



Deep Learning - Recitation 1

# What is Amazon Web Services (AWS)

Cloud Based Computational Resource



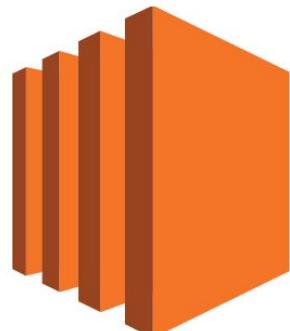
Google Cloud



# What does AWS offer?

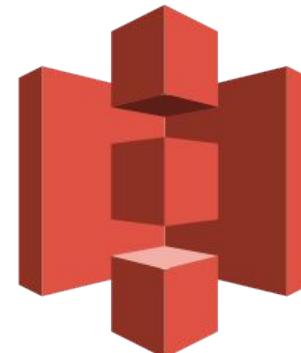
Many many things but here are the two main things to care about for DL...

EC2 - Compute Resources



Train the models

S3 - Data Storage



Store training data,  
models, etc

# EC2 - What kinds of machines are available?

Different types and different subtypes (you can mix and match what you want)...  
Here are the ones you may care about

## General Purpose:

T2 - Webservices

M3/M4 - Databases,  
Fileservers, etc

## Compute Optimized:

C2 - Multiplayer Gaming  
Servers, scientific  
computing

C3/C4 - Ad serving  
machines, MMO servers,  
etc

## GPU Optimized:

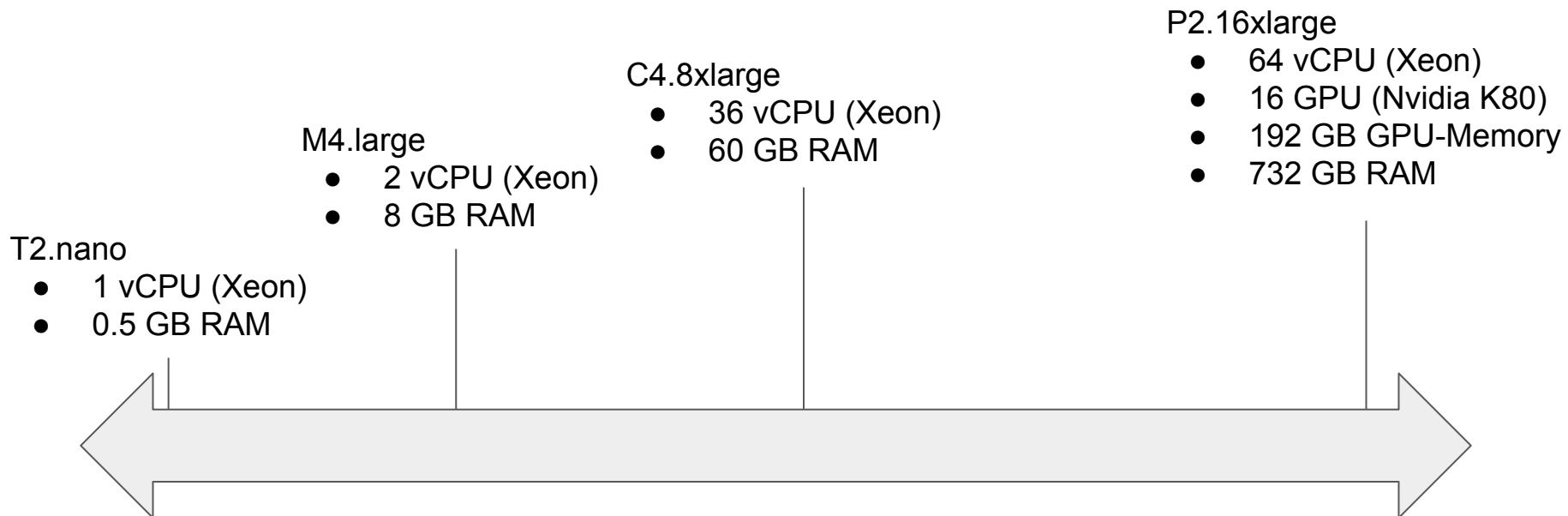
P3/P2 - Machine  
Learning

G3 - Fluid dynamics,  
graphics rendering, etc

Machine sizes - nano, micro, medium, large, xlarge, 2xlarge, ..., 16xlarge

# EC2 - What kinds of machines are available?

Different types and different subtypes (you can mix and match what you want)...



# EC2 - So what do we put on these machines?

## Amazon Machine Instances (AMIs)

- Virtual images of existing machines
  - You can create an image of your machine
    - Transfer it to a different machine
    - Save it as a backup
- Use cases
  - Software packages that are incredibly difficult to install
  - Need to create multiple different machines with the exact same data for parameters servers
  - Load balancing - create a new machine with the same AMI to be used in a different region depending on load

Now you know what AWS is and what you  
can do with it

Create An Account:

<https://aws.amazon.com>

Click “Sign Up” in the top right and follow the instructions

(If you already have an account you can skip this step)

[Menu](#)  [Contact Sales](#) [Products](#) [Solutions](#) [Pricing](#) [Getting Started](#) [Documentation](#) [More](#) [English](#) [My Account](#) [Create an AWS Account](#)

# Start Building on AWS Today

Whether you're looking for compute power, database storage, content delivery or other functionality, AWS has the services to help you build sophisticated applications with increased flexibility, scalability and reliability.

[Create A Free Account](#) [View AWS Free Tier Details »](#)

 [AWS re:Invent Live Streams](#)  
We'll send you a reminder with more information closer to the event  [Register Now](#) [X](#)

**Broad & Deep Platform**  
AWS has more than 90 services and is continually launching new features and functionality.  
[Learn more »](#)

**Customer Success**  
Explore how millions of active customers every month are innovating with AWS.  
[Learn more »](#)

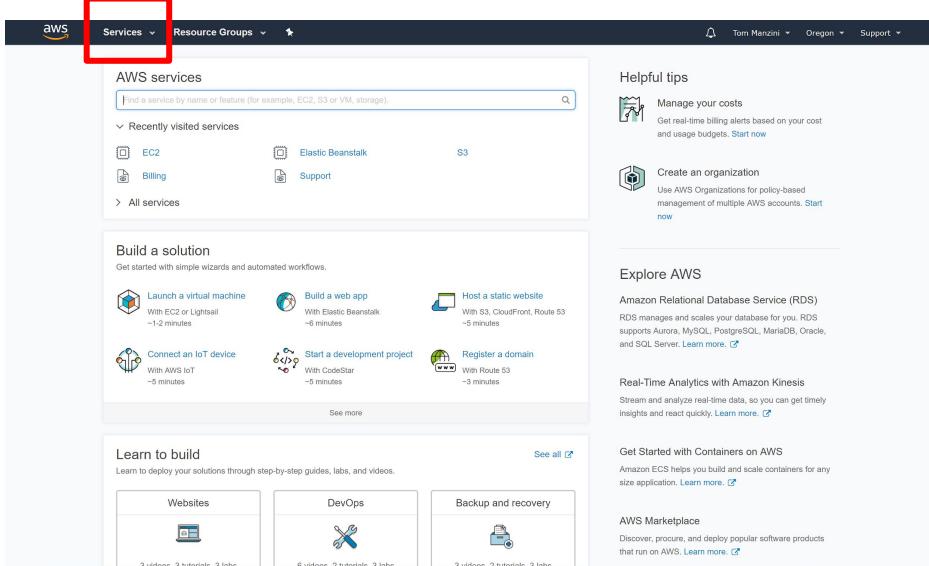
**Pace of Innovation**  
The AWS Cloud platform expands daily. Take a look at what we launched this week.  
[Learn more »](#)

**Global Infrastructure**  
AWS operates 43 Availability Zones within 16 geographic Regions around the world, with 11 more Availability Zones and 4 more Regions coming online soon.  
[Learn more »](#)

[https://aws.amazon.com/products/?nc2=h\\_qLp](https://aws.amazon.com/products/?nc2=h_qLp) [Learn with 10-Minute Tutorials](#) [Start Building with AWS](#)

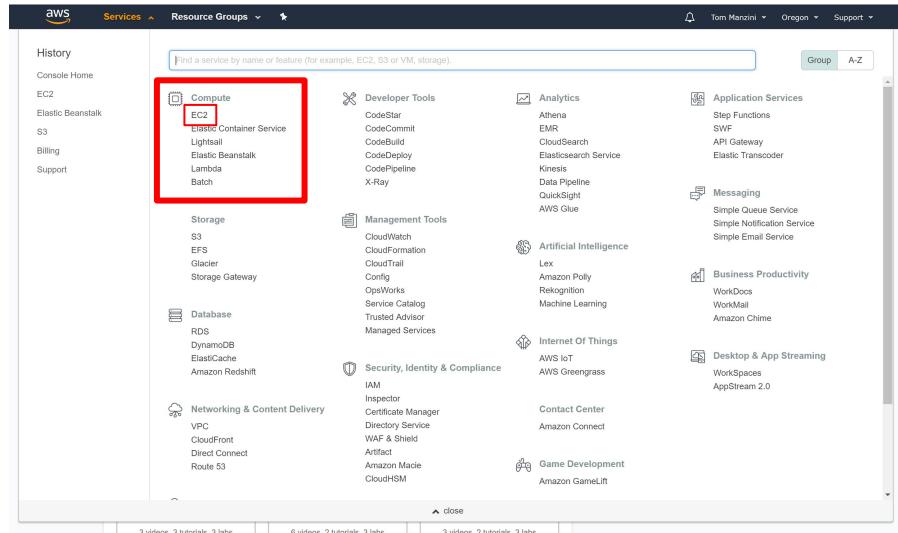
# Let's set up a basic machine



The screenshot shows the AWS Home Page. At the top left, there is a navigation bar with the AWS logo, a "Services" dropdown menu (which is highlighted with a red box), "Resource Groups", and links for "Tom Manzini", "Oregon", and "Support". Below the navigation bar, there is a search bar with placeholder text "Find a service by name or feature (for example, EC2, S3 or VM, storage)." and a "Find" button. The main content area is divided into several sections:

- AWS services**: A search bar and a list of recently visited services: EC2, Billing, Elastic Beanstalk, S3, Support, and All services.
- Helpful tips**: Includes "Manage your costs" (with a link to start alerts) and "Create an organization" (with a link to start now).
- Explore AWS**: Includes "Launch a virtual machine", "Build a web app", "Host a static website", "Connect an IoT device", "Start a development project", "Register a domain", "Amazon Relational Database Service (RDS)", "Real-Time Analytics with Amazon Kinesis", "Get Started with Containers on AWS", and "AWS Marketplace".
- Learn to build**: A section for deploying solutions through step-by-step guides, labs, and videos, featuring categories like Websites, DevOps, and Backup & recovery.

Click on “Services” in the top left



The screenshot shows the AWS Services Catalog. At the top, there is a navigation bar with the AWS logo, a "Services" dropdown menu, "Resource Groups", and links for "Tom Manzini", "Oregon", and "Support". Below the navigation bar, there is a search bar with placeholder text "Find a service by name or feature (for example, EC2, S3 or VM, storage)." and a "Group" button. The main content area is organized into several categories:

- History**: Includes "Console Home", "EC2", "Elastic Beanstalk", "S3", "Billing", and "Support".
- Compute**: Includes "EC2", "Lambda Container Service", "Lightsail", "Elastic Beanstalk", "Lambda", and "Batch". (This category is highlighted with a red box.)
- Storage**: Includes "S3", "EFS", "Glacier", and "Storage Gateway".
- Management Tools**: Includes "CloudWatch", "CloudFormation", "CloudTrail", "Config", "OpsWorks", "Service Catalog", "Trusted Advisor", and "Managed Services".
- Database**: Includes "RDS", "DynamoDB", "ElasticCache", and "Amazon Redshift".
- Analytics**: Includes "Athena", "EMR", "CloudSearch", "Elasticsearch Service", "Kinesis", "Data Pipeline", "QuickSight", and "AWS Glue".
- Developer Tools**: Includes "CodeStar", "CodeCommit", "CodeBuild", "CodeDeploy", "CodePipeline", and "X-Ray".
- Artificial Intelligence**: Includes "Lex", "Amazon Polly", "Rekognition", and "Machine Learning".
- Internet Of Things**: Includes "AWS IoT" and "AWS Greengrass".
- Networking & Content Delivery**: Includes "VPC", "CloudFront", "Direct Connect", and "Route 53".
- Security, Identity & Compliance**: Includes "IAM", "Inspector", "Certificate Manager", "Directory Service", "WAF & Shield", "Artifact", "Amazon Macie", and "CloudHSM".
- Business Productivity**: Includes "WorkDocs", "WorkMail", and "Amazon Chime".
- Desktop & App Streaming**: Includes "WorkSpaces" and "AppStream 2.0".
- Game Development**: Includes "Amazon GameLift".

Then, under “Compute”, select “EC2”

# Let's set up a basic machine

Then Click on  
“Running  
Instances”

The screenshot shows the AWS EC2 Dashboard. At the top, there is a navigation bar with the AWS logo, 'Services' dropdown, 'Resource Groups' dropdown, and user information ('Tom Manzini', 'Oregon', 'Support'). A red box highlights the 'Oregon' region selection. The main content area is titled 'Resources' and displays the following statistics:

| Category            | Value |
|---------------------|-------|
| 0 Running Instances |       |
| 0 Dedicated Hosts   |       |
| 1 Volumes           |       |
| 2 Key Pairs         |       |
| 0 Placement Groups  |       |
| 0 Elastic IPs       |       |
| 4 Snapshots         |       |
| 0 Load Balancers    |       |
| 9 Security Groups   |       |

A blue callout box at the bottom left of the dashboard area says: 'EC2 Spot. Save up to 90% off On-Demand Prices. Turbo Boost your Workloads. Get started with Amazon EC2 Spot Instances.' Below this, there is a 'Create Instance' section with a 'Launch Instance' button. To the right, there are sections for 'Service Health' and 'Scheduled Events'. The 'Service Health' section shows 'Service Status' for 'US West (Oregon)' with a green checkmark and the message 'This service is operating normally'. It also lists 'Availability Zone Status' for 'us-west-2a', 'us-west-2b', and 'us-west-2c', all of which are operating normally. The 'Scheduled Events' section shows 'US West (Oregon)' with a green checkmark and the message 'No events'. On the far right, there is an 'Account Attributes' sidebar with sections for 'Supported Platforms' (VPC), 'Default VPC' (vpc-f6d6a891), 'Resource ID length management', 'Additional Information' (Getting Started Guide, Documentation, All EC2 Resources, Forums, Pricing, Contact Us), and 'AWS Marketplace' (free software trial products from the AWS Marketplace). A red box highlights the 'Oregon' region selection in the top navigation bar.

Make sure you  
are in the “US  
West (Oregon)”  
region

# Let's set up a basic machine

The screenshot shows the AWS EC2 Dashboard. On the left, a sidebar lists various services: EC2 Dashboard, Events, Tags, Reports, Limits, Instances (selected), Spot Requests, Reserved Instances, Scheduled Instances, Dedicated Hosts, Images (AMIs), Bundle Tasks, Elastic Block Store (Volumes, Snapshots), Network & Security (Security Groups, Elastic IPs, Placement Groups, Key Pairs, Network Interfaces), and Load Balancing (Load Balancers, Target Groups). The main area has tabs for Launch Instance, Connect, and Actions. A search bar is present. The Instances section shows a table with columns: Name, Instance ID, Instance Type, Availability Zone, Instance State, Status Checks, Alarm Status, and Public DNS (IPv4). One instance is listed: i-02d40dd9b0a197529, t2.micro, us-west-2a, stopped, None, and None. Below this, a modal window displays detailed information for the selected instance (i-02d40dd9b0a197529). The modal has tabs for Description, Status Checks, Monitoring, and Tags. The Description tab shows the following details:

| Attribute              | Value   |
|------------------------|---|
| Instance ID            | i-02d40dd9b0a197529                               |
| Instance state         | stopped   |
| Instance type          | t2.micro  |
| Elastic IPs            | -   |
| Availability zone      | us-west-2a  |
| Security groups        | launch-wizard-1 . view inbound rules              |
| Scheduled events       | -   |
| AMI ID                 | Loading ami-92f1420ea...                          |
| Platform               | -   |
| IAM role               | -   |
| Key pair name          | tmanzini-personal                                 |
| Owner                  | 180942078756                                      |
| Launch time            | November 7, 2017 at 11:35:49 PM UTC-5 (448 hours) |
| Termination protection | -   |
| Lifecycle              | normal  |
| Monitoring             | basic   |
| Alarm status           | None  |
| Kernel ID              | -   |

On the right side of the instance details, there is a list of network interfaces:

| Attribute             | Value                                     |
|-----------------------|---|
| Public DNS (IPv4)     | -   |
| IPv4 Public IP        | -   |
| IPv6 IPs              | -   |
| Private DNS           | ip-172-31-45-0.us-west-2.compute.internal |
| Private IPs           | 172.31.45.0                               |
| Secondary private IPs | -   |
| VPC ID                | vpc-f6d6a891                              |
| Subnet ID             | subnet-18ed4551                           |
| Network interfaces    | eth0                                      |
| Source/dest. check    | True                                      |

At the bottom of the modal, there are sections for EBS-optimized (False), Root device type (ebs), Root device (/dev/sda1), Block devices (/dev/sda1), Elastic GPU (-), Elastic GPU type (-), and Elastic GPU status (-).

Here you see your current instances

(I have 1 that is stopped, you shouldn't have any)

Here are the details of that instance

# Let's set up a basic machine

Click “Launch Instance”

The screenshot shows the AWS EC2 Instances page. On the left, there's a navigation sidebar with various services like EC2 Dashboard, Events, Tags, Reports, and Limits. Under the 'INSTANCES' section, 'Instances' is selected, showing a single instance named 'i-02d40dd9b0a197529' which is currently 'stopped'. A red box highlights the 'Launch Instance' button at the top of the main content area. The main content area also includes tabs for 'Description', 'Status Checks', 'Monitoring', and 'Tags', and displays detailed information about the instance.

| Attribute              | Value   |
|------------------------|---|
| Instance ID            | i-02d40dd9b0a197529                               |
| Instance state         | stopped   |
| Instance type          | t2.micro  |
| Elastic IPs            | -   |
| Availability zone      | us-west-2a  |
| Security groups        | launch-wizard-1-. view inbound rules              |
| Scheduled events       | -   |
| AMI ID                 | Loading ami-92f1420ea...                          |
| Platform               | -   |
| IAM role               | -   |
| Key pair name          | tmanzini-personal                                 |
| Owner                  | 180942078756                                      |
| Launch time            | November 7, 2017 at 11:35:49 PM UTC-5 (448 hours) |
| Termination protection | -   |
| Lifecycle              | normal  |
| Monitoring             | basic   |
| Alarm status           | None  |
| Kernel ID              | -   |
| Public DNS (IPv4)      | -   |
| IPv4 Public IP         | -   |
| IPv6 IPs               | -   |
| Private DNS            | ip-172-31-45-0.us-west-2.compute.internal         |
| Private IPs            | 172.31.45.0                                       |
| Secondary private IPs  | -   |
| VPC ID                 | vpc-f6d6a891                                      |
| Subnet ID              | subnet-18ed4551                                   |
| Network interfaces     | eth0  |
| Source/dest. check     | True  |
| EBS-optimized          | False   |
| Root device type       | ebs   |
| Root device            | /dev/sda1   |
| Block devices          | /dev/sda1   |
| Elastic GPU            | -   |
| Elastic GPU type       | -   |
| Elastic GPU status     | -   |

Feedback English (US) © 2008 - 2017, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use

# Let's set up a basic machine

AWS Services Resource Groups ★

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

Step 1: Choose an Amazon Machine Image (AMI)

of commercial databases. [Learn more about RDS](#)

[Launch a database using RDS](#)

|  |   |
|--|---|
|  <b>Ubuntu Server 16.04 LTS (HVM), SSD Volume Type</b> - ami-076e276d85f524150 (64-bit x86) / ami-05e1b2aec3b47890f (64-bit Arm)<br><small>Ubuntu Server 16.04 LTS (HVM), EBS General Purpose (SSD) Volume Type. Support available from Canonical (<a href="http://www.ubuntu.com/cloud/services">http://www.ubuntu.com/cloud/services</a>).</small><br><small>Free tier eligible</small><br>Root device type: ebs Virtualization type: hvm ENA Enabled: Yes  | <a href="#">Select</a><br><input checked="" type="radio"/> 64-bit (x86)<br><input type="radio"/> 64-bit (Arm) |
|  <b>Microsoft Windows Server 2016 Base</b> - ami-019e99815e07ceb49<br><small>Microsoft Windows 2016 Datacenter edition. [English]</small><br><small>Free tier eligible</small><br>Root device type: ebs Virtualization type: hvm ENA Enabled: Yes   | <a href="#">Select</a><br>64-bit (x86)  |
|  <b>Deep Learning AMI (Ubuntu) Version 20.0</b> - ami-0d0ff0945ae093aea<br><small>With latest deep learning frameworks pre-installed: MXNet, TensorFlow, PyTorch, Keras, Chainer, Caffe/2, Theano &amp; CNTK, configured with NVIDIA CUDA, cuDNN, NCCL &amp; Intel MKL-DNN. For a fully managed experience, check: <a href="https://aws.amazon.com/sagemaker">https://aws.amazon.com/sagemaker</a></small><br>Root device type: ebs Virtualization type: hvm ENA Enabled: Yes                       | <a href="#">Select</a><br>64-bit (x86)  |
|  <b>Deep Learning AMI (Amazon Linux) Version 20.0</b> - ami-0305a0d7a68489e58<br><small>With latest deep learning frameworks pre-installed: MXNet, TensorFlow, PyTorch, Keras, Chainer, Caffe/2, Theano &amp; CNTK, configured with NVIDIA CUDA, cuDNN, NCCL &amp; Intel MKL-DNN. For a fully managed experience, check: <a href="https://aws.amazon.com/sagemaker">https://aws.amazon.com/sagemaker</a></small><br>Amazon Linux<br>Root device type: ebs Virtualization type: hvm ENA Enabled: Yes | <a href="#">Select</a><br>64-bit (x86)  |
|  <b>Deep Learning Base AMI (Ubuntu) Version 14.0</b> - ami-015eb46ac552e435f<br><small>Comes with foundational platform of Nvidia CUDA, cuDNN, NCCL, GPU Drivers, Intel MKL-DNN and other system libraries to deploy your own custom deep learning environment. For a fully managed experience, check: <a href="https://aws.amazon.com/sagemaker">https://aws.amazon.com/sagemaker</a></small><br>Root device type: ebs Virtualization type: hvm ENA Enabled: Yes                                   | <a href="#">Select</a><br>64-bit (x86)  |

[Cancel and Exit](#)

# Let's set up a basic machine

The screenshot shows the AWS EC2 instance creation wizard at Step 2: Choose an Instance Type. The user has selected the t2.micro instance type, which is highlighted with a red box. The t2.micro row is also highlighted with a red box. The 'Free tier eligible' status is visible in the table. At the bottom right, the 'Review and Launch' button is highlighted with a red box, and the text 'Next: Configure Instance Details' is displayed above it.

Step 2: Choose an Instance Type

Amazon EC2 provides a wide selection of instance types optimized to fit different use cases. Instances are virtual servers that can run applications. They have varying combinations of CPU, memory, storage, and networking capacity, and give you the flexibility to choose the appropriate mix of resources for your applications. [Learn more](#) about instance types and how they can meet your computing needs.

Filter by: All instance types ▾ Current generation ▾ Show/Hide Columns

Currently selected: t2.micro (Variable ECUs, 1 vCPUs, 2.5 GHz, Intel Xeon Family, 1 GiB memory, EBS only)

|                                     | Family          | Type                           | vCPUs | Memory (GiB) | Instance Storage (GB) | EBS-Optimized Available | Network Performance | IPv6 Support |
|-------------------------------------|-----------------|--------------------------------|-------|--------------|-----------------------|-------------------------|---------------------|--------------|
| <input type="checkbox"/>            | General purpose | t2.nano                        | 1     | 0.5          | EBS only              | -                       | Low to Moderate     | Yes          |
| <input checked="" type="checkbox"/> | General purpose | t2.micro<br>Free tier eligible | 1     | 1            | EBS only              | -                       | Low to Moderate     | Yes          |
| <input type="checkbox"/>            | General purpose | t2.small                       | 1     | 2            | EBS only              | -                       | Low to Moderate     | Yes          |
| <input type="checkbox"/>            | General purpose | t2.medium                      | 2     | 4            | EBS only              | -                       | Low to Moderate     | Yes          |
| <input type="checkbox"/>            | General purpose | t2.large                       | 2     | 8            | EBS only              | -                       | Low to Moderate     | Yes          |
| <input type="checkbox"/>            | General purpose | t2.xlarge                      | 4     | 16           | EBS only              | -                       | Moderate            | Yes          |
| <input type="checkbox"/>            | General purpose | t2.2xlarge                     | 8     | 32           | EBS only              | -                       | Moderate            | Yes          |
| <input type="checkbox"/>            | General purpose | m4.large                       | 2     | 8            | EBS only              | Yes                     | Moderate            | Yes          |
| <input type="checkbox"/>            | General purpose | m4.xlarge                      | 4     | 16           | EBS only              | Yes                     | High                | Yes          |
| <input type="checkbox"/>            | General purpose | m4.2xlarge                     | 8     | 32           | EBS only              | Yes                     | High                | Yes          |
| <input type="checkbox"/>            | General purpose | m4.4xlarge                     | 16    | 64           | EBS only              | Yes                     | High                | Yes          |

Cancel Previous Review and Launch Next: Configure Instance Details

Feedback English (US) © 2008 - 2017, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use

Select the t2-micro because it is “free tier eligible”

Select Next

# Let's set up a basic machine

The screenshot shows the AWS Step 3: Configure Instance Details page. The top navigation bar includes the AWS logo, Services dropdown, Resource Groups dropdown, a bell icon, user name 'Tom Manzini', location 'Oregon', and Support dropdown. Below the navigation, a progress bar shows steps 1 through 7: Choose AMI, Choose Instance Type, Configure Instance (highlighted in orange), Add Storage, Add Tags, Configure Security Group, and Review.

**Step 3: Configure Instance Details**

Configure the instance to suit your requirements. You can launch multiple instances from the same AMI, request Spot instances to take advantage of the lower pricing, assign an access management role to the instance, and more.

**Number of instances:** 1 [Launch into Auto Scaling Group](#)

**Purchasing option:**  Request Spot instances

**Network:** vpc-f6d6a891 (default) [Create new VPC](#)

**Subnet:** No preference (default subnet in any Availability Zone) [Create new subnet](#)

**Auto-assign Public IP:** Use subnet setting (Enable)

**IAM role:** None [Create new IAM role](#)

**Shutdown behavior:** Stop

**Enable termination protection:**  Protect against accidental termination

**Monitoring:**  Enable CloudWatch detailed monitoring  
Additional charges apply.

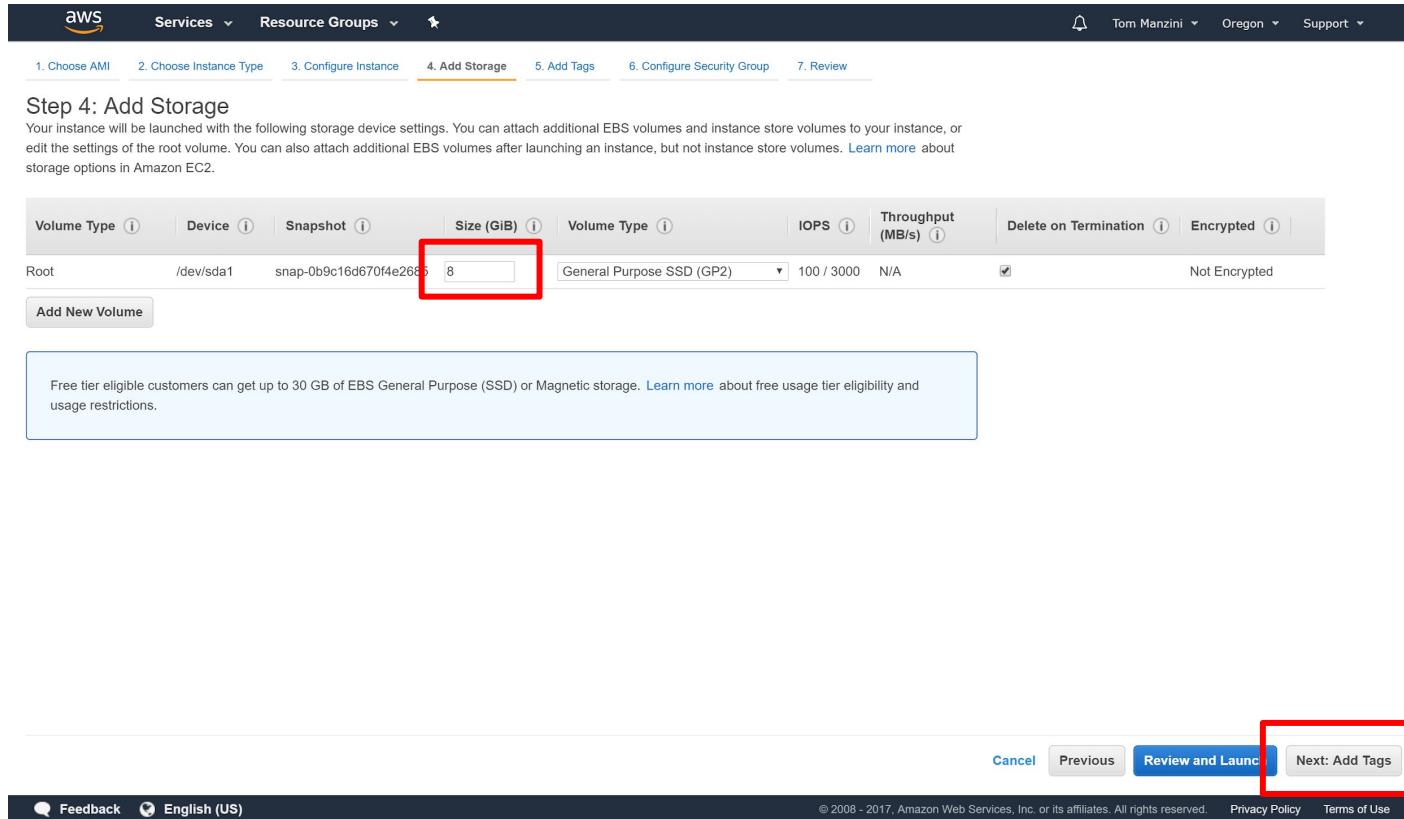
**Tenancy:** Shared - Run a shared hardware instance  
Additional charges will apply for dedicated tenancy.

[Advanced Details](#)

At the bottom, there are buttons: Cancel, Previous, **Review and Launch**, and **Next: Add Storage** (which is highlighted with a red box).

Just select next

# Let's set up a basic machine



Step 4: Add Storage

Your instance will be launched with the following storage device settings. You can attach additional EBS volumes and instance store volumes to your instance, or edit the settings of the root volume. You can also attach additional EBS volumes after launching an instance, but not instance store volumes. [Learn more](#) about storage options in Amazon EC2.

| Volume Type | Device    | Snapshot               | Size (GiB) | Volume Type               | IOPS       | Throughput (MB/s) | Delete on Termination               | Encrypted     |
|-------------|-----------|------------------------|------------|---------------------------|------------|-------------------|-------------------------------------|---------------|
| Root        | /dev/sda1 | snap-0b9c16d670f4e2685 | 8          | General Purpose SSD (GP2) | 100 / 3000 | N/A               | <input checked="" type="checkbox"/> | Not Encrypted |

Add New Volume

Free tier eligible customers can get up to 30 GB of EBS General Purpose (SSD) or Magnetic storage. [Learn more](#) about free usage tier eligibility and usage restrictions.

Cancel Previous Review and Launch Next: Add Tags

Make sure you choose 8 GB of SSD Storage Space

Select next

# Let's set up a basic machine

The screenshot shows the AWS EC2 Launch Instance wizard at Step 5: Add Tags. The top navigation bar includes the AWS logo, Services dropdown, Resource Groups dropdown, a bell icon, user Tom Manzini, location Oregon, and Support dropdown. Below the navigation is a progress bar with steps 1 through 7: Choose AMI, Choose Instance Type, Configure Instance, Add Storage, Add Tags (which is highlighted in orange), Configure Security Group, and 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 (i) | Volumes (i)

*This resource currently has no tags*

Choose the Add tag button or [click to add a Name tag](#). Make sure your [IAM policy](#) includes permissions to create tags.

Add Tag (Up to 50 tags maximum)

Cancel Previous Review and Launch Next: Configure Security Group

<https://docs.aws.amazon.com/console/ec2/launchinstance/tags/iam>

© 2008 - 2017, Amazon Web Services, Inc. or its affiliates. All rights reserved. [Privacy Policy](#) [Terms of Use](#)

Select next

# Let's set up a basic machine

Step 6: Configure Security Group

A security group is a set of firewall rules that control the traffic for your instance. On this page, you can add rules to allow specific traffic to reach your instance. For example, if you want to set up a web server and allow Internet traffic to reach your instance, add rules that allow unrestricted access to the HTTP and HTTPS ports. You can create a new security group or select from an existing one below. [Learn more about Amazon EC2 security groups.](#)

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

Security group name:

Description:

| Type       | Protocol | Port Range | Source             | Description                |
|------------|----------|------------|--------------------|----------------------------|
| SSH        | TCP      | 22         | Custom 0.0.0.0/0   | e.g. SSH for Admin Desktop |
| Custom TCP | TCP      | 8888       | Anywhere 0.0.0.0/0 | e.g. SSH for Admin Desktop |

Add Rule

**Warning**  
Rules with source of 0.0.0.0/0 allow all IP addresses to access your instance. We recommend setting security group rules to allow access from known IP addresses only.

Cancel Previous **Review and Launch**

Make sure you have an SSH rule set (This should be default) & a TCP rule for 8888 so you can connect your Ipython Notebook

You can set HTTP or other rules here too if you want

Select Review and Launch

# Let's set up a basic machine

The screenshot shows the AWS Instance Launch Wizard at Step 7: Review Instance Launch. The top navigation bar includes the AWS logo, Services dropdown, Resource Groups dropdown, a notification bell, user Tom Manzini, location Oregon, and Support dropdown. Below the navigation, a progress bar shows steps 1 through 7, with step 7 being the current active step.

**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.

**AMIs**

Ubuntu Server 16.04 LTS (HVM), SSD Volume Type - ami-0a00ce72

Free tier eligible

Ubuntu Server 16.04 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

**Instance Type**

| Instance Type | ECUs     | vCPUs | Memory (GiB) | Instance Storage (GB) | EBS-Optimized Available | Network Performance |
|---------------|----------|-------|--------------|-----------------------|-------------------------|---------------------|
| t2.micro      | Variable | 1     | 1            | EBS only              | -                       | Low to Moderate     |

**Security Groups**

Security group name: launch-wizard-4  
Description: launch-wizard-4 created 2017-11-26T17:20:35.221-05:00

| Type | Protocol | Port Range | Source    | Description |
|------|----------|------------|-----------|-------------|
| SSH  | TCP      | 22         | 0.0.0.0/0 |             |

**Instance Details**

Cancel Previous **Launch**

Take one last look to make sure you are happy with everything...

Select Launch

# Let's set up a basic machine

Enter a name  
for your private  
key

Select an existing key pair or create a new key pair X

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](#).

Create a new key pair

Key pair name   

Download Key Pair

You have to download the **private key file** (\*.pem file) before you can continue. **Store it in a secure and accessible location.** You will not be able to download the file again after it's created.

Cancel Launch Instances

This window  
allows your to  
create a private  
key to access  
your  
machine...  
when you ssh  
you will need it

Download your  
key and save it

Make sure it's  
in a place  
where it won't  
get lost - this  
key is the only  
way to connect  
to this specific  
instance!

Now you have a basic AWS machine up  
and running

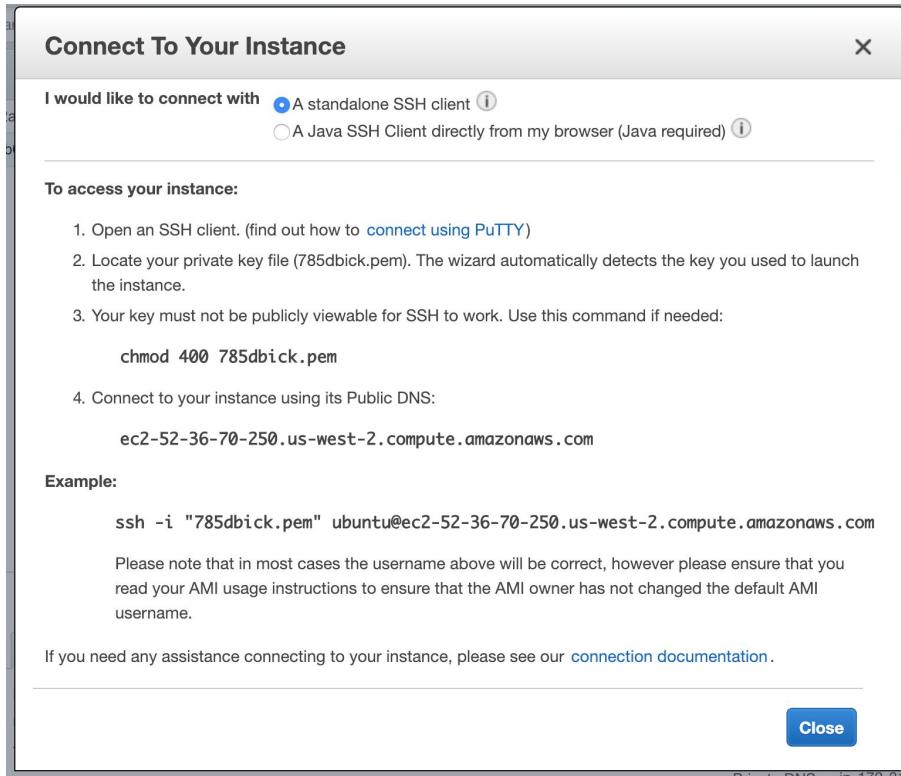
# Now let's connect to it

Your instance is now identified by its IP address. You can ssh to your instance by using this public IP address and your address key.

Let's start by changing the instance type. For student accounts, t2.micro is free-tier eligible, but let's still start on a t2.nano instance.

Remember your instance's IP address changes every time you restart it

# Connect to launched instance



Right-click and click connect

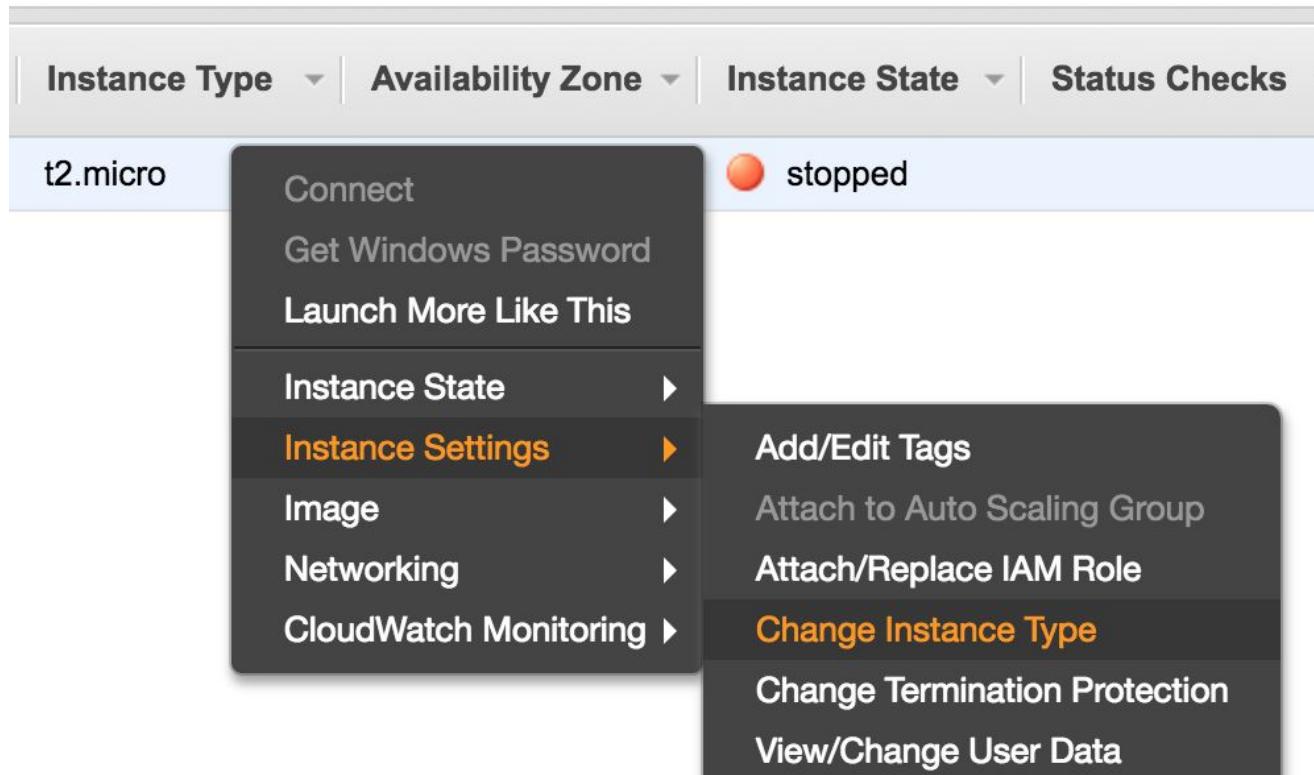
The ssh command shown includes the currently assigned IP address.

This will change at each launch!

Note - the ssh command will use the path to the key associated with this instance.

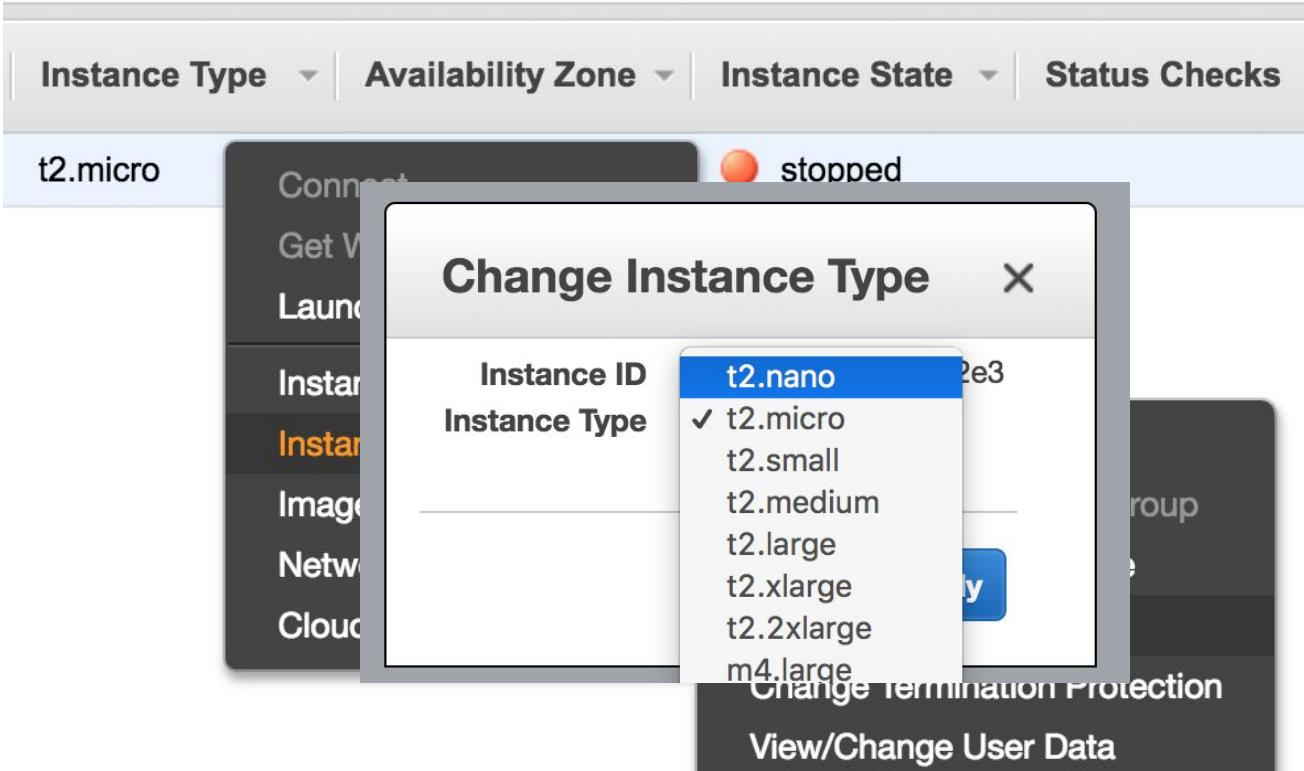
**Important** - will need to run chmod command every time you have a new key

# Set instance type



Right-click on your instance under the instance tab, go to instance settings and change instance type to t2.nano

# Set instance type



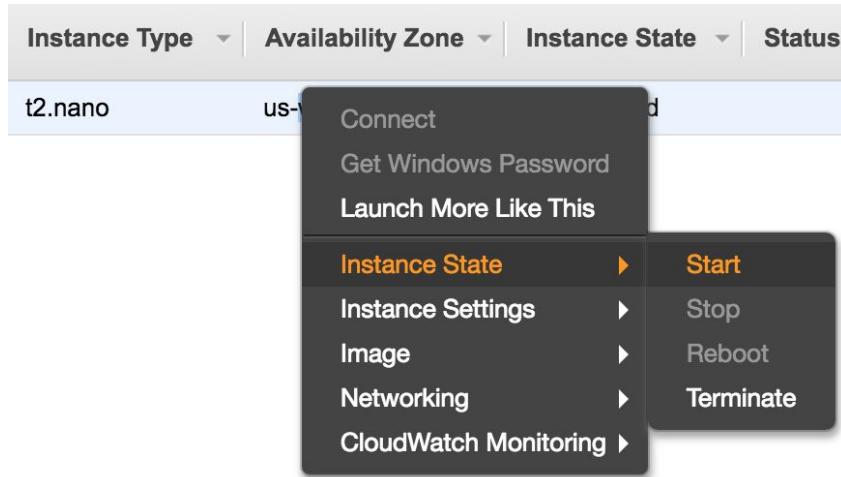
Right-click on your instance under the instance tab, go to instance settings and change instance type to t2.nano

# Note

You can spin up a t2.micro for development (which allows access to all the packages in the AMI and can put all your data on the instance), and then change the instance type to a p2.xlarge for GPU access (training).

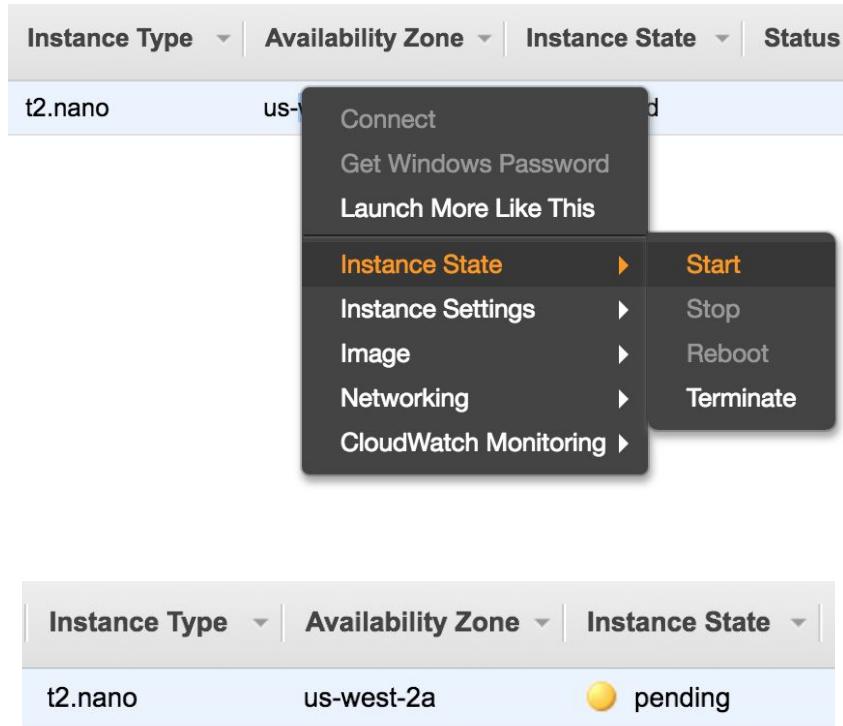
p2.xlarge is what I used for all my training, about \$1 per hour, so with 3 credits of \$50 that's more than enough hours of training for almost all cases.

# Launch instance



Right-click and Start your instance.

# Launch instance



Right-click and Start your instance.

This will start the process of allocating resources to your instance. Once this is completed, your instance will be running and you can connect to it.

Stopping this instance removes the compute associated with the current session.

Do NOT terminate! This will wipe the slate clean. Terminate ONLY when you are sure of 'throwing away' the data.

Now you have a running AWS machine  
and you can connect

Now you know how to use machines on  
EC2

Let's run a trivial little problem

# A Toy Problem to Test your AWS Instance

Will Test:

- Working instance
- Good Python installation
- Working PyTorch installed
- GPU / CUDA support
- Your ability to login and execute code on AWS

```
import torch

def main():
    GPU = torch.cuda.is_available()
    mat_size = (100, 100)
    cpu_mat_0 = torch.zeros(size = mat_size)
    cpu_mat_1 = torch.ones(size = mat_size)
    gpu_mat_0 = torch.zeros(size = mat_size)
    gpu_mat_1 = torch.ones(size = mat_size)

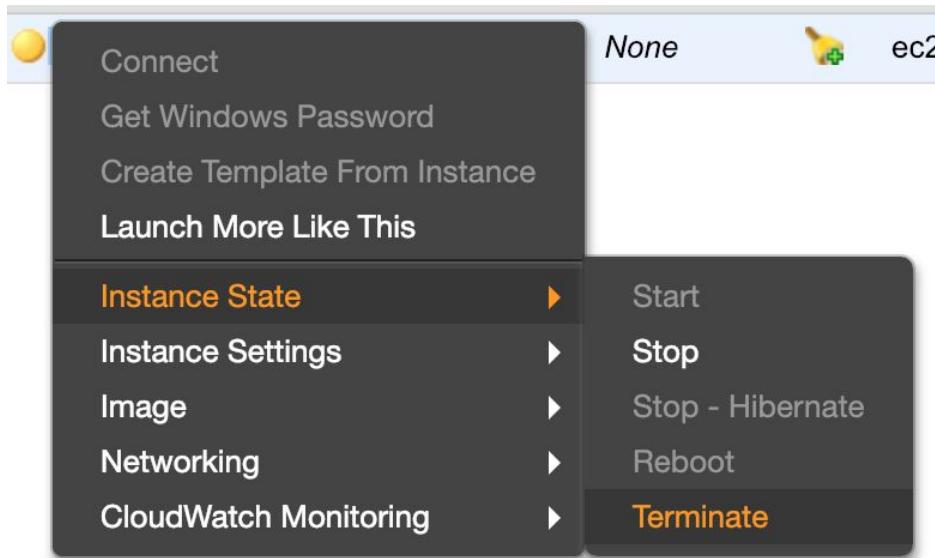
    if GPU:
        gpu_mat_0 = gpu_mat_0.cuda()
        gpu_mat_1 = gpu_mat_1.cuda()
        print("Using GPU")

    cpu_res = cpu_mat_0 + cpu_mat_1
    gpu_res = gpu_mat_0 + gpu_mat_1

    try:
        print(bool(torch.all(cpu_res == gpu_res)))
        return 0
    except:
        print("If using GPU, should be here")
        gpu_res = gpu_res.detach().cpu() # detach is for gradient computations
        print(bool(torch.all(cpu_res == gpu_res)))
        return 0

main()
```

# And let's shut it down so you don't get billed



If you stop the instance you can just start the instance and resume whenever you want.

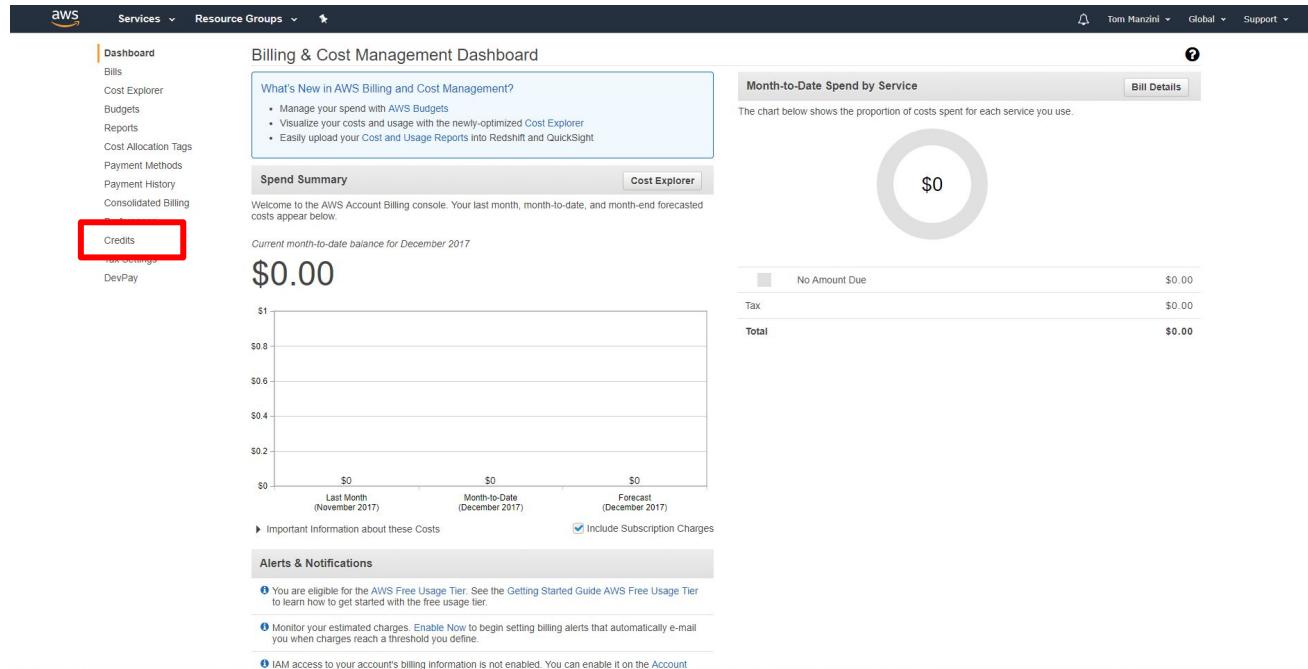
If you are done, terminate.

# Now lets redeem some AWS credits

The screenshot shows the AWS Home page with several sections:

- AWS services:** A search bar and a list of recently visited services (EC2, IAM, Billing, Support, Elastic Beanstalk). A red box highlights the "Billing" link.
- Build a solution:** A section with six cards: Launch a virtual machine, Build a web app, Host a static website, Connect an IoT device, Start a development project, and Register a domain. A red box highlights the "Build a web app" card.
- Learn to build:** A section with categories: Websites, DevOps, Backup and recovery, Big data, Databases, and Mobile. A red box highlights the "Websites" card.
- Helpful tips:** Manage your costs, Create an organization, and Explore AWS.
- Explore AWS:** Amazon Relational Database Service (RDS), Real-Time Analytics with Amazon Kinesis, Get Started with Containers on AWS, and AWS Marketplace.
- User Profile:** Shows Tom Manzini, Oregon, with a red box highlighting the "My Billing Dashboard" link in the dropdown menu.

# Now lets redeem some AWS credits



The screenshot shows the AWS Billing & Cost Management Dashboard. On the left sidebar, under the 'Services' dropdown, the 'Billing' option is selected. A red box highlights the 'Credits' link in the sidebar menu. The main content area displays the 'Billing & Cost Management Dashboard' with the following sections:

- What's New in AWS Billing and Cost Management?**:
  - Manage your spend with AWS Budgets
  - Visualize your costs and usage with the newly-optimized Cost Explorer
  - Easily upload your Cost and Usage Reports into Redshift and QuickSight
- Spend Summary**:
  - Welcome to the AWS Account Billing console. Your last month, month-to-date, and month-end forecasted costs appear below.
  - Current month-to-date balance for December 2017**: \$0.00
  - A chart showing the month-to-date balance for December 2017, with values for Last Month (November 2017), Month-to-Date (December 2017), and Forecast (December 2017) all at \$0.00.
  - Important Information about these Costs**: Includes links for 'Include Subscription Charges' and 'AWS Free Usage Tier'.
- Alerts & Notifications**:
  - You are eligible for the AWS Free Usage Tier. See the Getting Started Guide AWS Free Usage Tier to learn how to get started with the free usage tier.
  - Monitor your estimated charges. Enable Now to begin setting billing alerts that automatically e-mail you when charges reach a threshold you define.
  - IAM access to your account's billing information is not enabled. You can enable it on the Account page.
- Month-to-Date Spend by Service**: A donut chart showing the proportion of costs spent for each service used. The chart is currently empty, displaying '\$0'.
- Bill Details**: A table showing the current month-to-date balance breakdown.

| Category      | Amount Due |
|---------------|------------|
| No Amount Due | \$0.00     |
| Tax           | \$0.00     |
| Total         | \$0.00     |

# Now lets redeem some AWS credits

Screenshot of the AWS Credits page showing the redemption process.

The page title is "Credits". A red box highlights the "Promo Code" input field. Below it is a CAPTCHA image showing the characters "yb8nd5" and a "Refresh Image" link. A red box highlights the "Redeem" button. A note below the CAPTCHA states: "By clicking "Redeem" you indicate that you have read and agree to the terms of the AWS Promotional Credit Terms & Conditions located [here](#)".

Below the form, a message says: "Below are all the credits you have redeemed with AWS. Credits will automatically be applied to your bill. Only credits that apply to a specific service can be used."

| Expiration Date | Credit Name                   | Credits Used | Credits Remaining | Applicable Products               |
|-----------------|-------------------------------|--------------|-------------------|-----------------------------------|
| 2018-03-31      | EDU_ENG_FY2017_Q1_1_CMU_50USD | \$44.22      | \$5.78            | <a href="#">See complete list</a> |
| 2018-03-31      | EDU_ENG_FY2017_Q1_1_CMU_50USD | \$45.05      | \$4.95            | <a href="#">See complete list</a> |

Total Amount of Credits Remaining: \$10.73

Feedback English (US) © 2008 - 2017, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use

Now you know how to use machines on  
EC2 and you can afford to use the  
expensive ones

# Finally, some useful tidbits

Remember to refer back here later in the semester, likely  
won't remember all this right now

# Editing Volume

If you run out of space, you can adjust your volume without shutting off the instance, happened to me on hw2p2

<https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ebs-modify-volume.html>

<https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/console-modify.html>

EC2 Dashboard

Events

Tags

Reports

Limits

## INSTANCES

Instances

Launch Templates

Spot Requests

Reserved Instances

Dedicated Hosts

Scheduled Instances

Capacity Reservations

## IMAGES

AMIs

Bundle Tasks

## ELASTIC BLOCK STORE

**Volumes**

Snapshots

Lifecycle Manager

## NETWORK &amp; SECURITY

Security Groups

Elastic IPs

Placement Groups

Key Pairs

Network Interfaces

## LOAD BALANCING

**Create Volume****Actions** ▾**Modify Volume****Create Snapshot**

Delete Volume

Attach Volume

**Detach Volume**

Force Detach Volume

Change Auto-Enable IO Setting

Add/Edit Tags

| Volume Type | IOPS | Snapshot         | Created              | Availability Zone | State  | Alarm Status | Attachment Information | Monitoring |
|-------------|------|------------------|----------------------|-------------------|--------|--------------|------------------------|------------|
| gp2         | 225  | snap-014390f4... | December 21, 2018... | us-west-2a        | in-use | None         | i-08d9e0db1b613edb...  |            |

Volumes: vol-0adb44e78d8f4ab21

**Description**

Status Checks

Monitoring

Tags

Volume ID vol-0adb44e78d8f4ab21

Alarm status None

Size 75 GiB

Snapshot snap-014390f4a84518e09

Created December 21, 2018 at 12:37:08 AM UTC-5

Availability Zone us-west-2a

State in-use

Encrypted Not Encrypted

Attachment information i-08d9e0db1b613edb5:/dev/sda1 (attached)

KMS Key ID

Volume type gp2

KMS Key Aliases



EC2 Dashboard

Events

Tags

Reports

Limits

INSTANCES

Instances

Launch Templates

Spot Requests

Reserved Instances

Dedicated Hosts

Scheduled Instances

Capacity Reservations

IMAGES

AMIs

Bundle Tasks

ELASTIC BLOCK STORE

Volumes

Snapshots

Lifecycle Manager

NETWORK &amp; SECURITY

Security Groups

Elastic IPs

Placement Groups

Key Pairs

Network Interfaces

LOAD BALANCING

Create Volume

Actions ▾

Filter by tags and attributes or search by keyword

? &lt; 1 to 1 of 1 &gt;

| Name | Volume ID      | Size   | Volume Type | IOPS | Snapshot         | Created              | Availability Zone | State  | Alarm Status | Attachment Information | Monitoring |
|------|----------------|--------|-------------|------|------------------|----------------------|-------------------|--------|--------------|------------------------|------------|
|      | vol-0adb44e... | 75 GiB | gp2         | 225  | snap-014390f4... | December 21, 2018... | us-west-2a        | in-use | None         | i-08d9e0db1b613edb...  |            |

## Modify Volume



Volume ID vol-0adb44e78d8f4ab21

Volume Type General Purpose SSD (gp2)

Size 75 (Min: 1 GiB, Max: 16384 GiB)

IOPS 225 / 3000   
(Baseline of 3 IOPS per GiB with a minimum of 100 IOPS, burstable to 3000 IOPS)

Volumes: vol-0adb44e78d8f4ab21

Description

Status Checks

Mo

Cancel

Modify

|                        |  |             |        |         |                                |       |        |                 |                        |
|------------------------|--|-------------|--------|---------|--------------------------------|-------|--------|-----------------|------------------------|
| Volume ID              | vol-0adb44e78d8f4ab21                    | Size        | 75 GiB | Created | December 21, 2018 10:45:10 UTC | State | in-use | Encrypted       | None                   |
| Attachment information | i-08d9e0db1b613edb5 /dev/sda1 (attached) | Volume type | gp2    |         |                                |       |        | KMS Key ID      | snap-014390f4a84518e09 |
|                        |  |             |        |         |                                |       |        | KMS Key Aliases | us-west-2a             |
|                        |  |             |        |         |                                |       |        |                 | Not Encrypted          |

# Here's How to Connect to Jupyter Notebook

Call ssh -N -L localhost:8888:localhost:8887 -i ~/path/to/key ubuntu@ip

Create new terminal window, ssh into instance

In ssh window, source activate pytorch\_p36

Call jupyter notebook --no-browser --port=8887

Go to browser, enter localhost:8888

# Parting wisdom

- Remember to shut down your machines
  - Just because you don't have an open ssh connection doesn't mean your machine is off
  - When you're not running code, Stop; when you can get rid of the data, Terminate
- Use PyTorch on the previously given AMI
  - After you ssh into the instance, run command 'source activate pytorch\_p36'
  - Without this you cannot import torch
  - Unless you are already very comfortable with tensorflow, it is much easier to get help from TAs on PyTorch

# Parting wisdom

- You need permission to launch a GPU instance
  - You will need to create a support ticket to launch GPU Machines
  - If you attempt to launch on you will be guided through the process
  - You are typically only allowed to launch 1 GPU machine at a time
- Only launch an expensive instance when it is time to train, not develop
  - Launch a basic instance with everything you need to develop and test your code
  - When it comes time to train your system for real, then launch a decked out instance
  - Saves you money

# Parting wisdom

If you want to use a local IDE instead of VIM on your ssh window, you can develop locally and run an scp to transfer the file to the aws instance

```
scp -i ~/path/to/key ~/path/to/file/ ubuntu@ec2...
```

If you use jupyter notebook connected to aws, no need to worry about this.

While writing code, I kept a note with common commands that I could just change the IP address on, such as an ssh command, and scp of my local code

# Requesting Instances

Must request access to instances from AWS in the Support Center section of AWS. Check the latest status on AWS, as they may have changed something or upgraded offerings. Currently, should request EC2 instance limit increase, but be sure that it has not changed.

# How to Request

## Instructions to apply for GPU access on AWS

Starting from [aws.amazon.com](https://aws.amazon.com)

Top right corner, "My Account" -> "Account Settings"

At top right corner click "Support" -> Support Center

Click "Create case"

Click "Service Limit increase"

Limit type -> EC2 Instance

If Pitt, select Region US East (Ohio). If SV, select Region US West (Oregon)

New limit value -> 1 Instance type -> p2.xlarge (can also request other instance types)

Use case description -> Describe that you are in this class at CMU and you need a GPU to train deep learning models for the homeworks

Specify your contact method