



Deep Learning - Recitation (Fall 2018)

What is Amazon Web Services (AWS)

Cloud Based Computational Resource



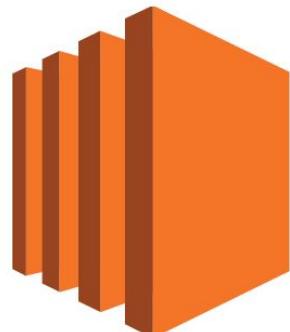
Google Cloud



What does AWS offer?

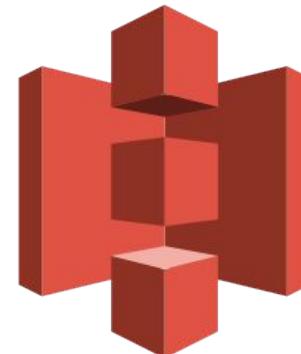
Many many things but here are the two main things to care about for DL...

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:

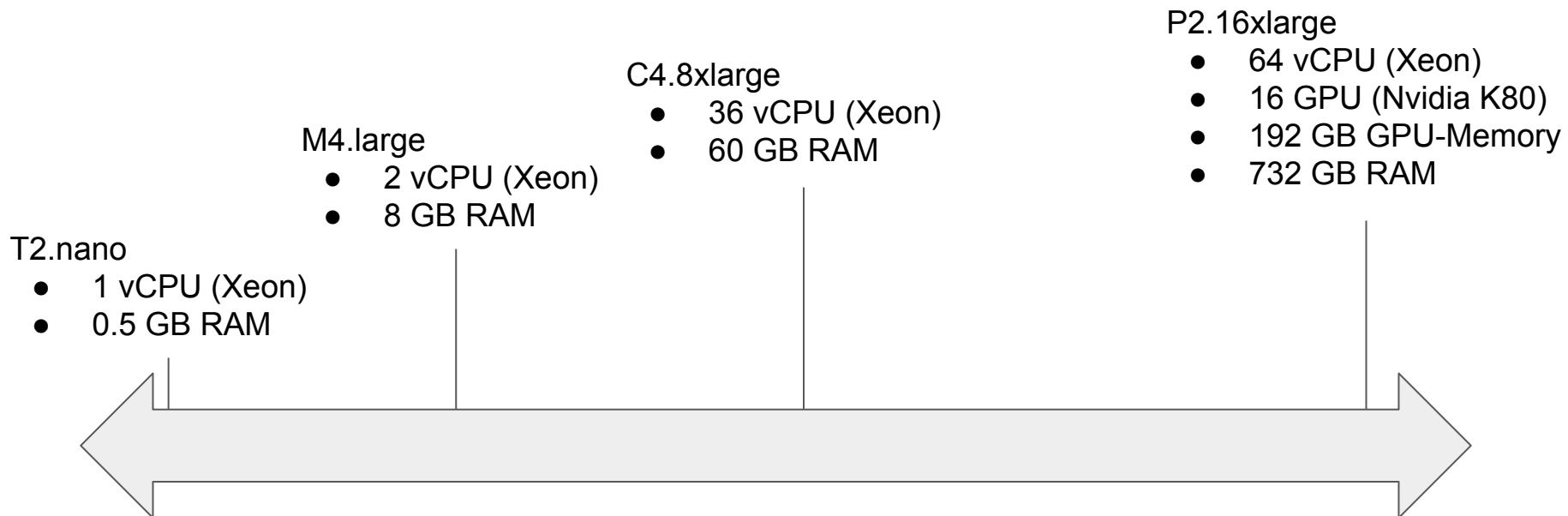
P3/P2 - Machine
Learning

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

Who needs to know how to do this and why

- Anyone who doesn't have a NVIDIA GPU
- Anyone who doesn't have their GPU configured to be correctly used by Pytorch/Tensorflow/etc.
- GPU is good for training not processing data or if the code is not configured to leverage GPU
- PS: There are screenshots for each step that go with this guide (steps are numbered)

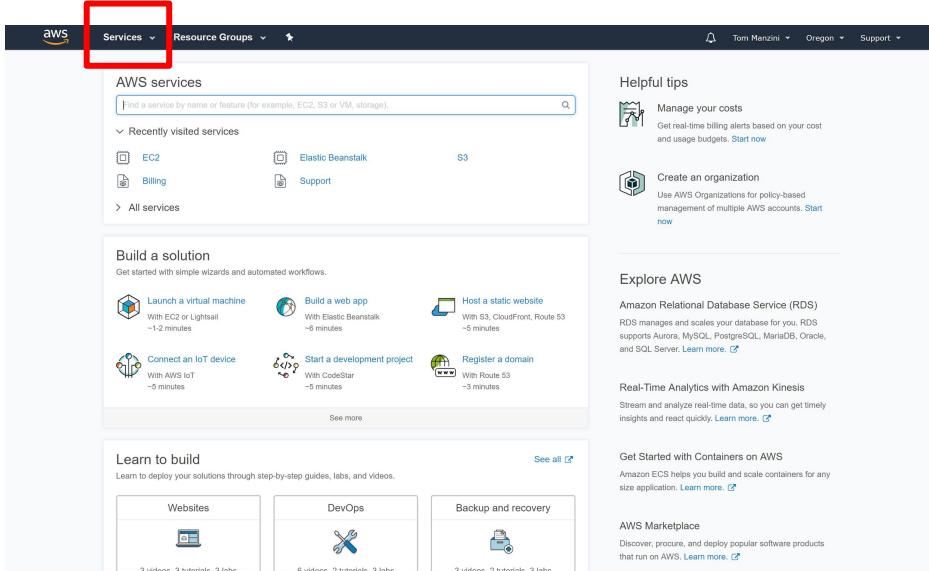
Steps you need to get started

1. Create/sign in to your AWS account (aws.amazon.com)
2. Click on the region in the top right and then click either US East (N. Virginia) or US West (Oregon) GPU instances are only available in certain regions
3. Click on EC2 on the right
4. Click on running instances
5. Click on launch instances (top right)
6. Choose the Deep Learning AMI (ami-0faada77)
7. Pick a p2.xlarge instance (you can explore other instances if you want but these are the cheapest and most appropriate for personal use IMO)

Steps cont.

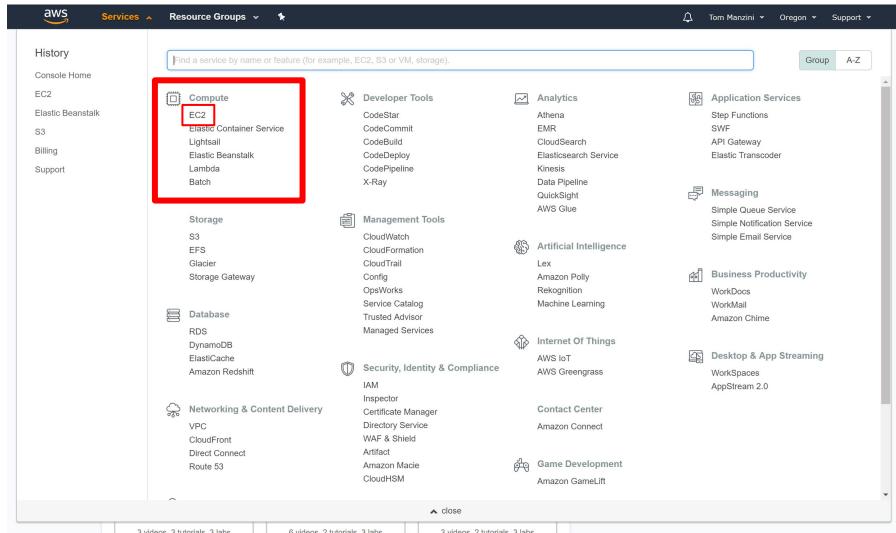
8. Do nothing at the Configure Instance Details
9. Configure the instance to have at least 75GB storage
10. Do nothing at add tags
11. Do nothing at security group (they will create a SSH security group for you)(If you want to set up a Jupyter notebook you will have to add a couple of ports to open)
12. AWS Step 7 Review and Launch (click launch)

Step 1,2



The screenshot shows the AWS Home Page. At the top, there's a navigation bar with the AWS logo, a "Services" dropdown (which is highlighted with a red box), "Resource Groups", and "Support". Below the navigation is a search bar with placeholder text "Find a service by name or feature (for example, EC2, S3 or VM, storage)." Under the search bar, there's a section titled "Recently visited services" with links to EC2, Billing, Elastic Beanstalk, and S3. A "Helpful tips" section follows, featuring "Manage your costs" and "Create an organization". The main content area is titled "Explore AWS" and includes sections like "Launch a virtual machine", "Build a web app", "Host a static website", "Connect an IoT device", "Start a development project", "Register a domain", "Get Started with Containers on AWS", and "AWS Marketplace". At the bottom, there's a "Learn to build" section with links to "Websites", "DevOps", and "Backup and recovery".

Click on “Services” in the top left



The screenshot shows the AWS Services Catalog. The top navigation bar includes the AWS logo, a "Services" dropdown, "Resource Groups", and user information (Tom Manzini, Oregon, Support). A search bar at the top right says "Find a service by name or feature (for example, EC2, S3 or VM, storage)." The main content area is organized into categories: History, Compute, Storage, Database, Management Tools, Analytics, Developer Tools, Artificial Intelligence, Internet Of Things, Security, Identity & Compliance, Networking & Content Delivery, and Game Development. The "Compute" section is expanded, showing EC2 and other services like Lambda, Batch, Lightsail, and Elastic Beanstalk. EC2 is also highlighted with a red box. The "Developer Tools" section is also highlighted with a red box.

Then, under “Compute”, select “EC2”

Step 3

Then Click on
“Running
Instances”

The screenshot shows the AWS EC2 Dashboard. On the left, there's a sidebar with various navigation options like EC2 Dashboard, Events, Tags, Reports, Limits, Instances, AMIs, Bundle Tasks, Elastic Block Store, Security Groups, Placement Groups, Network Interfaces, Load Balancers, Target Groups, and Auto Scaling. The main content area has a heading 'Resources' and a message: 'You are using the following Amazon EC2 resources in the US West (Oregon) region:'. It lists 0 Running Instances, 0 Dedicated Hosts, 1 Volumes, 2 Key Pairs, 0 Placement Groups, 0 Elastic IPs, 4 Snapshots, 0 Load Balancers, and 9 Security Groups. Below this is a promotional banner for EC2 Spot instances. To the right, there are sections for 'Create Instance' (with a 'Launch Instance' button), 'Service Health' (showing service status for US West (Oregon) and availability zones us-west-2a, us-west-2b, and us-west-2c, all operating normally), and 'Scheduled Events' (showing no events for the US West (Oregon) region). At the top right, there are account attributes (Supported Platforms: VPC; Default VPC: vpc-f6d6a891; Resource ID length management), additional information links (Getting Started Guide, Documentation, All EC2 Resources, Forums, Pricing, Contact Us), and an AWS Marketplace section. A red box highlights the 'Oregon' region selection in the top right corner.

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

Feedback English (US)

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

Step 4,5

The screenshot shows the AWS EC2 Dashboard. On the left, there's a sidebar with navigation links for EC2 Dashboard, Events, Tags, Reports, Limits, Instances, 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. The main area has tabs for Launch Instance, Connect, and Actions. A search bar says "Filter by tags and attributes or search by keyword". Below it is a table with columns: Name, Instance ID, Instance Type, Availability Zone, Instance State, Status Checks, Alarm Status, and Public DNS (IPv4). One row is selected, showing "i-02d40dd9b0a197529", "t2.micro", "us-west-2a", "stopped", "None", and a yellow icon. A red box highlights this table. Below the table, a modal window titled "Instance: i-02d40dd9b0a197529 Private IP: 172.31.45.0" shows detailed information. It has tabs for Description, Status Checks, Monitoring, and Tags. The Description tab is active, displaying fields like Instance ID, Instance state, Instance type, Elastic IPs, Availability zone, Security groups, Scheduled events, AMI ID, Platform, IAM role, Key pair name, Owner, Launch time, Termination protection, Lifecycle, Monitoring, Alarm status, and Kernel ID. To the right, another set of fields includes Public DNS (IPv4), IPv4 Public IP, IPv6 IPs, Private DNS, Private IPs, Secondary private IPs, VPC ID, Subnet ID, Network interfaces, Source/dest. check, EBS-optimized, Root device type, Root device, Block devices, Elastic GPU, Elastic GPU type, and Elastic GPU status. A red box highlights this modal window.

Here you see your current instances

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

Here are the details of that instance

Step 6

The screenshot shows the AWS EC2 Management Console interface for launching a new instance. The top navigation bar includes 'Services', 'Resource Groups', and tabs for '1. Choose AMI', '2. Choose Instance Type', '3. Configure Instance', '4. Add Storage', '5. Add Tags', '6. Configure Security Group', and '7. Review'. The main content area is titled 'Step 1: Choose an Amazon Machine Image (AMI)'. It lists several AMI options:

- SUSE Linux Enterprise Server 12 SP3 (HVM), SSD Volume Type - ami-81e29cf9 (Free tier eligible)
- Ubuntu Server 16.04 LTS (HVM), SSD Volume Type - ami-db710fa3 (Free tier eligible)
- Amazon RDS (Launch a database using RDS button)
- Microsoft Windows Server 2016 Base - ami-c9deaf1 (Free tier eligible)
- Deep Learning AMI (Ubuntu) Version 9.0 - ami-0faada77 (Free tier eligible)
- Deep Learning AMI (Amazon Linux) Version 9.0 - ami-94a9d9ec

The 'Deep Learning AMI (Ubuntu) Version 9.0' option is highlighted with a red border.

Select the Deep Learning AMI. It should have CUDA and all of the ML frameworks installed (ami-0faada77)

If you can't find it just search Deep Learning AMI (Ubuntu)

<https://aws.amazon.com/marketplace/pp/B077GCH38C>

Step 7

The screenshot shows the AWS EC2 Management Console Launch Instance Wizard, Step 2: Choose an Instance Type. The table lists the following instance types:

Compute optimized	c4.large	2	3.75	EBS only	Yes	Moderate	Yes
Compute optimized	c4.xlarge	4	7.5	EBS only	Yes	High	Yes
Compute optimized	c4.2xlarge	8	15	EBS only	Yes	High	Yes
Compute optimized	c4.4xlarge	16	30	EBS only	Yes	High	Yes
Compute optimized	c4.8xlarge	36	60	EBS only	Yes	10 Gigabit	Yes
FPGA instances	f1.2xlarge	8	122	1 x 470 (SSD)	Yes	Up to 10 Gigabit	Yes
FPGA instances	f1.16xlarge	64	976	4 x 940 (SSD)	Yes	25 Gigabit	Yes
GPU graphics	g3.4xlarge	16	122	EBS only	Yes	Up to 10 Gigabit	Yes
GPU graphics	g3.8xlarge	32	244	EBS only	Yes	10 Gigabit	Yes
GPU graphics	g3.16xlarge	64	488	EBS only	Yes	25 Gigabit	Yes
GPU instances	g2.2xlarge	8	15	1 x 60 (SSD)	Yes	High	-
GPU instances	g2.8xlarge	32	60	2 x 120 (SSD)	-	10 Gigabit	-
GPU compute	p2.xlarge	4	61	EBS only	Yes	High	Yes
GPU compute	p2.8xlarge	32	488	EBS only	Yes	10 Gigabit	Yes
GPU compute	p2.16xlarge	64	732	EBS only	Yes	25 Gigabit	Yes
GPU compute	p3.2xlarge	8	61	EBS only	Yes	Up to 10 Gigabit	Yes

At the bottom right, a red box highlights the "Review and Launch" button.

Select the t2-micro for the recitation because it is “free tier eligible”, we will explain GPU access in the later slides

Hit Next after selecting the instance

Step 8

The screenshot shows the AWS EC2 Management Console Launch Instance Wizard at Step 3: Configure Instance Details. The URL is https://us-west-2.console.aws.amazon.com/ec2/v2/home?region=us-west-2#LaunchInstanceWizard:.

The page displays various configuration options:

- Number of instances:** 1 (highlighted)
- Purchasing option:** Request Spot instances (unchecked)
- Network:** vpc-cb9090b2 (default) (highlighted)
- Subnet:** No preference (default subnet in any Availability Zone)
- Auto-assign Public IP:** Use subnet setting (Enable)
- Placement group:** Add instance to placement group (unchecked)
- IAM role:** None (highlighted)
- Shutdown behavior:** Stop
- Enable termination protection:** Protect against accidental termination (unchecked)
- Monitoring:** Enable CloudWatch detailed monitoring (Additional charges apply)
- EBS-optimized instance:** Launch as EBS-optimized instance (Additional charges apply)
- Tenancy:** Shared - Run a shared hardware instance (Additional charges will apply for dedicated tenancy)

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

We don't really use this page just hit next

Hit next

Step 9

EC2 Management Console

Secure | https://us-west-2.console.aws.amazon.com/ec2/v2/home?region=us-west-2#LaunchInstanceWizard:

David Oregon Support

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

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-0f680d89f3b5e5284	75	General Purpose SSD (GP2)	225 / 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**

Feedback English (US)

© 2008 - 2018, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use

To use the Deep Learning AMI you need at least 75GB so just leave this alone for now, in some later assignments you're probably going to have to go bigger

Hit Next

Step 10

The screenshot shows the AWS EC2 Management Console with the URL <https://us-west-2.console.aws.amazon.com/ec2/v2/home?region=us-west-2#LaunchInstanceWizard>. The page is titled "Step 5: Add Tags". The navigation bar at the top includes "Services", "Resource Groups", and "Support". Below the navigation, a progress bar shows steps 1 through 7: 1. Choose AMI, 2. Choose Instance Type, 3. Configure Instance, 4. Add Storage, 5. Add Tags (which is highlighted in red), 6. Configure Security Group, and 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)

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

Feedback English (US) © 2008 - 2018, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use

We don't really use tags just hit next don't bother with this page

Hit next

Step 11

The screenshot shows the AWS EC2 Management Console. The URL is https://us-west-2.console.aws.amazon.com/ec2/v2/home?region=us-west-2#LaunchInstanceWizard:6. The top navigation bar includes 'Services' and 'Resource Groups'. Below the navigation, a progress bar shows steps 1 through 7. Step 6, 'Configure Security Group', is highlighted.

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: launch-wizard-1

Description: launch-wizard-1 created 2018-06-02T12:31:49.774-07:00

Type	Protocol	Port Range	Source	Description
SSH	TCP	22	Custom 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**

Feedback English (US) © 2008 - 2018, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use

Leave this page alone unless you know what you are doing. To get a Jupyter notebook to run on your AWS instance you need to add a couple more ports to open, if we have time that this step will covered for that.

Hit Review and Launch

Step 12

The screenshot shows the AWS EC2 Management Console Launch Instance Wizard at Step 7: Review Instance Launch. The URL is https://us-west-2.console.aws.amazon.com/ec2/v2/home?region=us-west-2#LaunchInstanceWizard:.

The navigation bar includes AWS logo, Services (selected), Resource Groups, David (user), Oregon (region), Support, and a bell icon.

The steps are: 1. Choose AMI, 2. Choose Instance Type, 3. Configure Instance, 4. Add Storage, 5. Add Tags, 6. Configure Security Group, 7. Review (selected).

Step 7: Review Instance Launch

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.

AMI Details

Deep Learning AMI (Ubuntu) Version 9.0 - ami-0faada77
Free tier eligible

Comes with latest binaries of deep learning frameworks pre-installed in separate virtual environments: MXNet, TensorFlow, Caffe, Caffe2, PyTorch, Keras, Chainer, Theano and CNTK. Fully-configured with Nvidia CUDA, cuDNN and NCCL as well as Intel MKL-DNN.

Root Device Type: ebs Virtualization type: hvm

Instance Type

Instance Type	ECUs	vCPUs	Memory (GiB)	Instance Storage (GB)	EBS-Optimized Available	Network Performance
p2.xlarge	11.75	4	61	EBS only	Yes	High

Security Groups

Type	Protocol	Port Range	Source	Description
SSH	TCP	22	0.0.0.0/0	launch-wizard-1

Instance Details

Storage

Tags

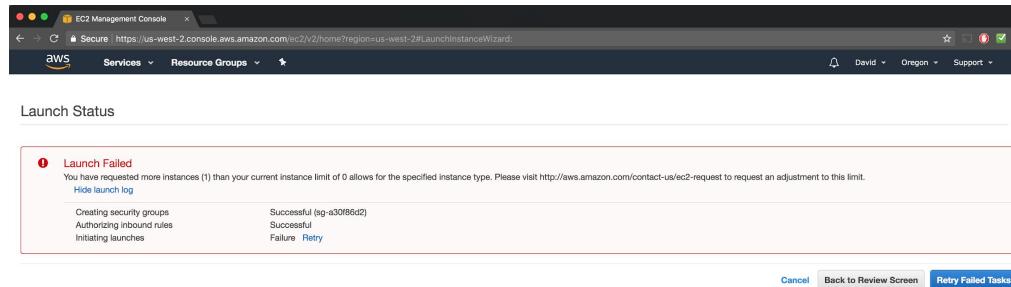
Buttons: Cancel, Previous, **Launch** (highlighted with a red box), Next, Done.

Bottom Bar: Feedback, English (US), © 2008 - 2018, Amazon Web Services, Inc. or its affiliates. All rights reserved., Privacy Policy, Terms of Use.

Hit launch

You need permission to get a GPU

- You will run into an error launching this instance unless you have already requested an AWS GPU instance before
- Looks like this:



Follow the instructions

- Visit the link provided by the error page in the previous slide
<https://aws.amazon.com/contact-us/ec2-request>
- Request a reasonable number of p2.xlarge instances
- Give them a reason why, give them a way to contact you
- Finish and wait a couple days for them to give you permission

Screenshots for p2.xlarge request

The screenshot shows the 'Create Case' page in the AWS EC2 Management Console. The 'Regarding' section is set to 'Service Limit Increase'. The 'Limit Type' dropdown is set to 'EC2 Instances'. A red box highlights the 'Request 1' section where the primary instance type is specified as 'p2.xlarge'. The 'New limit value*' field contains the number '5'. The 'Basic Support Plan' is selected.

EC2 Management Console AWS Support Dashboard

Secure | https://console.aws.amazon.com/support/v1#/case/create?issueType=service-limit-increase&limitType=service-code-ec2-instances&serviceLimitIncreaseType=ec2-instances&type=service_limit_increase

David Global Support

Support Center Account Number [REDACTED]

Dashboard Create Case Case History

Create Case

Basic Support Plan Change

Name [REDACTED]

Account [REDACTED]

Regarding:

- Account and Billing Support
- Service Limit Increase
- Technical Support
Unavailable under the Basic Support Plan

Limit Type*

EC2 Instances

Request 1

Region* US West (Oregon)

Primary Instance Type* p2.xlarge

Limit* Instance Limit

New limit value* 5

Add another request

Feedback English (US) © 2008 - 2018, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use

Screenshot 2 for p2.xlarge

EC2 Management Console AWS Support Dashboard

Secure | https://console.aws.amazon.com/support/v1#/case/create?issueType=service-limit-increase&limitType=service-code-ec2-instances&serviceLimitIncreaseType=ec2-instances&type=service_limit_increase

AWS Services Resource Groups David Global Support Account Number [REDACTED]

Support Center

Dashboard Create Case Case History

Add another request

Use Case Description* I need GPUs for some deep learning projects that I do on the side and I don't have a GPU of my own. Will be using it for Kaggle competitions.

Support Language* English Please choose your preferred correspondence language for this case.

Contact method* Web Phone

Country/Region* United States

Phone Number* Ext.
Please enter digits only (1234567890).
Please do not include the country code.

Submit

Feedback English (US) © 2006 - 2018, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use

Billing, AWS coupons... i.e. don't go broke

- Amazon charges you for a lot of services, the most significant is having running instances (GPU time is NOT cheap) **MAKE SURE YOU STOP YOUR INSTANCE WHEN YOU ARE DONE USING IT**
- Stopping an instance is basically shutting down a computer; the saved files persist, etc...
- Terminating an instance deletes the entire machine only do this when you completely done with whatever you are doing with your AWS instance
- To stop/terminate an instance simply go to step 4 from tutorial then right click on the instance and under the option instance state, hit stop/terminate, the next slide will show you how
- We will provide you with AWS tokens, you need to redeem them from Billing, the next next slide will show you how

Stopping the Instance

The screenshot shows the AWS EC2 Instances page. On the left, there's a sidebar with various navigation options like EC2 Dashboard, Events, Tags, Reports, Limits, Instances, Launch Templates, Spot Requests, Reserved Instances, Dedicated Hosts, Scheduled Instances, Images, AMIs, Bundle Tasks, Elastic Block Store, Volumes, Snapshots, Lifecycle Manager, Network & Security, Security Groups, Elastic IPs, Placement Groups, Key Pairs, Network Interfaces, Load Balancing, Target Groups, and Auto Scaling.

The main area displays a single instance named "i-0696d90375614bbd0" with a Private IP of 172.31.38.210. The instance is currently in a "stopped" state, indicated by a red circle icon. A context menu is open over this instance, with a red box highlighting the "Actions" dropdown. The dropdown menu includes options: Connect, Get Windows Password, Launch More Like This, Instance State (with "stopped" selected), Instance Settings, Image, Networking, and CloudWatch Monitoring.

Description	Value
Instance ID	i-0696d90375614bbd0
Instance state	stopped
Instance type	p2.xlarge
Elastic IPs	-
Availability zone	us-west-2b
Security groups	launch-wizard-1, view inbound rules, view outbound rules
Public DNS (IPv4)	-
IPv4 Public IP	-
IPv6 IPs	-
Private DNS	ip-172-31-38-210.us-west-2.compute.internal
Private IPs	172.31.38.210
Secondary private IPs	-

When you're done
with your instance
**MAKE SURE YOU
STOP IT OR YOU
WILL RUN OUT
OF AWS
CREDITS**

Billing

The screenshot shows the AWS EC2 Dashboard. On the left, there's a sidebar with various navigation options like Services, Resource Groups, EC2 Dashboard, Instances, Launch Templates, Spot Requests, Reserved Instances, Dedicated Hosts, Scheduled Instances, Images, AMIs, and Auto Scaling. The main content area displays EC2 resource statistics for the US West (Oregon) region, including 0 Running Instances, 0 Elastic IPs, 0 Dedicated Hosts, 0 Snapshots, 1 Volumes, 0 Load Balancers, 1 Key Pairs, and 3 Security Groups. Below this is a 'Create Instance' section with a 'Launch Instance' button. To the right, there are sections for Service Health, Scheduled Events (showing 'No events'), and AWS Marketplace (listing products like Barracuda CloudGen Firewall for AWS - PAYG and Matillion ETL for Amazon Redshift). At the bottom, there are links for Feedback, English (US), and other AWS services.

Under your name hit
My Billing Dashboard

Billing/Redeeming Tokens

The screenshot shows the AWS Billing & Cost Management Dashboard. The left sidebar includes links for Dashboard, Bills, Cost Explorer, Budgets, Reports, Cost Allocation Tags, Payment Methods, Payment History, Consolidated Billing, Preferences, Credits (which is highlighted with a red box), and Tax Settings.

The main content area features a "Month-to-Date Spend by Service" donut chart showing \$0 spent. Below it is a "Spend Summary" table:

	No Amount Due	\$0.00
Tax	\$0.00	\$0.00
Total	\$0.00	\$0.00

A bar chart titled "Current month-to-date balance for August 2018" compares "Last Month (July 2018)" at \$1.71 and "Forecast (August 2018)" at \$3.59. A link "Important Information about these Costs" is present below the chart.

At the bottom, there's a "Top Free Tier Services by Usage" section with a "View all" button.

Page URL: https://console.aws.amazon.com/billing/home?region=us-west-2#/paymenthistory

This is where you can see your bill. Hit Credits if you want to redeem your AWS tokens

Redemption Page

The screenshot shows the AWS Billing Management Console interface. On the left, a sidebar lists various services: Dashboard, Bills, Cost Explorer, Budgets, Reports, Cost Allocation Tags, Payment Methods, Payment History, Consolidated Billing, Preferences, Credits (which is selected and highlighted in orange), and Tax Settings. The main content area is titled 'Credits' and contains instructions: 'Please enter your code below to redeem your credits.' Below this is a 'Promo Code' input field, a 'Security Check' CAPTCHA field containing 'm8wdwx' with a 'Refresh Image' link, and a text input field for typing the characters shown. A note below the CAPTCHA says: 'By clicking "Redeem" you indicate that you have read and agree to the terms of the AWS Promotional Credit Terms & Conditions located [here](#)'. At the bottom of this section is a blue 'Redeem' button. Below this, a table displays redeemed credits. The table has columns: Expiration Date, Credit Name, Amount Used, Amount Remaining, and Applicable Products. One row is shown: '2019-03-31' under 'Expiration Date', 'EDU_ENG_FY2018_IC_Q1_2_CMU_50USD' under 'Credit Name', '\$0.00' under 'Amount Used', '\$50.00' under 'Amount Remaining', and a 'See complete list' link under 'Applicable Products'. At the bottom of the page, a note states: 'Total Credit Amount Remaining (as of 2018-08-01): \$50.00'.

This is where you
redeem your tokens, just
follow the instructions
and enter your token

TMUX makes your life easier

- Use tmux if you don't want to wait on your computer like its your child
- After SSHing into an AWS instance you can type in the command tmux
- Then run anything as would normally
- Usually when you run something on SSH if your connection dies (ie when your computer sleeps/etc) the program you are running in the SSH session dies too
- tmux prevents this (you can close out the session whenever)
- To reconnect just SSH back in and then do: tmux ls
- And then you should see your session (your first unnamed session is 0)
- tmux attach-session -t 0

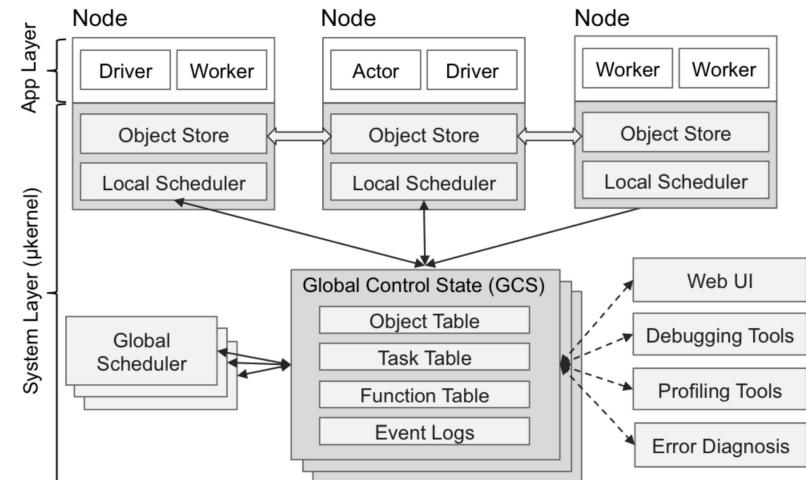
Using S3

- Instead of reading through more powerpoint you can just follow Amazon's tutorial (first one is for the GUI, second one is for the CLI)
- You would use S3 if you don't have enough storage on your local machine and you don't want to store it on EC2 storage (EBS is the EC2 HDD service) either
- Good for sharing large files has high bandwidth to AWS instances in the same region
- We might use it to share data files you with guys
- <https://docs.aws.amazon.com/AmazonS3/latest/gsg/GetStartedWithS3.html>
- <https://docs.aws.amazon.com/cli/latest/reference/s3/index.html>

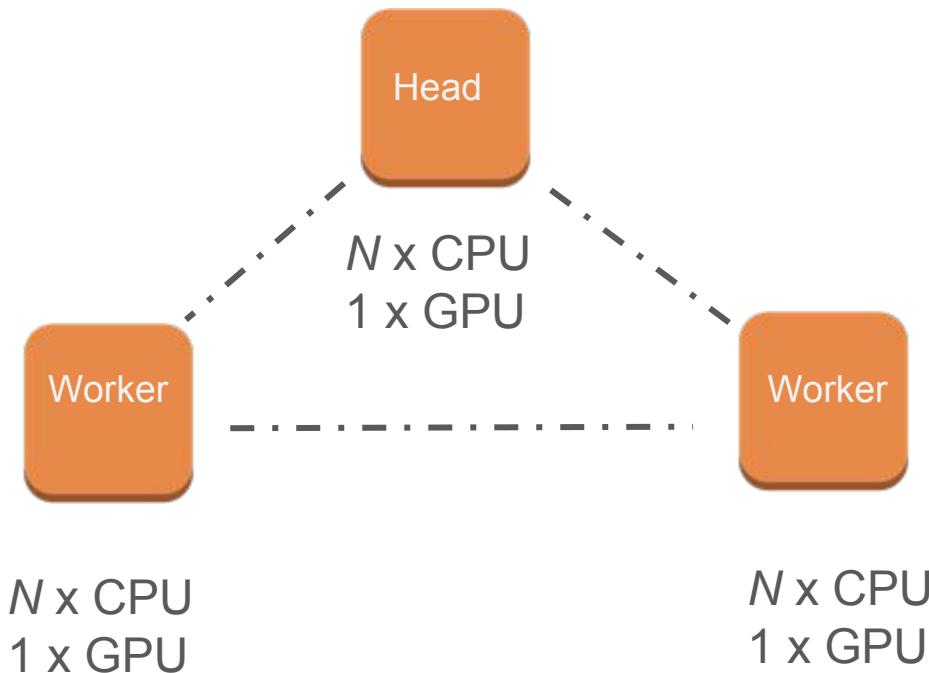
Can we develop scalable deep learning code that only relies on resource requirements and is agnostic to a specific system's topology?

Ray - A Framework for Distributed Deep Learning

- Write code with respect to available **resources** (CPUs, GPUs, etc).
- Don't worry about **placement** (local host, on-demand cloud instances, spot instances, etc).
- Write everything in a pythonic, declarative manner.
- Ray will take care of scheduling, serialization, and communication.



Ray - A Framework for Distributed Deep Learning



Simple Code Example 1

```
import time

def f1():
    time.sleep(1)

@ray.remote
def f2():
    time.sleep(1)

# The following takes ten seconds.
[f1() for _ in range(10)]

# The following takes one second (assuming the system has at least ten CPUs).
ray.get([f2.remote() for _ in range(10)])
```

Simple Code Example 2

```
@ray.remote
def f(x):
    return x + 1

x = f.remote(0)
y = f.remote(x)
z = f.remote(y)
ray.get(z) # 3
```

Install Ray

```
>>> pip install ray
```

Registering AWS credentials on your local machine

1. Open the [IAM console](#).
2. In the navigation pane of the console, choose **Users**.
3. Choose your IAM user name (not the check box).
4. Choose the **Security credentials** tab and then choose **Create access key**.
5. To see the new access key, choose **Show**. Your credentials will look something like this:
 - Access key ID: AKIAIOSFODNN7EXAMPLE
 - Secret access key: wJalrXUtnFEMI/K7MDENG/bPxRfiCYEXAMPLEKEY
6. To download the key pair, choose **Download .csv file**. Store the keys in a secure location.

Registering AWS credentials on your local machine

- You will need to provide AWS's CLI client with the credentials you just created.
- First install the awscli and boto3 packages with pip.
- Next add the `~/.aws/credentials` file with the following template

```
1 [default]
2 aws_access_key_id = [YOUR_ACCESS_KEY_ID]
3 aws_secret_access_key = [YOUR_SECRET_ACCESS_KEY]
```

Specify your cluster template

- Provide Ray your configuration preferences and let it do the EC2 setup work for you, even if you only need a single node in the cluster.
- Configuration template provided in the the recitation1 directory in the 11-785 tutorials repo on GitHub as `dev_template.yml`
 - Specify instance type, number of workers, AMI, volume size, and more

Launch the cluster

- Copy the template into a file you can edit.

```
$ cp dev_template.yml dev.yml
```

- Make your changes to the config and then launch the cluster.

```
$ ray create_or_update dev.yml
```

- The output will give you a command to log into the cluster with.

Stop the cluster

- Stop and terminate all cluster instances.

```
$ ray teardown dev.yml
```

WARNING: This will destroy all cluster instance and volumes. Make sure you copy your data back to your host machine first.

Always check the AWS console to ensure your instances have been stopped/terminated. You have been warned!