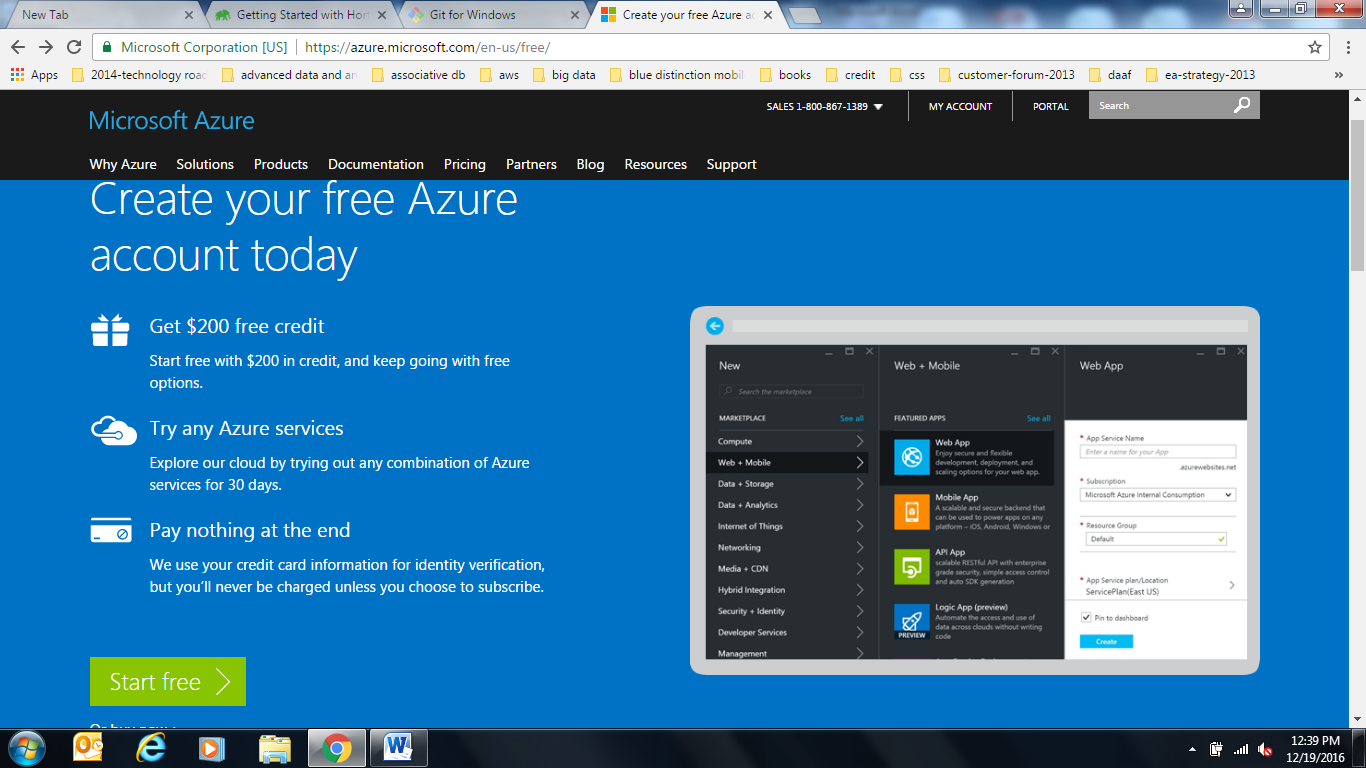
## Setting Up Your Hortonworks Sandbox

Note, I have included specific instructions below, but some of the following material is based on the “official” Hortonworks installation instructions. I have included a copy of the Hortonworks instructions on the blackboard in the “Free Books and Chapters” section. It is called “Deploying Hortonworks Sandbox on Microsoft Azure – Hortonworks.” If you encounter any issues with the following matrial check the Hortonworks file. But feel free to contact me for help.

### Step 1

Use an existing Azure account or establish a new one as follows:

1. Go to: <https://azure.microsoft.com/en-us/free/>

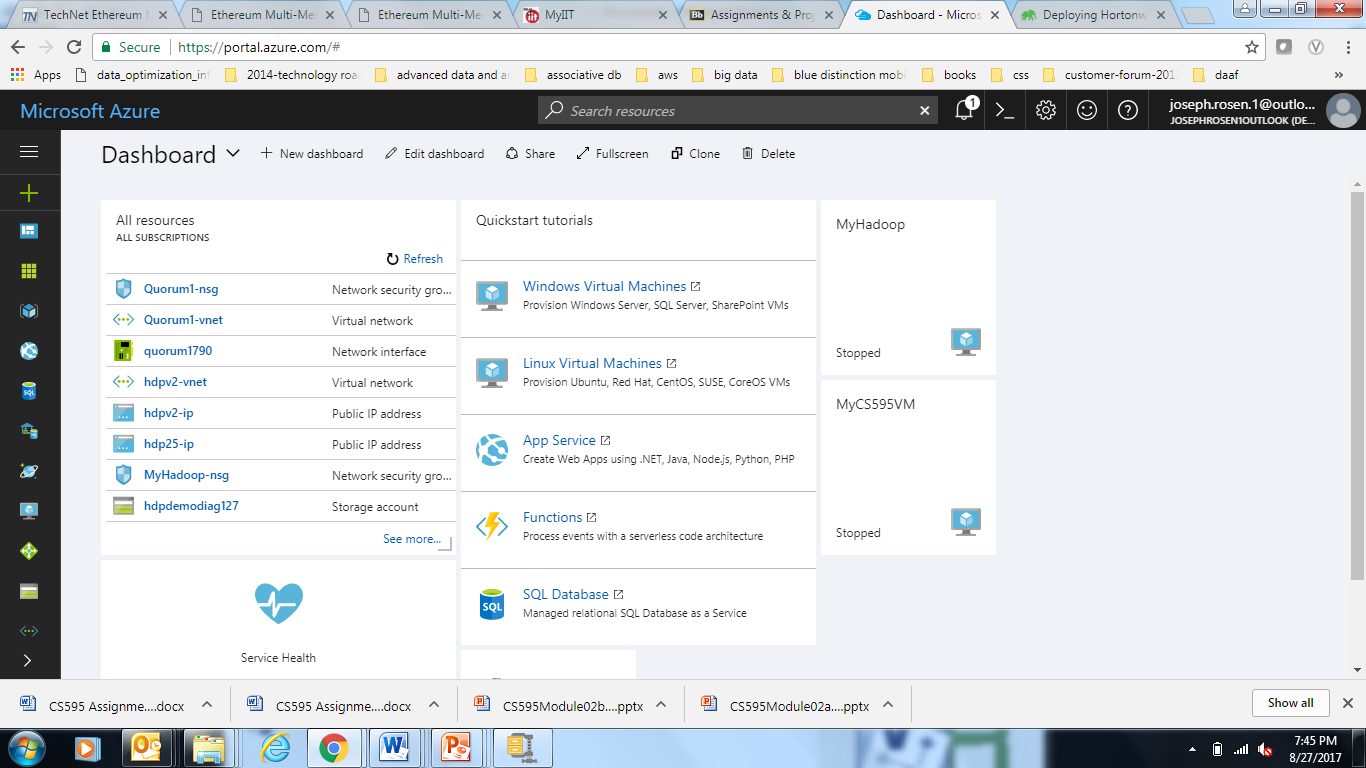


1. Set up a new account as per instructions. You will need a credit card. If you are careful the first 30 days of use will be free. The following months of our course should each cost some small amount AS LONG AS YOU STOP THE SANDBOX VIRTUAL MACHINE (vm) WHEN IT IS NOT IN USE.

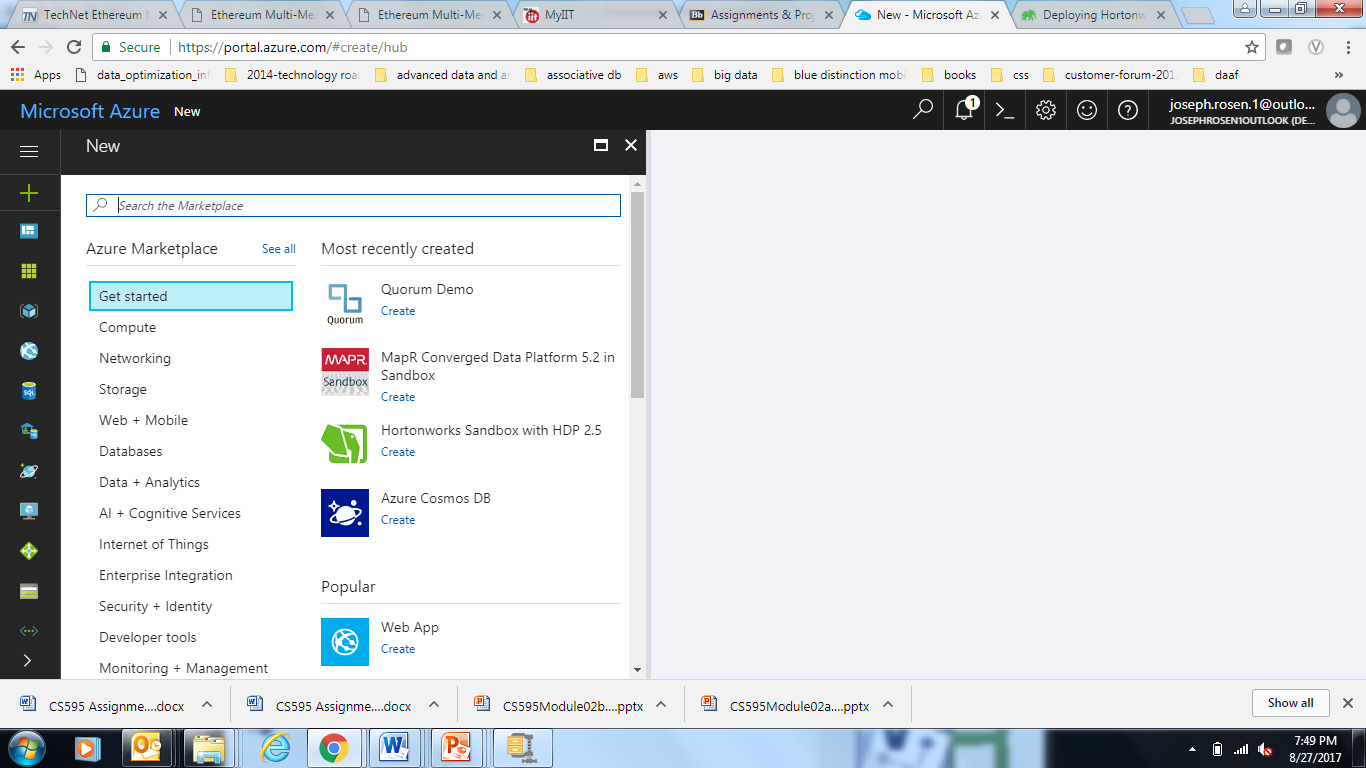
### Step 2: FIND HORTONWORKS SANDBOX ON AZURE MARKETPLACE

Log on to your Azure account and view the dashboard.

Click on the “+” in the top left corner (see below)

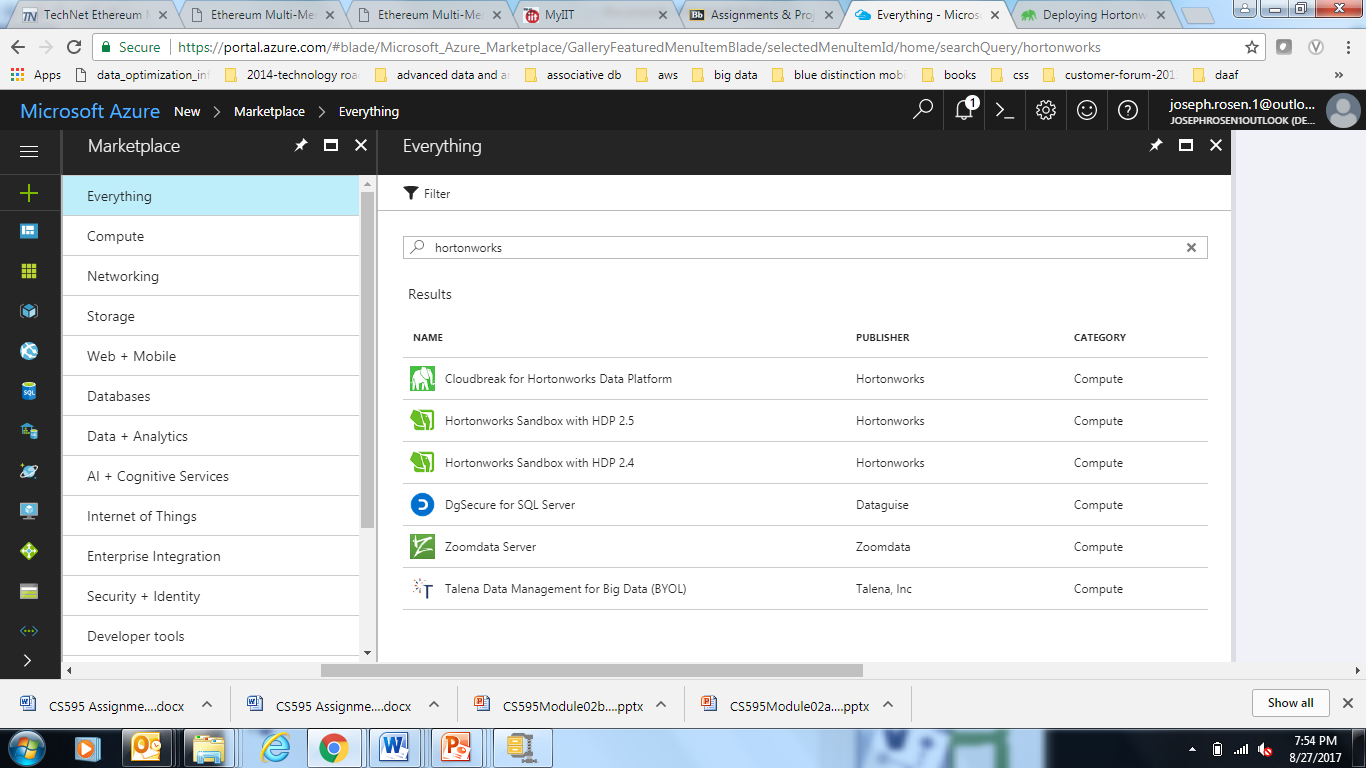


Now you should see the Azure marketplace.



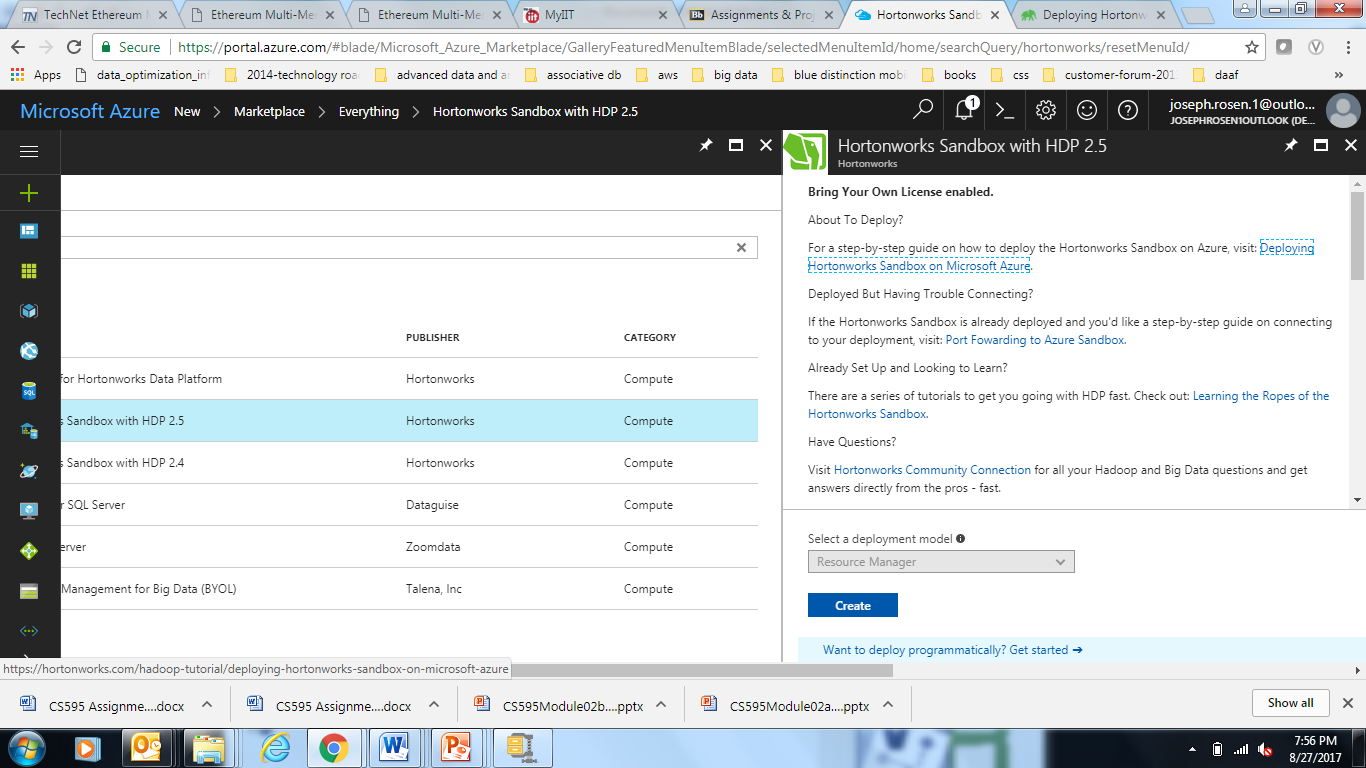
Enter “Hortonworks” into the search box at the top of the marketplace screen.

Now you should see a selection of Hortonworks products from which to choose as depicted below:



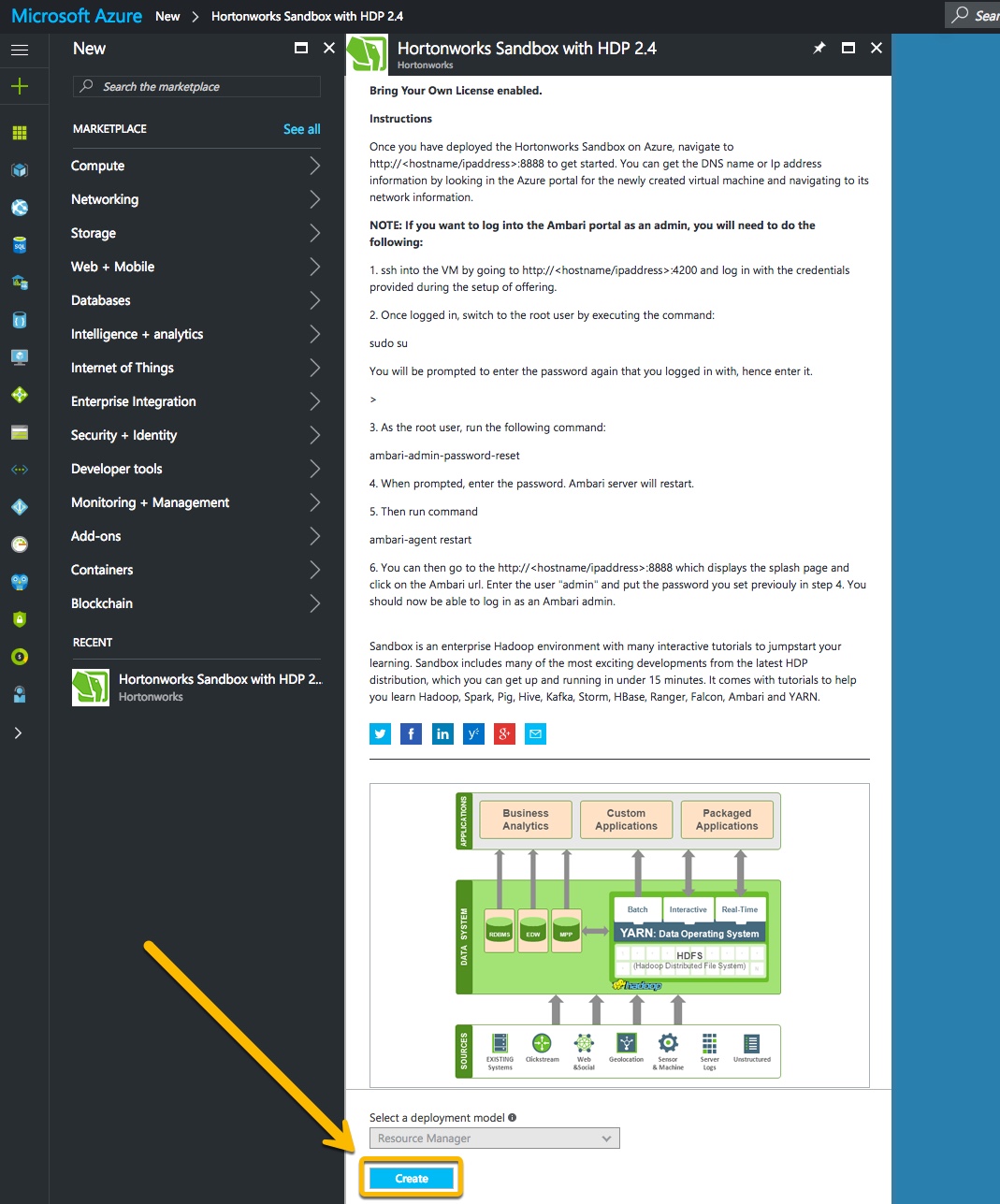
Now select the product “Hortonworks Sandbox with HDP 2.5”

At this point you should see:

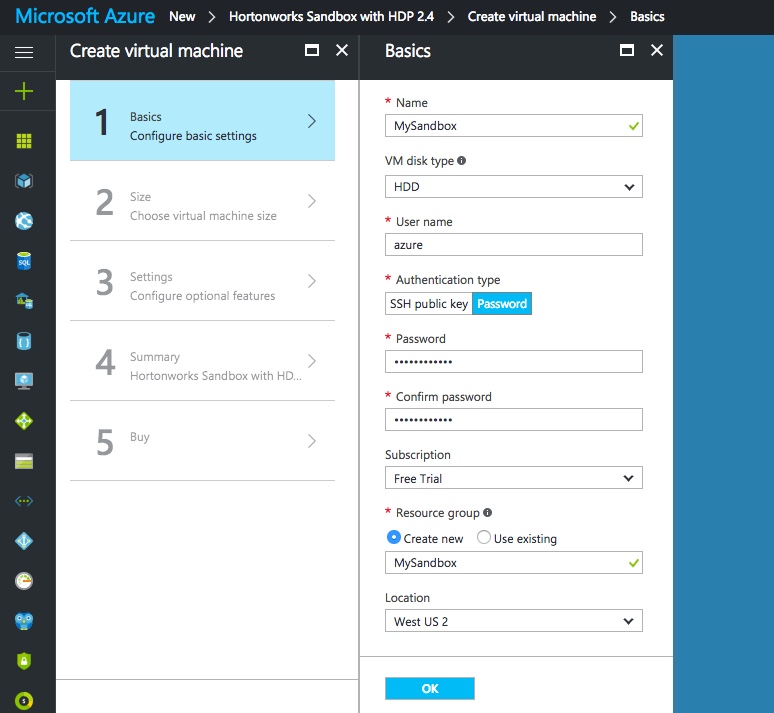


### Step 3: CREATING THE SANDBOX

An explanation of the Hortonworks Sandbox will come on the screen. When ready to begin the deployment process, select **Create** from the bottom of the screen.



Fill out some basic sandbox configuration settings. Scroll down for an explanation of the different fields.



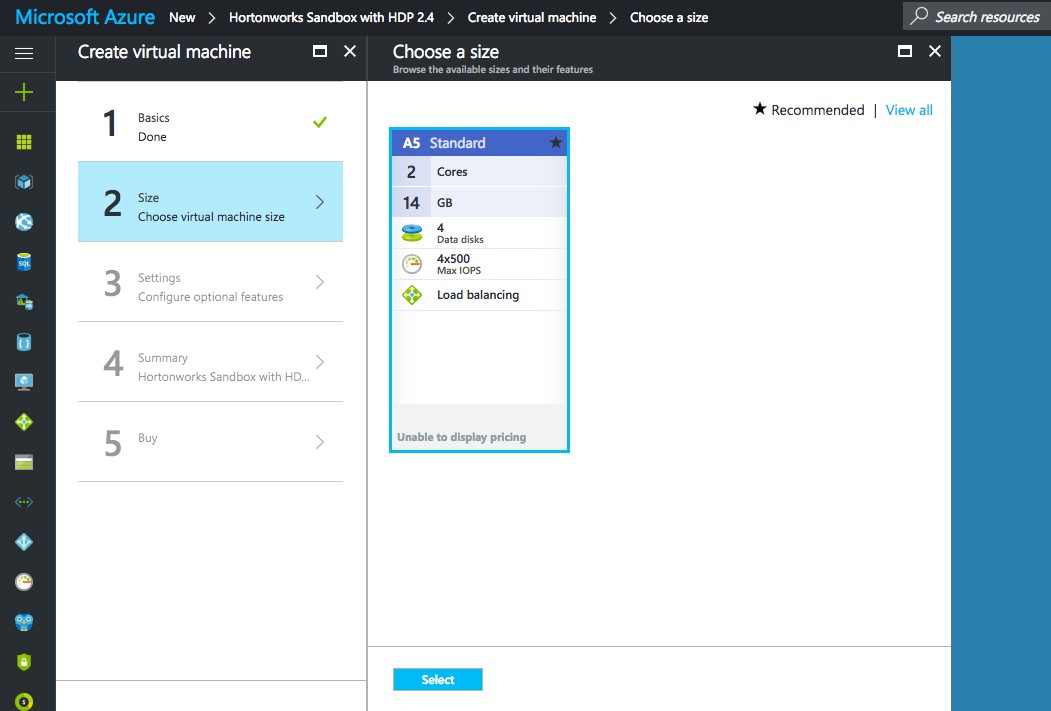
The different fields:

* **Name**: This is the name you want to use to reference the machine. In the example above, we use the name “MySandbox”
* **User name**: The name of the user account that will be used to log into the machine. Throughout these tutorials, we will use azure as the user name.
* **Authentication type**: By default, the machine will be deployed and allow you to connect via SSH key or password. In this example, we opt to use a password.
* **Subscription**: The subscription to deploy the machine under. Select one already in your list.
* **Resource group**: The name of the resource group to create, or use an existing one. Here, we create a new one with the same name as the machine itself.
* **Location**: Which region in the Azure offering to deploy the machine to.

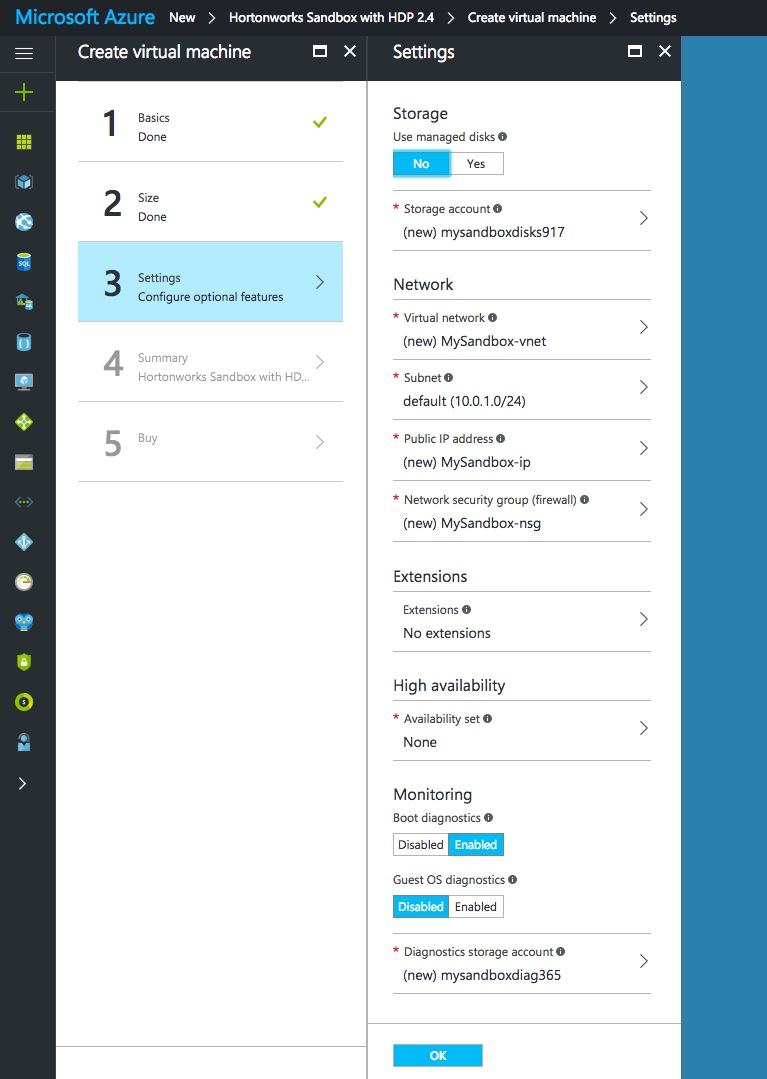
Note: Make sure to write down or remember your username and password. If using SSH, ensure you have the corresponding private key. Otherwise, you will not be able to log in to the machine.

The next step is to choose a size for the virtual machine. It is recommended to use a machine with A4 specifications, or higher.

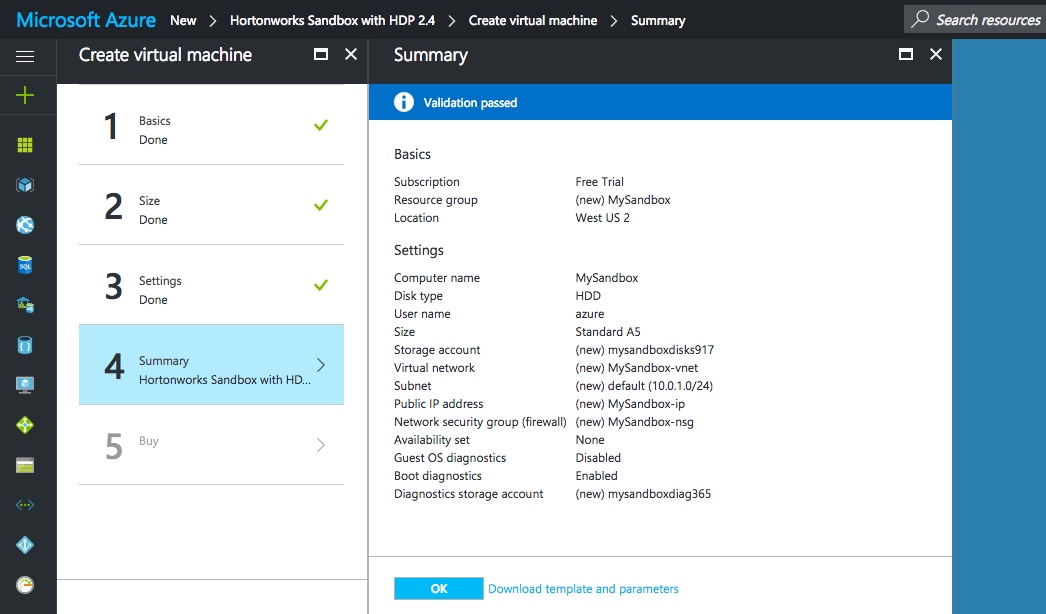
Make sure to select “View All.” Then look for a machine with 2 Cores or more, 14 Gb of RAM or more, and 28GB or more of local disk. This could be “D3\_V2, D11\_V2, D12\_V2” or similar.



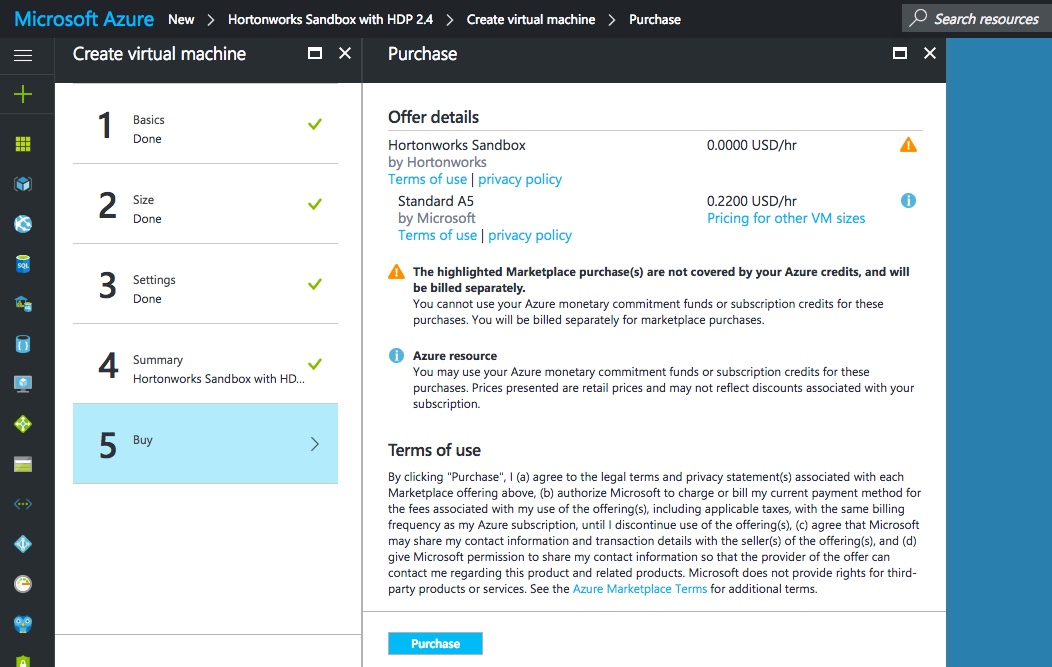
Optional settings to configure. You can safely leave these at their defaults.



Look over the summary and continue when ready.



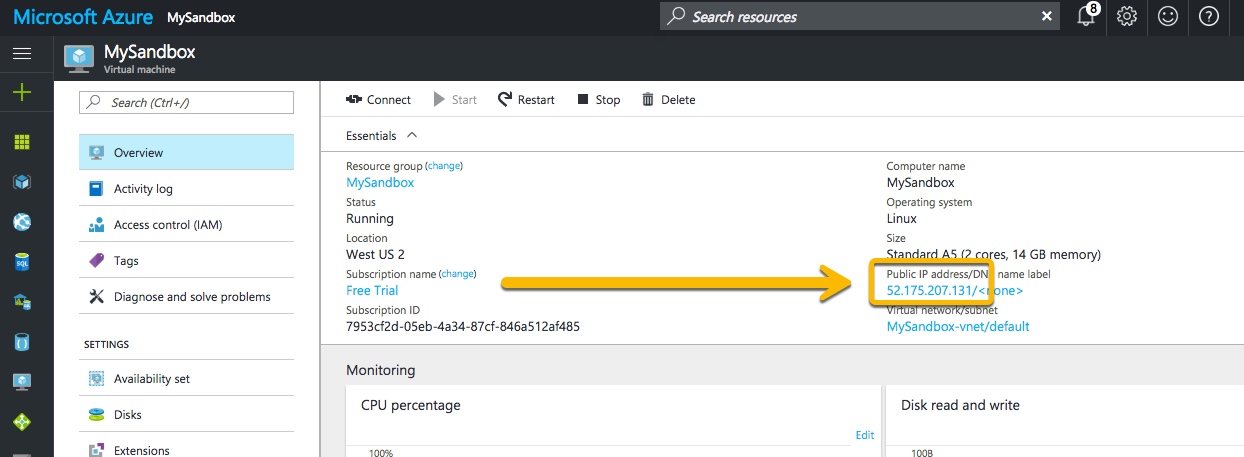
Alright, we’re ready to deploy! Review the offer details and purchase when ready.



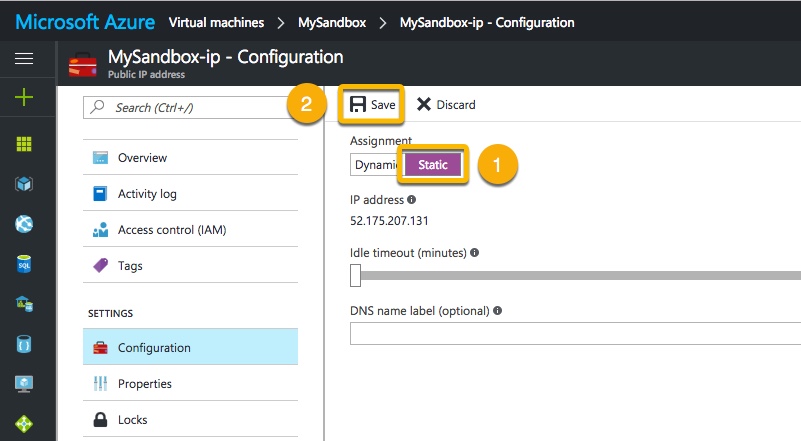
Once the offer is submitted by selecting **Purchase**, the sandbox will take a few minutes to set up and deploy. After deployment is complete, we can move on to connecting to the sandbox.

### Step 4: SET A STATIC IP

Once the machine is deployed, it’s overview will appear on the screen. Find the sandbox’s public IP address and click on it.



Clicking on the IP address will bring up the IP configuration panel. Select **Static** as the Assignment, and then make sure to save your changes. This will keep the sandbox from changing IP addresses each time it’s rebooted.



### Step 5: CONFIGURE SSH TUNNELING

SSH tunneling allows us a way to port forward securely, without actually opening the machine’s ports for the entire world to access. Follow these steps to access the endpoints of your Azure deployment from your computer.

USING SSH

If you have completed assignment two then you have a config file in your “.ssh” directory. Save a copy in case of emergency to “config.old” Now edit the original “config” as indicated below.

Use your favorite editor and edit your ~/.ssh/config file. For example:

vi ~/.ssh/config

Enter the following configuration, replacing the **HostName** IP with the public IP of your instance and also replace the **User** with the user name you configured above for the sandbox.

Note: Spacing and capitalization is important.

Host azureSandbox (or any other host alias)

Port 22

User <your-specified-azure-username-here>

HostName <your-azure-public-ip-here>

LocalForward 8080 127.0.0.1:8080

LocalForward 8888 127.0.0.1:8888

LocalForward 9995 127.0.0.1:9995

LocalForward 9996 127.0.0.1:9996

LocalForward 8886 127.0.0.1:8886

LocalForward 10500 127.0.0.1:10500

LocalForward 4200 127.0.0.1:4200

LocalForward 2222 127.0.0.1:2222

Save and close the file. Now SSH into the Azure machine by using the **Host** alias we just created, by using the command below. This will connect automatically using the IP address specified in the config file.

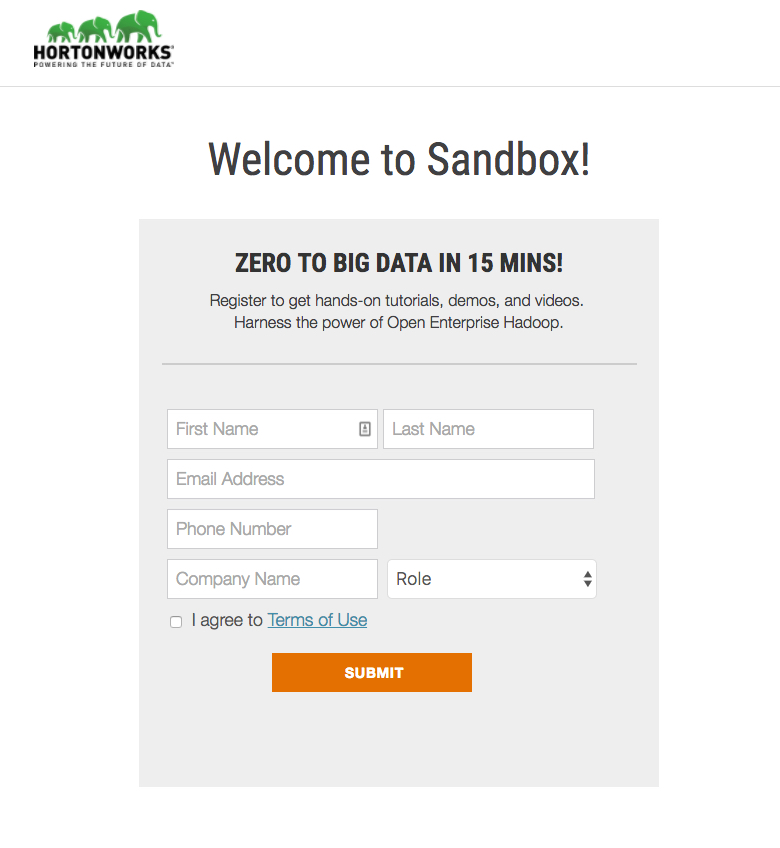
ssh azureSandbox

You’ll be asked for a password, which is the one you set during initial configuration on Azure.

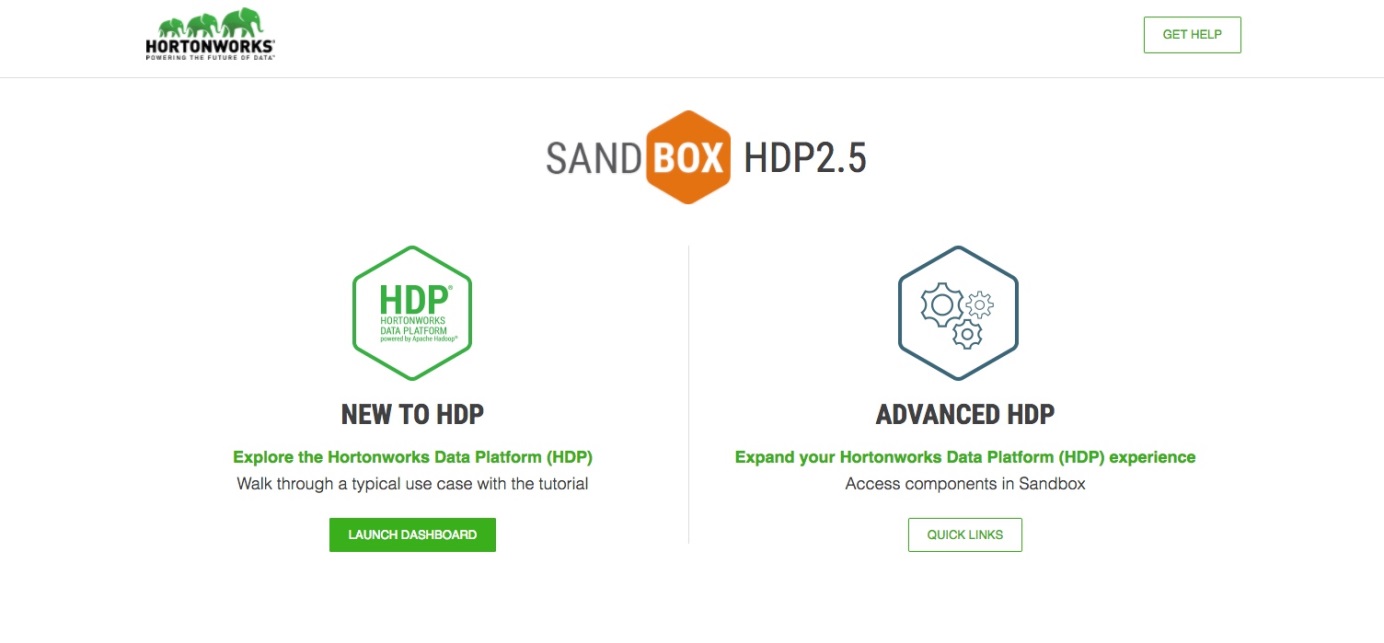
That’s it! Keep this SSH connection open for the duration of your interaction with the sandbox on Azure.

### Step 6: SPLASH SCREEN

Now that you’ve port forwarded by following the tutorial linked above, you can explore the sandbox as you see fit. Point your browser to [http://localhost:8888](http://localhost:8888/) for the sandbox’s splash screen.



Fill out the form and hit **Submit** to access the sandbox.



That’s it! Keep this SSH connection open for the duration of your interaction with the sandbox on Azure.

You can now access all forwarded ports by pointing a browser to [http://localhost:portNumber](http://localhost:portNumber/). For example: [http://localhost:8888](http://localhost:8888/) will connect to the Azure machine and sandbox over port 8888.

SSH tunneling allows us a way to port forward securely, without actually opening the machine’s ports for the entire world to access.

### Step 7:

Now you can access your Hadoop sandbox as before by establishing an ssh session to the maria\_dev account as with assignment #2.