



DSA Course Instance Guide

DS201 - DataStax Enterprise Foundations of Apache Cassandra™



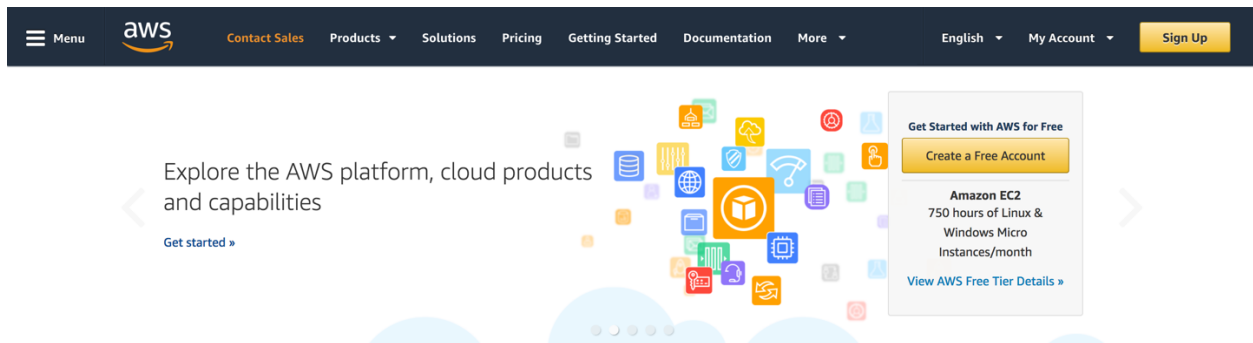
Table of Contents

OBTAINING AN AWS ACCOUNT	3
STARTING AN INSTANCE	4
AWS Management Console	4
Launching an Instance	6
Choose an Instance Type	7
Configure Instance Details	7
Add Storage	7
Add Tags	7
Configure Security Group	8
Required ports	8
Example security group setting	9
SSH Key Pair	10
Finding the IP address of the instance	10
TERMINATING AN INSTANCE	11

Obtaining an AWS Account

You will need your own AWS account, or have one provided from your company or organization in order to start up an instance.

To sign up for an AWS account, go to <https://aws.amazon.com> and click on **Create a Free Account** and follow the instructions there. Although some usage of AWS may be free, DataStax Academy instances do not run on the AWS Free Tier and will cost money to run.



You may find additional documentation about AWS here: <https://aws.amazon.com/documentation/>

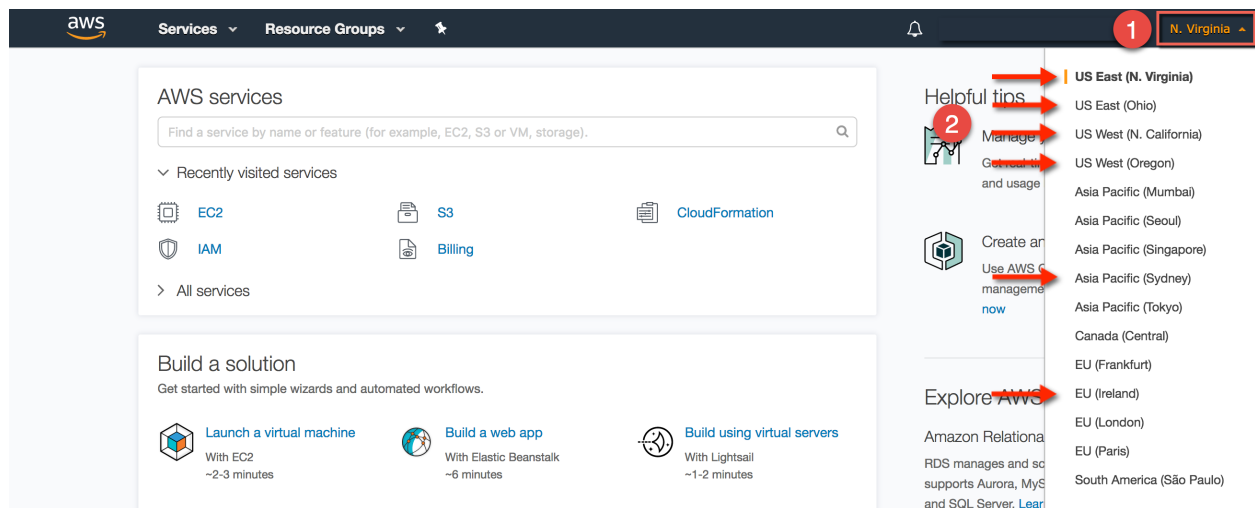
Starting an Instance

AWS Management Console

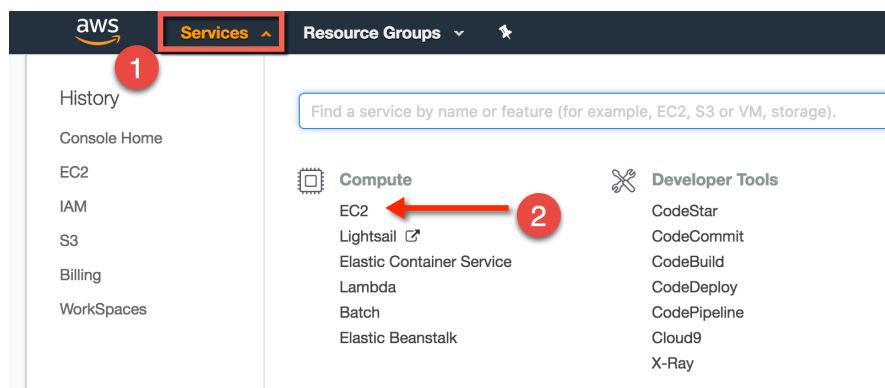
To start an instance, you will first need to sign into the AWS Management Console at <https://aws.amazon.com/console>.

After you sign in, make sure to select one of the supported regions, whichever region is closest to you:

US East (N. Virginia) – us-east-1
US East (Ohio) – us-east-2
US West (N. California) – us-west-1
US West (Oregon) – us-west-2
Asia Pacific (Sydney) – ap-southeast-2
EU (Ireland) – eu-west-1



Afterwards click on **Services > EC2**.



You should now be in the *EC2 Dashboard*.

aws

Services

Resource Groups

EC2 Dashboard

Events

Tags

Reports

Limits

INSTANCES

Instances

Launch Templates

Spot Requests

Reserved Instances

Dedicated Hosts

IMAGES

AMIs

Bundle Tasks

ELASTIC BLOCK STORE

Resources

You are using the following Amazon EC2 resources in the Asia Pacific (Seoul) region:

0 Running Instances

0 Elastic IPs

0 Dedicated Hosts

0 Snapshots

0 Volumes

0 Load Balancers

0 Key Pairs

2 Security Groups

0 Placement Groups

Create Instance

To start using Amazon EC2 you will want to launch a virtual server, known as an Amazon EC2 instance.

Launch Instance

Note: Your instances will launch in the Asia Pacific (Seoul) region

Service Health

Scheduled Events

Launching an Instance

From the *EC2 Dashboard*, click on the button to **Launch Instance**. Click on the link **Community AMIs**. Using the table below, get the AMI ID for region you are starting the instance in and search for that.

Step 1: Choose an Amazon Machine Image (AMI) [Cancel and Exit](#)

An AMI is a template that contains the software configuration (operating system, application server, and applications) required to launch your instance. You can select an AMI provided by AWS, our user community, or the AWS Marketplace; or you can select one of your own AMIs.

Quick Start 2 Search community AMIs 1 to 50 of 104,324 AMIs

My AMIs

AWS Marketplace

Community AMIs 1

Operating system

- ☐ Amazon Linux
- ☐ Cent OS
- ☐ Debian

amzn-ami-hvm-2018.03.0.20180412-x86_64-gp2 - ami-467ca739 **Select**

Amazon Linux AMI 2018.03.0.20180412 x86_64 HVM GP2 64-bit

Root device type: ebs Virtualization type: hvm ENA Enabled: Yes

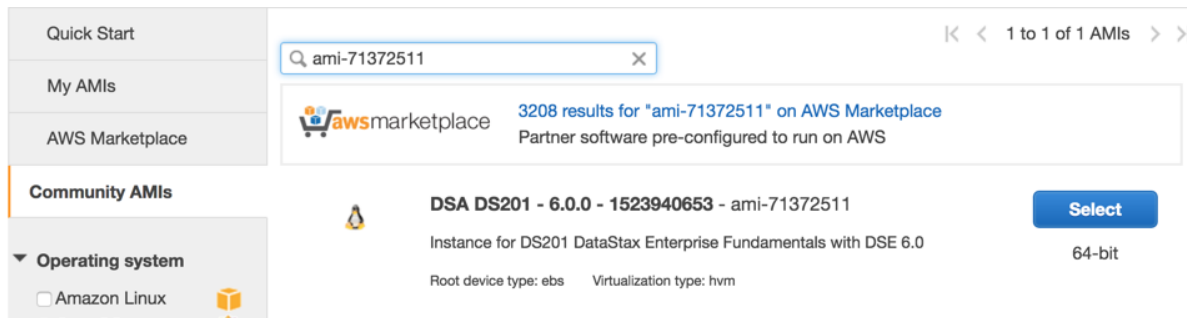
amzn2-ami-hvm-2017.12.0.20180328.1-x86_64-gp2 - ami-f973ab84 **Select**

Amazon Linux 2 LTS Candidate AMI 2017.12.0.20180328.1 x86_64 HVM GP2 64-bit

Root device type: ebs Virtualization type: hvm ENA Enabled: Yes

Region	AMI ID
N. Virginia (us-east-1)	ami-f160c08e
Ohio (us-east-2)	ami-c10635a4
N. California (us-west-1)	ami-71372511
Oregon (us-west-2)	ami-73ddb00b
Ireland (eu-west-1)	ami-5bfadc22
Sydney (ap-southeast-2)	ami-a18942c3

The search should display only one AMI, which you can then choose with the **Select** button.



Choose an Instance Type

In Step 2, select an appropriate instance type to launch. The instance will need at least 4 GB of RAM, with the recommended instance type being **t2.medium**. Once the instance type is selected, click on the button **Next: Configure Instance Details**.

Note: You will be charged to start up an instance and to keep it running. Each instance type has a different cost. Please make sure you choose an instance type appropriate for the amount you are willing to spend. The same instance types in different regions may also be priced differently. A list of instance types and their costs can be found here: <https://aws.amazon.com/ec2/pricing/on-demand>

Configure Instance Details

In Step 3, you should not need to make any changes here. The course DS201 will need only 1 instance in order to do all of the exercises. After setting the number of instances, click on **Next: Add Storage**.

Add Storage

There is nothing that you are required to do in Step 4, so click on **Next: Add Tags** to proceed if you do not need to add any additional storage.

Add Tags

In Step 5, tags are not required but can be useful to keep track of the instances that you launch. It is recommended to add at least one tag to name your instance.

Click on the button **Add Tag**. Set the **Key** as *Name* and the **Value** to something descriptive, like *DS201 Instance*. You can add additional tags if desired, but once you're done click on the button **Next: Configure Security Group**.

1. Choose AMI 2. Choose Instance Type 3. Configure Instance 4. Add Storage 5. Add Tags 6. Configure Security Group 7. Review

Step 5: Add Tags

A tag consists of a case-sensitive key-value pair. For example, you could define a tag with key = Name and value = Webserver.

A copy of a tag can be applied to volumes, instances or both.

Tags will be applied to all instances and volumes. [Learn more](#) about tagging your Amazon EC2 resources.

Key (127 characters maximum)	Value (255 characters maximum)	Instances ⓘ	Volumes ⓘ
Name	DS201 Instance	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input type="button" value="Add another tag"/> (Up to 50 tags maximum)			

Configure Security Group

The security group sets the firewall rules for the instance you are launching. The rules need to be configured to allow you to interact with the instance, whether it is through SSH, DataStax Studio, OpsCenter, or some other application. You can also set additional firewall rules if you want to interact with the instance in a way outside the scope of the exercises.

Required ports

The following table lists all of the possible ports used throughout all of the different DSA course exercises.

Port	Description
22	SSH
80	HTTP
443	HTTPS
4040	Spark Application Web UI
7000	DSE Inter-node Cluster Communication Port
7001	DSE Inter-node Cluster Communication Port (SSL)
7080	Spark Master Web UI
7081	Spark Worker Web UI
7199	JMX
8888	DataStax OpsCenter
8983	Solr Admin Web UI
9042	Native Transport Protocol
9142	Native Transport Protocol (SSL)
9091	DataStax Studio

Example security group setting

Type	Protocol	Port Range	Source
SSH	TCP	22	0.0.0.0/0
HTTP	TCP	80	0.0.0.0/0
HTTPS	TCP	443	0.0.0.0/0
All ICMP – IPV4	ICMP	0 - 65535	<i>Restrict to security group</i>
Custom TCP Rule	TCP	4040	0.0.0.0/0
Custom TCP Rule	TCP	7000	<i>Restrict to security group</i>
Custom TCP Rule	TCP	7001	<i>Restrict to security group</i>
Custom TCP Rule	TCP	7080	0.0.0.0/0
Custom TCP Rule	TCP	7081	0.0.0.0/0
Custom TCP Rule	TCP	7199	0.0.0.0/0
Custom TCP Rule	TCP	8888	0.0.0.0/0
Custom TCP Rule	TCP	8983	0.0.0.0/0
Custom TCP Rule	TCP	9042	0.0.0.0/0
Custom TCP Rule	TCP	9142	0.0.0.0/0
Custom TCP Rule	TCP	9091	0.0.0.0/0

Once a security group is selected or created, click on the button **Review and Launch**.

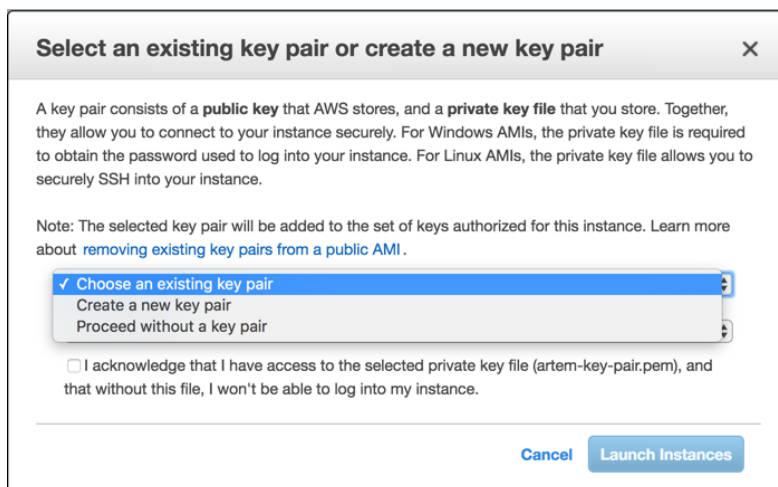
SSH Key Pair

In Step 7, you can review the instance settings to check if it is configured properly. If everything looks ok, click on the button **Launch**. A window will pop up to have you select an existing key pair or to create a new key pair.

If you already have an existing key pair set up, and have access to the private key file, then you use that key pair for your instance(s).

If you do not have an existing key pair, select the option **Create a new key pair** and then enter a name for the new key pair. Click on the button **Download Key Pair** and save the private key file to a location that you can remember; you will need that file to SSH into the instance later.

Warning: You should never select the option **Proceed without a key pair**, as this will prevent you from being able to access your instance.



The screenshot shows a modal dialog box titled "Select an existing key pair or create a new key pair" with a close button (X) in the top right corner. The dialog contains the following text:

A key pair consists of a **public key** that AWS stores, and a **private key file** that you store. Together, they allow you to connect to your instance securely. For Windows AMIs, the private key file is required to obtain the password used to log into your instance. For Linux AMIs, the private key file allows you to securely SSH into your instance.

Note: The selected key pair will be added to the set of keys authorized for this instance. Learn more about [removing existing key pairs from a public AMI](#).

Below the note is a dropdown menu with three options: "Choose an existing key pair" (selected with a checkmark), "Create a new key pair", and "Proceed without a key pair".

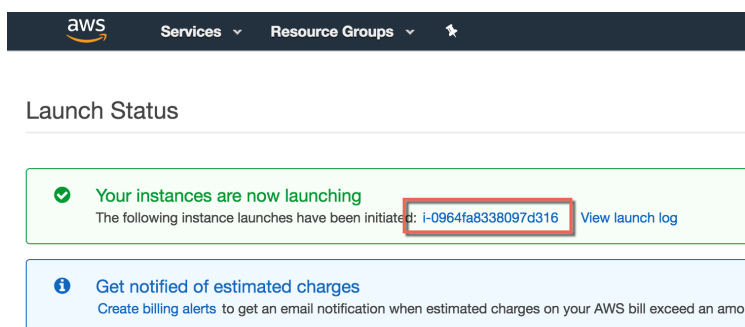
Below the dropdown is a checkbox labeled "I acknowledge that I have access to the selected private key file (artem-key-pair.pem), and that without this file, I won't be able to log into my instance." The checkbox is currently unchecked.

At the bottom right of the dialog are two buttons: "Cancel" and "Launch Instances".

Once the key pair is selected or created, click on the button **Launch Instances**.

Finding the IP address of the instance

When your instance(s) launches, you will be transferred to the **Launch Status** page. Click on the link with the instance ID to check out the instance details.



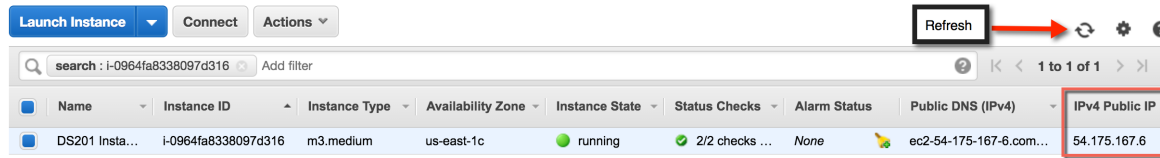
The screenshot shows the AWS "Launch Status" page. At the top is the AWS navigation bar with the logo and "Services" and "Resource Groups" dropdown menus. Below the bar, the page title "Launch Status" is displayed.

The main content area has a green background with a checkmark icon and the text "Your instances are now launching". Below this, it says "The following instance launches have been initiated: i-0964fa8338097d316" with a red box highlighting the instance ID "i-0964fa8338097d316". To the right of the instance ID is a link "View launch log".

Below the green box is a blue box with an information icon and the text "Get notified of estimated charges". Below this, it says "Create billing alerts to get an email notification when estimated charges on your AWS bill exceed an amount".

Find the value for **IPv4 Public IP**, which is the IP address of the instance that you will use to SSH, or access files or applications from the web browser.

If the instance is still starting up, the IP address may not be available yet. Wait a few minutes, and then refresh the instance details to see if the IP is available.

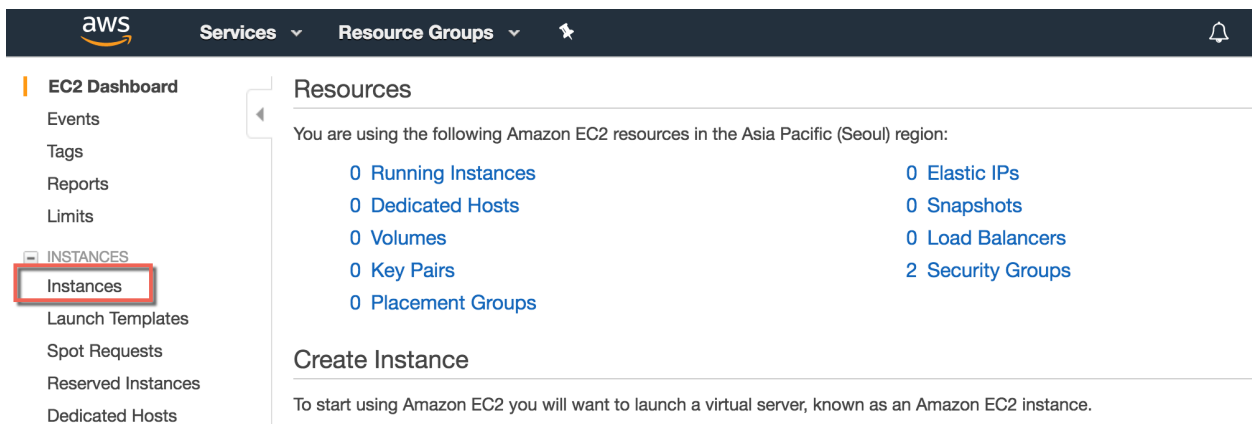


With the IP address, you can SSH to the instance once the instance has reached the **running** state.

Terminating an Instance

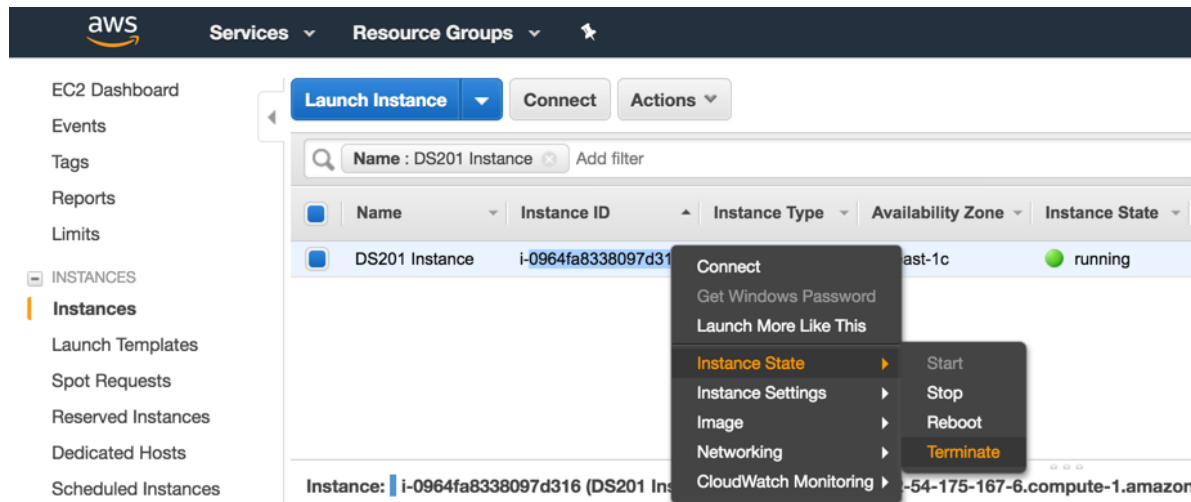
Instances incur a cost while running. When you are done with the exercises, or wish to stop for a period of time, we recommend that you terminate the instance.

In the EC2 page, click on the link **Instances** to go to the list of instances.



From the list of instances, find the instance you want to terminate. Right-click on that instance, and select **Instance State > Terminate**. The instance will change to the state *shutting-down* and will be completely terminated once the state is *terminated*.

Alternatively, you can also select the instance(s) and then click on the button **Actions > Instance State > Terminate**.



Note that terminating the instance will delete all data on the instance. If there is anything that you wish to save, please transfer that data from the instance before terminating.

There is another action called *Stop* that will also shut down the instance but save the EBS disk and allows the instance to start again with the same disk. In general, we do not recommend using this action for the following reasons:

- You are still charged for storage of the EBS disk, although not for the instance itself
- The ephemeral disk is not saved, which is usually where DSE data is stored
- The stopped instance may be assigned a different IP address when you start it again, and will potentially break your DSE configuration