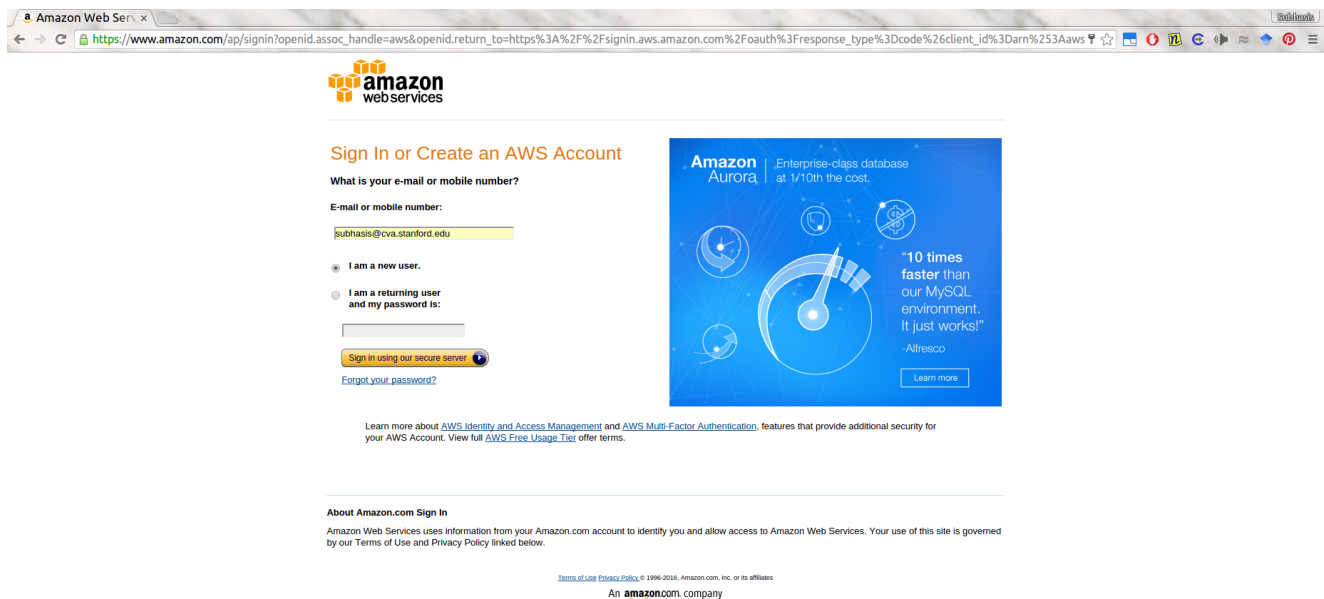


AWS Tutorial

For GPU instances, we also have an Amazon Machine Image (AMI) that you can use to launch GPU instances on Amazon EC2. This tutorial goes through how to set up your own EC2 instance with the provided AMI. **We do not currently distribute AWS credits to CS231N students but you are welcome to use this snapshot on your own budget.**

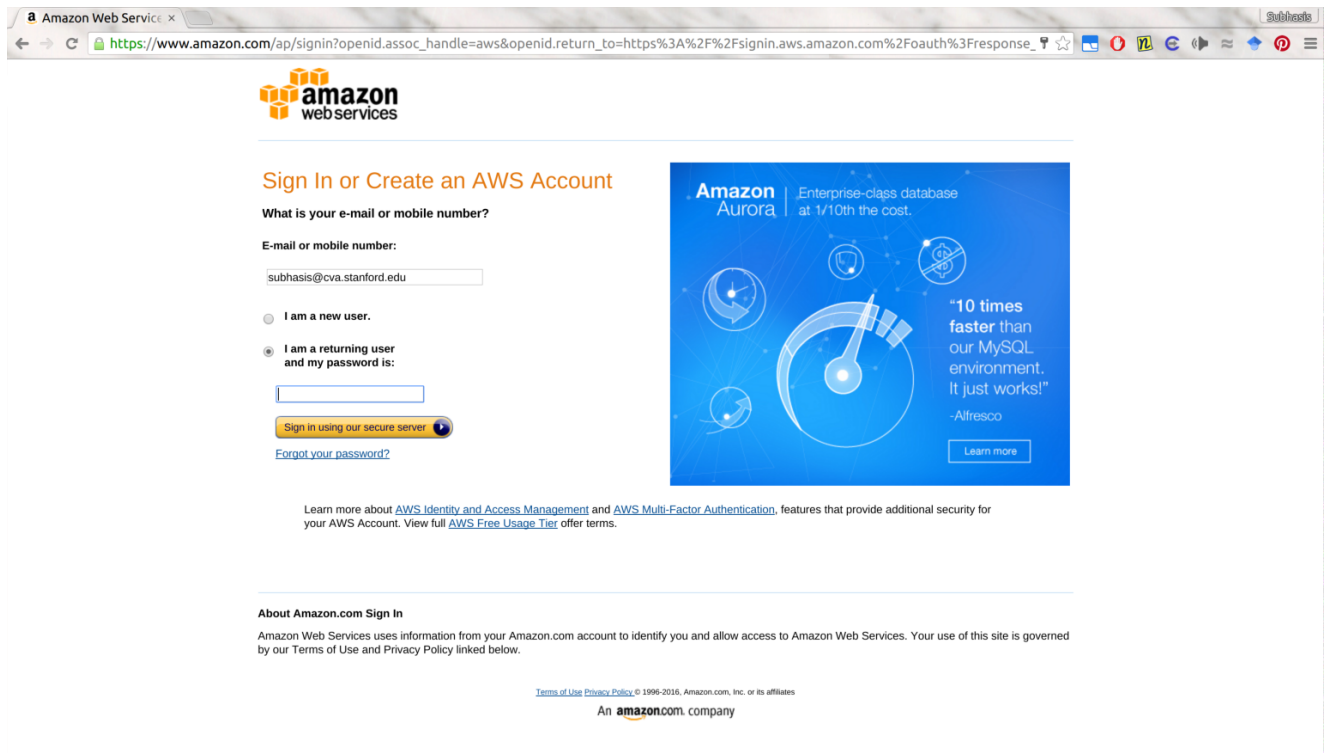
TL;DR for the AWS-savvy: Our image is `cs231n_caffe_torch7_keras_lasagne_v2`, AMI ID: `ami-125b2c72` in the us-west-1 region. Use a `g2.xlarge` instance. Caffe, Torch7, Theano, Keras and Lasagne are pre-installed. Python bindings of caffe are available. It has CUDA 7.5 and CuDNN v3.

First, if you don't have an AWS account already, create one by going to the [AWS homepage](#), and clicking on the yellow "Sign In to the Console" button. It will direct you to a signup page which looks like the following.



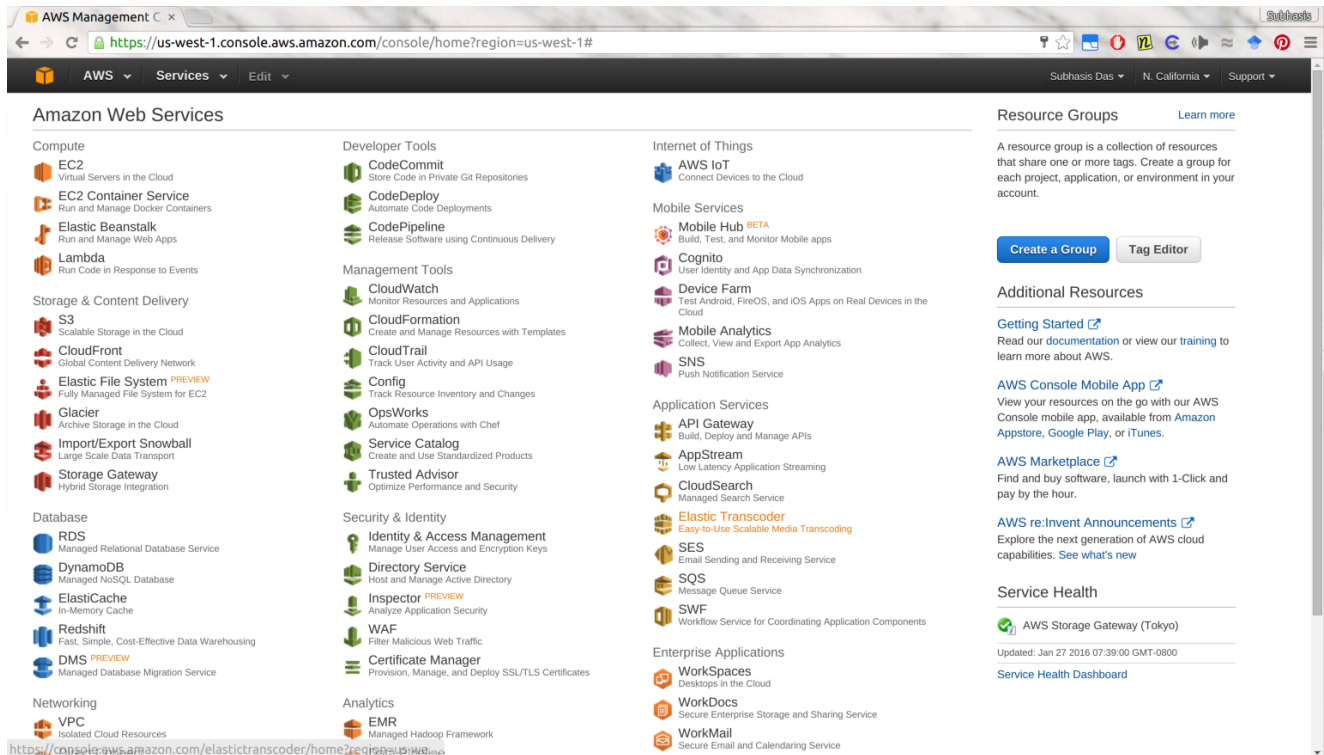
Select the “I am a new user” checkbox, click the “Sign in using our secure server” button, and follow the subsequent pages to provide the required details. They will ask for a credit card information, and also a phone verification, so have your phone and credit card ready.

Once you have signed up, go back to the [AWS homepage](#), click on “Sign In to the Console”, and this time sign in using your username and password.



The screenshot shows the Amazon Web Services sign-in page in a web browser. The browser's address bar displays the URL: https://www.amazon.com/ap/signin?openid.assoc_handle=aws&openid.return_to=https%3A%2F%2Fsignin.aws.amazon.com%2Foauth%3Fresponse_. The page features the Amazon Web Services logo at the top left. Below the logo, the heading "Sign In or Create an AWS Account" is displayed. Underneath, the text "What is your e-mail or mobile number?" is followed by a text input field containing the email address "subhasis@cva.stanford.edu". Below the input field, there are two radio button options: "I am a new user." and "I am a returning user and my password is:". The "I am a returning user" option is selected. Below this, there is a password input field. A yellow button with the text "Sign in using our secure server" and a right-pointing arrow is positioned below the password field. To the left of the button is a link that says "Forgot your password?". To the right of the sign-in form is a blue promotional banner for Amazon Aurora. The banner includes the text "Amazon Aurora | Enterprise-class database at 1/10th the cost." and a quote: "10 times faster than our MySQL environment. It just works!" attributed to -Alfresco. A "Learn more" button is located at the bottom right of the banner. Below the banner, there is a line of text: "Learn more about [AWS Identity and Access Management](#) and [AWS Multi-Factor Authentication](#), features that provide additional security for your AWS Account. View full [AWS Free Usage Tier](#) offer terms." At the bottom of the page, there is a section titled "About Amazon.com Sign In" with the text: "Amazon Web Services uses information from your Amazon.com account to identify you and allow access to Amazon Web Services. Your use of this site is governed by our Terms of Use and Privacy Policy linked below." Below this text are links for "Terms of Use" and "Privacy Policy". At the very bottom, it says "© 1996-2016, Amazon.com, Inc. or its affiliates" and "An amazon.com company".

Once you have signed in, you will be greeted by a page like this:



Make sure that the region information on the top right is set to N. California. If it is not, change it to N. California by selecting from the dropdown menu there.

(Note that the subsequent steps requires your account to be “Verified” by Amazon. This may take up to 2 hrs, and you may not be able to launch instances until your account verification is complete.)

Next, click on the EC2 link (first link under the Compute category). You will go to a dashboard page like this:

EC2 Management Console

Resources

You are using the following Amazon EC2 resources in the US West (N. California) region:

- 0 Running Instances
- 0 Elastic IPs
- 0 Dedicated Hosts
- 0 Snapshots
- 0 Volumes
- 0 Load Balancers
- 0 Key Pairs
- 1 Security Groups
- 0 Placement Groups

Need fast, reliable, scalable, fully-managed message queuing? Try Amazon Simple Queue Service.

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 US West (N. California) region

Service Health

Service Status:

US West (N. California):

This service is operating normally

Availability Zone Status:

us-west-1a:

Availability zone is operating normally

us-west-1c:

Availability zone is operating normally

Scheduled Events

US West (N. California):

No events

Account Attributes

Supported Platforms

VPC

Default VPC

vpc-90bc22f5

Resource ID length management

Additional Information

Getting Started Guide

Documentation

All EC2 Resources

Forums

Pricing

Contact Us

AWS Marketplace

Find free software trial products in the AWS Marketplace from the EC2 Launch Wizard. Or try these popular AMIs:

Tableau Server (10 users)

Provided by Tableau

Rating ★★★★★

Pay by the hour for Tableau software and AWS usage

View all Business Intelligence

SAP HANA One 244GiB

Provided by SAP America, Inc

Click the blue “Launch Instance” button, and you will be redirected to a page like the following:

EC2 Management Console

Step 1: Choose an Amazon Machine Image (AMI)

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

My AMIs

AWS Marketplace

Community AMIs

Free tier only

Amazon Linux AMI 2015.09.1 (HVM), SSD Volume Type - ami-d5ea86b5

The Amazon Linux AMI is an EBS-backed, AWS-supported image. The default image includes AWS command line tools, Python, Ruby, Perl, and Java. The repositories include Docker, PHP, MySQL, PostgreSQL, and other packages.

Root device type: ebs Virtualization type: hvm

Select

64-bit

Red Hat Enterprise Linux 7.2 (HVM), SSD Volume Type - ami-d1315fb1

Red Hat Enterprise Linux version 7.2 (HVM), EBS General Purpose (SSD) Volume Type

Root device type: ebs Virtualization type: hvm

Select

64-bit

SUSE Linux Enterprise Server 12 SP 1 (HVM), SSD Volume Type - ami-6d701b0d

SUSE Linux Enterprise Server 12 Service Pack 1 (HVM), EBS General Purpose (SSD) Volume Type. Public Cloud, Advanced Systems Management, Web and Scripting, and Legacy modules enabled.

Root device type: ebs Virtualization type: hvm

Select

64-bit

Ubuntu Server 14.04 LTS (HVM), SSD Volume Type - ami-06116566

Ubuntu Server 14.04 LTS (HVM), EBS General Purpose (SSD) Volume Type. Support available from Canonical (http://www.ubuntu.com/cloud/services).

Root device type: ebs Virtualization type: hvm

Select

64-bit

Microsoft Windows Server 2012 R2 Base - ami-e7e78d87

Select

Click on the “Community AMIs” link on the left sidebar, and search for “cs231n” in the search box. You should be able to see the AMI `cs231n_caffe_torch7_keras_lasagne_v2` (AMI ID:

ami-125b2c72). Select that AMI, and continue to the next step to choose your instance type.

EC2 Management Console

Step 1: Choose an Amazon Machine Image (AMI)

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

My AMIs

AWS Marketplace

Community AMIs

Operating system

- ☐ Amazon Linux
- ☐ Cent OS
- ☐ Debian
- ☐ Fedora
- ☐ Gentoo
- ☐ OpenSUSE
- ☐ Other Linux
- ☐ Red Hat
- ☐ SUSE Linux
- ☐ Ubuntu
- ☐ Windows

Architecture

- ☐ 32-bit

Search: cs231n

cs231n_caffe_torch7_keras_lasagne_v2 - ami-125b2c72

Image for CS231N. Contains Caffe, Torch7, Keras, Lasagne. CUDA v7.5, CUDNN v3.

Root device type: ebs Virtualization type: hvm

Select

64-bit

Feedback English

© 2008 - 2016, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use

Choose the instance type `g2.2xlarge`, and click on "Review and Launch".

EC2 Management Console

Step 2: Choose an Instance Type

	Instance Type	VCpus	Memory (GiB)	Storage (SSD)	Network (Gigabit)	Price (USD)	
<input type="checkbox"/>	Compute optimized	c3.xlarge	4	7.5	2 x 40 (SSD)	Yes	Moderate
<input type="checkbox"/>	Compute optimized	c3.2xlarge	8	15	2 x 80 (SSD)	Yes	High
<input type="checkbox"/>	Compute optimized	c3.4xlarge	16	30	2 x 160 (SSD)	Yes	High
<input type="checkbox"/>	Compute optimized	c3.8xlarge	32	60	2 x 320 (SSD)	-	10 Gigabit
<input checked="" type="checkbox"/>	GPU instances	g2.2xlarge	8	15	1 x 60 (SSD)	Yes	High
<input type="checkbox"/>	GPU instances	g2.8xlarge	32	60	2 x 120 (SSD)	-	10 Gigabit
<input type="checkbox"/>	Memory optimized	r3.large	2	15	1 x 32 (SSD)	-	Moderate
<input type="checkbox"/>	Memory optimized	r3.xlarge	4	30.5	1 x 80 (SSD)	Yes	Moderate
<input type="checkbox"/>	Memory optimized	r3.2xlarge	8	61	1 x 160 (SSD)	Yes	High
<input type="checkbox"/>	Memory optimized	r3.4xlarge	16	122	1 x 320 (SSD)	Yes	High
<input type="checkbox"/>	Memory optimized	r3.8xlarge	32	244	2 x 320 (SSD)	-	10 Gigabit
<input type="checkbox"/>	Storage optimized	i2.xlarge	4	30.5	1 x 800 (SSD)	Yes	Moderate
<input type="checkbox"/>	Storage optimized	i2.2xlarge	8	61	2 x 800 (SSD)	Yes	High
<input type="checkbox"/>	Storage optimized	i2.4xlarge	16	122	4 x 800 (SSD)	Yes	High

Cancel Previous Review and Launch Next: Configure Instance Details

Feedback English

© 2008 - 2016, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use

In the next page, click on Launch.

EC2 Management Console

https://us-west-1.console.aws.amazon.com/ec2/v2/home?region=us-west-1#LaunchInstanceWizard:

AWS Services Edit

Subhash Das N. California Support

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

Step 7: Review Instance Launch

Please review your instance launch details. You can go back to edit changes for each section. Click **Launch** to assign a key pair to your instance and complete the launch process.

Improve your instances' security. Your security group, launch-wizard-1, is open to the world.

Your instances may be accessible from any IP address. We recommend that you update your security group rules to allow access from known IP addresses only. You can also open additional ports in your security group to facilitate access to the application or service you're running, e.g., HTTP (80) for web servers. [Edit security groups](#)

Your instance configuration is not eligible for the free usage tier

To launch an instance that's eligible for the free usage tier, check your AMI selection, instance type, configuration options, or storage devices. Learn more about [free usage tier](#) eligibility and usage restrictions.

[Don't show me this again](#)

AMI Details [Edit AMI](#)

cs231n_caffe_torch7_keras_lasagne_v2 - ami-125b2c72

Image for CS231N. Contains Caffe, Torch7, Keras, Lasagne, CUDA v7.5, CUDNN v3.

Root Device Type: ebs Virtualization type: hvm

Instance Type [Edit instance type](#)

Instance Type	ECUs	vCPUs	Memory (GiB)	Instance Storage (GB)	EBS-Optimized Available	Network Performance
g2.xlarge	26	8	15	1 x 60	Yes	High

Security Groups [Edit security groups](#)

Security group name: launch-wizard-1

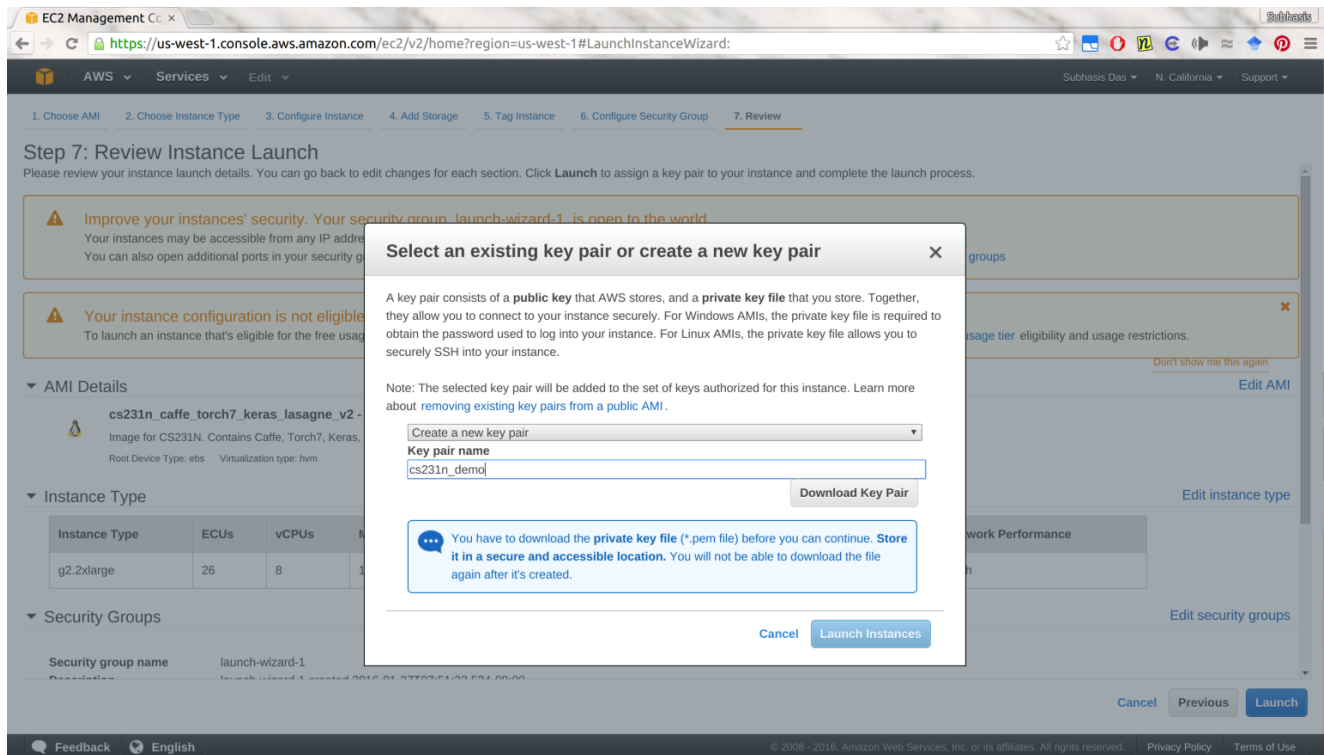
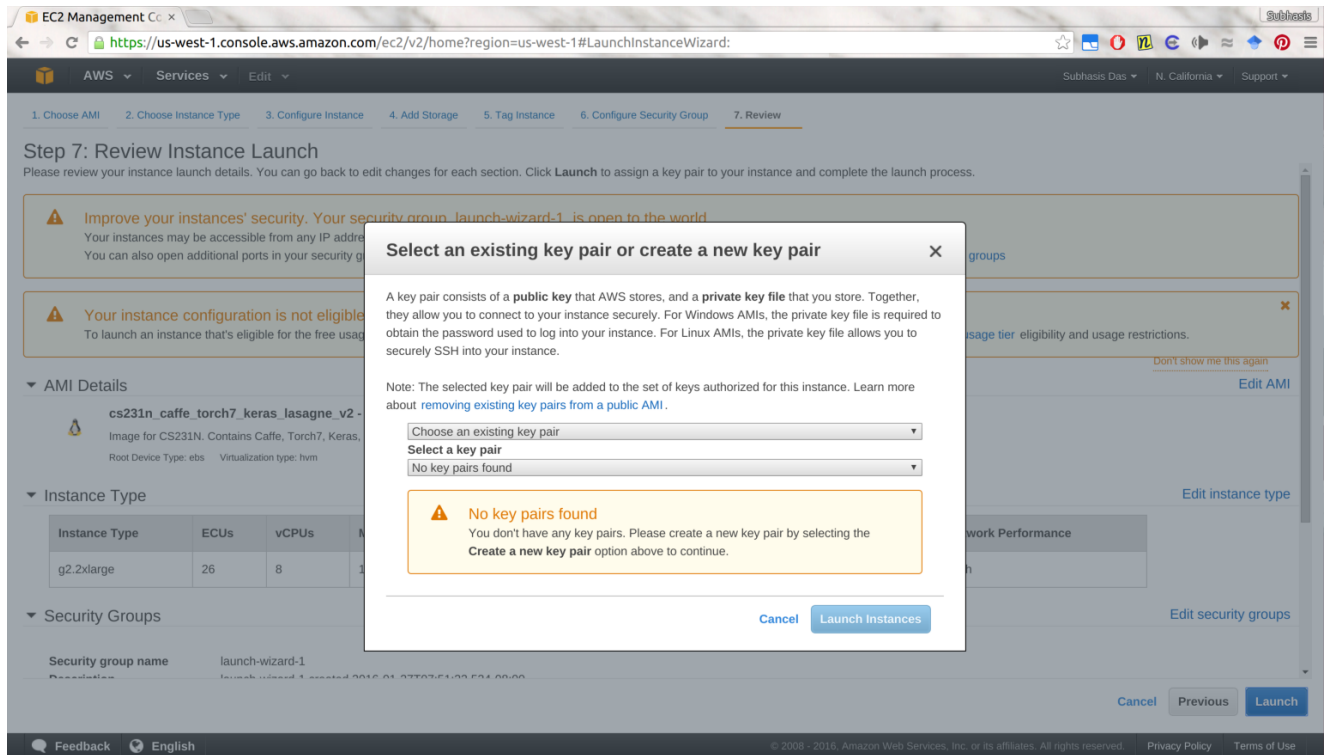
Description: launch-wizard-1 created 2016-01-27T03:51:03-08:00

[Cancel](#) [Previous](#) [Launch](#)

Feedback English

© 2008 - 2016, Amazon Web Services, Inc. or its affiliates. All rights reserved. [Privacy Policy](#) [Terms of Use](#)

You will be then prompted to create or use an existing key-pair. If you already use AWS and have a key-pair, you can use that, or alternately you can create a new one by choosing “Create a new key pair” from the drop-down menu and giving it some name of your choice. You should then download the key pair, and keep it somewhere that you won’t accidentally delete. Remember that there is **NO WAY** to get to your instance if you lose your key.

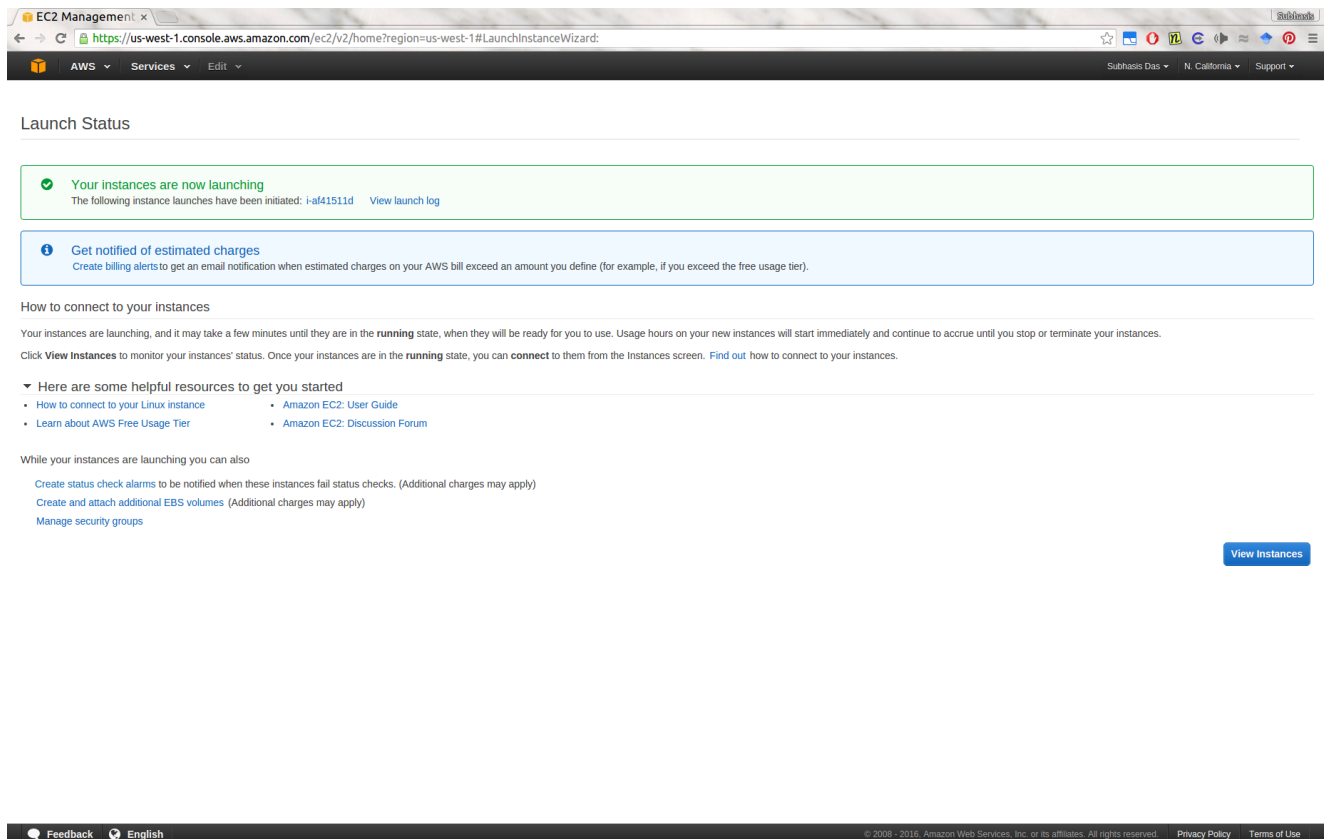


Once you download your key, you should change the permissions of the key to user-only RW, In Linux/OSX you can do it by:

```
$ chmod 600 PEM_FILENAME
```

Here `PEM_FILENAME` is the full file name of the .pem file you just downloaded.

After this is done, click on “Launch Instances”, and you should see a screen showing that your instances are launching:



The screenshot shows the AWS Management Console interface for the 'Launch Status' of EC2 instances. The browser address bar shows the URL: `https://us-west-1.console.aws.amazon.com/ec2/v2/home?region=us-west-1#LaunchInstanceWizard`. The console header includes the AWS logo, navigation tabs (AWS, Services, Edit), and user information (Subhasis Das, N. California, Support).

Launch Status

✓ Your instances are now launching
The following instance launches have been initiated: `i-a441511d` [View launch log](#)

i Get notified of estimated charges
[Create billing alerts](#) to get an email notification when estimated charges on your AWS bill exceed an amount you define (for example, if you exceed the free usage tier).

How to connect to your instances

Your instances are launching, and it may take a few minutes until they are in the **running** state, when they will be ready for you to use. Usage hours on your new instances will start immediately and continue to accrue until you stop or terminate your instances. Click **View instances** to monitor your instances' status. Once your instances are in the **running** state, you can **connect** to them from the Instances screen. [Find out](#) how to connect to your instances.

▼ Here are some helpful resources to get you started

- [How to connect to your Linux instance](#)
- [Amazon EC2: User Guide](#)
- [Learn about AWS Free Usage Tier](#)
- [Amazon EC2: Discussion Forum](#)

While your instances are launching you can also

- [Create status check alarms](#) to be notified when these instances fail status checks. (Additional charges may apply)
- [Create and attach additional EBS volumes](#) (Additional charges may apply)
- [Manage security groups](#)

[View Instances](#)

Feedback English © 2008 - 2016, Amazon Web Services, Inc. or its affiliates. All rights reserved. [Privacy Policy](#) [Terms of Use](#)

Click on “View Instances” to see your instance state. It should change to “Running” and “2/2 status checks passed” as shown below within some time. You are now ready to ssh into the instance.

The screenshot shows the AWS Management Console for the EC2 service. The left sidebar contains navigation links for various AWS services. The main content area displays a table of EC2 instances. The instance i-af41511d is highlighted, and its details are shown in a modal window. The details include the instance ID, state (running), type (g2.xlarge), availability zone (us-west-1c), and public IP (54.183.60.80).

Name	Instance ID	Instance Type	Availability Zone	Instance State	Status Checks	Alarm Status	Public DNS	Public IP
	i-af41511d	g2.xlarge	us-west-1c	running	2/2 checks ...	None	ec2-54-183-60-80.us-w...	54.183.60.80

Instance: i-af41511d		Public DNS: ec2-54-183-60-80.us-west-1.compute.amazonaws.com	
Description			
Instance ID	i-af41511d	Public DNS	ec2-54-183-60-80.us-west-1.compute.amazonaws.com
Instance state	running	Public IP	54.183.60.80
Instance type	g2.xlarge	Elastic IP	-
Private DNS	ip-172-31-0-241.us-west-1.compute.internal	Availability zone	us-west-1c
Private IPs	172.31.0.241	Security groups	launch-wizard-3, view rules
Secondary private IPs		Scheduled events	No scheduled events
VPC ID	vpc-90bc22f5	AMI ID	cs231n_caffe_torch7_keras_lasagne_v2 (ami-125b2c72)
Subnet ID	subnet-7f199526	Platform	-

First, note down the Public IP of the instance from the instance listing. Then, do:

```
ssh -i PEM_FILENAME ubuntu@PUBLIC_IP
```

Now you should be logged in to the instance. You can check that Caffe is working by doing:

```
$ cd caffe
$ ./build/tools/caffe time --gpu 0 --model examples/mnist/lenet.prototxt
```

We have Caffe, Theano, Torch7, Keras and Lasagne pre-installed. Caffe python bindings are also available by default. We have CUDA 7.5 and CuDNN v3 installed.

If you encounter any error such as

```
Check failed: error == cudaSuccess (77 vs. 0) an illegal memory access
```

you might want to terminate your instance and start over again. I have observed this rarely, and I am not sure what causes this.

About how to use these instances:

- The root directory is only 12GB, and only ~ 3GB of that is free.
- There should be a 60GB `/mnt` directory that you can use to put your data, model checkpoints, models etc.
- Remember that the `/mnt` directory won't be persistent across reboots/terminations.
- Stop your instances when are done for the day to avoid incurring charges. GPU instances are costly. Use your funds wisely. Terminate them when you are sure you are done with your instance (disk storage also costs something, and can be significant if you have a large disk footprint).
- Look into creating custom alarms to automatically stop your instances when they are not doing anything.
- If you need access to a large dataset and don't want to download it every time you spin up an instance, the best way to go would be to create an AMI for that and attach that AMI to your machine when configuring your instance (before launching but after you have selected the AMI).

 [cs231n](#)

 [cs231n](#)

karpathy@cs.stanford.edu