**Creating a Spot Amazon EC2 instance**

An**EC2 instance** is nothing but a virtual server in Amazon Web services terminology. It stands for**Elastic Compute Cloud.** It is a web service where an AWS subscriber can request and provision a compute server in AWS cloud.

A spot instance is an offering from AWS; it allows an AWS business subscriber to bid on unused AWS compute capacity. The hourly price for a Spot instance is decided by AWS, and it fluctuates depending on the supply and demand for Spot instances.

Your Spot instance runs whenever your bid exceeds the current market price. The price of a spot instance varies based on the instance type and the Availability Zone in which the instance can be provisioned.

When your bid price exceeds the market spot price of the instance called as the ‘spot price,' your instance stays running. When the spot price overshoots the bid price, AWS will terminate your instance automatically. Therefore, it is necessary to plan the spot instances in your application architecture carefully.

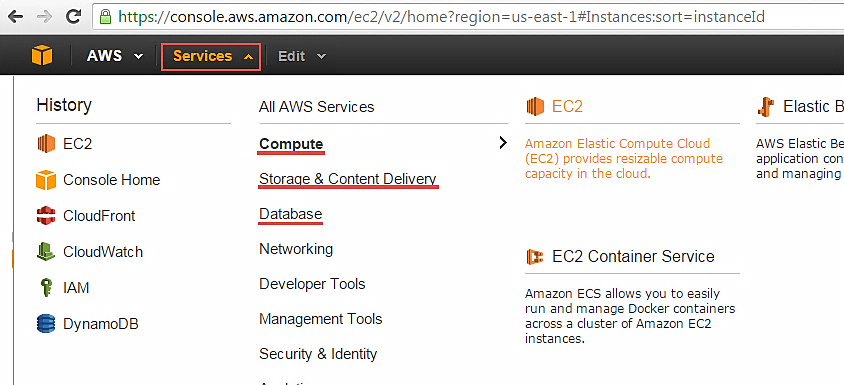
In this tutorial, you will learn-

* Login and access to AWS services
* Creating a Spot Request
* Find Instance Types
* Configure the Spot instance
* Reviewing your Spot instance
* Connecting to the spot instance

### Login and access to AWS services

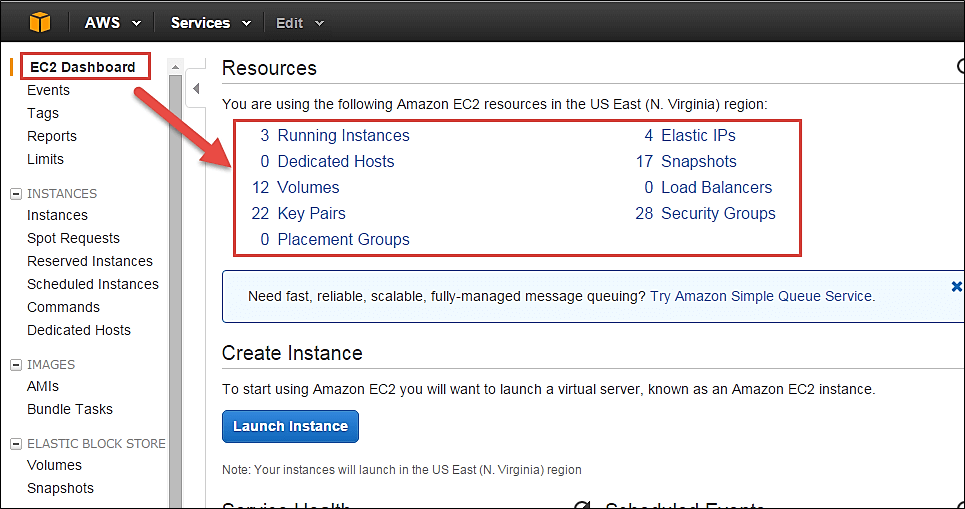
**Step 1)**Login to your AWS account and go to the AWS Services tab at the top left corner.

Here, you will see all of the AWS Services categorized as per their area viz. Compute, Storage, Database, etc. For creating an EC2 instance, we have to choose Computeà EC2 as in the next step.



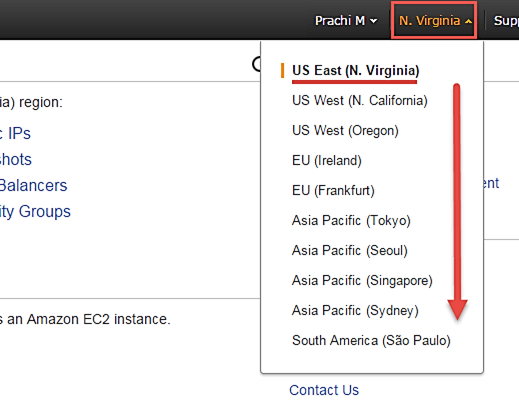
Open all the services and click on EC2 under Compute services. This will launch the dashboard of EC2.

Here is the EC2 dashboard. Here you will get all the information in gist about the AWS EC2 resources running.



**Step 2)** On the top right corner of the EC2 dashboard, choose the AWS Region in which you want to provision the EC2 server.

Here we are selecting N. Virginia. AWS provides 10 Regions all over the globe.



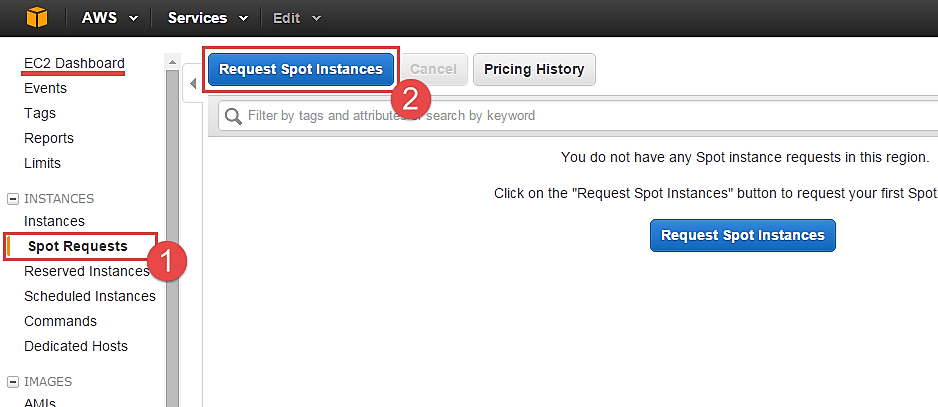
**Step 3)**Once your desired Region is selected, come back to the EC2 Dashboard.

### Creating a Spot Request

In order to launch a spot instance, you have to first create a Spot Request.

Follow the steps below to create a Spot Request.

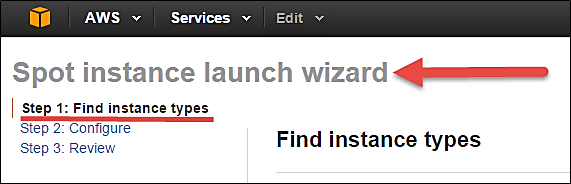
1. On the EC2 Dashboard select 'Spot Requests' from the left pane under Instances.
2. Click on the button 'Request Spot Instances" as shown below.



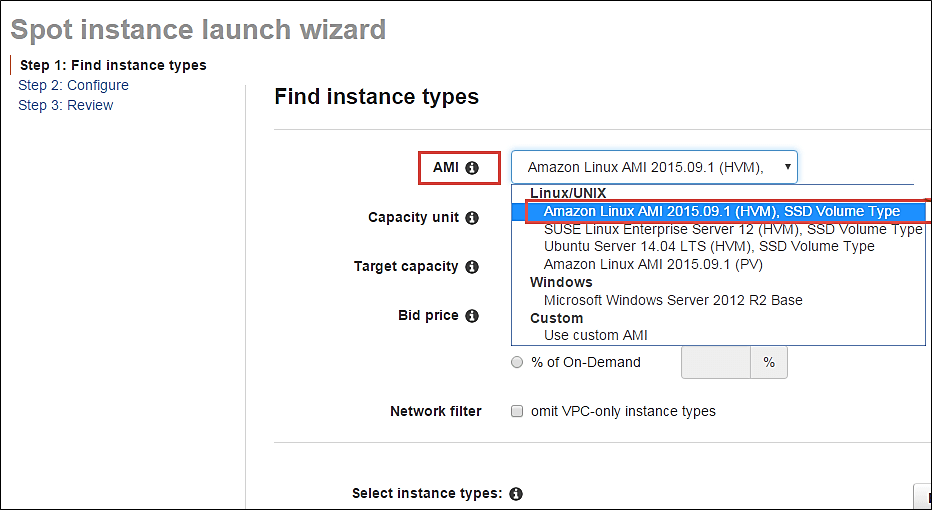
Spot instance launch wizard will open up. You can now go ahead with selecting the parameters and the instance configuration.

#### Find Instance Types

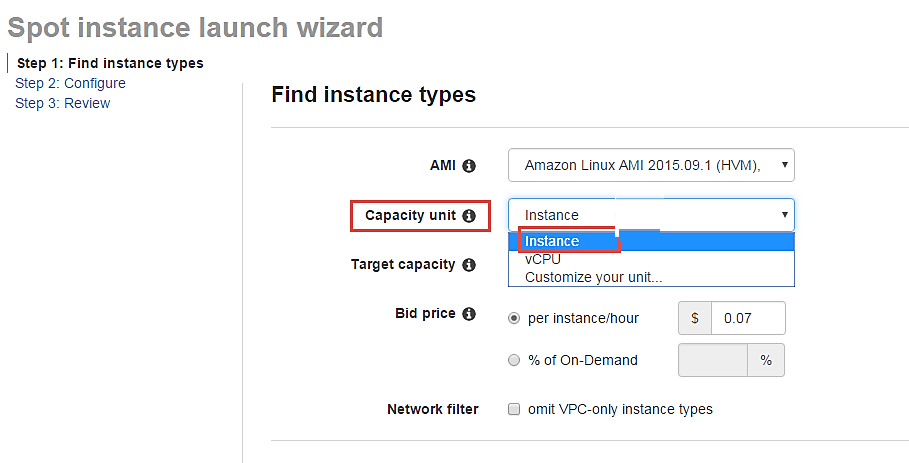
The first step for spot instance is to "Find instance types."



**Step 1)** Select an AMI- an AMI is a template consisting of the OS platform and software to be installed in the instance. Select your desired AMI from the existing list. We are selecting Amazon Linux AMI for this tutorial.

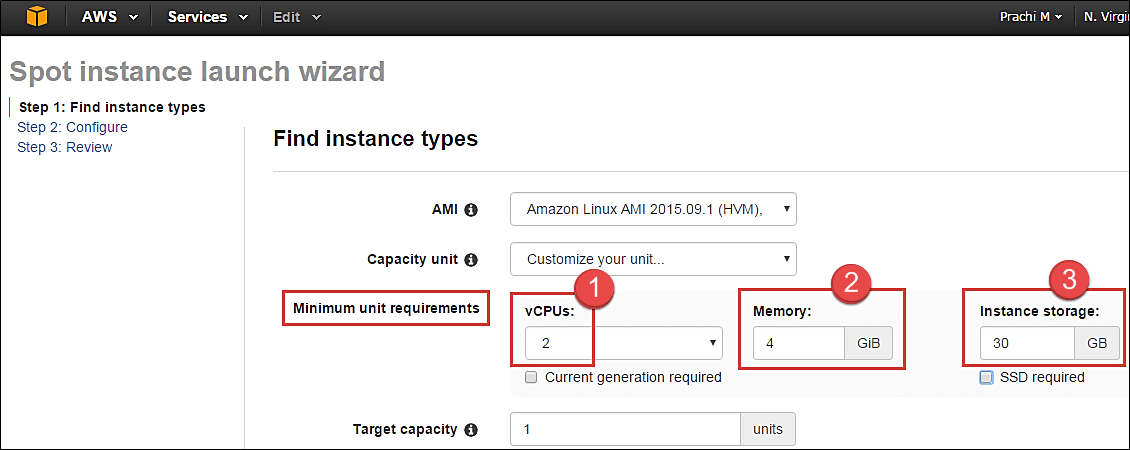


**Step 2)** Capacity Unit- a Capacity Unit is your application requirement. You may decide to launch an instance based on the instance type, vCPU or custom configuration like your choice of vCPU/memory/storage requirements. Here we are selecting an Instance.

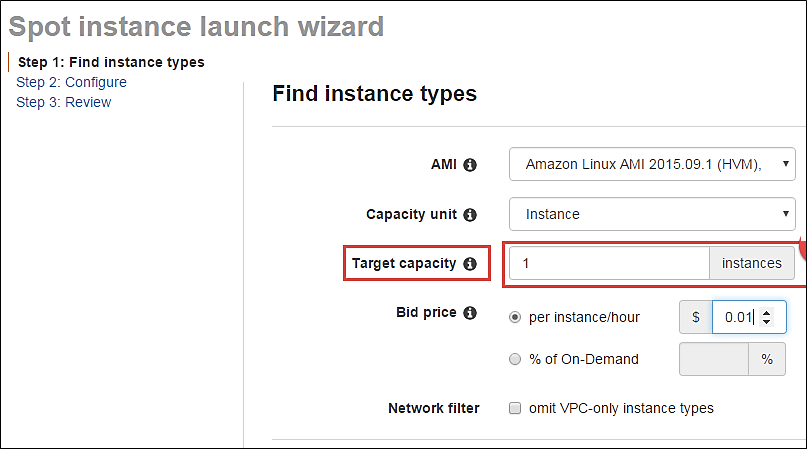


If you wish to customize the capacity, you can add your choice of

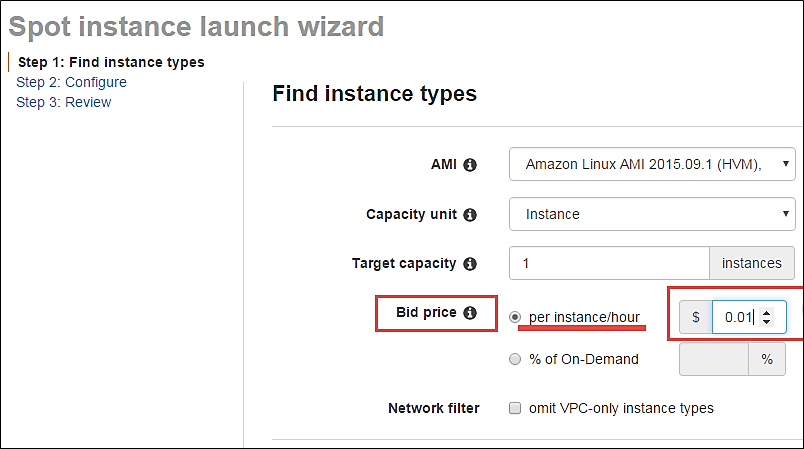
1. vCPU,
2. Memory and
3. Instance storage as below.



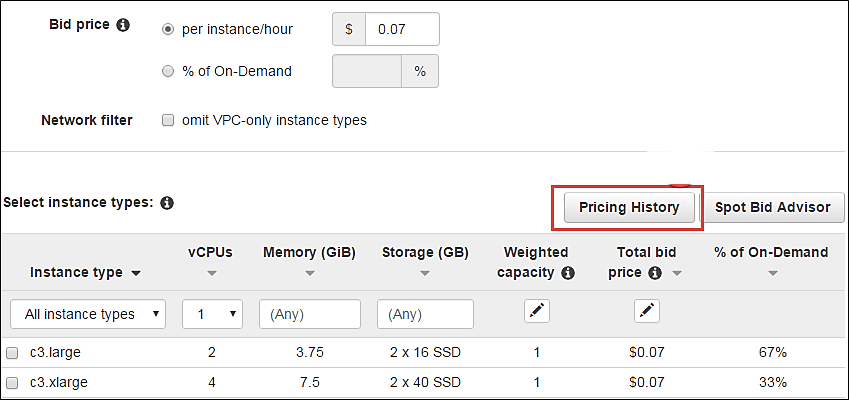
**Step 3)** Target Capacity depicts how many spot instances you wish to maintain in your request. Here we are selecting one.



**Step 4)** Bid Price – this is the maximum price we are ready to pay for the instance. We are going to set a particular price per instance/hour. This is the simplest to calculate based on our business requirement. We will see ahead how we should determine the bid price so that our bid price always remains high and doesn't exceed the spot price so that our instance keeps running.



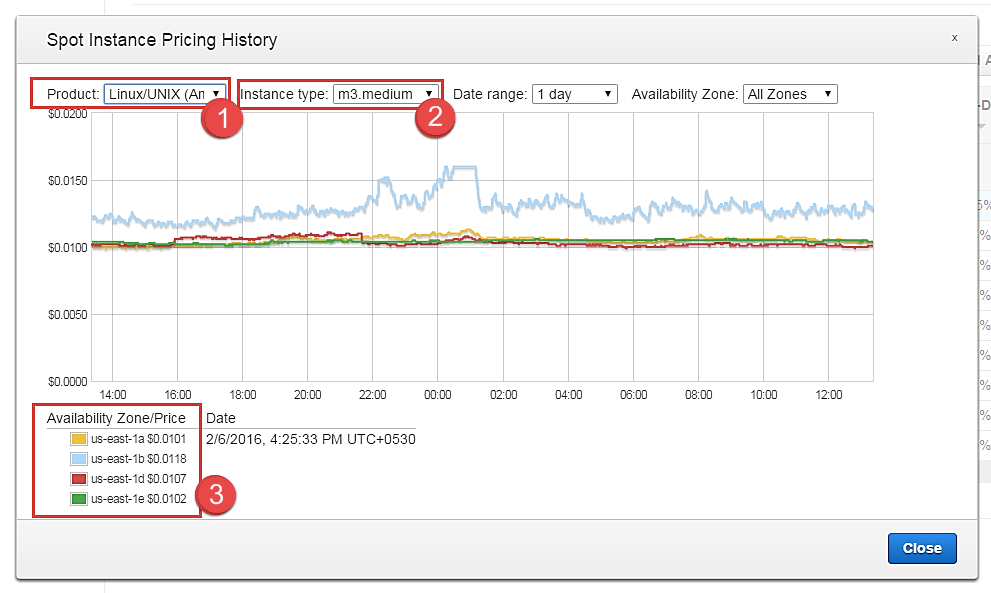
**j**ust below the bid price you can see a button of Pricing History. Click on that as shown below.



Here in Pricing History, we can see a graph depicting instance pricing trends with historical data. You can select the parameters and get an idea of the pricing of our desired instance over a period of time.

1. Select the product. We have selected our Linux AMI.
2. Select the instance type. We have selected m3.medium.
3. Note the average prices for over a day here.

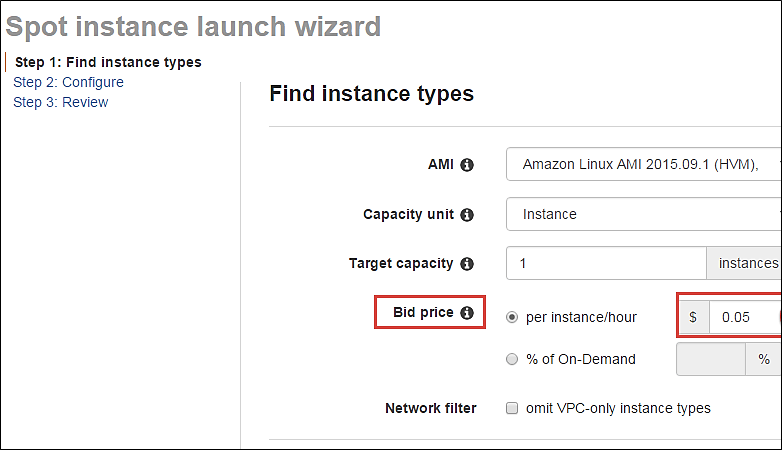
Thus, from the chart below, we can see that the instance type that we are planning to provision lies in the pricing range of $0.01xx, and it seems that Availability Zone 'us-east 1a' has the lowest price.



cont. to step 4.

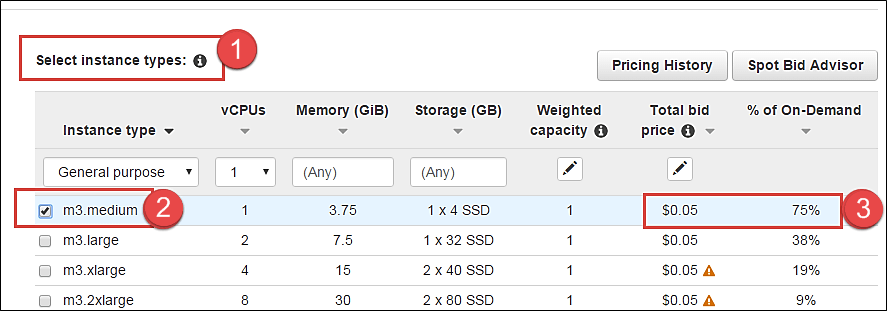
So let's come back to our step of quoting a bid price.

For the sake of maintaining our instance always available and if it falls within our budget, we can quote a higher bid price. Here we have quoted a slightly higher price of $0.05.

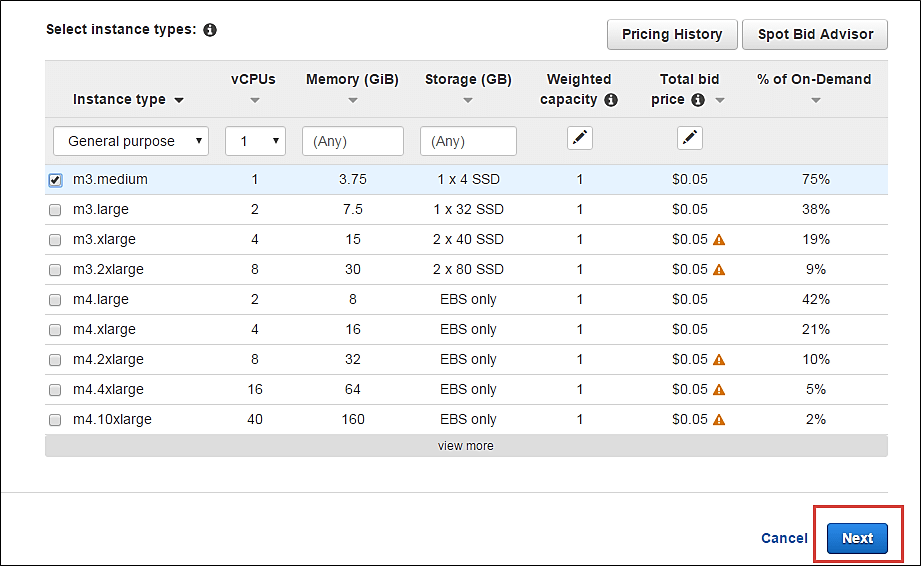


You can see some trends in the wizard itself.

1. Note the instance types section
2. Select the instance type that we are planning to provision
3. Note the price that we are planning to bid. % of on-demand shows us that our quoted price is 75% of the on-demand price for the same instance type. This means we are saving 25% per hour as compared to an on-demand instance. You can further lower the price and save costs drastically.



**Step 5)** Once we are done looking at the trends and quoting our bid price, click on next.



#### Configure the Spot instance

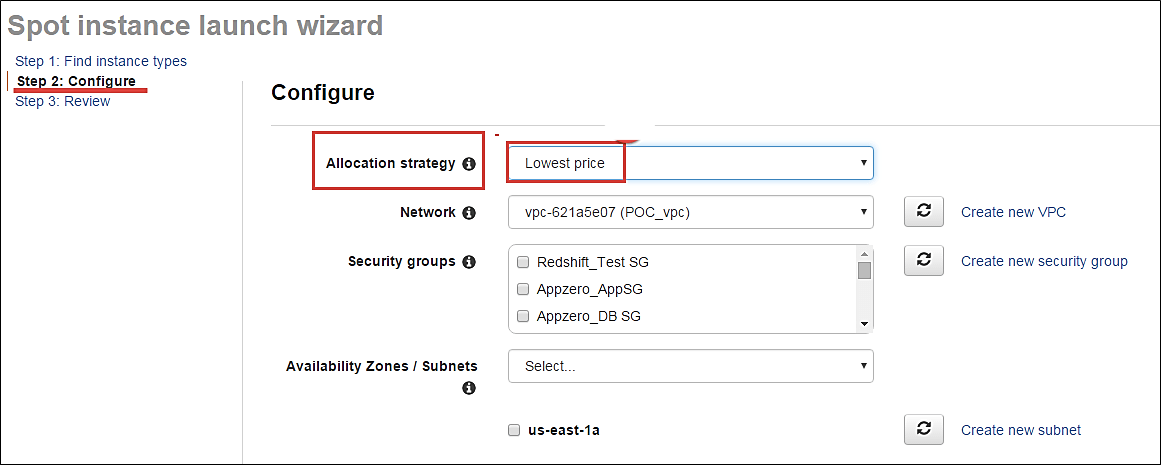
Our next step is to configure the instance, in this step of the wizard, we'll configure instance parameters like VPC, subnets, etc.

Let's take a look.

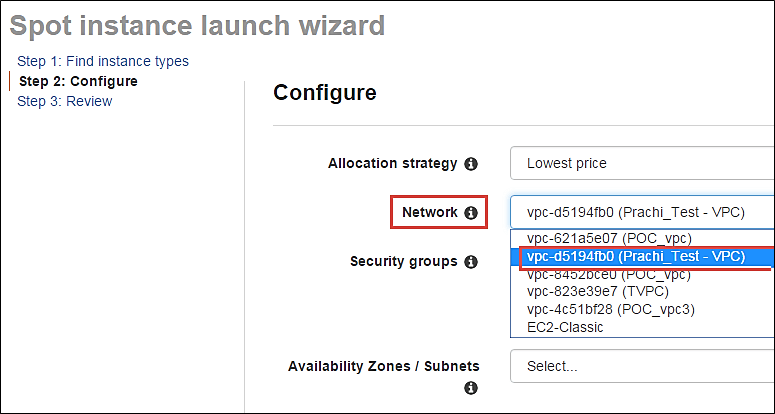
**Step 1)**Allocation Strategy – it determines how your spot request is fulfilled from the AWS's spot pools. There are two types of strategies:

* Diversified – here, spot instances are balanced across all the spot pools
* Lowest price – here, spot instances are launched from the pool which has lowest price offers

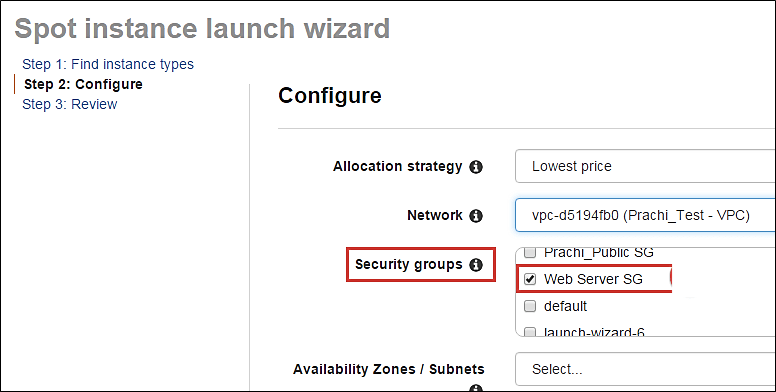
For this tutorial, we'll select Lowest Price as our allocation strategy.



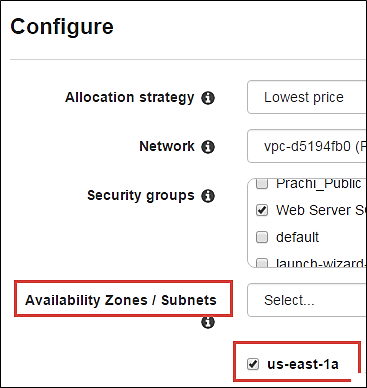
**Step 2)**Select the VPC- we'll select from the list of available VPCs that we have created earlier. We can also create a new VPC in this step.



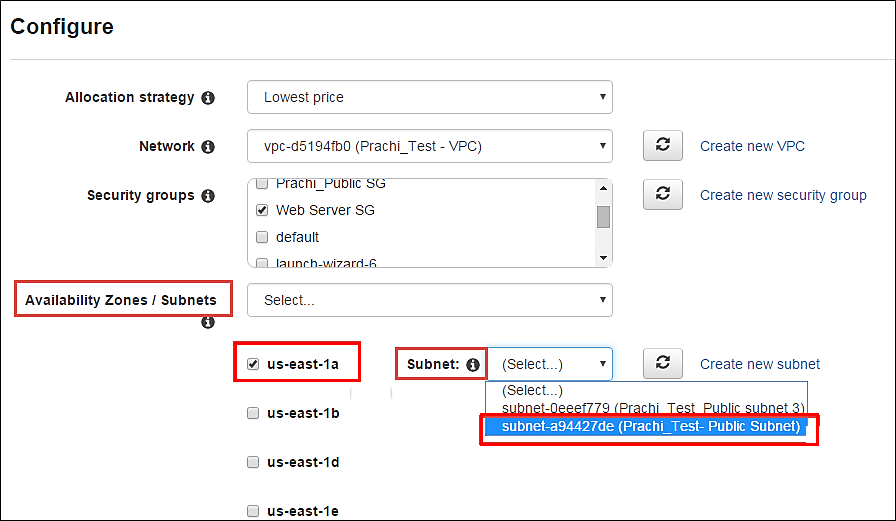
**Step 3)** Next we'll select the security group for the instance. We can select an already existing SG or create a new one.



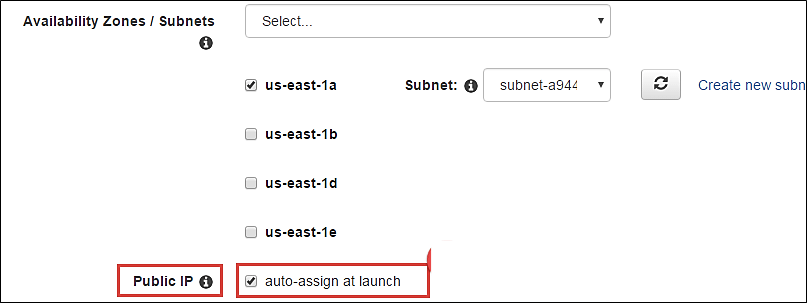
**Step 4)** Availability Zone- we'll select the AZ where we want to place our instance based on our application architecture. We are selecting AZ- us-east-1a.



**Step 5)** Subnets- we are going to select the subnet from our list of already available list.



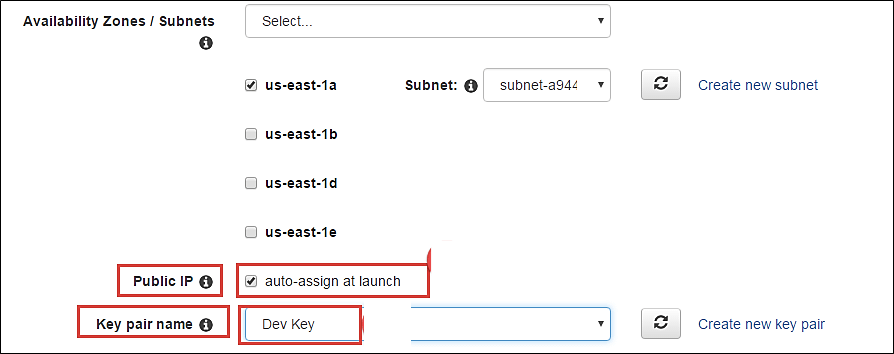
**Step 6)** Public IP- we'll choose to assign the instance a public IP as soon as it launches. In this step, you can choose if you want AWS to assign it an IP automatically, or you want to do it manually later. You can enable/ disable 'Auto assign Public IP' feature here likewise.



**Step 7)** Key pair- A key pair is a set of public-private keys.

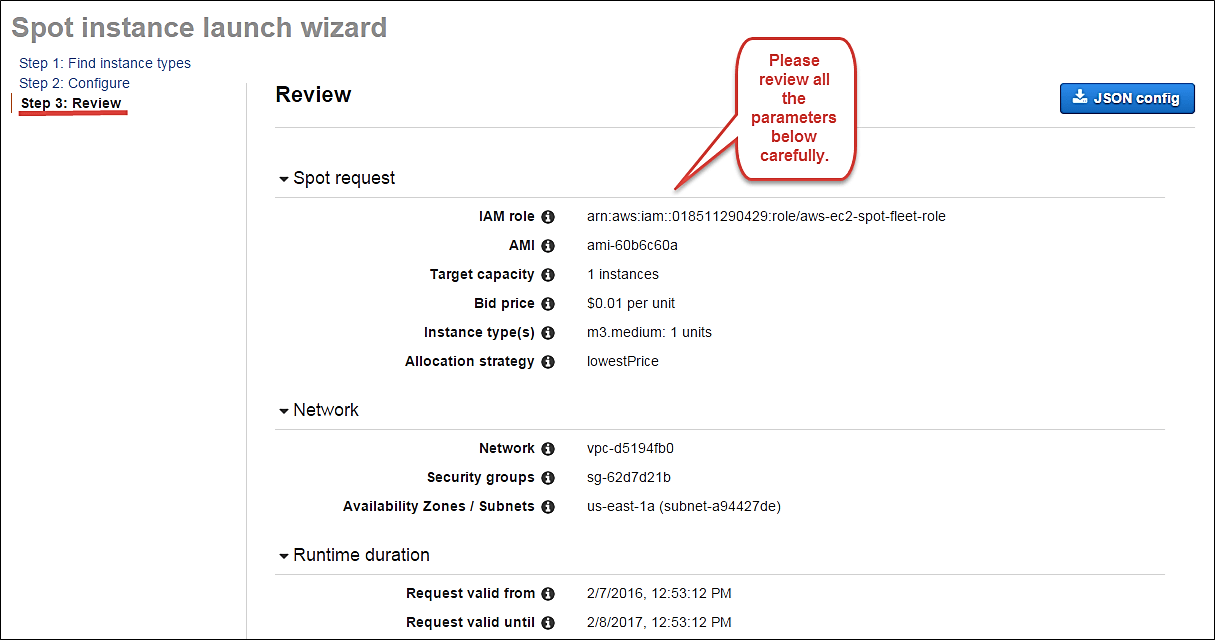
AWS stores the private key in the instance, and you are asked to download the public key. Make sure you download the key and keep it safe and secured**; if it is lost you cannot download it again.**

After selecting public IP, here we are selecting a key which we already have created in our last tutorial.

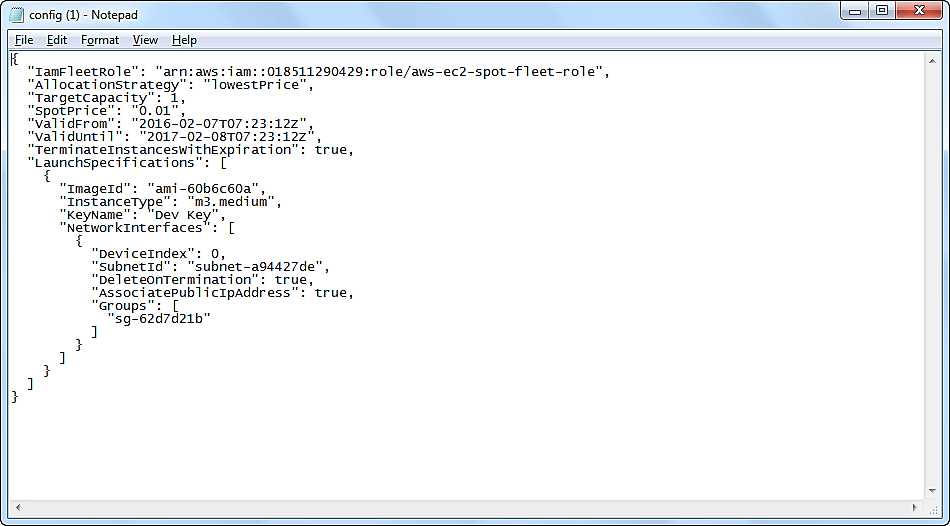


#### Reviewing your Spot instance

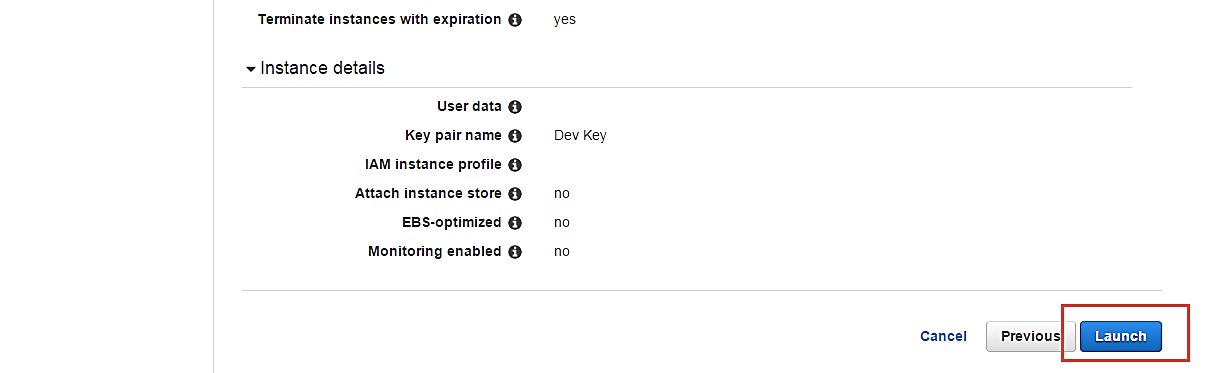
Once we are done configuring our spot instance request in the 2 steps earlier in our wizard, we'll take a look at the overall configuration.



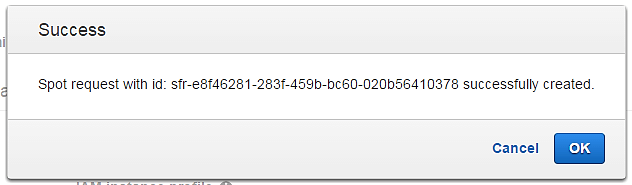
1. **We can also download a JSON file with all the configurations. Below is our JSON file.**



After we are done reviewing, we can proceed with the launching by clicking the Launch button as shown below.

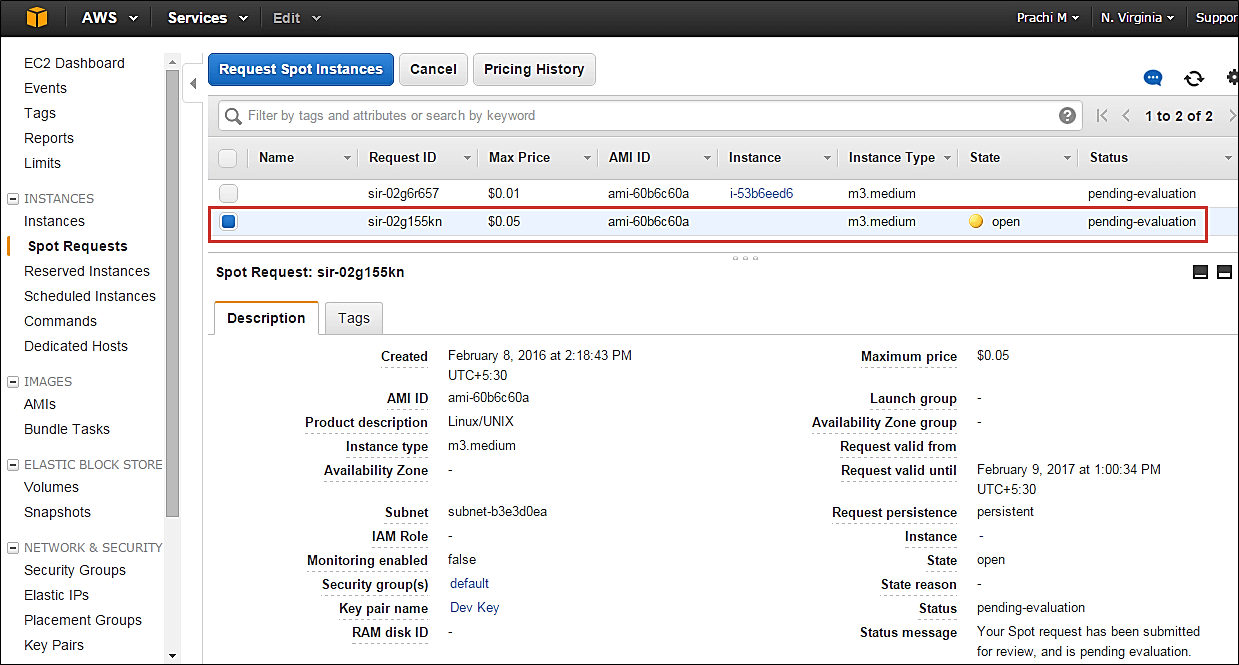


Once we select Launch, we can see a notification about the request getting created.

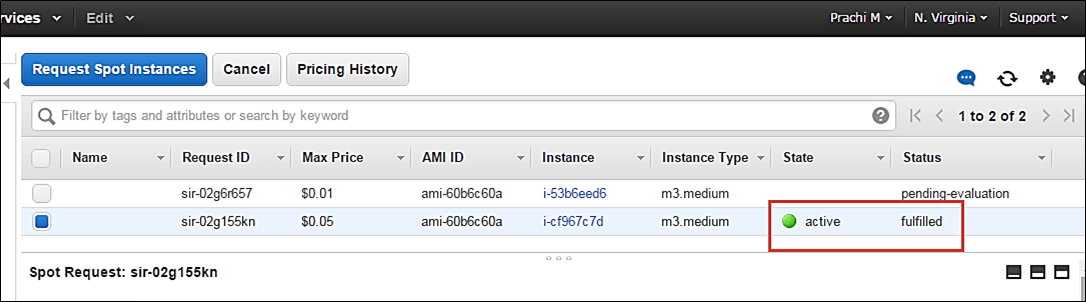


The spot request creation wizard will close, and the page will automatically direct back to the EC2 Dashboard.

You can see as shown below that the State of our request is 'open' which means that it is getting evaluated from the AWS's side. AWS EC2 will check if the required instance is available in its spot pool.



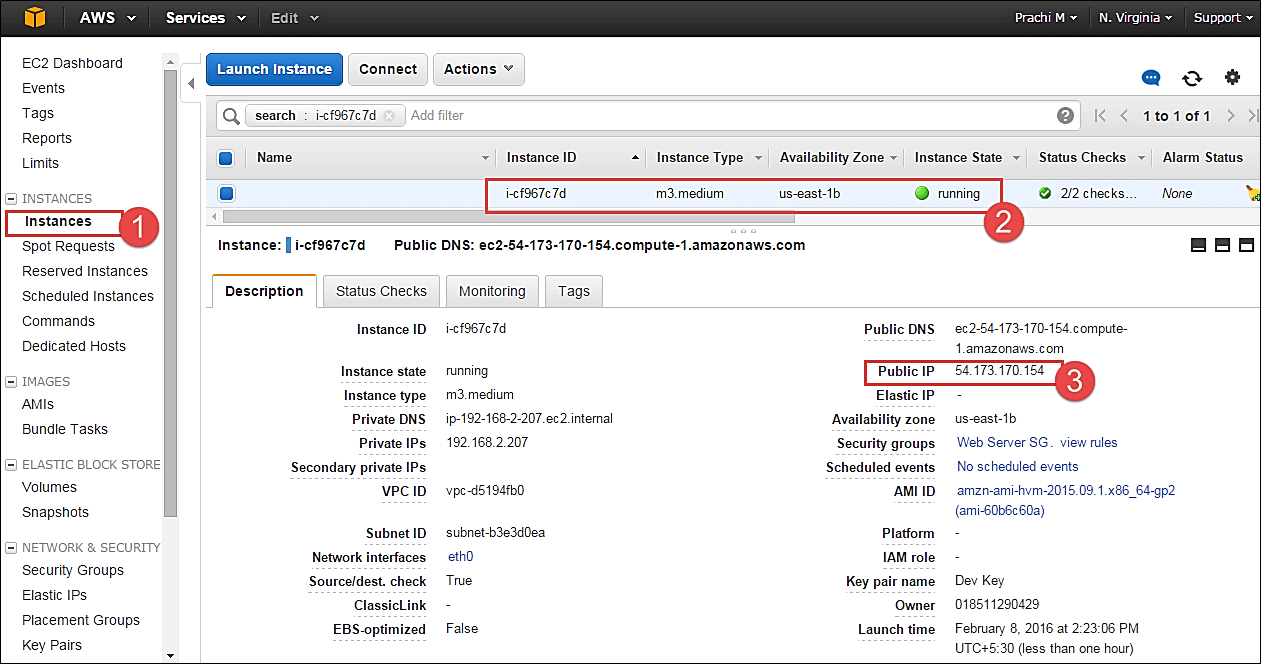
After a couple of minutes, you can see that the state is changed to 'active', and now our spot request is successfully fulfilled. You can note the configuration parameters below.



### Connecting to the spot instance

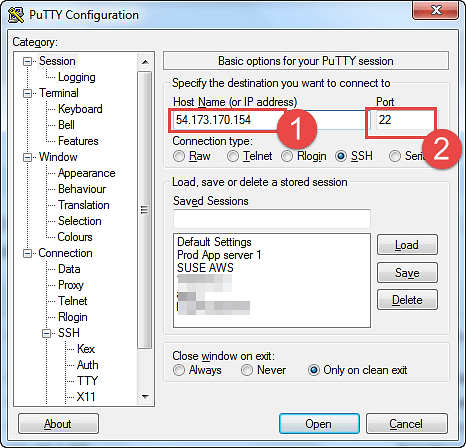
**Step 1)** In this step, you will notice following things

1. on the EC2 Dashboard, click on Instances in the left pane
2. Here you can see that the spot instance is running
3. Note down its Public IP



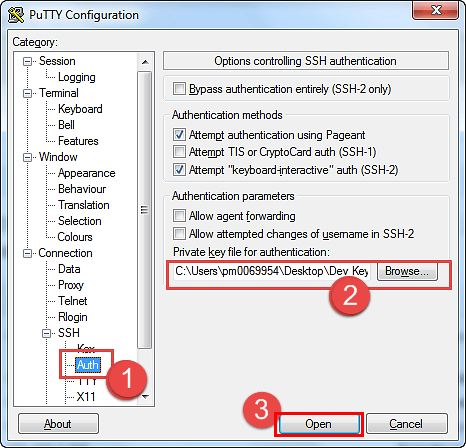
**Step 2)** Open putty in your computer.

1. Connect via the said Public IP that we noted above
2. Assign the connecting port as 22 for ssh access



**Step 3)** In this step,

1. go to Auth Browse and
2. add our .ppk key to connect to the instance
3. Click open button



You can now connect to your instance.

