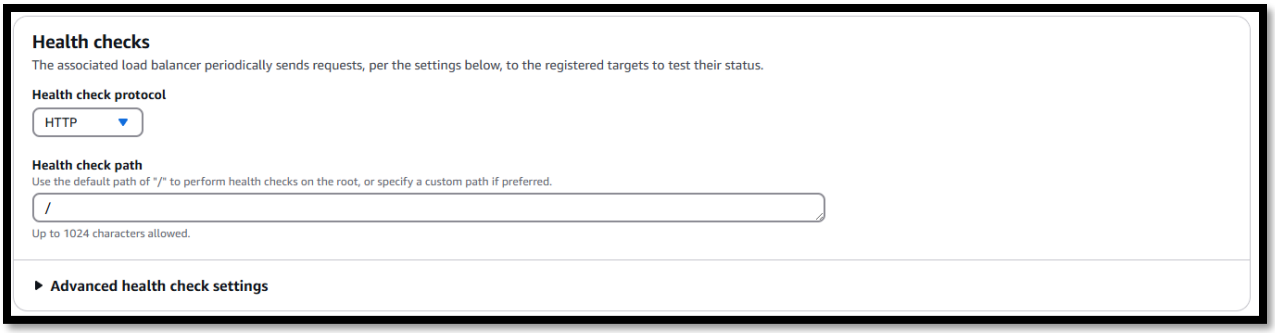
5) Next go ahead and click the Create target group button. Even though we haven't created our EC2 instances yet, we are going to create this group.



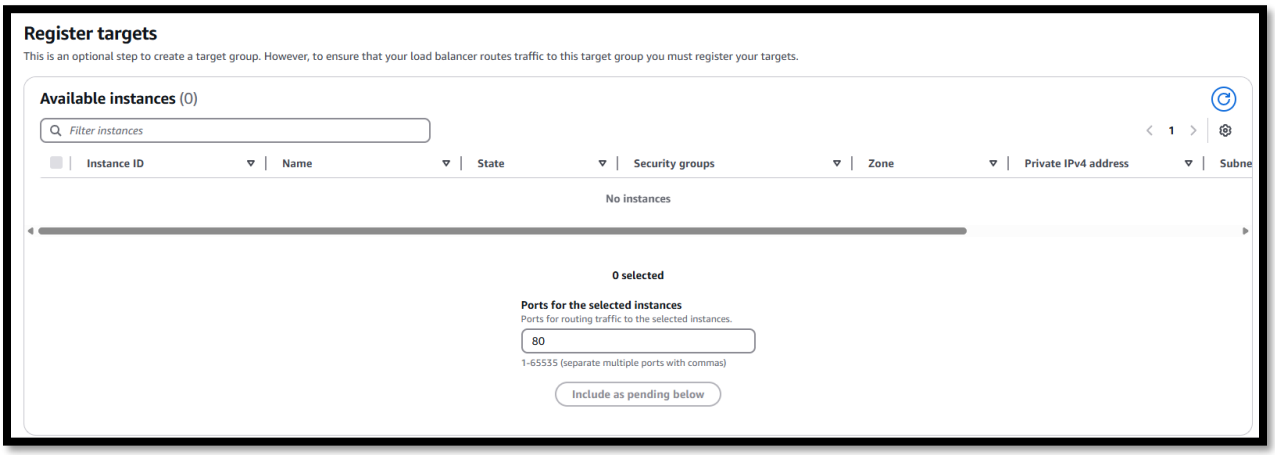
6) Give your target group a name and leave almost everything else default.


Basic configuration
Settings in this section can't be changed after the target group is created.
Choose a target type
Instances
Supports load balancing to instances within a specific VPC.
Facilitates the use of Amazon EC2 Auto Scaling to manage and scale your EC2 capacity.
IP addresses
Supports load balancing to VPC and on-premises resources.
Facilitates routing to multiple IP addresses and network interfaces on the same instance.
Offers flexibility with microservice based architectures, simplifying inter-application communication.
Supports IPv6 targets, enabling end-to-end IPv6 communication, and IPv4-to-IPv6 NAT.
Lambda function
Facilitates routing to a single Lambda function.
Accessible to Application Load Balancers only.
Application Load Balancer
Offers the flexibility for a Network Load Balancer to accept and route TCP requests within a specific VPC.
Facilitates using static IP addresses and PrivateLink with an Application Load Balancer.
Target group name
lab-tg
A maximum of 32 alphanumeric characters including hyphens are allowed, but the name must not begin or end with a hyphen.
Protocol : Port
Choose a protocol for your target group that corresponds to the Load Balancer type that will route traffic to it. Some protocols now include anomaly detection for the targets and you can set mitigation options once your target group is created. This choice cannot be changed after creation
HTTP
80

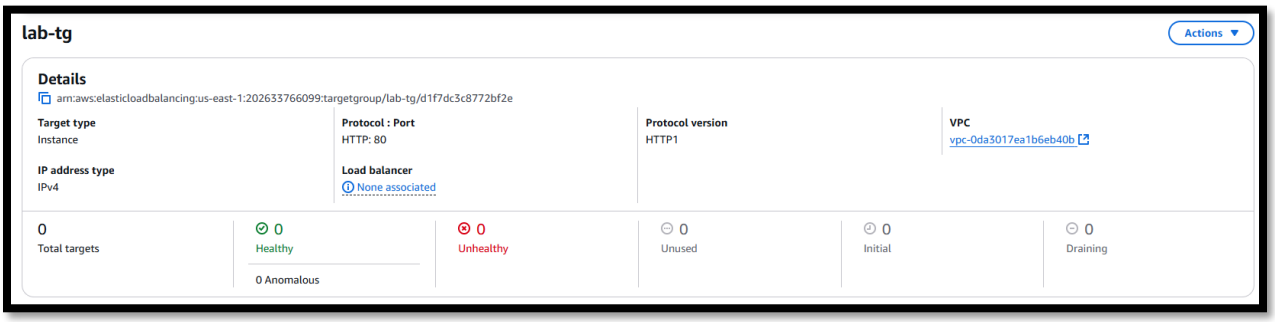

Protocol : Port
Choose a protocol for your target group that corresponds to the Load Balancer type that will route traffic to it. Some protocols now include anomaly detection for the targets and you can set mitigation options once your target group is created. This choice cannot be changed after creation
HTTP
80
1-65535
IP address type
Only targets with the indicated IP address type can be registered to this target group.
IPv4
Each instance has a default network interface (eth0) that is assigned the primary private IPv4 address. The instance's primary private IPv4 address is the one that will be applied to the target.
IPv6
Each instance you register must have an assigned primary IPv6 address. This is configured on the instance's default network interface (eth0). Learn more
VPC
Select the VPC with the instances that you want to include in the target group. Only VPCs that support the IP address type selected above are available in this list.
lb-web-lab-vpc
vpc-0da3017ea1b6eb40b
IPv4 VPC CIDR: 10.0.0.0/22
Protocol version
HTTP1
Send requests to targets using HTTP/1.1. Supported when the request protocol is HTTP/1.1 or HTTP/2.
HTTP2
Send requests to targets using HTTP/2. Supported when the request protocol is HTTP/2 or gRPC, but gRPC-specific features are not available.
gRPC
Send requests to targets using gRPC. Supported when the request protocol is gRPC.



4) On the next page you will be asked to select targets. You have none right now and we will come back to add our instances once we create them. Continue to create the target group.



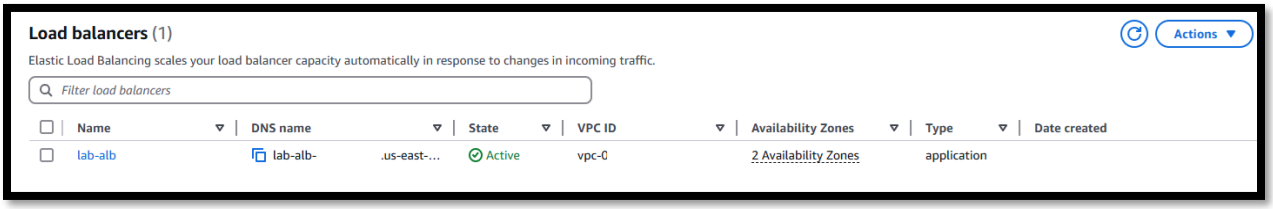
5) Your target group is now created and empty.

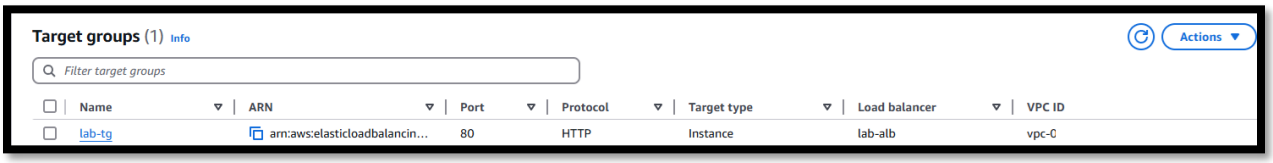


6) Go back to your "Create application load balancer" tab and click the Create load balancer button. 

7) With your new target group selected you should be able to create this.

8) When finished you should have a load balancer and target group configured similarly to what is seen below:





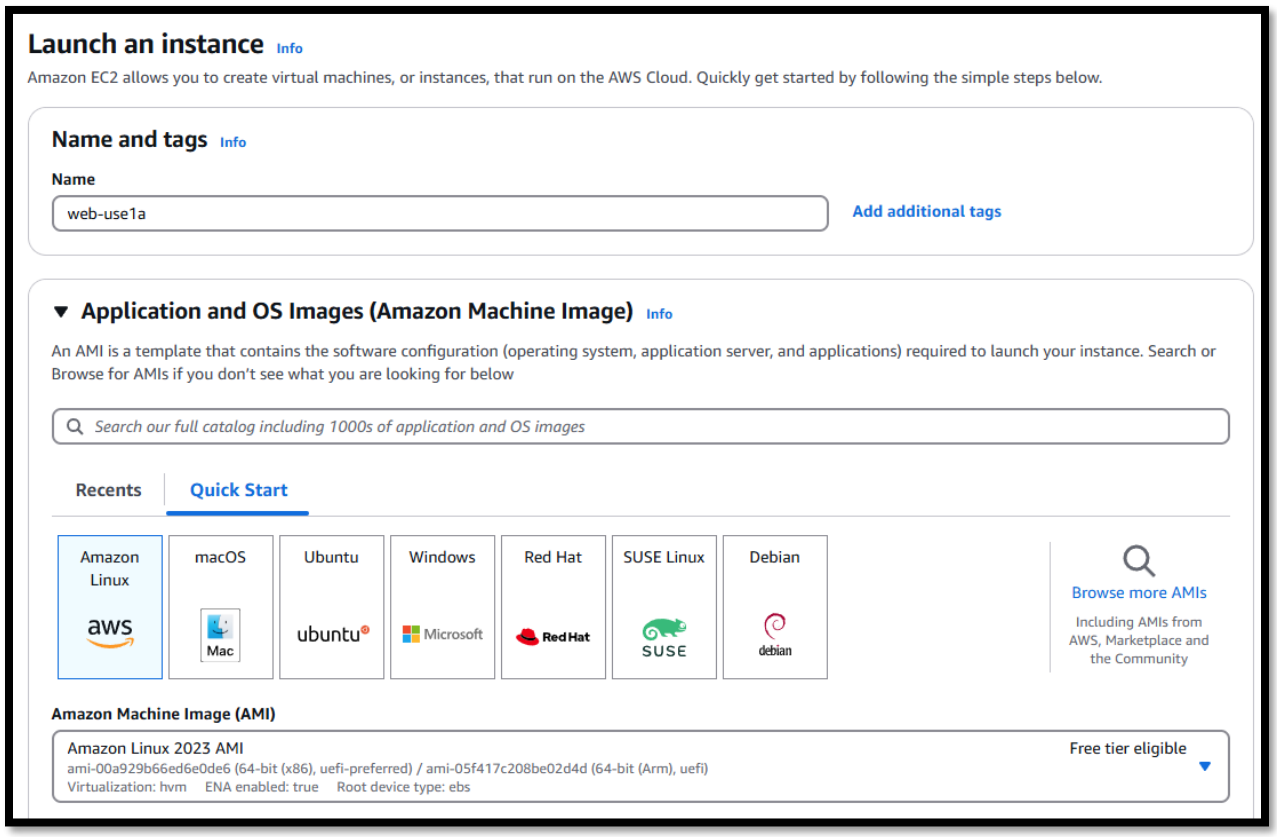
EC2 Instances

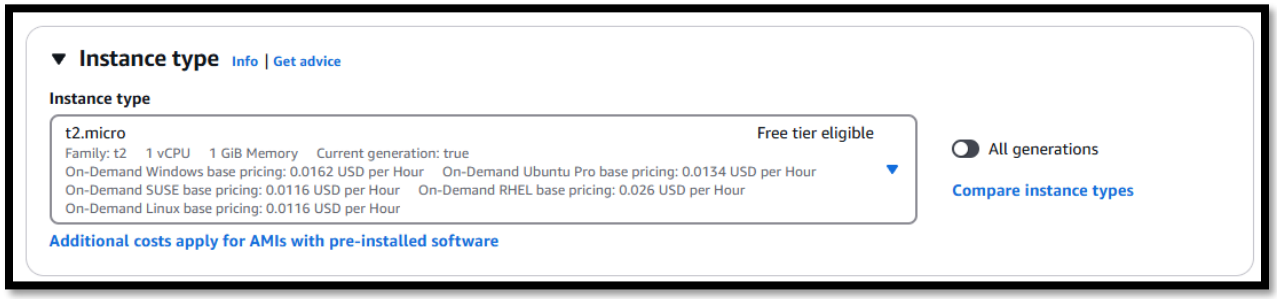
In this section we will create our two web servers and configure them as basic HTTP servers. I will keep everything free-tier so you can follow along.

1) Go to EC2 and Launch an instance.

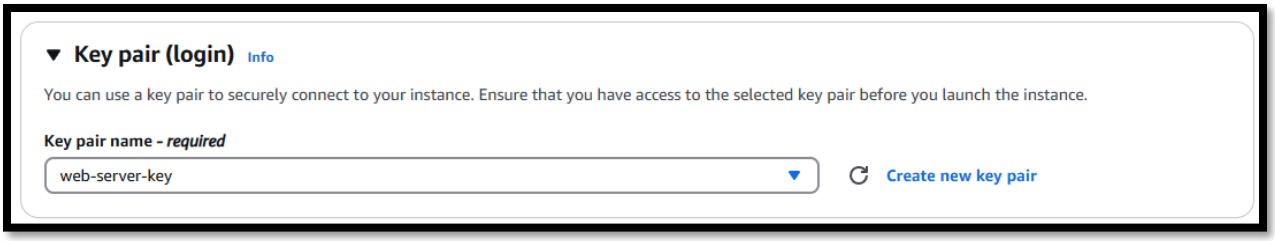
a. I am going to use Amazon Linux and keep most settings default.

b. Give it an appropriate name.





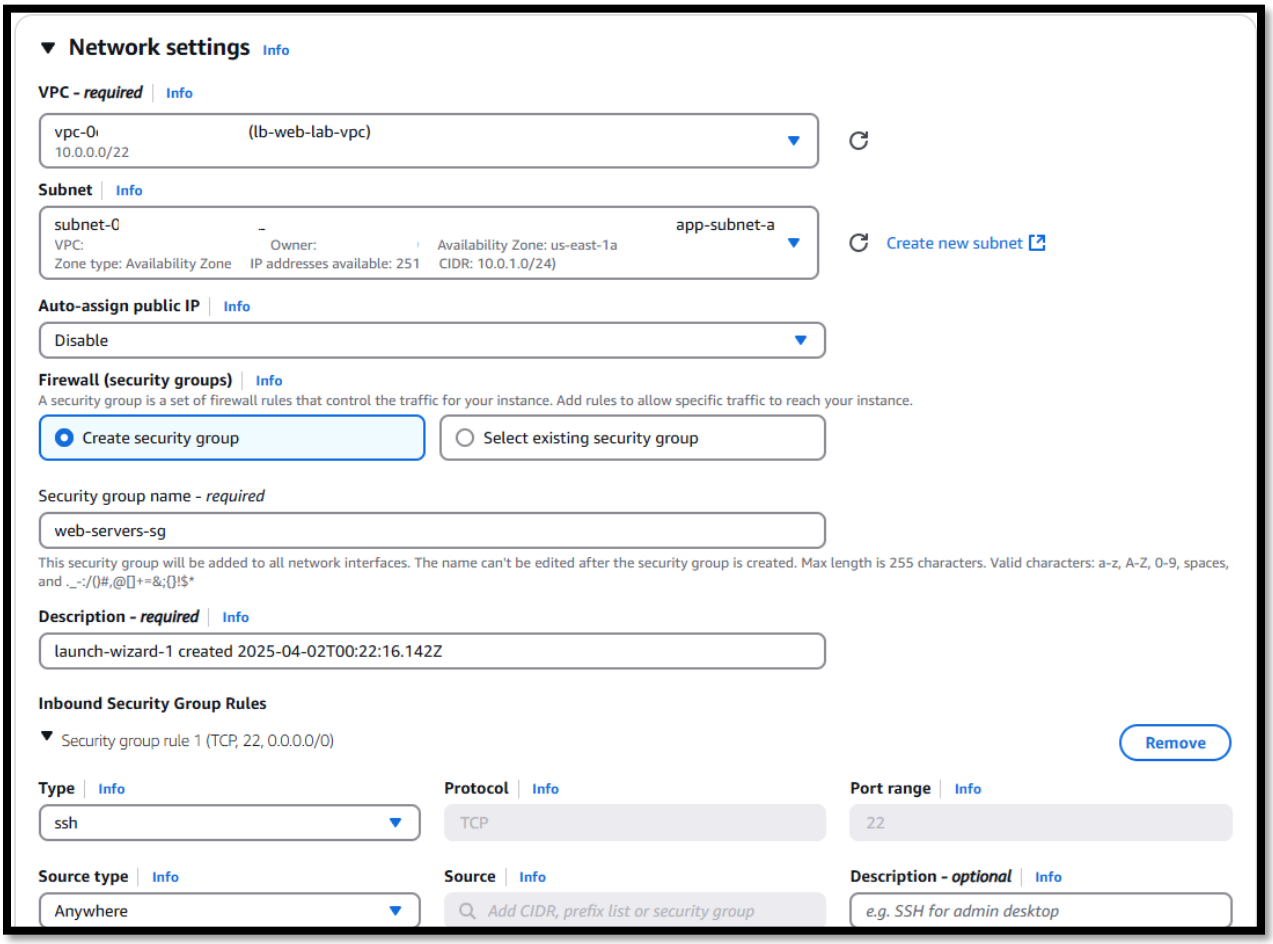
2) Create a new SSH key pair or use an existing one you've already created.



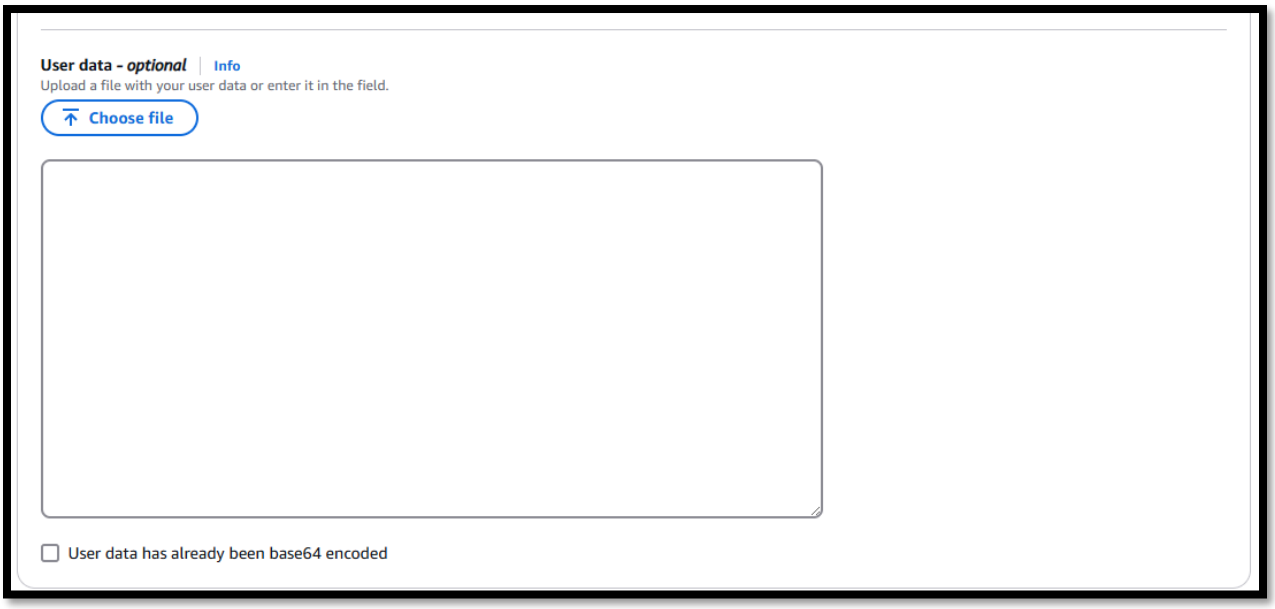
3) Ensure you have the correct VPC and subnet selected.

4) Use the default security group and allow SSH for now from anywhere. Give it a different name to identify it by.

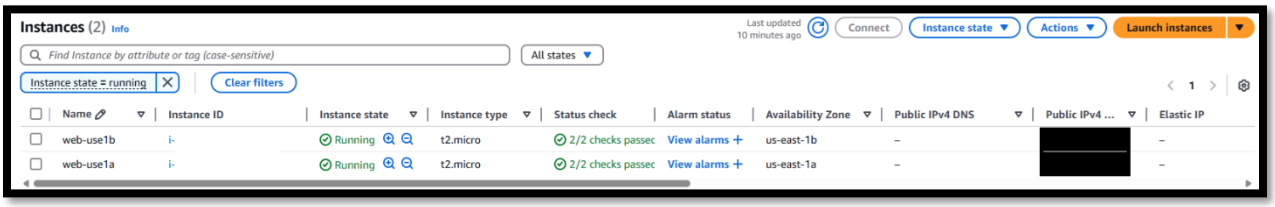
5) Give your server a public IP by changing Auto-assign public IP to enable.



6) I want you to go to Advanced options and look for user data. This is data you can pass directly to the instance when it is launched. This allows for quick bootstrapping and automation. We are not going to use this in this lab because we're going to configure our servers manually. I want to point this out because it's something to note of importance. All major cloud providers offer this.



7) Refresh your EC2 dashboard and you should have two public instances with public Ips.



Recap

Before moving on to the next section ensure you've created and configured the following:

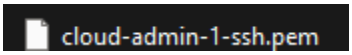
- ✓ A load balancer.
- ✓ An empty target group.
- ✓ (2) EC2 free-tier instances running Amazon Linux, one in each availability zone.
- ✓ (2) public IPs, one attached to each instance created.

Connecting to EC2 Instances

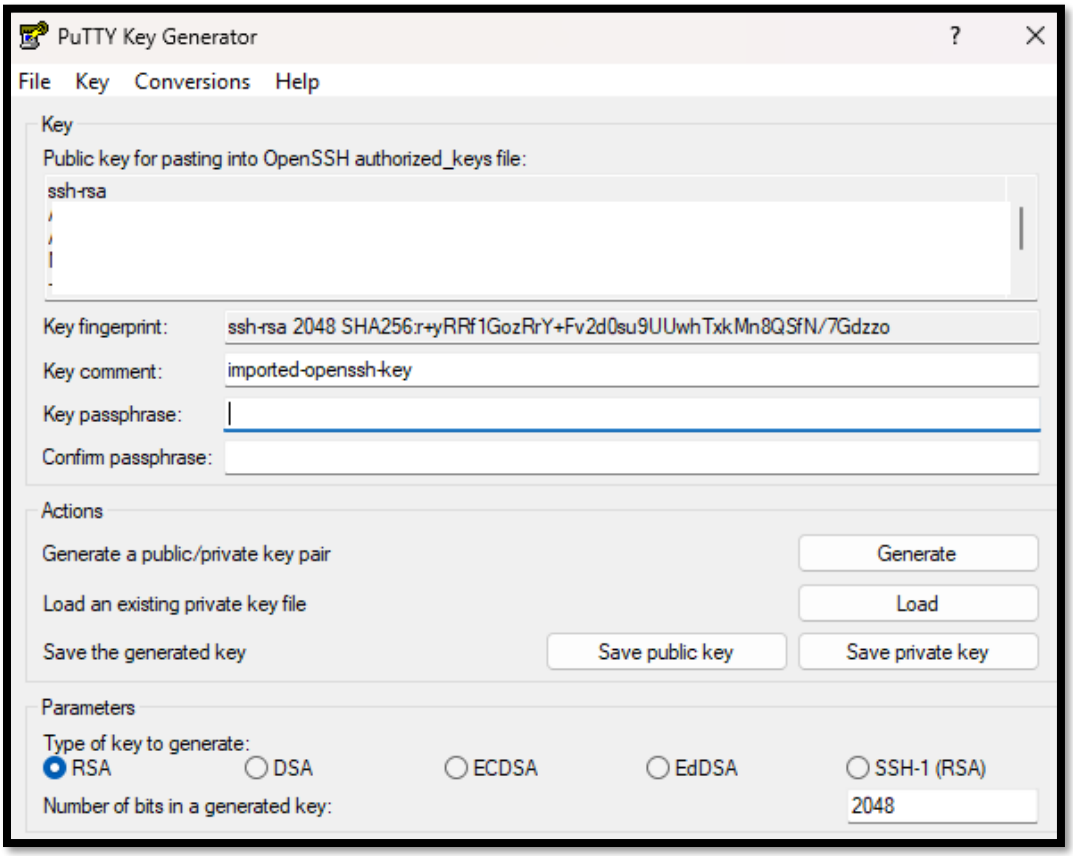
This section will walk-through how to connect to our EC2 instances initially and how to setup a simple HTTP server on our EC2 instances. Then we will finalize our load balancer config and test it all out.

PuTTY and PuTTYgen

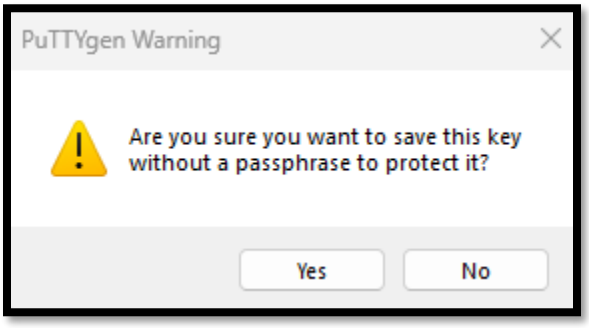
Before we get started you will want PuTTY for Windows and PuTTYgen (comes with PuTTY) to create a .ppk file from our .pem key we downloaded earlier when we created our EC2 instances. My private key is still in my downloads folder as seen below.



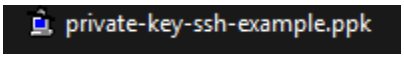
1) First, open PuTTYgen and click File, then Load private key.



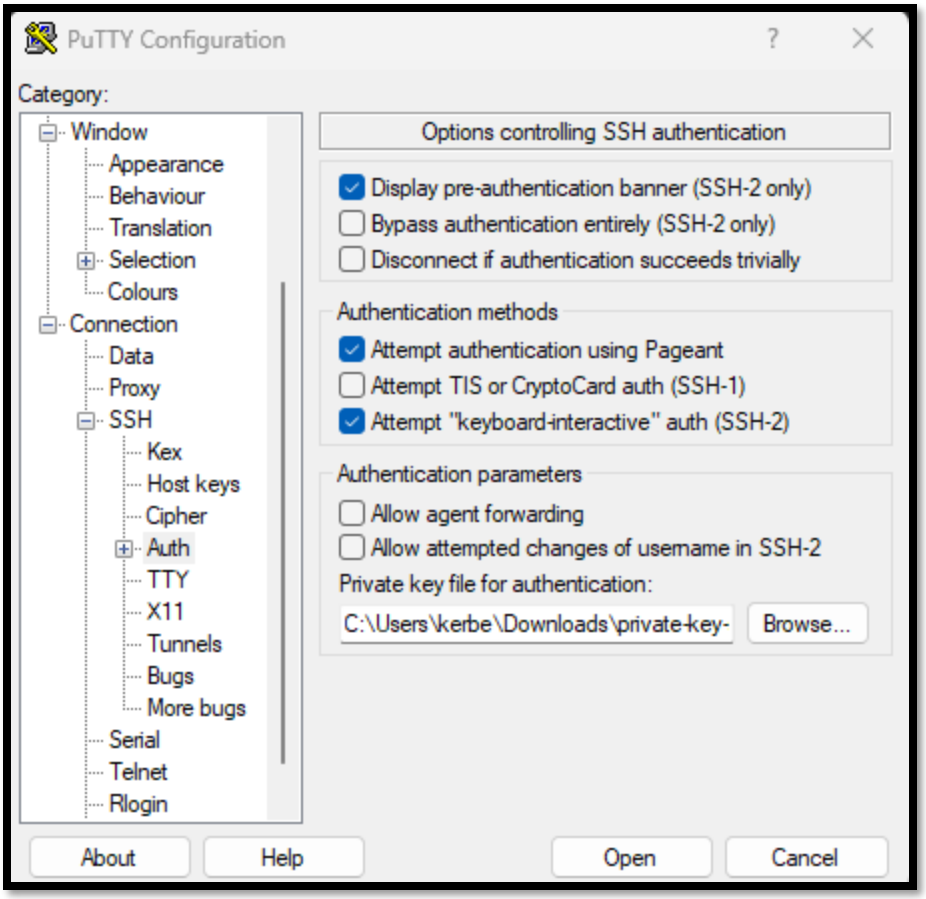
2) Next, I am going to just hit save private key and I will get a warning, and I hit yes. I don't care about passwords because this is a lab. Obviously, in the real world you would want to utilize a passphrase.



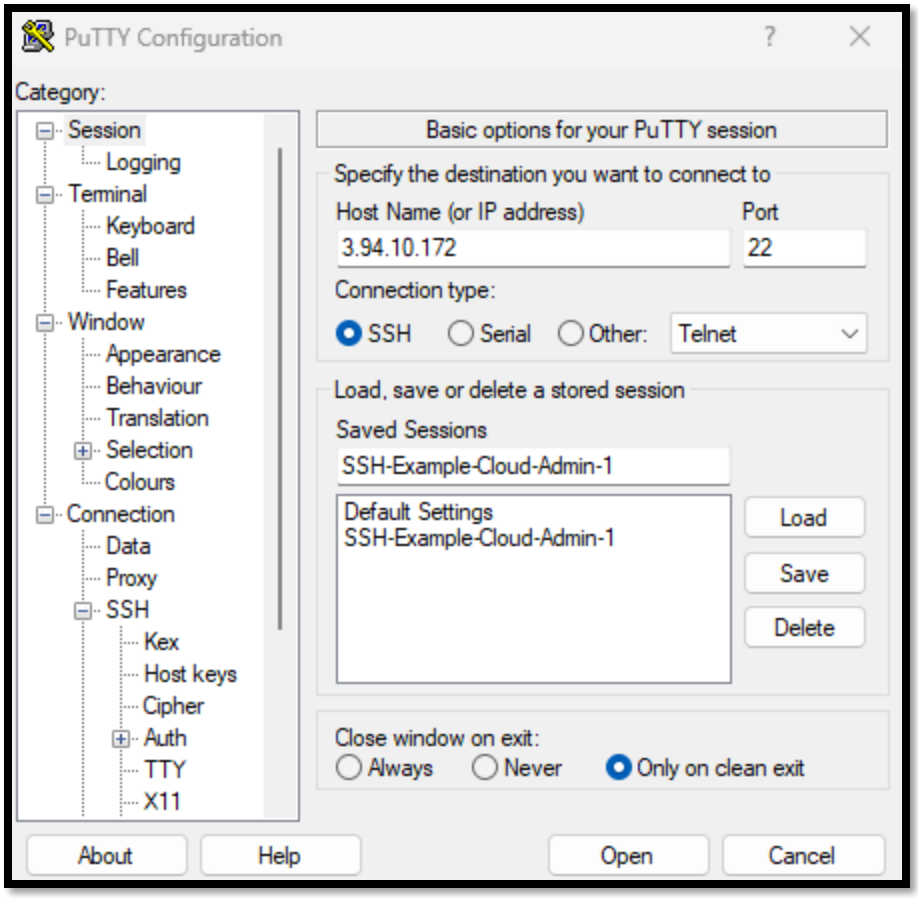
3) I save this as a .ppk and now I have my key to connect to my instance.



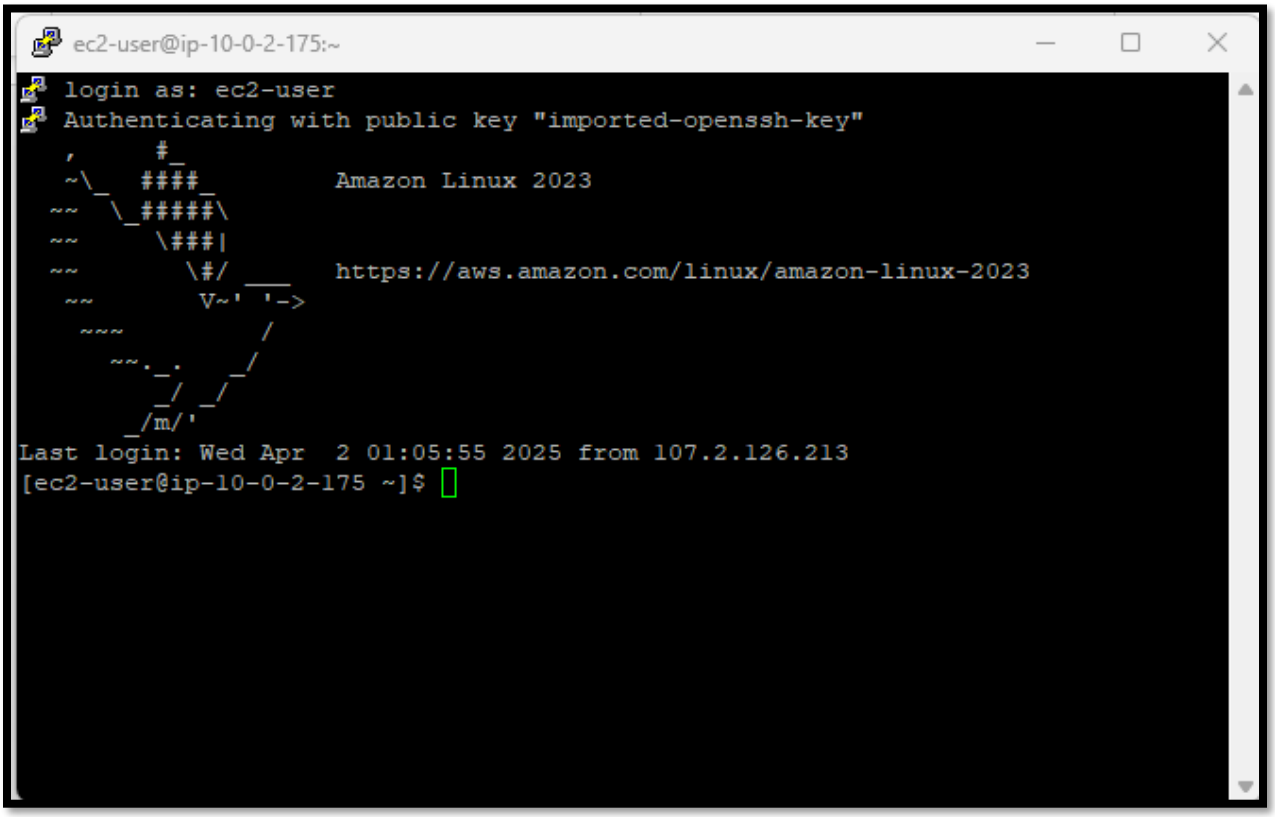
4) Now we will open PuTTY and drill to Connection, SSH, Auth, Private key file for authentication:



5) Provide your newly created .ppk then go back to Session and add your IP address and click SSH. I also recommend saving these settings so you don't have to configure this every time. I saved mine as "SSH-Example-Cloud-Admin-1"



6) You will now be asked for a user, and you will type **ec2-user** exactly as you see it here.



You are now successfully connected to your EC2 instances. Ensure you can connect to both!

Recap

Before moving on to the next section ensure you've created and configured the following:

- ✓ You can successfully connect and log into both EC2 instances
- ✓ Both EC2 instances can successfully ping 8.8.8.8 or any external IP

Configure Web Servers

Now that you can log into both servers, we need to configure them to be actual servers. For this lab we are going to use python and a useful library for python called simple http server. You will create a simple configuration on both machines with only one difference, the names on the html pages.

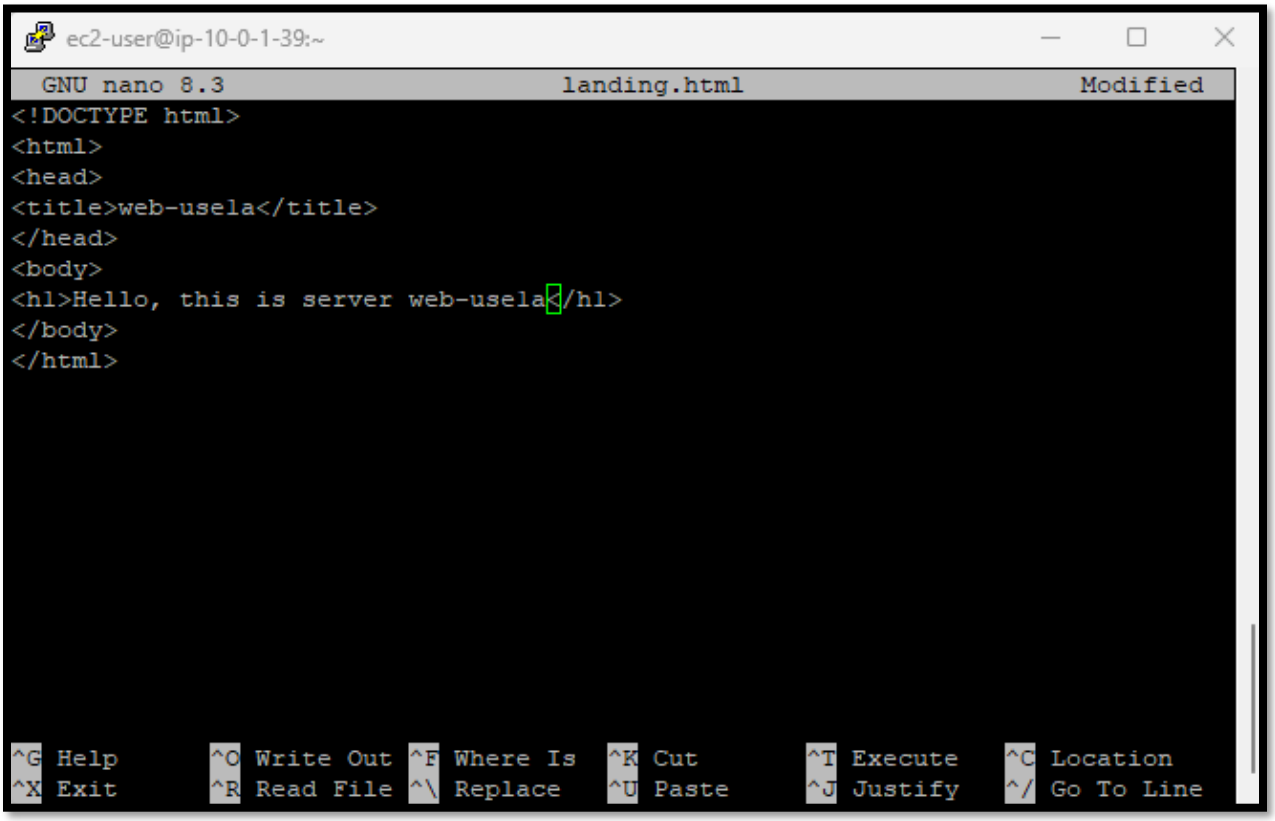
1) Run the following commands on each EC2 instance.

```
sudo yum install python3 -y
sudo yum install python3-pip -y
python3 -m pip install simple_http_server
```

2) In your user directly create a basic html file and call it landing.html. You can use nano or vim (I prefer nano). You can do this by running the following command.

```
nano landing.html
```

3) Copy and paste the following HTML code and make sure you modify the name for which server page you're creating. When finished hit CTRL + O key to save then CTRL + X



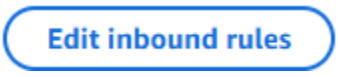
4) Lastly, run the following command from the directory your file is in (it should be your user directory).

```
sudo python3 -m http.server 80
```

5) You should now see a running web server service in the CLI.

6) To test this we need to allow HTTP 80- from anywhere in the world. Go to VPC then Security groups in the AWS consol.

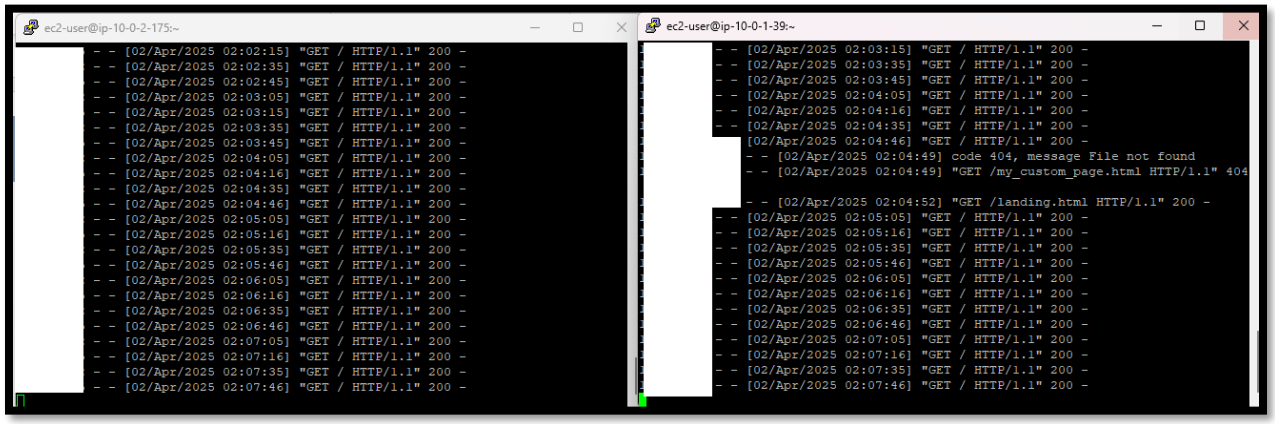
7) Find your security group from earlier and click the Edit inbound rules button.



8) Create a new policy to allow TCP 80 from anywhere inbound.

9) Open your web browser and put the public IP address of your EC2 instance in, starting with http://x.x.x.x/landing.html. You should now see both pages displayed! Congratulations, you've got two successful HTTP servers.

10) What is also great about using python simple http server is that it also allows you to confirm your traffic. If you look at the CLI of your PuTTY sessions, you'll see the GET Requests as they come in.



Recap

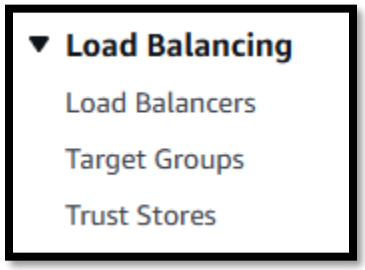
This has been a monster of a lab, but we are not done yet. To move to the final part of this lab you need to ensure the following:

- ✓ You have two EC2 instances running python3 and simple http server
- ✓ Both EC2 instances are successfully serving http server pages

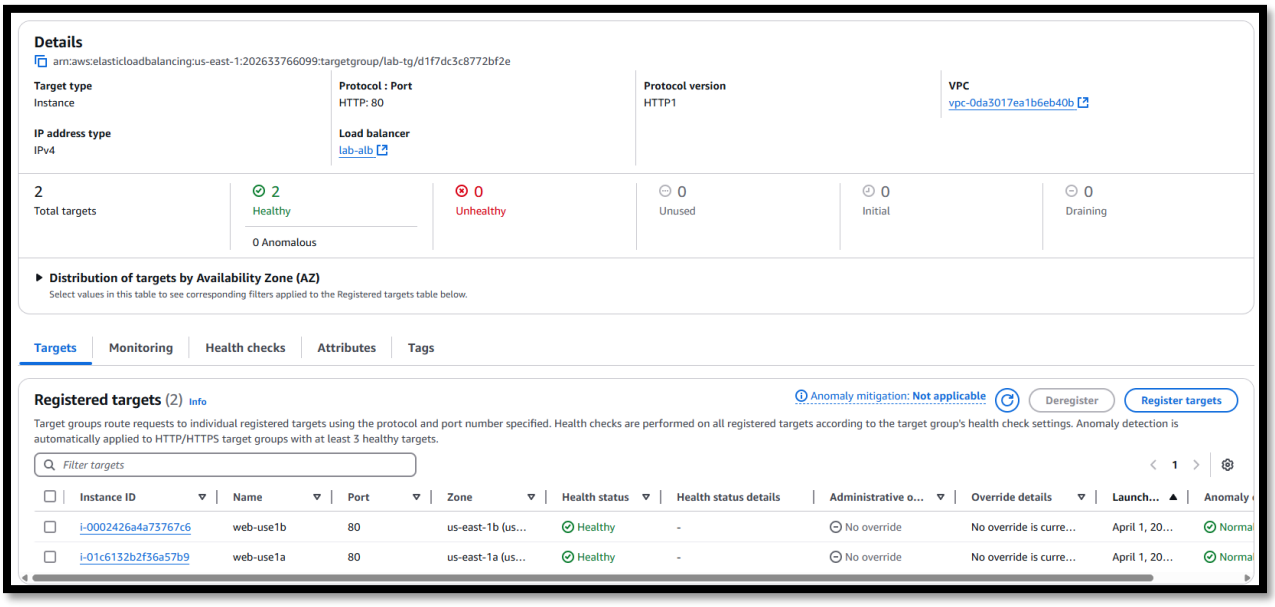
Finalize and Test Deployment

We've completed almost every single piece of this lab, and it has been done in small phases. We initially created our load balancer and target pool but left it empty. Now that we have confirmed our web servers are behaving accordingly, we can go back and finish the load balancer config.

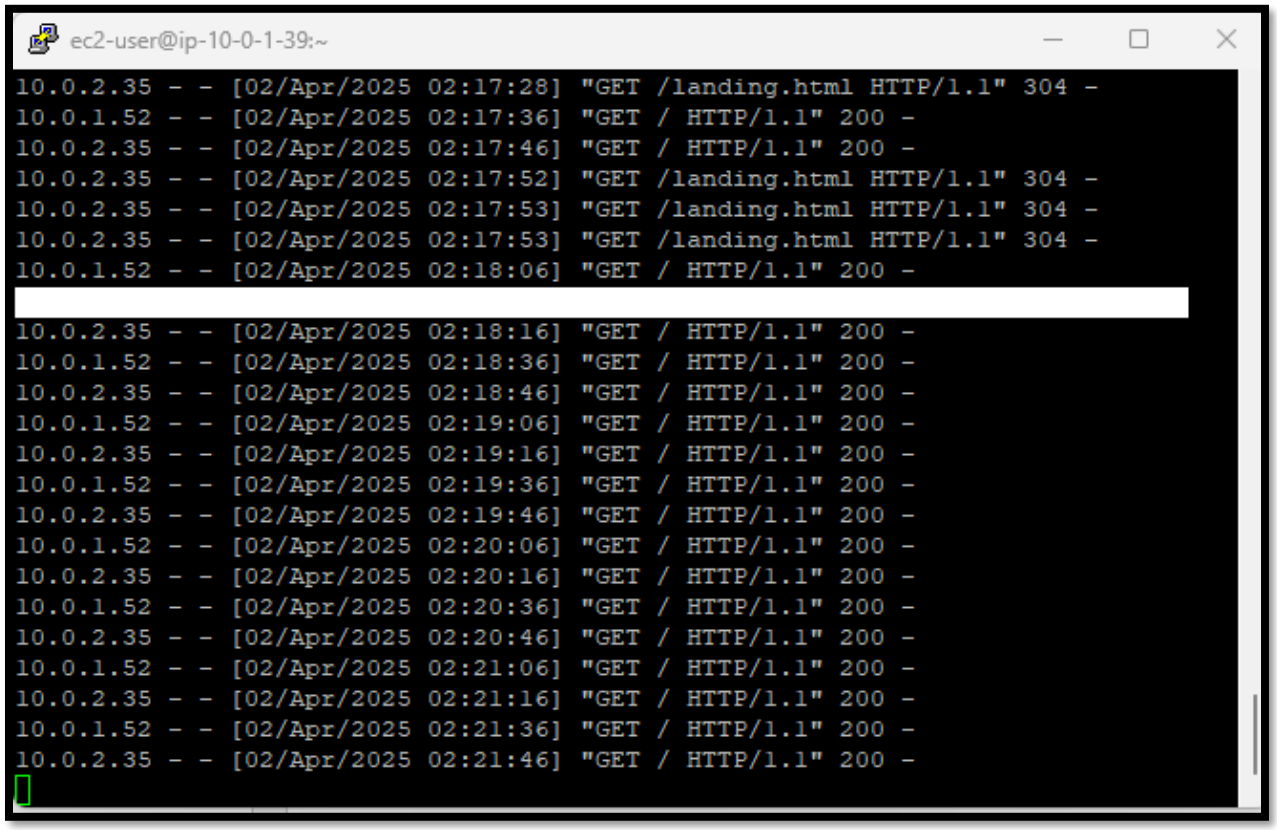
1) First, we need to update our target group with our new healthy web servers. We can do this by going to EC2 on the AWS console, then Load Balancing, and finally Target Groups



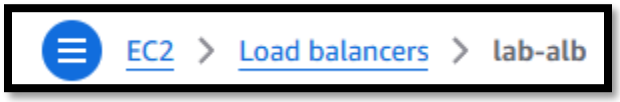
2) Add our new EC2 instances as targets then register them at the bottom of the page. Once completed and after a few minutes, you should see something like the output below:



3) Open your PuTTY connection and notice you will start seeing lots of HTTP GET requests from an internal IP. This is the AWS load balancer checking the health of your servers now.

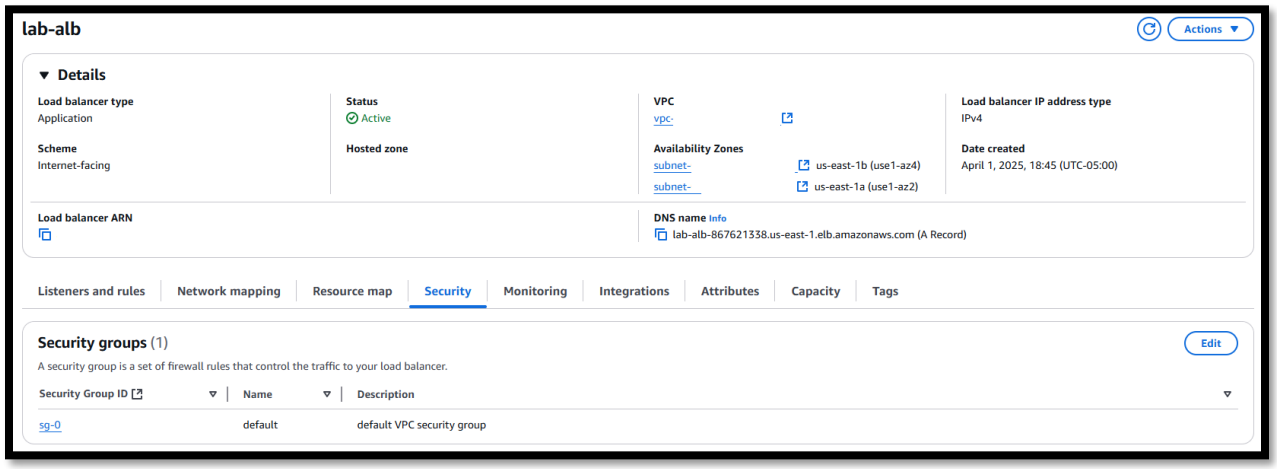


4) Next go to the Load Balancers tab in the AWS Console under EC2 and select your ALB.

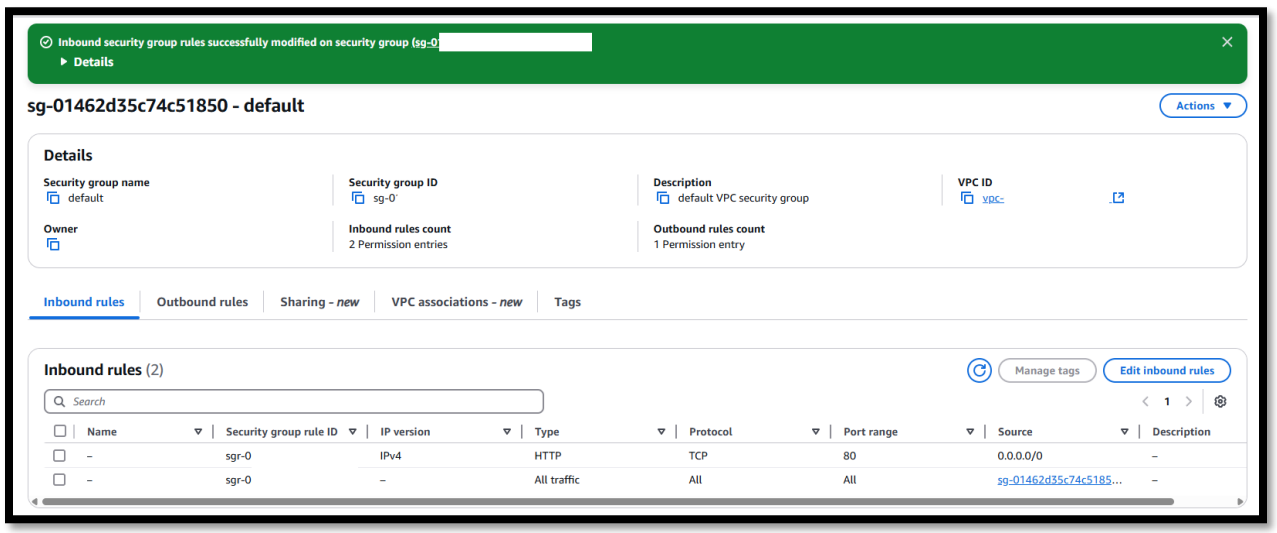


5) Find its DNS name, it should look like mine below. You will want to copy this and paste it into your browser momentarily. This is your ALB frontend DNS name. We can use it to connect to our backend web servers.

6) Before you do this, you need to allow TCP 80 traffic to your ALB from anywhere in the world. In your ALB configuration click on security and then click on the security group.



7) Create one last security group policy like the one seen below.



8) Finally, we can test our full deployment with the DNS name of the ALB we grabbed from step 5. Ensure you put /landing.html behind the name and it will look like mine below.

Summary

Congratulations on finishing this lab! This lab is by no means short or easy. Keep in mind this lab is meant to be a large introduction to utilizing AWS and building real-world based solutions. If you find this lab to be confusing or not making any sense, you may want to head over to the references on the next page and check out some of those links. If you have additional questions or have found errors, typos, and would like to contact me, please do so with the email address below:

taylor.kerber@hennepintech.edu

References

URL:	Description:
What is an Application Load Balancer?	An in-depth AWS technical document to explain Application Load Balancers.
Python SimpleHTTPServer	A quick tutorial of Python Simple HTTP Server commands and how to use it.
Connect to your Linux instance using PuTTY	A quick guide from AWS on how to connect to a Linux EC2 intance.
AWS Free Tier	An AWS official document for the types of services and limitations of free tier.
Enable internet access for a VPC using an internet gateway	Official AWS documentation on how to configure an AWS internet gateway.

Revisioning

Revision:	Editor:	Description:
1.0	Taylor Kerber	Initial release of lab.