



Lesson 2: Cloud Computing



Launch an Instance

SEARCH



RESOURCES



CONCEPTS



- ✓ 1. Overview
- ✓ 2. Create an AWS Account
- ✓ 3. Get Access to GPU Instances
- ✓ 4. Apply Credits
- ✓ 5. Launch an Instance
- ✓ 6. Login to the Instance
- ✓ 7. More Resources



Knowledge

Get support and stay on track

Launch an Instance

Once AWS approves your GPU Limit Increase Request, you can start the process of launching your instance.

Visit the [EC2 Management Console](#), and click on the "Launch Instance" button.

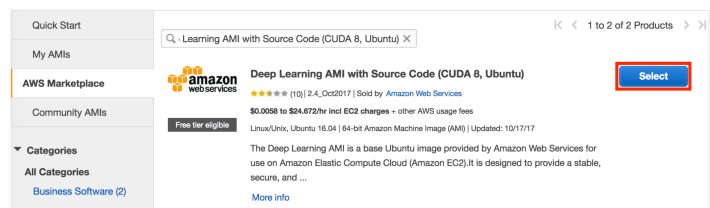
Create Instance

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

[Launch Instance](#)

Next, you must choose an AMI (Amazon Machine Image) which defines the operating system for your instance, as well as any configurations and pre-installed software.

Click on **AWS Marketplace**, and search for **Deep Learning AMI with Source Code (CUDA 8, Ubuntu)**. Once you find the appropriate AMI, click on the "Select" button.



This [Amazon Machine Image \(AMI\)](#) contains all the environment files and drivers for you to train on a GPU. It has [cuDNN](#), and many the other packages required for this course. Any additional packages required for specific

Lesson 2: Cloud Computing



Launch an Instance

SEARCH



RESOURCES



CONCEPTS



- ✓ 1. Overview
- ✓ 2. Create an AWS Account
- ✓ 3. Get Access to GPU Instances
- ✓ 4. Apply Credits
- ✓ 5. Launch an Instance
- ✓ 6. Login to the Instance
- ✓ 7. More Resources

Select the Instance Type

You must next choose an instance type, which is the hardware on which the AMI will run.

Filter the instance list to only show "GPU compute":

Filter by: **GPU compute** **Current generation**

Currently selected: p2.xlarge (11.75 ECUs, 4 vCPUs, 2.7 GHz,

	Family	Type
<input checked="" type="checkbox"/>	GPU compute	p2.xlarge
<input type="checkbox"/>	GPU compute	p2.8xlarge
<input type="checkbox"/>	GPU compute	p2.16xlarge

Select the p2.xlarge instance type:

Filter by: **GPU compute** **Current generation** [Show/Hide Columns](#)

Currently selected: p2.xlarge (11.75 ECUs, 4 vCPUs, 2.7 GHz, E5-2686v4, 61 GiB memory, EBS only)
Note: The vendor recommends using a **p2.xlarge** instance (or larger) for the best experience with this p

	Family	Type	vCPUs	Memory (GiB)	Instance Storage (GB)
<input checked="" type="checkbox"/>	GPU compute	p2.xlarge	4	61	EBS only
<input type="checkbox"/>	GPU compute	p2.8xlarge	32	488	EBS only
<input type="checkbox"/>	GPU compute	p2.16xlarge	64	732	EBS only

Finally, click on the "Review and Launch" button:

[Cancel](#) [Previous](#) **Review and Launch** [Next: Configure Instance Details](#)

Configure the Security Group



Lesson 2: Cloud Computing



Launch an Instance

SEARCH



RESOURCES



CONCEPTS



- ✓ 1. Overview
- ✓ 2. Create an AWS Account
- ✓ 3. Get Access to GPU Instances
- ✓ 4. Apply Credits
- ✓ 5. Launch an Instance
- ✓ 6. Login to the Instance
- ✓ 7. More Resources

By default, AWS restricts access to most ports on an EC2 instance. In order to access the Jupyter notebook, you must configure the AWS Security Group to allow access to port 8888.

Click on "Edit security groups".

Security Groups

Edit security groups

On the "Configure Security Group" page:

1. Select "Create a **new** security group"
2. Set the "Security group name" (i.e. "Jupyter")
3. Click "Add Rule"
4. Set a "Custom TCP Rule"
 1. Set the "Port Range" to "8888"
 2. Select "Anywhere" as the "Source"
5. Click "Review and Launch" (again)

Assign a security group: ☒ Create a new security group ☐ Select an existing security group

Security group name:

Description:

Type	Protocol	Port Range	Source
SSH	TCP	22	Custom 0.0.0.0/0
Custom TCP Rule	TCP	8888	Anywhere 0.0.0.0/0

Launch the Instance

Click on the "Launch" button to launch your GPU instance.

Cancel

Previous

Launch

Create an Authentication Key Pair



Lesson 2: Cloud Computing



Launch an Instance

SEARCH	Q
RESOURCES	^
CONCEPTS	∨

- ✓ 1. Overview
- ✓ 2. Create an AWS Account
- ✓ 3. Get Access to GPU Instances
- ✓ 4. Apply Credits
- ✓ 5. Launch an Instance
- ✓ 6. Login to the Instance
- ✓ 7. More Resources

order to access your instance, so you select "Create a new key pair" and click the "Download Key Pair" button. This will download a .pem file, which you'll need to be able to access 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](#).

Create a new key pair

Key pair name

YourKeyName

Download Key Pair

Move the .pem file to a secure and easily remembered location on your computer; you'll need to access your instance through the location you select.

After the .pem file has been downloaded, click the "Launch Instances" button.

Click the "View Instances" button to go to the EC2 Management Console and watch your instance boot.

Be Careful!

From this point on, AWS will charge you for running this EC2 instance. You can find the details on the [EC2 On-Demand Pricing page](#).

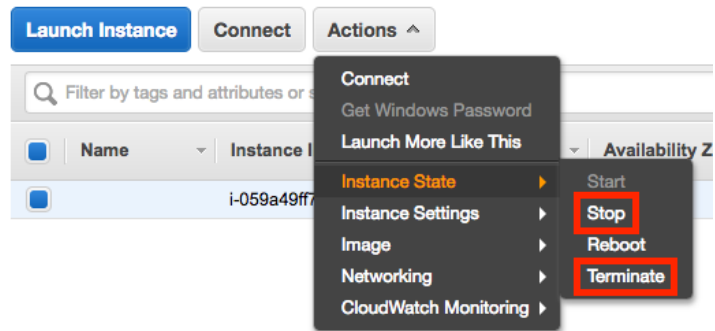
Most importantly, remember to "stop" (i.e. shutdown) your instances when you are not using them. Otherwise, your instances might run for a day, week, month, or longer without you remembering, and you'll wind up with a large bill!

< Lesson 2: Cloud Computing



Launch an Instance

stop the instance. However, there are smaller storage charges that continue to accrue until you “terminate” (i.e. delete) the instance.



There is no way to limit AWS to only a certain budget and have it auto-shutdown when it hits that threshold. However, you can set [AWS Billing Alarms](#).

[NEXT](#)

- ✓ 1. Overview
- ✓ 2. Create an AWS Account
- ✓ 3. Get Access to GPU Instances
- ✓ 4. Apply Credits
- ✓ 5. Launch an Instance
- ✓ 6. Login to the Instance
- ✓ 7. More Resources