Lab 4: Amazon EC2 and Load Balancing

Tutorial and Instructions

Type your answers and paste your screenshots directly in Lab4_Submission_File.docx in the indicated spots. Save the file with a filename in this format Lab4_*FirstName_LastName*.docx and upload it to D2L.

When asked to paste screenshots, please format screenshot so that your toolbar and system time are visible.

1. Setting up with Amazon EC2

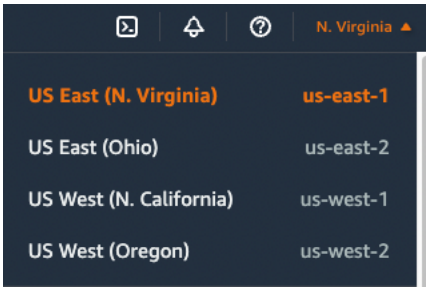
A. Create a Key Pair

AWS uses public-key cryptography to secure the login information for your instance. You specify the name of the key pair when you launch your instance, then provide the private key to obtain the administrator password for your Windows instance so you can log in using RDP (Remote Desktop Protocol).

If you haven't created a key pair already, you can create one using the Amazon EC2 console.

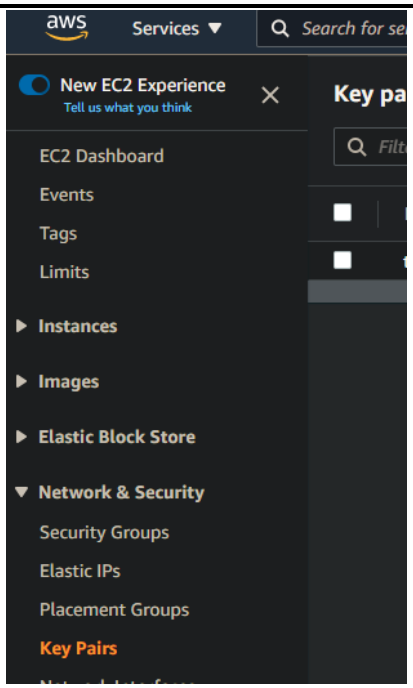
<u>To Create a Key Pair</u>

1. Sign in to AWS using your AWS Academy Account.
2. From the AWS dashboard, choose **EC2** to open the Amazon EC2 console.
3. From the navigation bar, select a region for the key pair. You can select any region that's available to you, regardless of your location (Our AWS Academy account may only allow US East N. Virginia us-east-1). The key pairs are specific to a region; for example, if you plan to launch an instance in the US East (N. Virginia) Region, you must create a key pair for the instance in the US East (N. Virginia) Region. Go to top right corner and you will see the below screenshot.

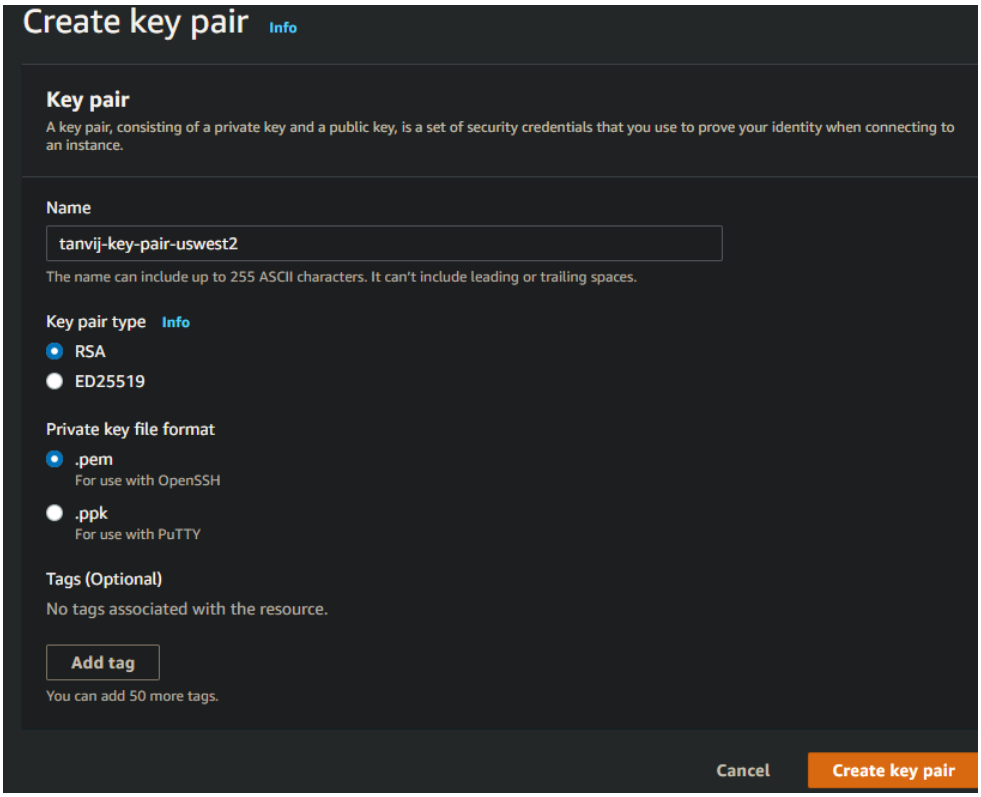


4. In the navigation pane, under **NETWORK & SECURITY**, click **Key Pairs**.

   Tip
   The navigation pane is on the left side of the console. If you do not see the pane, it might be minimized; click the arrow to expand the pane. You may have to scroll down to see the **Key Pairs** link.

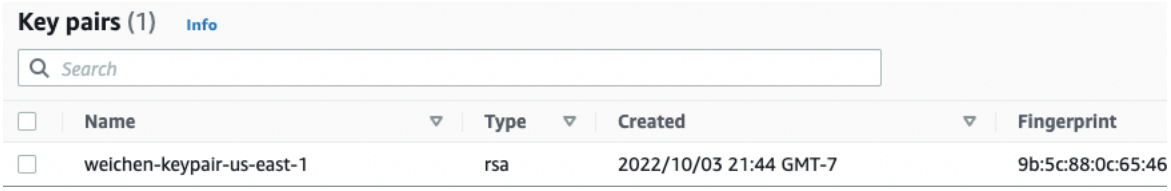


5. Click **Create Key Pair**.
6. Enter a name for the new key pair in the **Key pair name** field of the **Create Key Pair** dialog box, and then click **Create**. Choose a name that is easy for you to remember, such as your <NetID>, followed by `-key-pair`, plus the region name. For example, *netid*-key-pair-*useast1*.



7. The private key file is automatically downloaded by your browser. The base file name is the name you specified as the name of your key pair, and the file name extension is". **pem" (.cer on Mac)**. Save the private key file in a safe place.

   Important
   This is the only chance for you to save the private key file. You'll need to provide the name of your key pair when you launch an instance and the corresponding private key each time you connect to the instance.



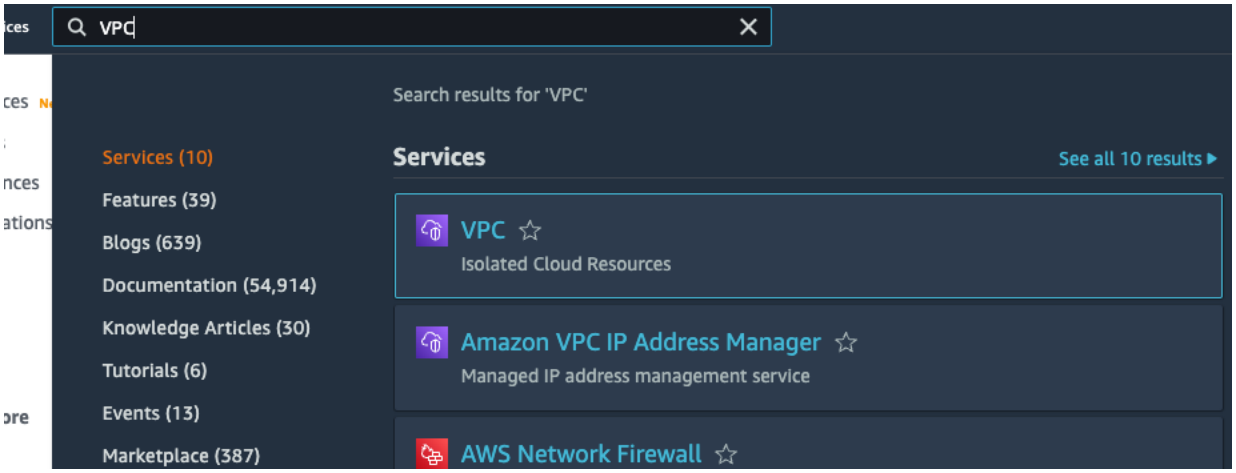
Paste the Screenshot showing Key Pair Generated
<<INSERT SCREENSHOT>>

B. Create a Virtual Private Cloud (VPC)

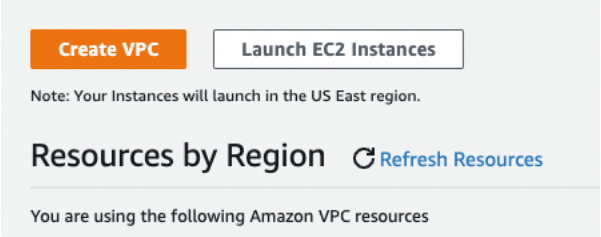
Amazon VPC enables you to launch AWS resources into a virtual network that you've defined. A Default VPC is already created, but we are creating a manual VPC.

To create a nondefault VPC

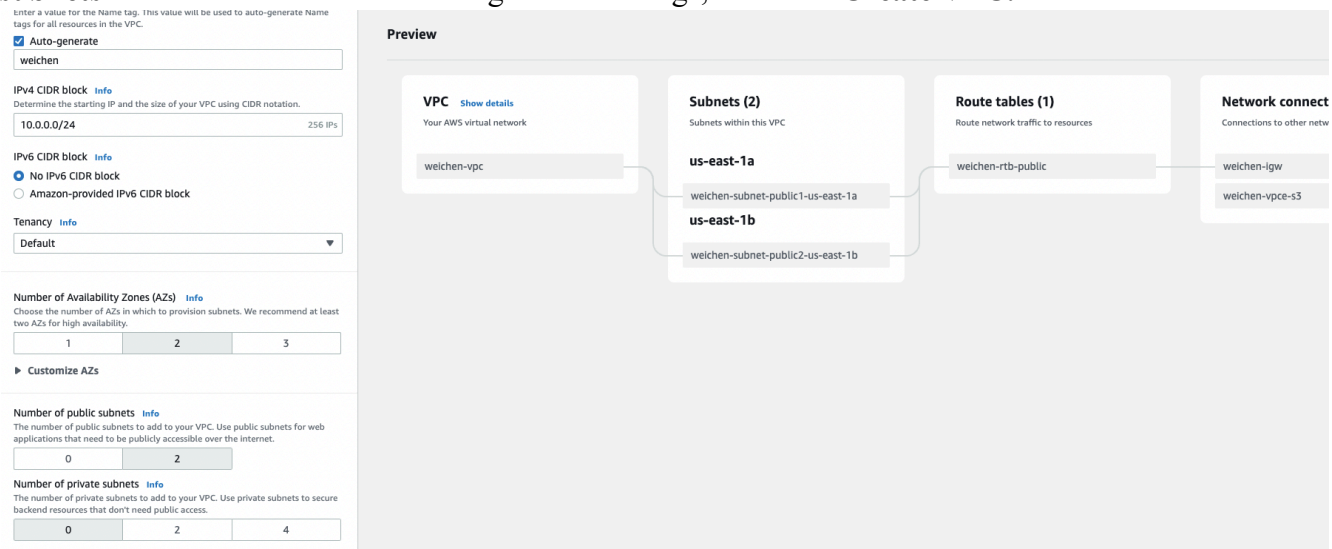
1. Open the Amazon VPC console by searching for VPC in the search bar.



2. From the navigation bar, select a region for the VPC. **VPCs are specific to a region, so you should select the same region in which you created your key pair.**
3. On the VPC dashboard, click **Create VPC**.



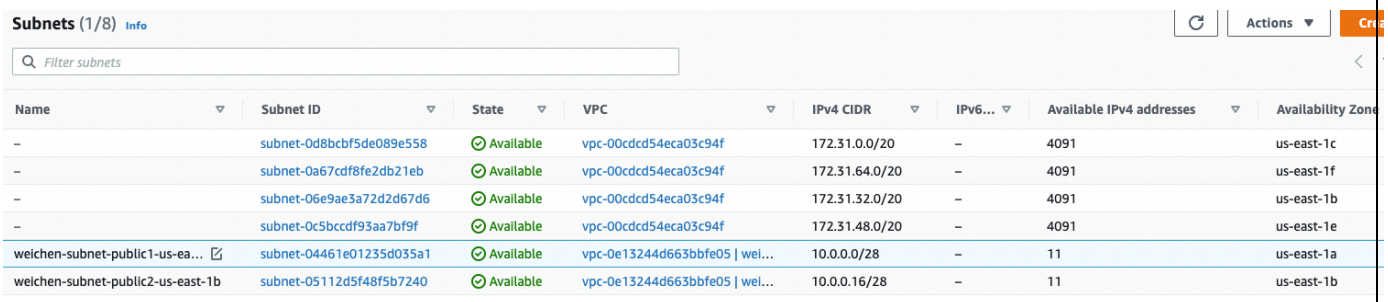
4. On the **Create VPC** page, ensure that **VPC and more** is selected, and enter your netid for the **Auto-generate** name tag.
5. Change your **IPv4 CIDR block** to 10.0.0.0/24, and ensure that we have 2 **public subnets**, and 0 **private subnets**. Leave the other default configuration settings, and click **Create VPC.**



Paste the Screenshot showing VPC
<<INSERT SCREENSHOT>>

Check subnets in different availability zones

1. Select Subnets in VPC Dashboard. You will find the two subnets 10.0.0.0/28 and 10.0.0.16/28. Note that the first one was created in the Availability Zone us-east-1a, and the second one was created in the Availability Zone us-east-1b.



Name	Subnet ID	State	VPC	IPv4 CIDR	IPv6...	Available IPv4 addresses	Availability Zone
–	subnet-0d8bcbf5de089e558	Available	vpc-00cdcd54eca03c94f	172.31.0.0/20	–	4091	us-east-1c
–	subnet-0a67cdf8fe2db21eb	Available	vpc-00cdcd54eca03c94f	172.31.64.0/20	–	4091	us-east-1f
–	subnet-06e9ae3a72d2d67d6	Available	vpc-00cdcd54eca03c94f	172.31.32.0/20	–	4091	us-east-1b
–	subnet-0c5bccdf93aa7bf9f	Available	vpc-00cdcd54eca03c94f	172.31.48.0/20	–	4091	us-east-1e
weichen-subnet-public1-us-ea...	subnet-04461e01235d035a1	Available	vpc-0e13244d663bbfe05 | wei...	10.0.0.0/28	–	11	us-east-1a
weichen-subnet-public2-us-east-1b	subnet-05112d5f48f5b7240	Available	vpc-0e13244d663bbfe05 | wei...	10.0.0.16/28	–	11	us-east-1b

Paste the Screenshot showing both your subnets in two different availability zones
<<INSERT SCREENSHOT>>

C. Create a Security Group

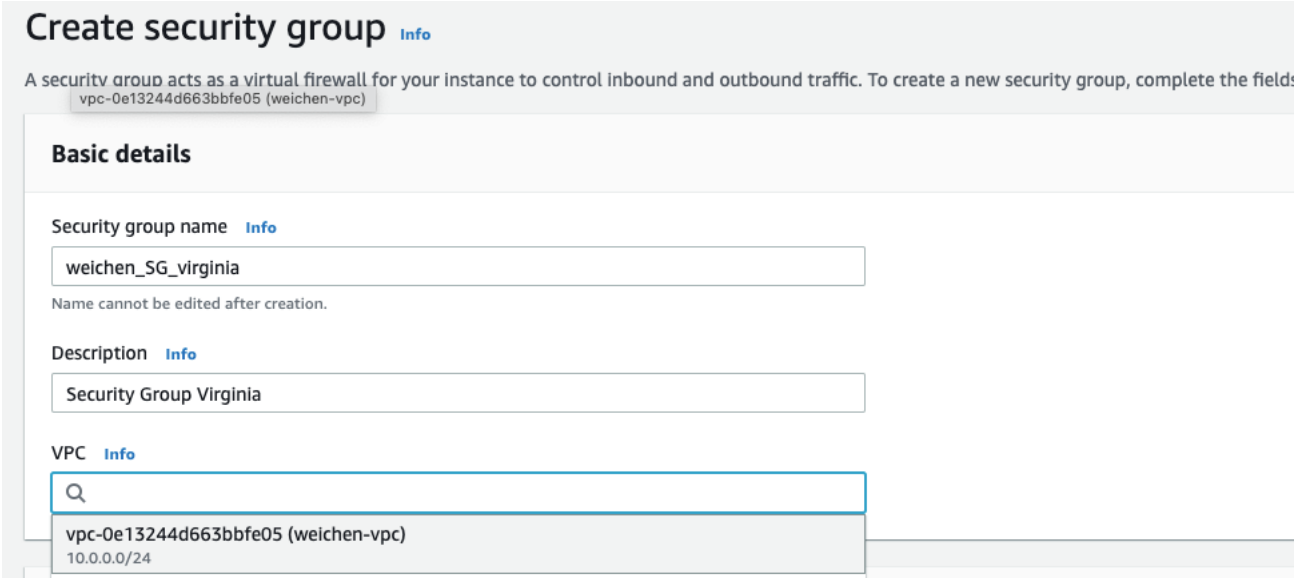
Security groups act as a firewall for associated instances, controlling both inbound and outbound traffic at the instance level. You must add rules to a security group that enable you to connect to your instance from your IP address using RDP (Remote Desktop Protocol). You can also add rules that allow inbound and outbound HTTP and HTTPS access from anywhere.

Prerequisites

You'll need the public IPv4 address of your local computer. The security group editor in the Amazon EC2 console can automatically detect the public IPv4 address for you. Alternatively, you can use the search phrase "what is my IP address" in an Internet browser.

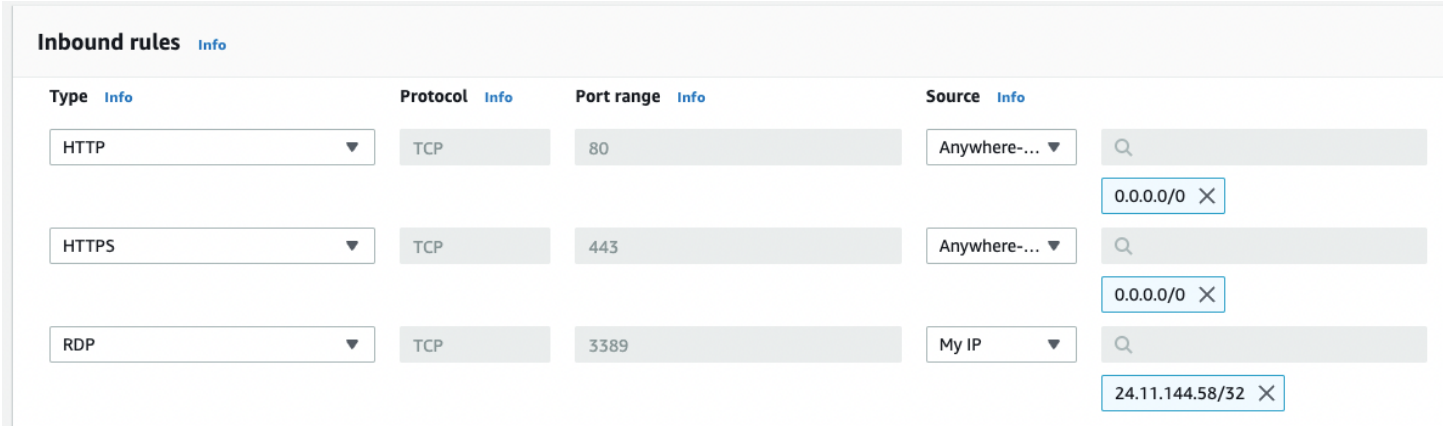
To create a security group with least privilege

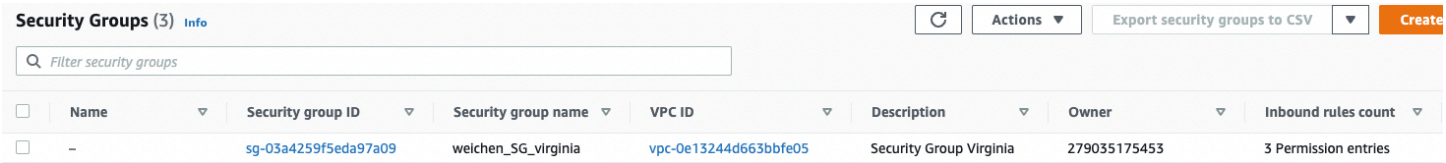
1. From the navigation bar, select a region for the security group. Security groups are specific to a region, so you should select the same region in which you created your key pair (e.g., us-east-1).
2. Click **Security Groups** in the navigation pane.
3. Click **Create Security Group**.
4. Enter a name for the new security group and a description. Choose a name that is easy for you to remember, such as your <NetID> name, followed by _SG_, plus the region name. For example, *netid*_SG_*virginia.*
5. In the **VPC** list, select your VPC.



6. Select the security group you created and on the **Inbound** tab, create the following rules (click **Add Rule** for each new rule), and then click **Create Security Group**:
   - Select **HTTP** from the **Type** list, and make sure that **Source** is set to **Anywhere-IPv4** (0.0.0.0/0).
   - Select **HTTPS** from the **Type** list, and make sure that **Source** is set to **Anywhere-IPv4** (0.0.0.0/0).

- Select **RDP** from the **Type** list. Choose **My IP** to specify your public IPv4 addresses in CIDR notation (You will need to update it just in case the next time your router obtains a different public IP).





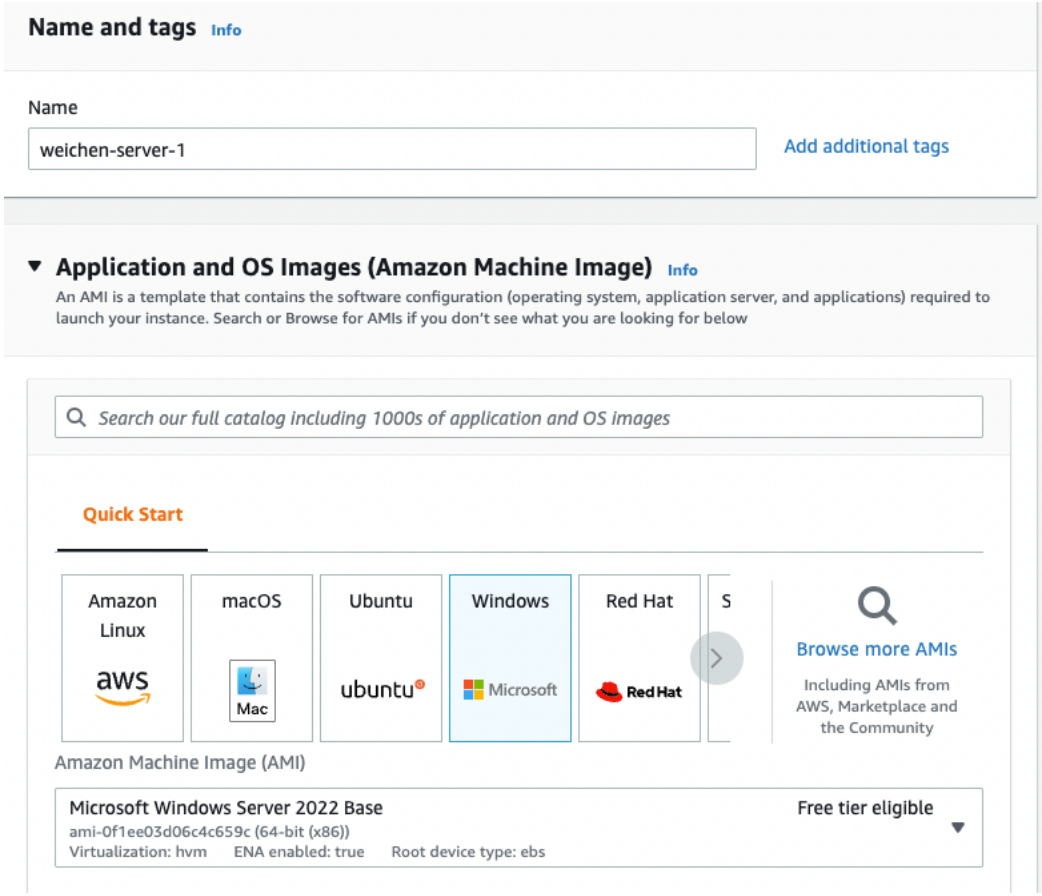
Paste the Screenshot showing security Group
<<INSERT SCREENSHOT>>

2. Launching an EC2 Instance

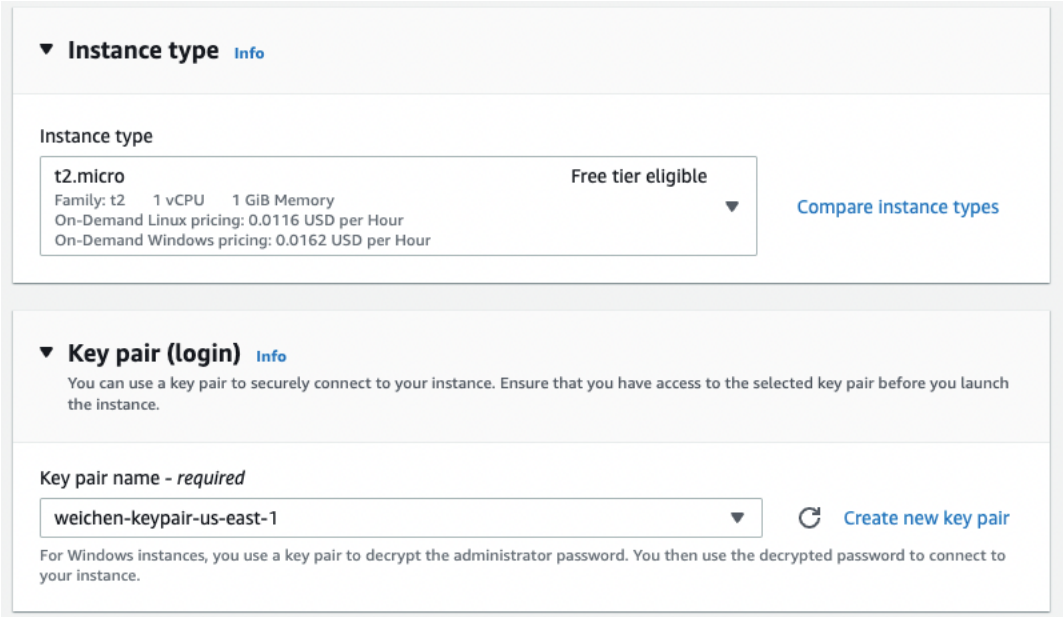
A. Launch an Instance

You can launch a Windows instance using the AWS Management Console as described in the following procedure.

1. Open the Amazon EC2 console through the Services menu or search bar.
2. From the console dashboard, choose **Launch Instance**.
3. Name it as <NetID>-server-1.
4. For the **Application and OS Images** (Amazon Machine Image), choose Windows, which will install Microsoft Windows Server 2022 Base. Notice that this AMI is marked "Free tier eligible."



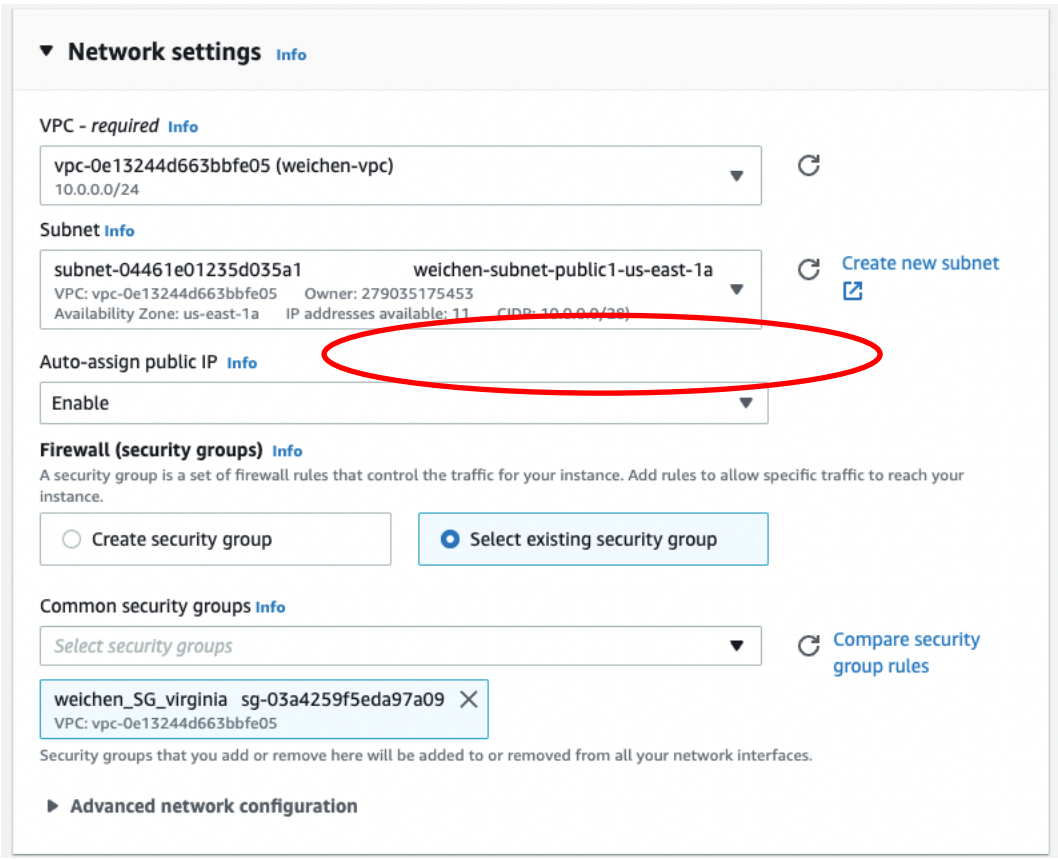
5. On the **Instance Type** dropdown menu, you can select the hardware configuration of your instance. Select the `t2.micro` type, which is selected by default.
6. On the **Key pair (login)** item, choose the key pair that you have created.



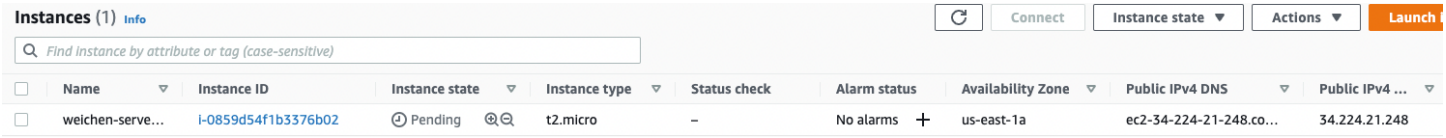
Caution

Don't select the **Proceed without a key pair** option. If you launch your instance without a key pair, then you can't connect to it.

7. Click **Edit** for the **Network settings**. Change VPC to your non-default VPC, which you have created earlier. Change "Subnet" to **the first subnet** (weichen-subnet-public1-us-east-1a in my case. Second instance to the second subnet). Change "Auto-Assign Public IP" to **Enable.**
8. Under **Firewall (security groups)**, choose **Select existing security group**, and select the security group that you have created earlier.

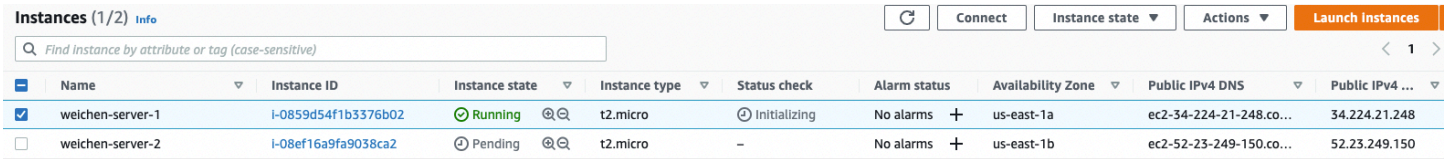


9. On the **Summary** pane, choose **Launch instance**.
10. A confirmation page lets you know that your instance is launching. Choose **View Instances** to close the confirmation page and return to the console.



11. On the **Instances** screen, you can view the status of the launch. It takes a short time for an instance to launch. When you launch an instance, its initial state is `pending`. After the instance starts, its state changes to `running` and it receives a public DNS name. (If the **Public DNS (IPv4)** column is hidden, choose the Show/Hide icon in the top right corner of the page and then select **Public DNS (IPv4)**.)

12. It can take up to 5 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.
13. Repeat the above steps to create a second instance. **Remember to create it in a different subnet.**



Paste the Screenshot showing EC2 Instances
<<INSERT SCREENSHOT>>

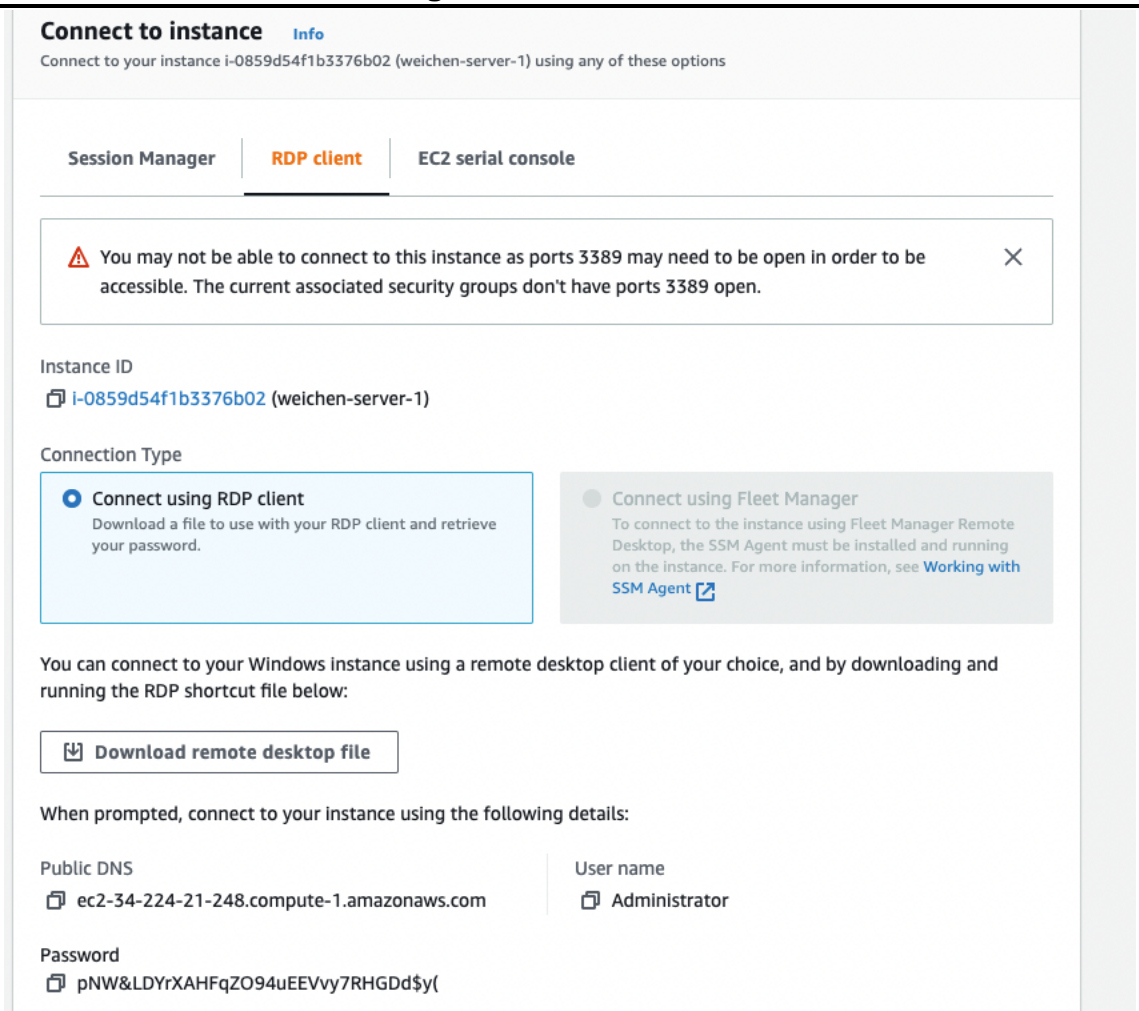
B. Connect to Your Instance

To connect to a Windows instance, you must retrieve the initial administrator password and then specify this password when you connect to your instance using Remote Desktop.

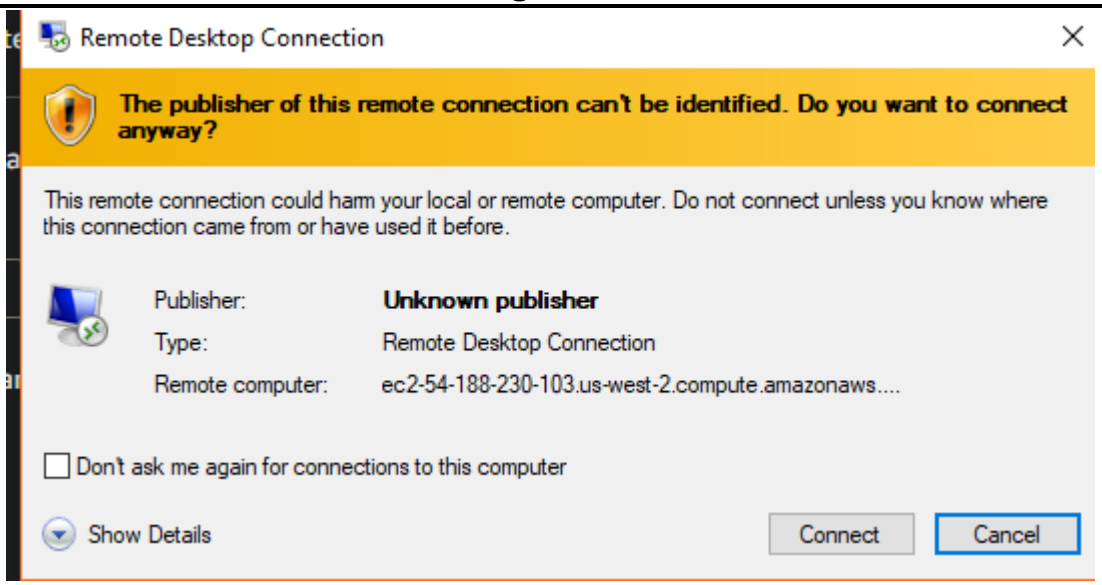
The name of the administrator account depends on the language of the operating system. For example, for English, it's Administrator, for French it's Administrateur, and for Portuguese it's Administrador.

To connect to your Windows instance using an RDP client

1. In the Amazon EC2 console, select the instance, and then choose **Connect**.
2. In the **Connect to Instance** page, choose the **RDP client** tab. If you are on a Mac, you will need to download the Microsoft Remote Desktop app from the App Store (it is free).
3. Click **Get password** at the bottom of the page (it will take a few minutes after the instance is launched before the password is available).
4. Choose **Upload private key file** and navigate to the private key file you created when you launched the instance. Select the file and choose **Open** to copy the entire contents of the file into the **Contents** field.
5. Choose **Decrypt Password**. The console displays the default administrator password for the instance in the **Connect to Your Instance** dialog box, replacing the link to **Get Password** shown previously with the actual password.
6. Record the default administrator password, or copy it to the clipboard. You need this password to connect to the instance.



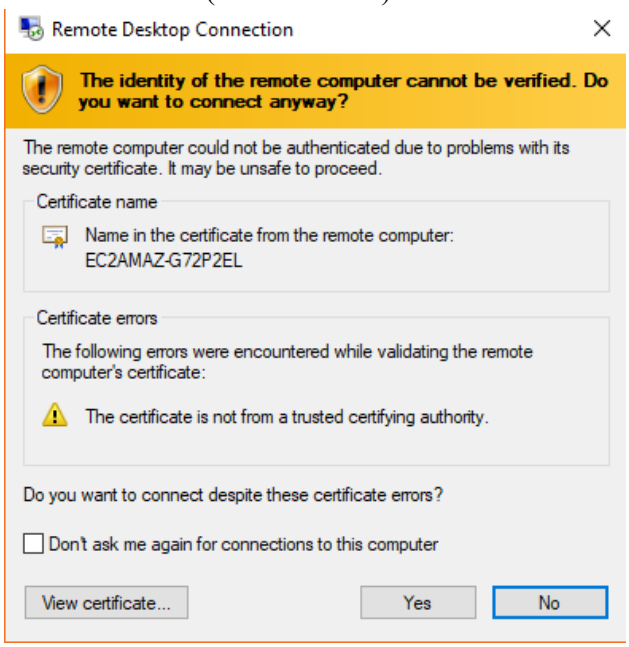
7. Choose **Download Remote Desktop File**. Your browser prompts you to either open or save the .rdp file. Either option is fine. When you have finished, you can choose **Close** to dismiss the **Connect to Your Instance** dialog box.
    - If you opened the .rdp file, you'll see the **Remote Desktop Connection** dialog box.
    - If you saved the .rdp file, navigate to your downloads directory, and open the .rdp file to display the dialog box.
8. You may get a warning that the publisher of the remote connection is unknown. If you are using **Remote Desktop Connection** from a Windows PC, choose **Connect** to connect to your instance. If you are using **Microsoft Remote Desktop** on a Mac, skip the next step.



9. When prompted, log in to the instance, using the administrator account for the operating system and the password that you recorded or copied previously. If your **Remote Desktop Connection** already has an administrator account set up, you might have to choose the **Use another account** option and enter the user name and password manually.

   Note
   Sometimes copying and pasting content can corrupt data. If you encounter a "Password Failed" error when you log in, try typing in the password manually.

10. Due to the nature of self-signed certificates, you may get a warning that the security certificate could not be authenticated. Use the following steps to verify the identity of the remote computer, or simply choose **Yes** or **Continue** to continue if you trust the certificate (recommended).

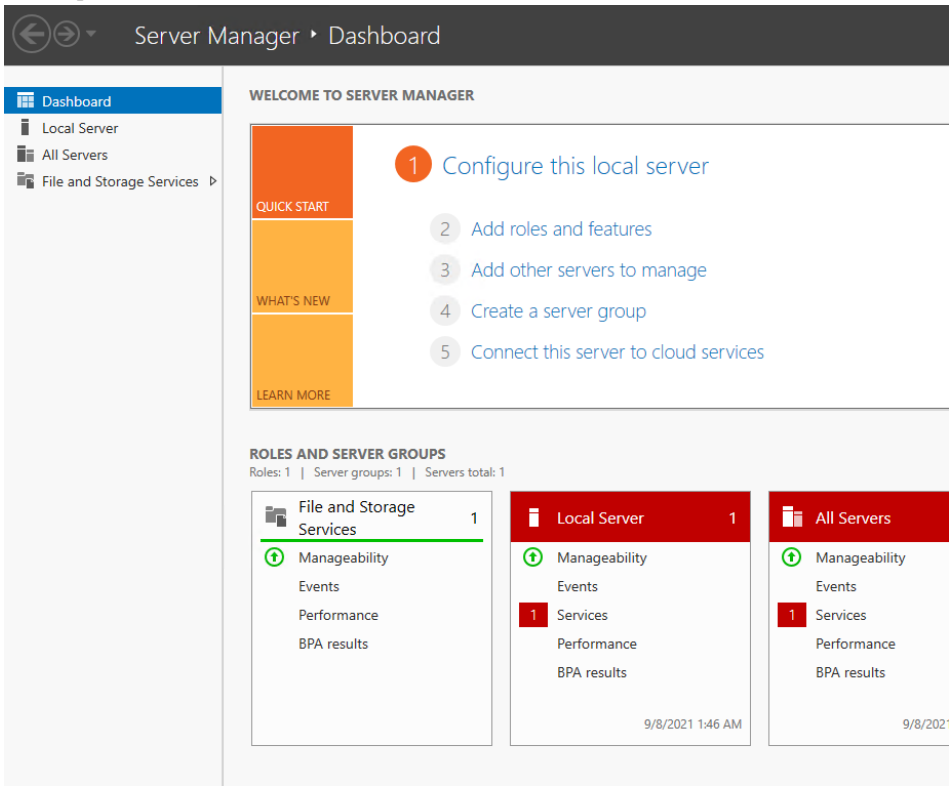


C. Setup the Internet Information Services (IIS)

You need to download Remote Desktop File for each instance. Here, I have explained the steps for 1st instance. You need to repeat the process for 2nd instance too.

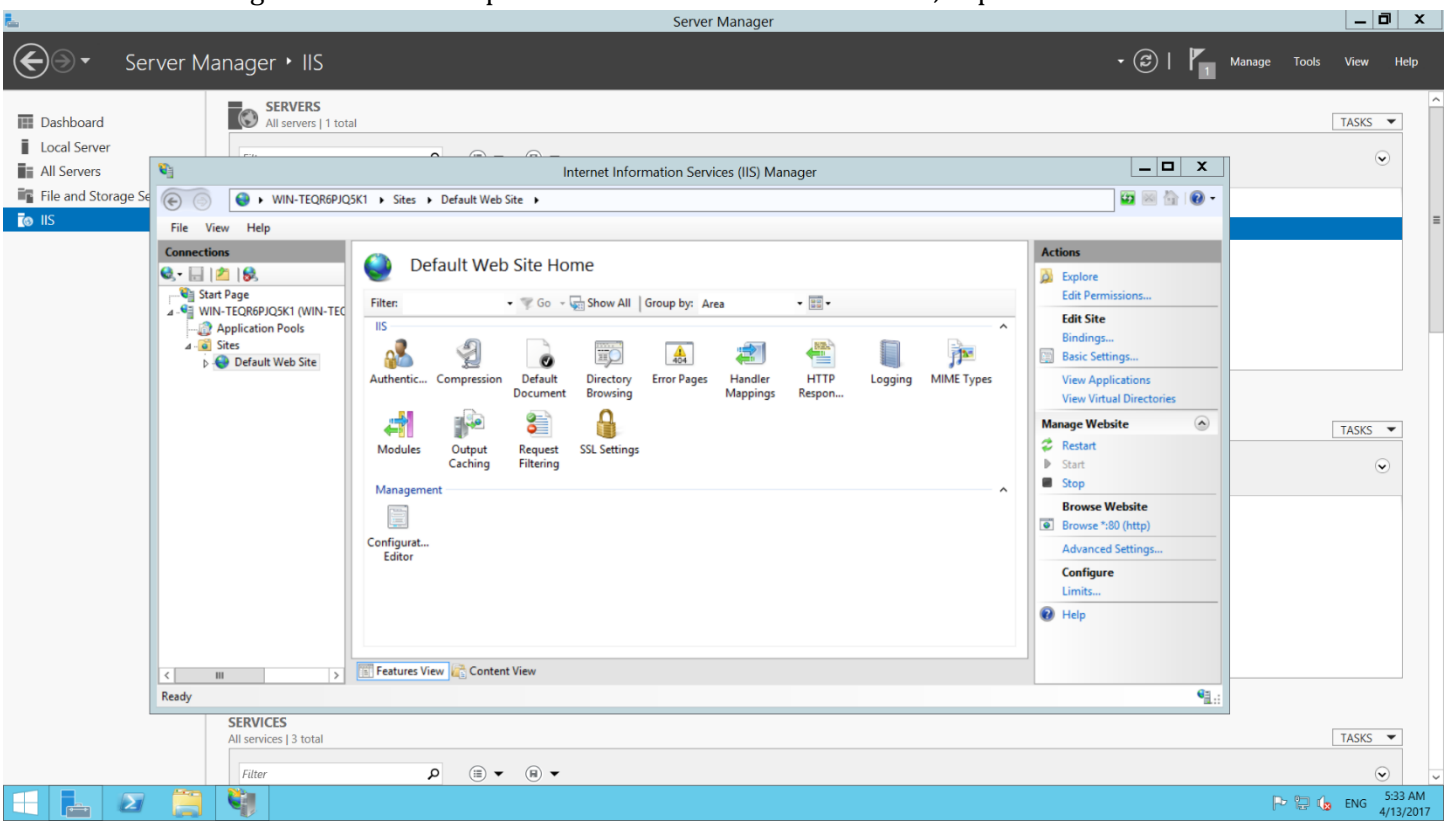
<u>These Steps would be repeated twice:</u>

1. Click on the Remote Desktop File. You might need to wait for few minutes till the RDP Client load its profile.
2. Open the "Start" Menu or click on , and search for "Server Manager." In the Top-Right, you will Click on "Manage," and then "Add Roles and Features."
   - You Might get a dialog box stating "Server Manager is collecting inventory data"(This happens because the Server Manager has not fully loaded). Click on OK. Wait for the Server Manager to load all its profile.



3. Click on Next in "Before you Begin." In Installation type, "Role based or feature-based installation" is selected. In Server Selection, leave the default value and Click on Next.
4. In Server Role, scroll down and Select Web Server (IIS) and Click Add Features.
5. Leave the default values as it is and Click on Next till you reach "Confirmation".
6. Click on Install to start the process.
7. After the installation is done, Click on IIS on the navigation pane.
8. Right Click on Server Name and Open Internet Information Services (IIS) Manager.

9. Go to the Navigation Pane and expand the content of IIS Server. Then, Expand the "Sites"

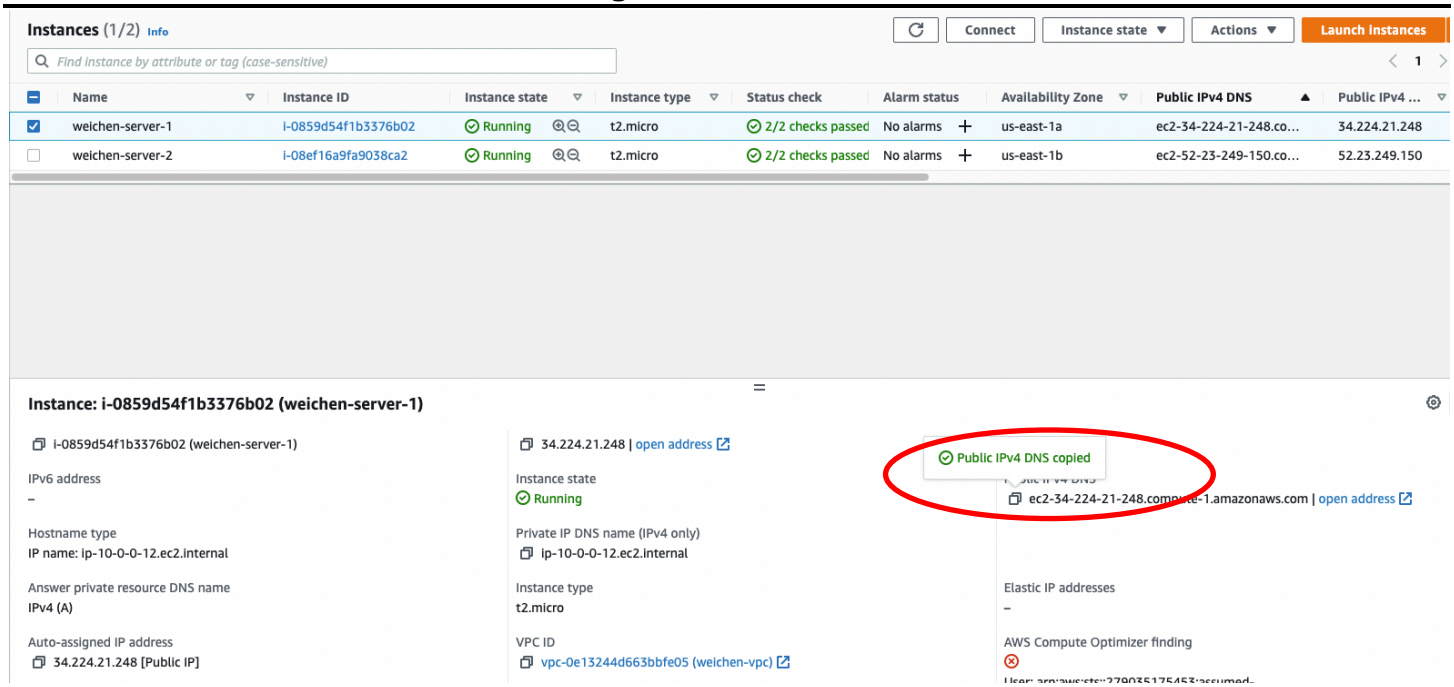


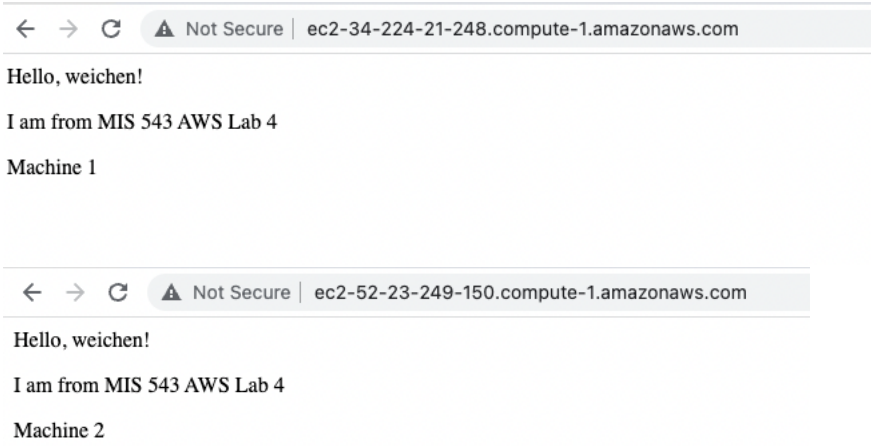
10. Right Click on **Default Web Site** and navigate to Explore**.** Delete the HTML Document "iisstart" and Copy paste the Index.html (Pay attention to the file extension. It may not be visible). Open the file in **Notepad.**
11. Replace the content with:

```
<html>
<body>
<p>Hello, <your_UA_Net_id>!</p>
<p>I am from MIS 543 AWS Lab 4</p>
<p>Machine 1 or 2</p>
</body>
</html>
```

Note: You will copy and paste "Index.html" for each Instance. Please specify for which instance this file belongs to. Is it for Machine 1(corresponds to first instance) or Machine 2(corresponds to second instance) (Highlighted in yellow). This step is necessary because load balancer routes the traffic to either 1st or 2nd instance (web page).

12. **Save** the changes.
13. Now, go back to your EC2 Instance Window, and copy the string from **Public DNS name** (for example, ec2-54-200-87-74.us-west-2.compute.amazonaws.com) and paste it into the address field of an Internet-connected web browser. If IIS is working, you see the default page of your server.





Paste the Screenshot of the default page for **both** of the machines
<<INSERT SCREENSHOT>>

3. Application Load Balancer

A. Create Application Load Balancer

1. Open the Amazon EC2 console.
2. On the navigation bar, choose a region for your load balancer. Be sure to select the same region that you selected for your EC2 instances.
3. On the navigation pane, under **LOAD BALANCING**, choose **Load Balancers**.
4. Choose **Create Load Balancer**.
5. Choose **Application Load Balancer**, and then choose **Create**.

B. Define Your Load Balancer

You must provide a basic configuration for your load balancer, such as a name, a network, and a listener.

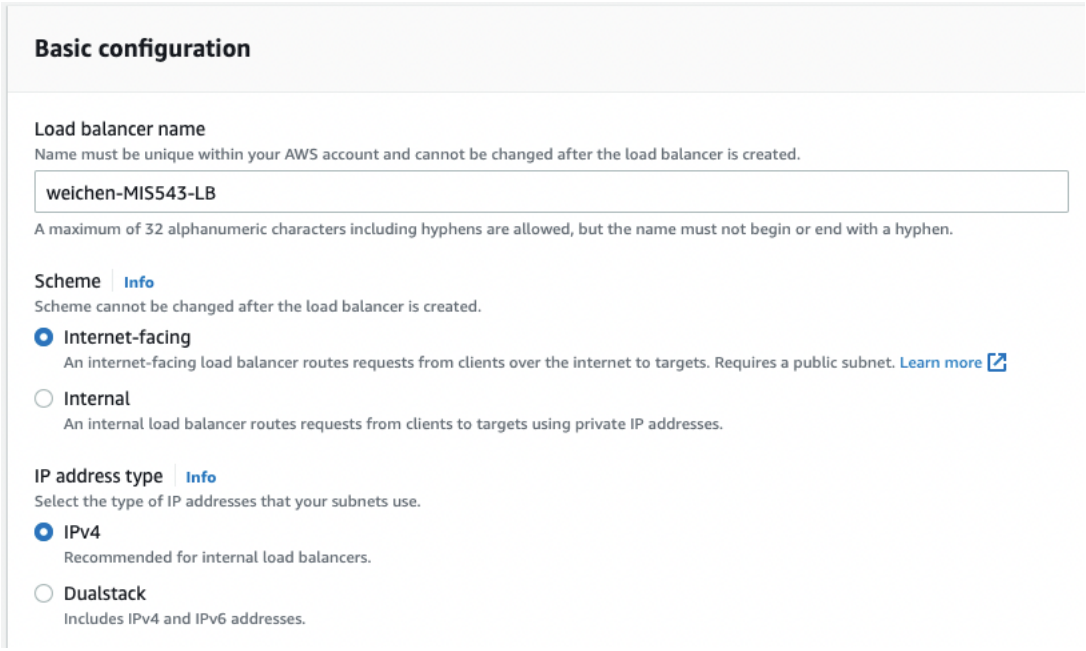
A listener is a process that checks for connection requests. It is configured with a protocol and a port for front-end (client to load balancer) connections and a protocol and a port for back-end (load balancer to instance) connections. In this, you configure a listener that accepts HTTP requests on port 80 and sends them to your instances on port 80 using HTTP.

To define your load balancer and listener

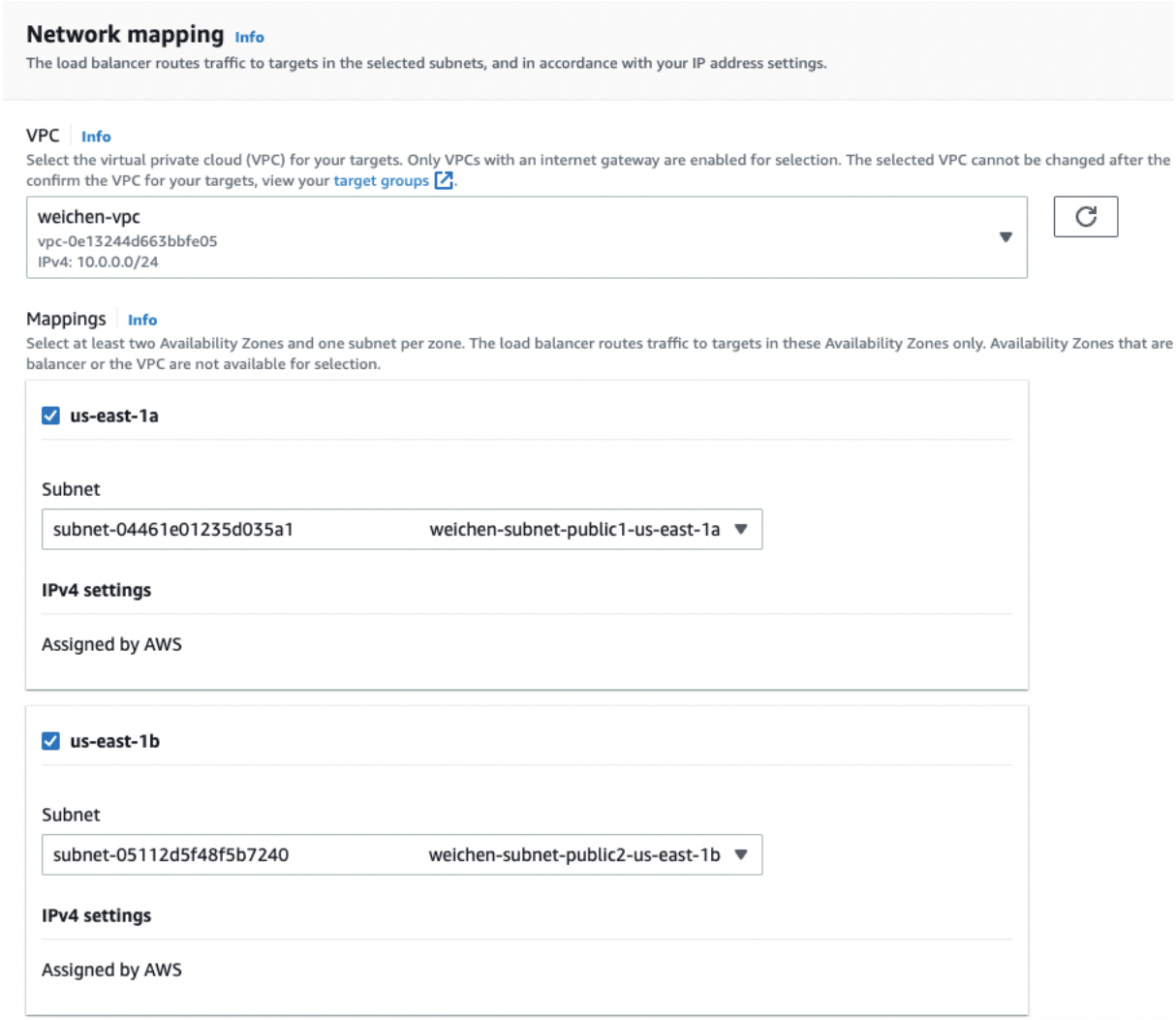
1. For **Load Balancer name**, use the name <NetID>-MIS543-LB for your load balancer.

   The name of your Classic Load Balancer must be unique within your set of Classic Load Balancers for the region, can have a maximum of 32 characters, can contain only alphanumeric characters and hyphens, and must not begin or end with a hyphen.

2. Leave the default Scheme and IP address type configuration.



3. For Network Mapping, select both subnet in your manually created VPC.

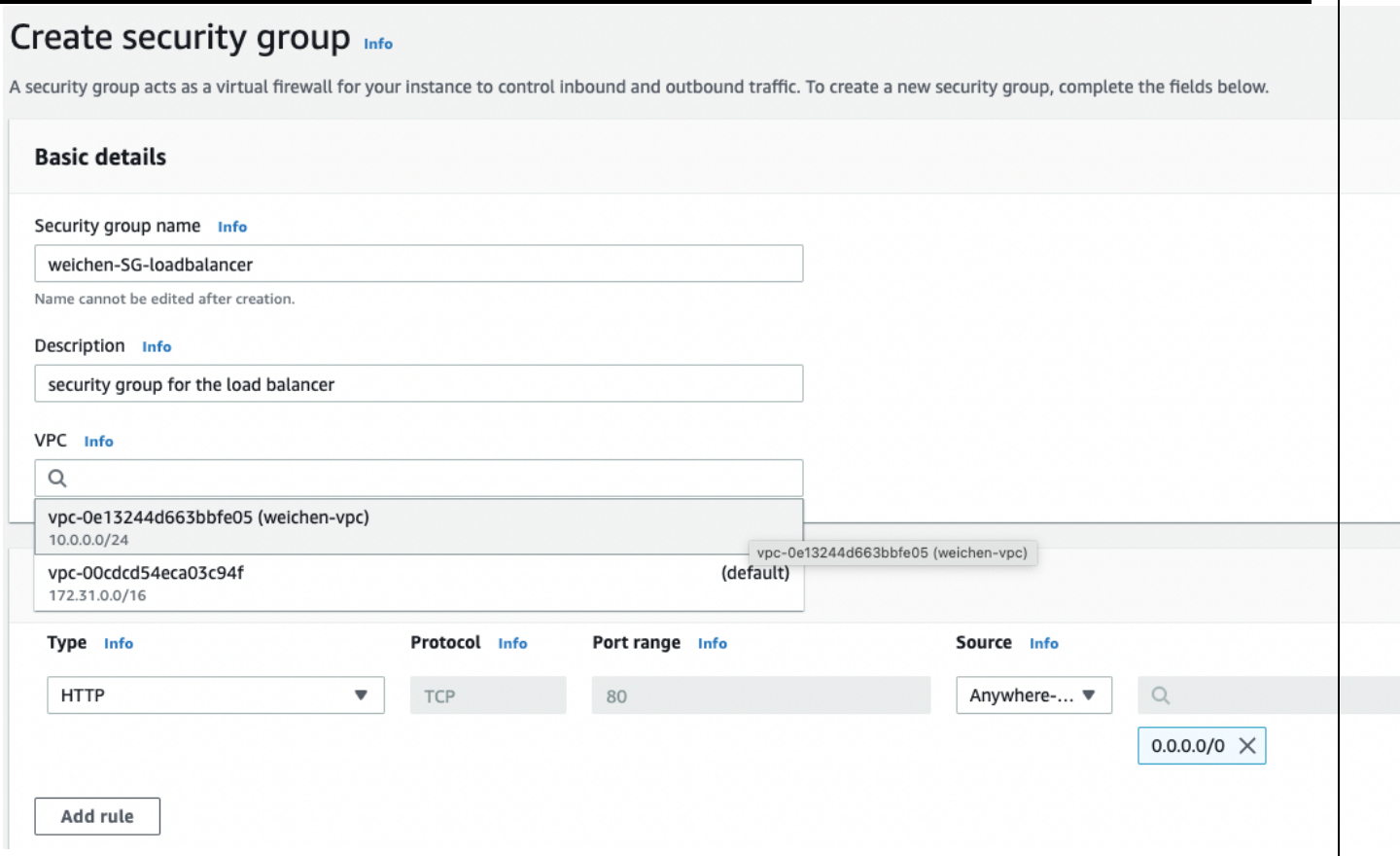


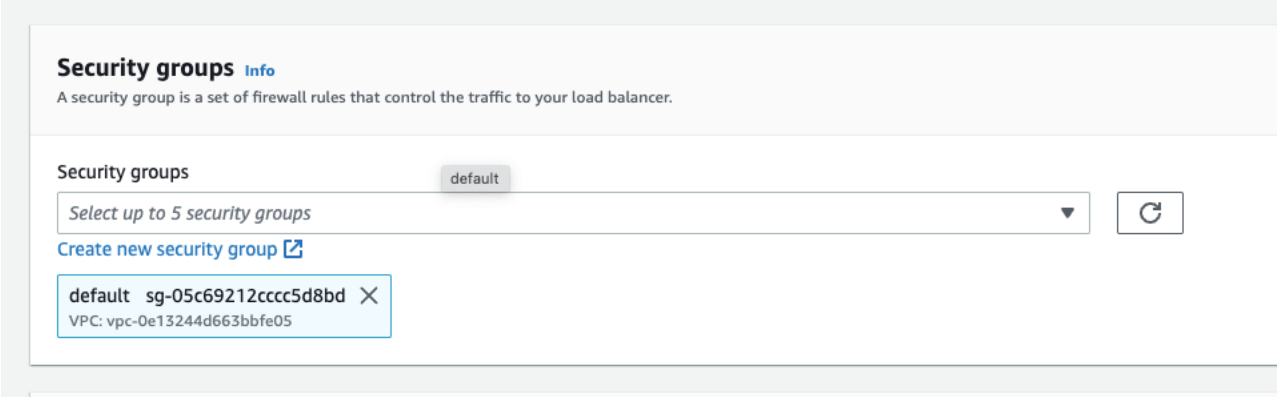
4. Next: **Security Groups**.

If you selected a VPC as your network, you must assign your load balancer a security group that allows inbound traffic to the ports that you specified for your load balancer and the health checks for your load balancer.

To assign security group to your load balancer

1. On the **Assign Security Groups** page, select **Create a new security group**.
2. Type a name and description for your security group. *For example,* Security group name: <NetID>-SG-loadbalancer. This new security group contains a rule that allows traffic to the port that you configured your load balancer to use. For the **Inbound rules**, change Type to **HTTP** and Source to **Anywhere**



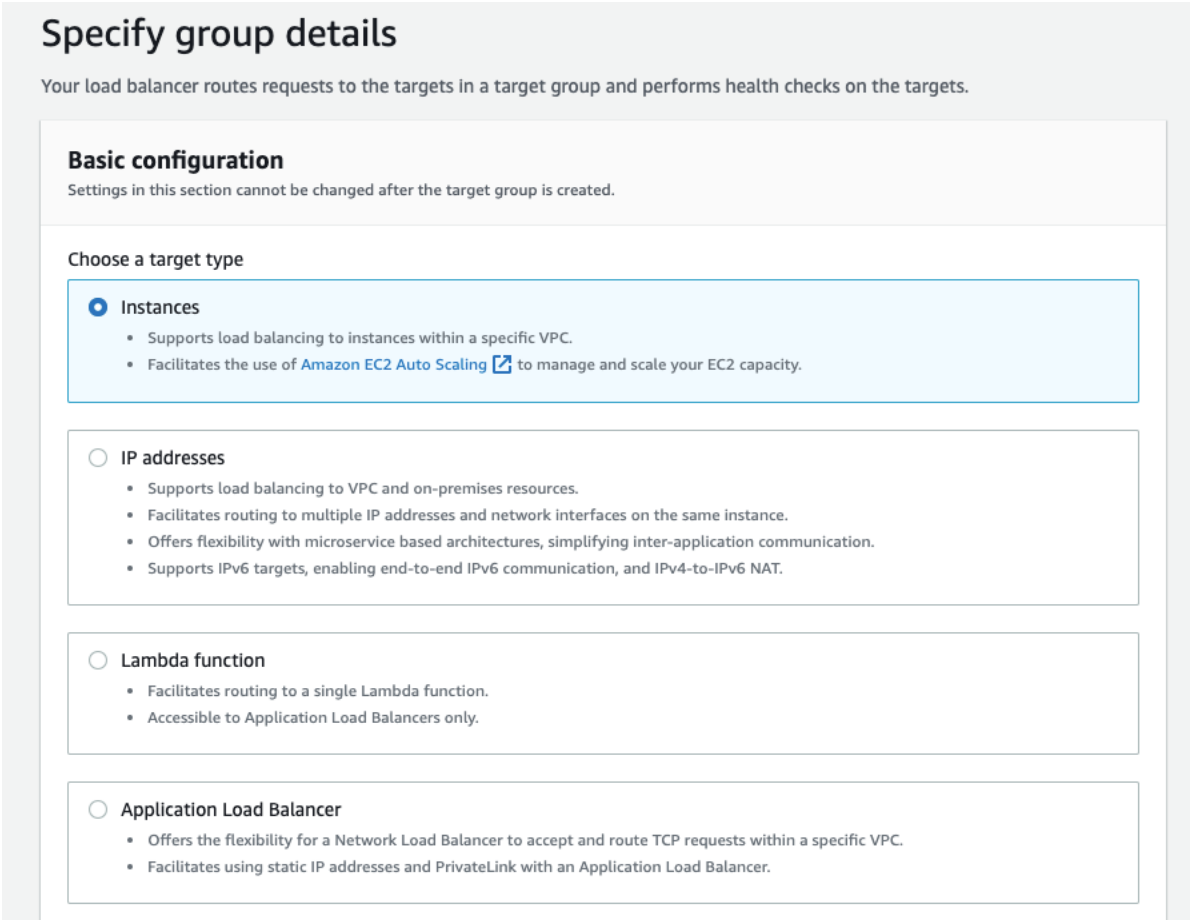


C. Configure Routing for your EC2 Instances

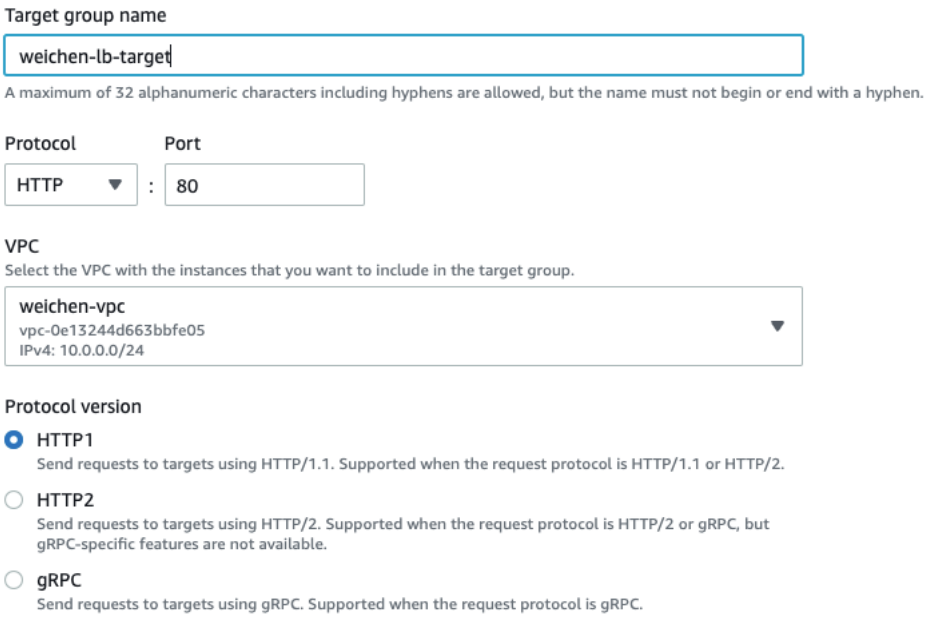
Elastic Load Balancing automatically checks the health of the EC2 instances for your load balancer. If Elastic Load Balancing finds an unhealthy instance, it stops sending traffic to the instance and reroutes traffic to healthy instances. In this step, you customize the health checks for your load balancer.

To configure target group and health checks for your instances

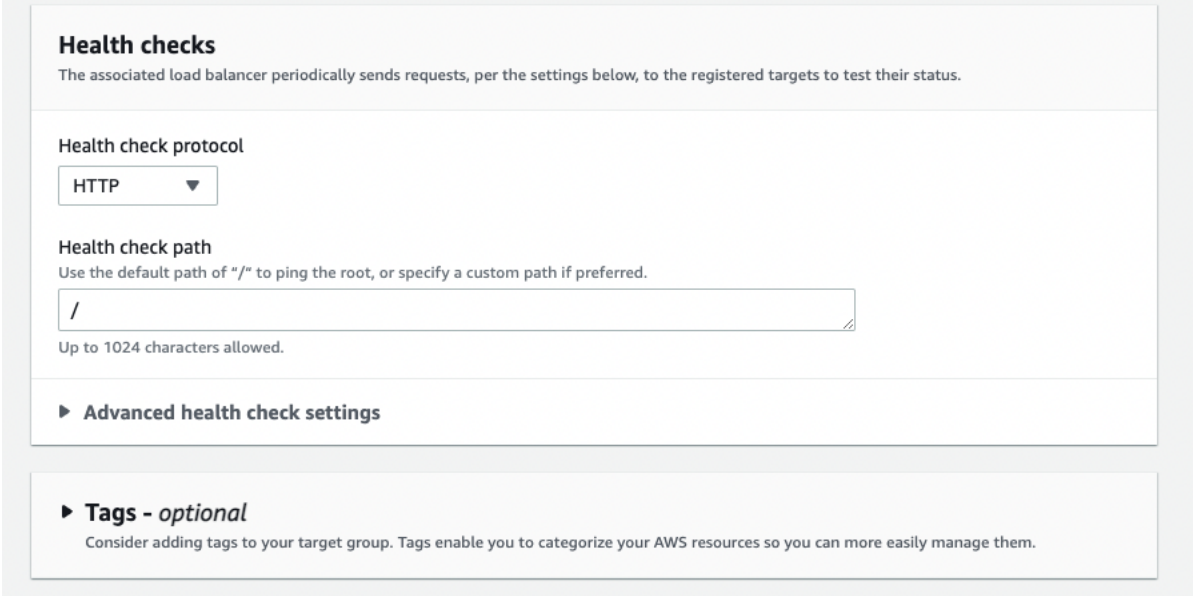
1. On the **Listerners and routing** section, leave the default values for HTTP and 80.
2. For Default action, click **Create target group**. Select Instances for the target type.

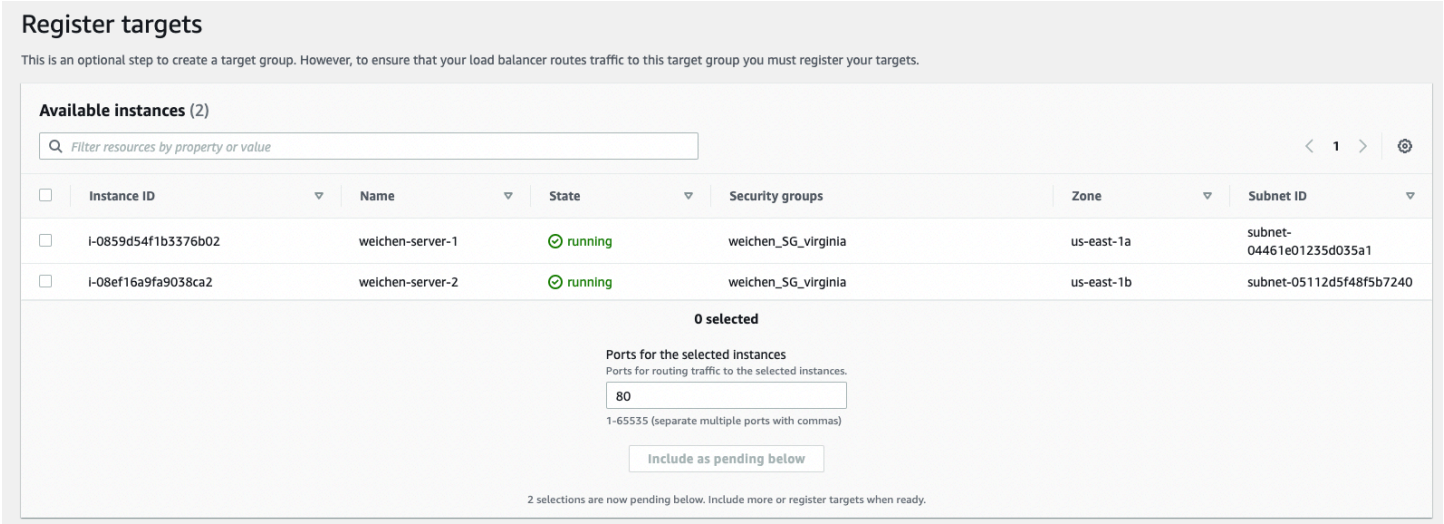


3. Change the target group name to <NetID>-lb-target, and ensure that your VPC is chosen.

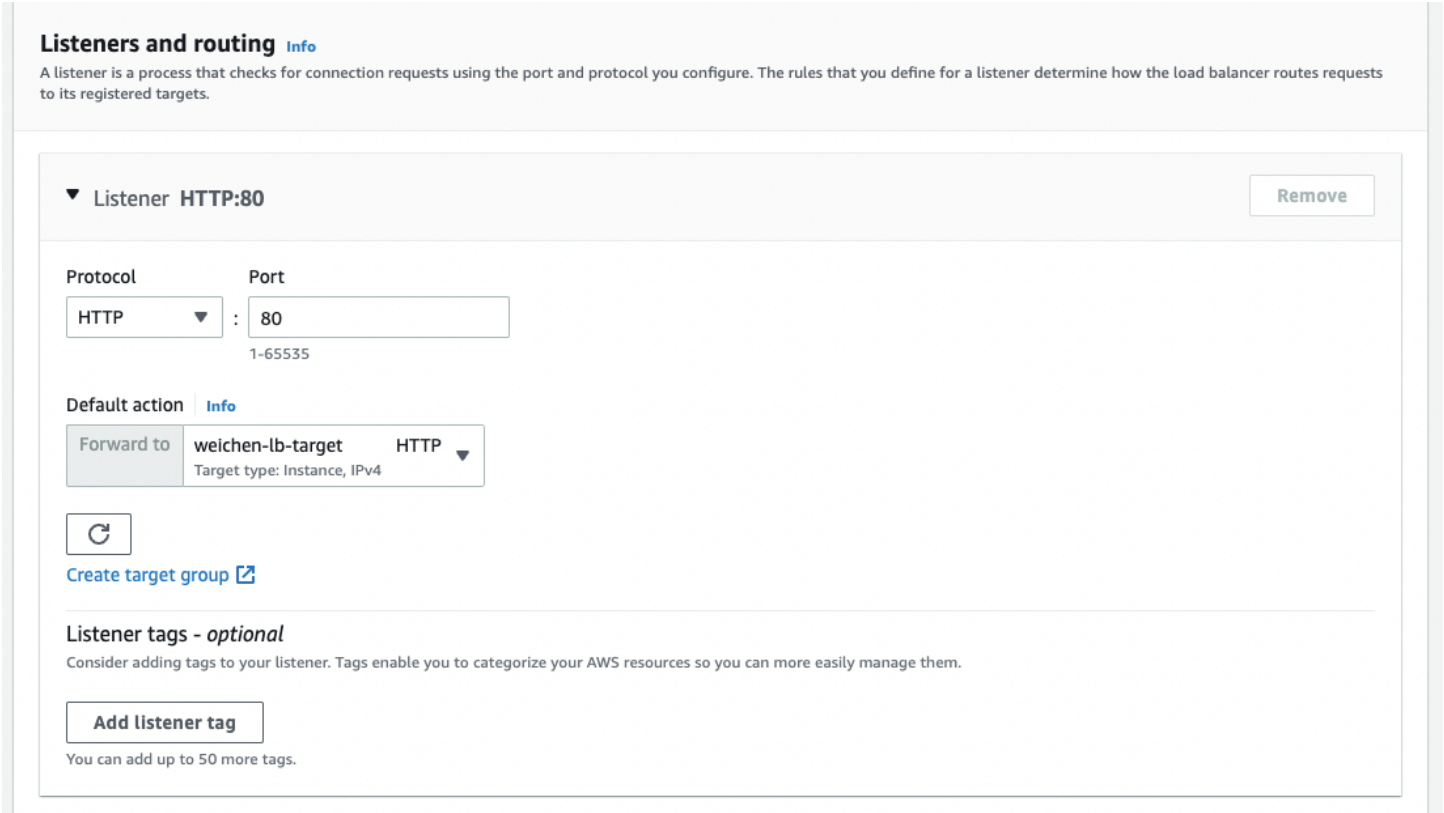


4. For Health checks, leave the default value, and Choose **Next: Register Targets**.





5. On the **Register Targets** page, select the both the instances, and click **Include as pending below**.
6. Choose **Create Target Group**
7. Add target group to the Load Balancer. Choose **Protocol: Port = HTTP:80 and** assign the newly created target group in the next field
8. Click **Create load balancer**.



D. Register EC2 Instances with Your Load Balancer

Your load balancer distributes traffic between the instances that are registered to it.

Note

When you register an instance with an elastic network interface (ENI) attached, the load balancer routes traffic to the primary IP address of the primary interface (eth0) of the instance.

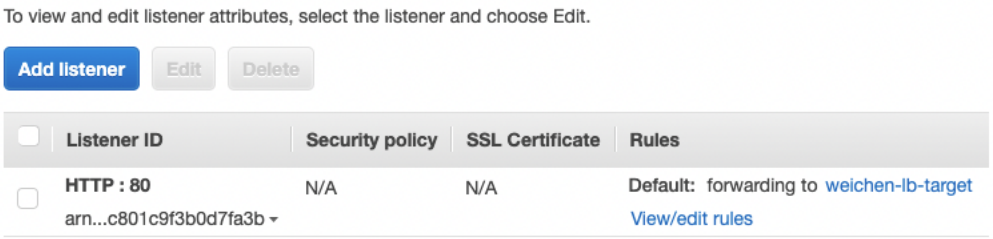
To register EC2 instances with your load balancer

E. Verify and Test Your Load Balancer

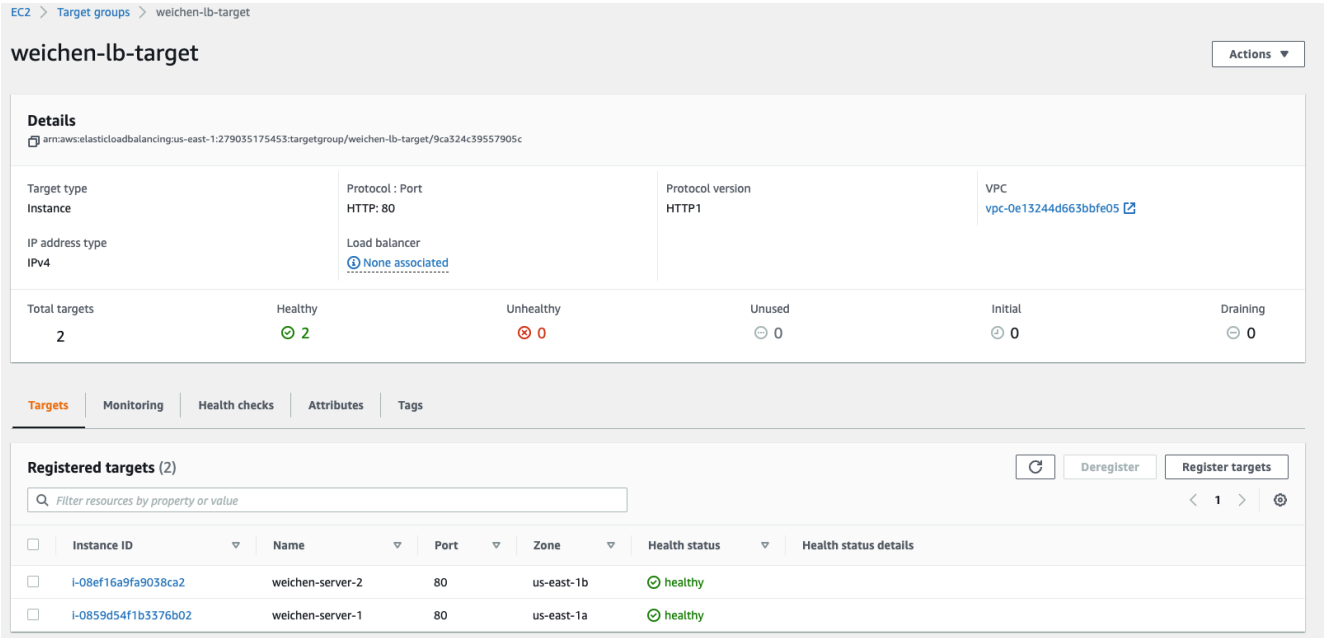
After creating the load balancer, you can verify that it's sending traffic to your EC2 instances.

To create and test your load balancer

1. Select your new load balancer.
2. As shown in below screenshot, click on the **Listeners** tab. You should see that this load balancer is listening on Port 80, and will forward to weichen-lb-target target group.

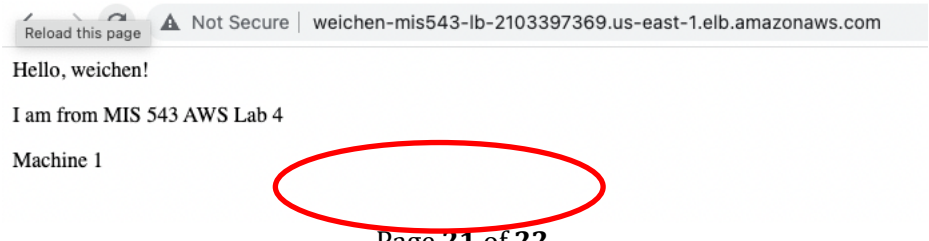


3. Click Target Groups in your navigation pane. You can see the details of your weichen-lb-target target group here.



4. If it indicates that some of your instances are not in service, it's probably because they are still in the registration process. If there is not registered then you can add them by Choosing Register Targets>> Select the machines from available instances>> Select Include as pending below>>Register the pending targets. Refer the below screenshot.

5. After at least one of your EC2 instances is in service, you can test your load balancer. Copy the string from **DNS name in Description Tab** (for example, weichen-mis543-lb-2103397369.us-east-1.elb.amazonaws.com) and paste it into the address field of an Internet-connected web browser. If your load balancer is working, you see the default page of your server. Refresh the page, you will see the content of both your Instances, that is, Machine 1 and Machine 2.



Paste the Screenshot of the Listener tab for the load balancer
<<INSERT SCREENSHOT>>

Paste the Screenshot of the Targets tab for the target groups
<<INSERT SCREENSHOT>>

Paste two Screenshots of visiting the load balancer DNS name in browser, one showing machine 1, and another showing machine 2
<<INSERT SCREENSHOT>>