Azure Guide for CS224n

- This guide will help you setup and use Azure Virtual Machines for your final project.
- Before we start, it cannot be stressed enough: **do not leave your machine running when you are not using it**.
- The expected time to complete the setup guide is **15 min** to **1 hour**, depending on which configuration you opt to take.

Contents

Your Azure subscription for this class

Microsoft has generously agreed to sponsor CS224n, and has provided us with Azure credit to distribute to CS224n students. We expect that there will be enough credit for teams to run as many experiments as they need for their projects. **However, it's very important for students to manage their credit carefully, so that they can get the most out of it (see next section).**

You should receive an email on Tuesday or Wednesday (Feb 13/14) with an invitation to claim your initial Azure credit of **$150**. Credit has been assigned per team (according to the teams you

gave us in your project proposal), with the same amount allocated regardless of team size. The $150 corresponds to about **126 hours**, or slightly over **5 days** on a NV6 machine.

The $150 is an initial allocation. If you use it up running *genuine* experiments, that's **perfectly OK and completely expected** – we expect that most teams will need more credit, and we have plenty more to give you. However, please don't use up your credit by leaving your machine running when you're not using it! Nor should you use up many hours of credit using your VM to write your code (see next section).

When you run out of credit (or before you run out), you can ask us for more on Piazza using the "azure" tag.

Best practices for managing your Azure credit

Azure virtual machines are charged at a flat rate, for each minute that they are turned on. This is irrespective of:

- whether you are ssh'd to the machine at that time
- whether you are running any processes on the machine at that time
- the computational intensity of the the processes you're running
- whether you're using GPUs

Therefore, the most important thing you need to do to, to manage your Azure credit, is to **carefully turn your VM on and off just when you need it**. If you are using a NV6 VM, it is charged at $855.60/month (i.e., approximately $1.19/hour) while it is turned on.

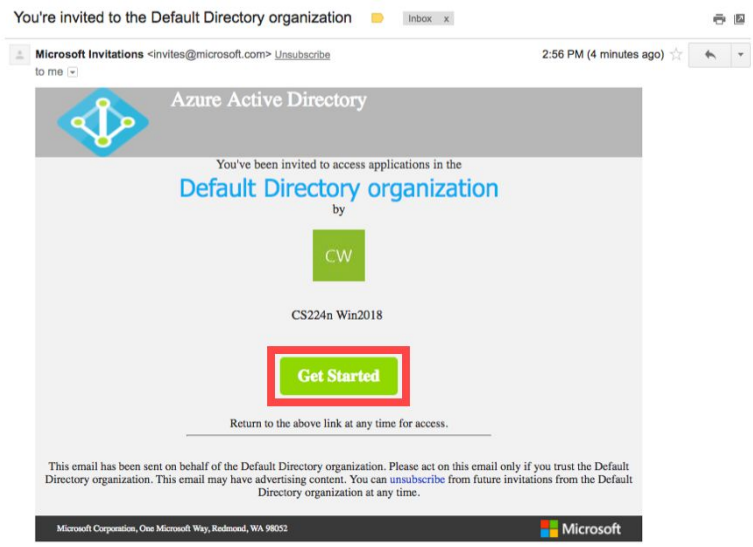
We advise you to **develop your code on your local machine** (for example your laptop with the CPU version of TensorFlow installed) for debugging (i.e., work on your new code until you are able to complete several training iterations without errors), then run your code on your Azure VM when it's time to train on a GPU.

Azure also has an auto-shutdown feature that allows you to specify a time when you want your VM to turn off - this allows you to turn off the machine at a time when you are unable to do it manually. For example, if you start an experiment at 9pm, and you want to stop it after 5 hours, you can set auto-shutdown to turn your VM off at 2am. This will prevent you spending credit that you would have otherwise spent until you woke up many hours later to turn off the VM.

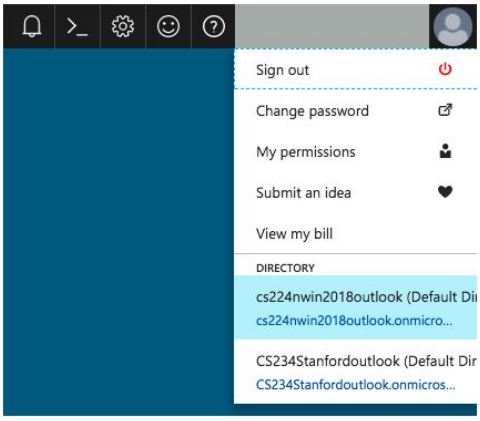
Configuring your Azure VM

Creating an Azure account (5 min)

You should have received an email with an invitation to join a subscription. Click Get Started and follow the instructions, using the email account that received this invitation. After completing the instructions, you will be redirected to your dashboard; if not, you can login to your account at portal.azure.com.



If you have multiple subscriptions (e.g. you're sharing the same email account for CS 224N with another course using Azure like CS 234 or CS 273B), then you must select CS 224N in the Account Menu in the top-right corner.



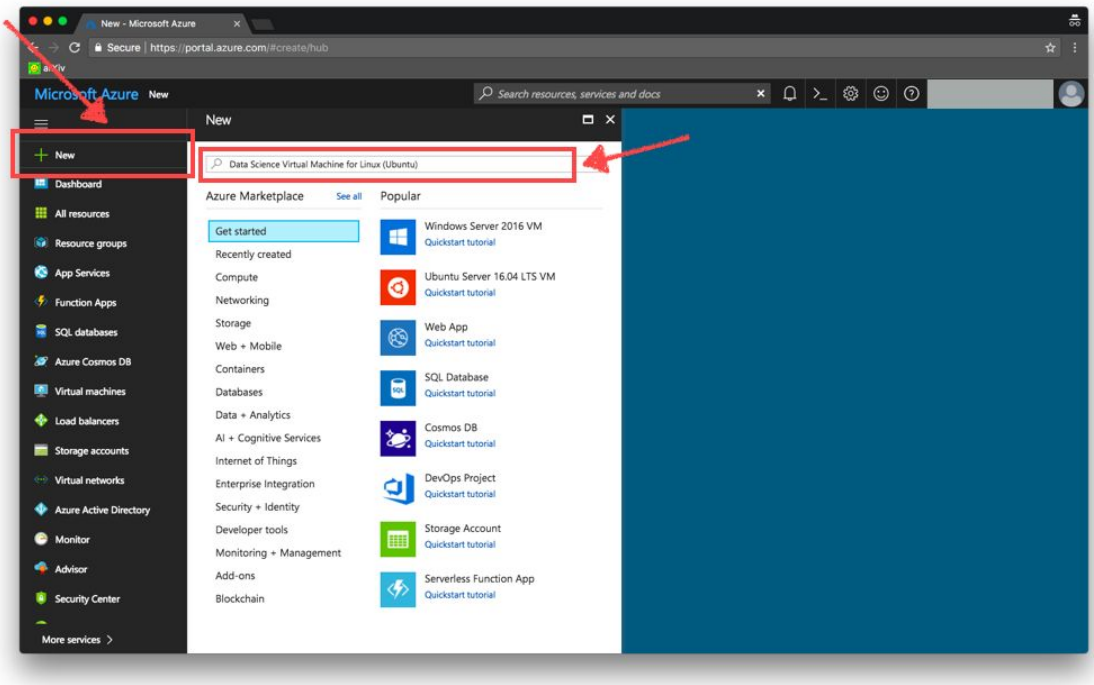
Creating a VM (15-45 min)

In this step, you will have a few different choices. You might prefer to use a predefined image if you want a more convenient setup (see: Using a predefined image), and do not need specific package versions. You might prefer to define your own image if you want more control over the configuration (see: Using your own image).

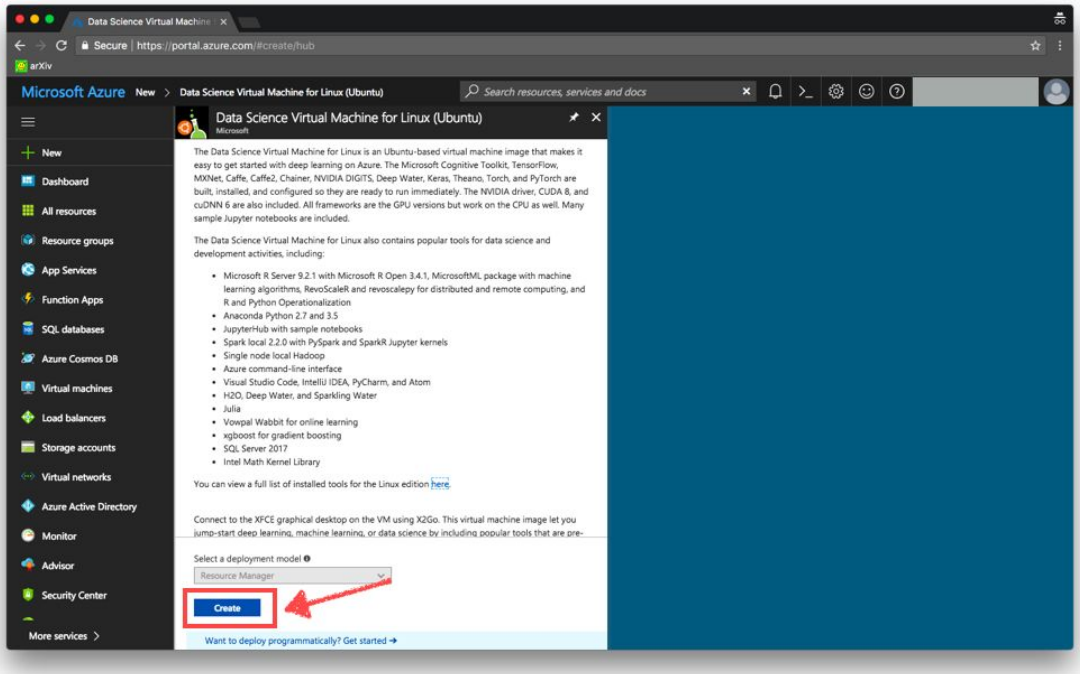
Using a predefined image (15 min)

If you use a predefined image, we recommend using the `Data Science Virtual Machine for Linux (Ubuntu)` image, which comes installed with Python 3.5, tensorflow-gpu, CUDA, and cuDNN.

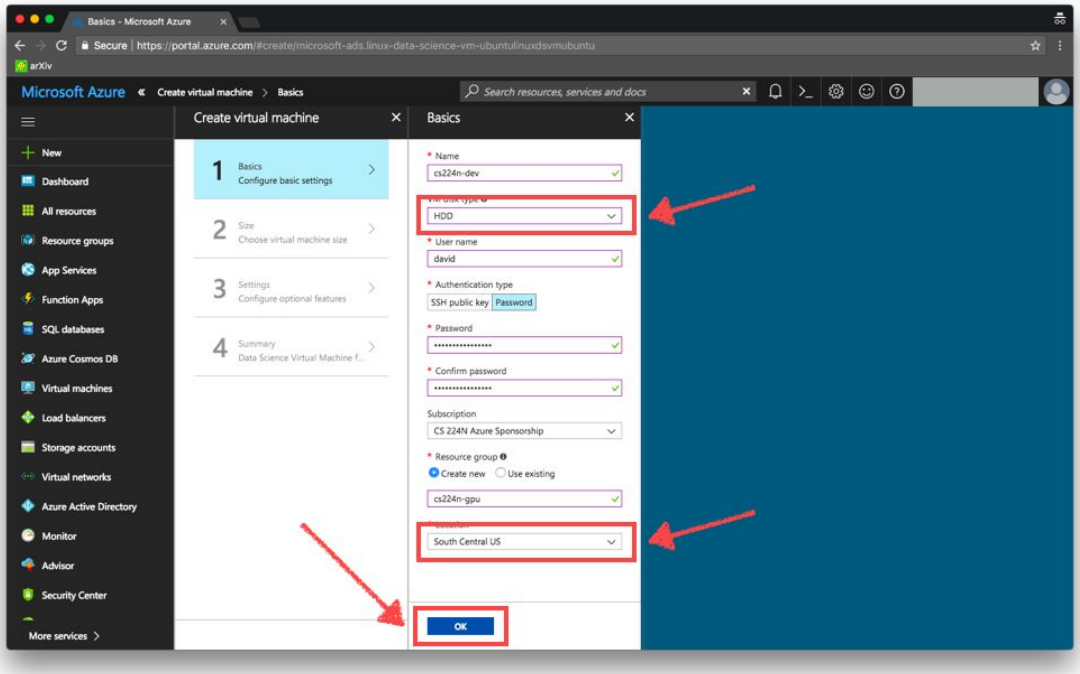
1. Click the `+` in the left sidebar menu and type in `Data Science Virtual Machine for Linux (Ubuntu)`. It's essential that you select the Ubuntu and **not** CentOS distribution.



2. Click `Create`.



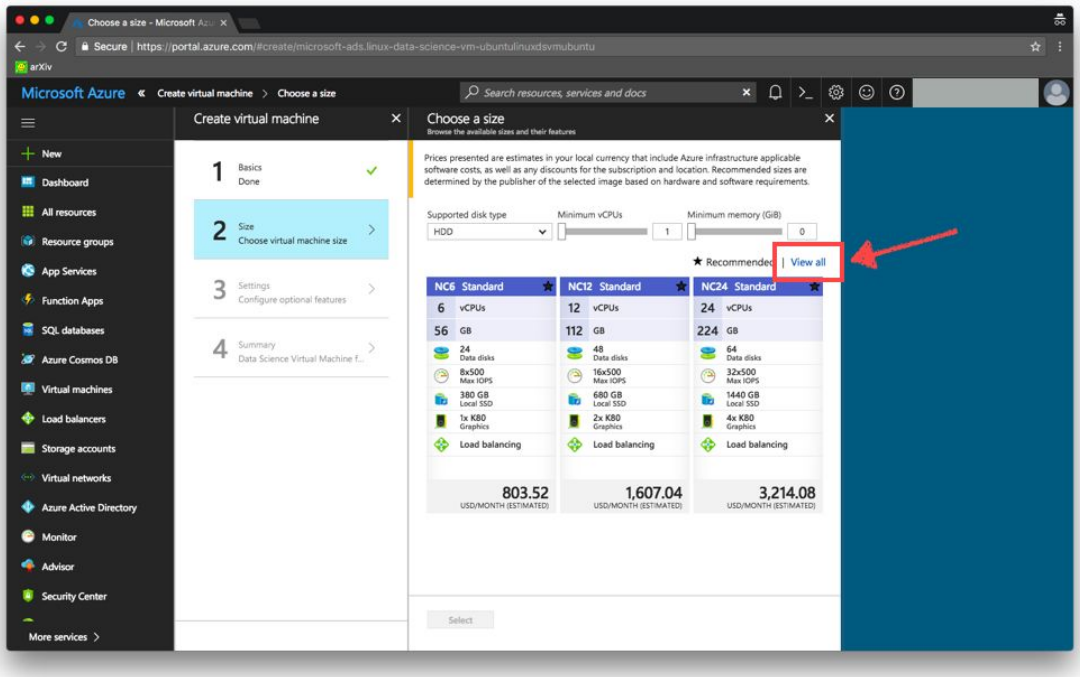
3. Fill in the the following fields … Then click `OK`.



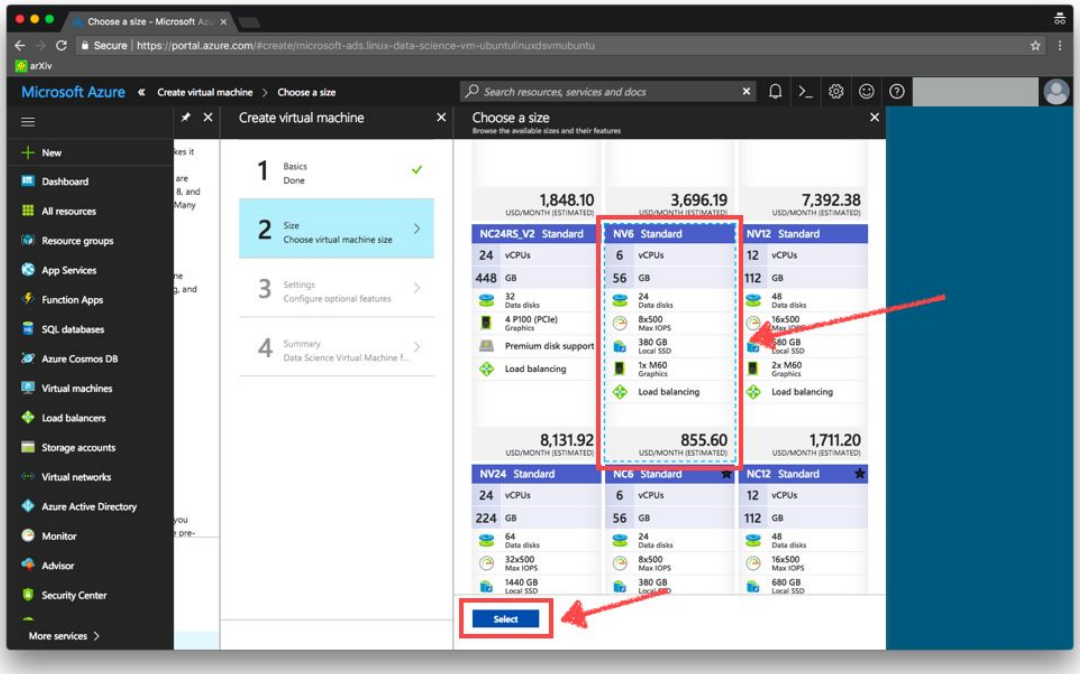
   - **Name.** This will be the name of your VM. You can name it whatever you want.
   - **VM disk type.** You **must** select `HDD`.
   - **User name.** This will be the username used on the VM. You can name yourself whatever you want.

- **Authentication type.** If you are not familiar with SSH keys, authenticate using password; otherwise, choose whichever you prefer.
- **Subscription.** Choose the only available option. If there are multiple options or none, post on Piazza for additional assistance.
- **Resource group.** If you create multiple VMs, those within the same resource group will share resources. Unless you create multiple VMs, this configuration does not matter, so click `Create New` and type `cs224n-gpu`.
- **Location.** You **must** select `South Central US` for Location; otherwise, there might not be any GPUs available to you at the next step.

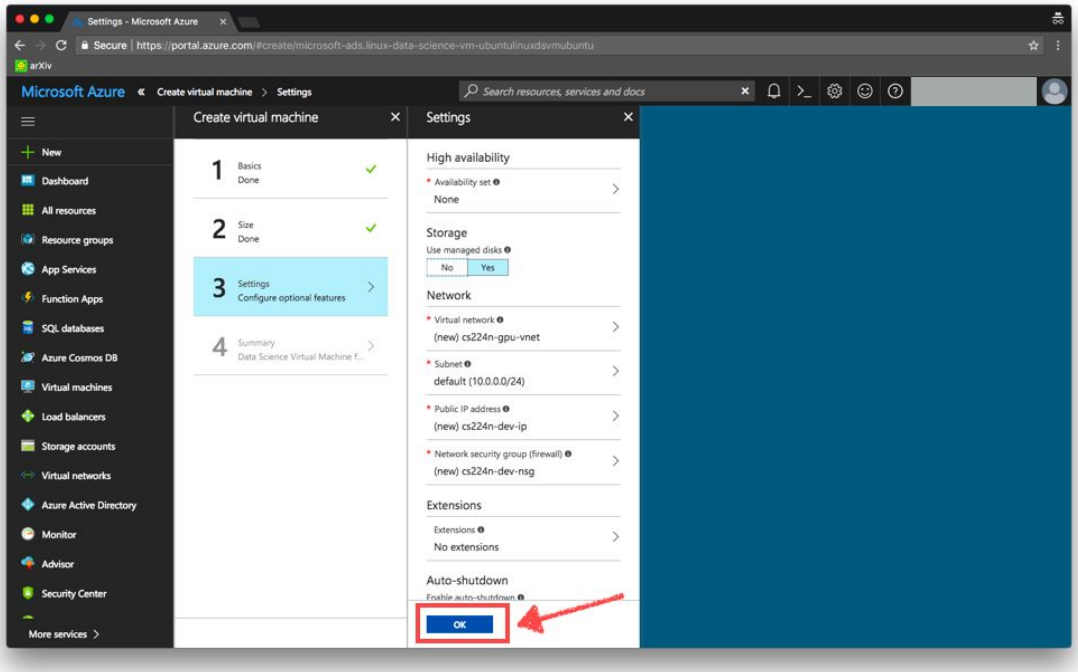
4. Click `View all`.



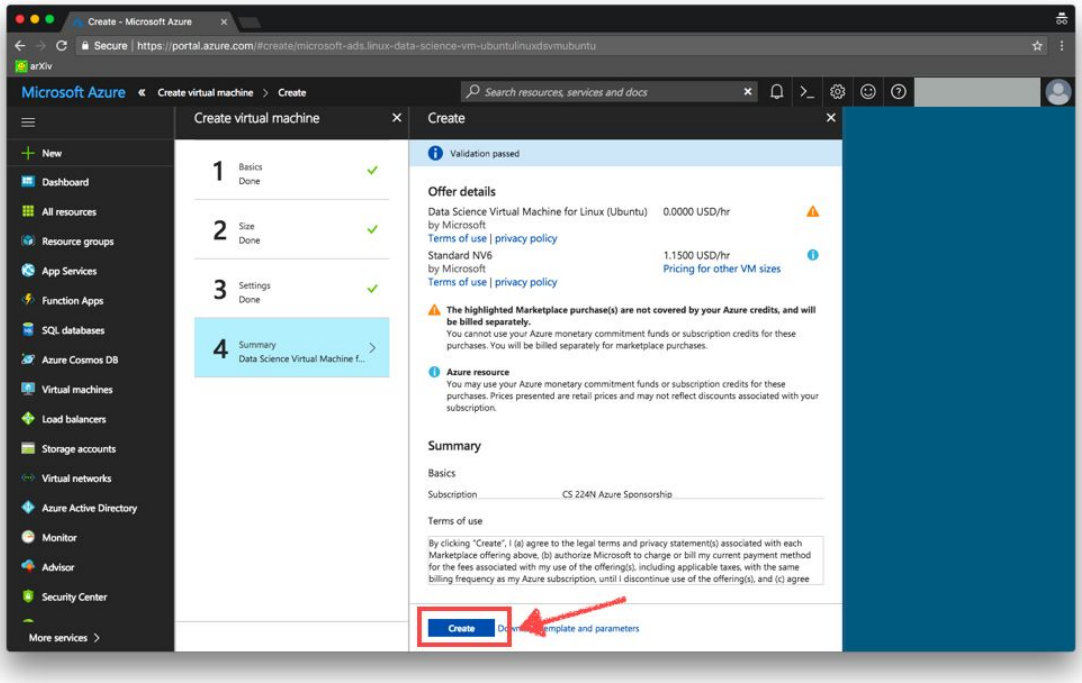
5. Select `NV6 Standard`. Click `Select`.



6. Accept the default configuration. Click `OK`.



7. Wait for the configuration to validate. Click `Create`.

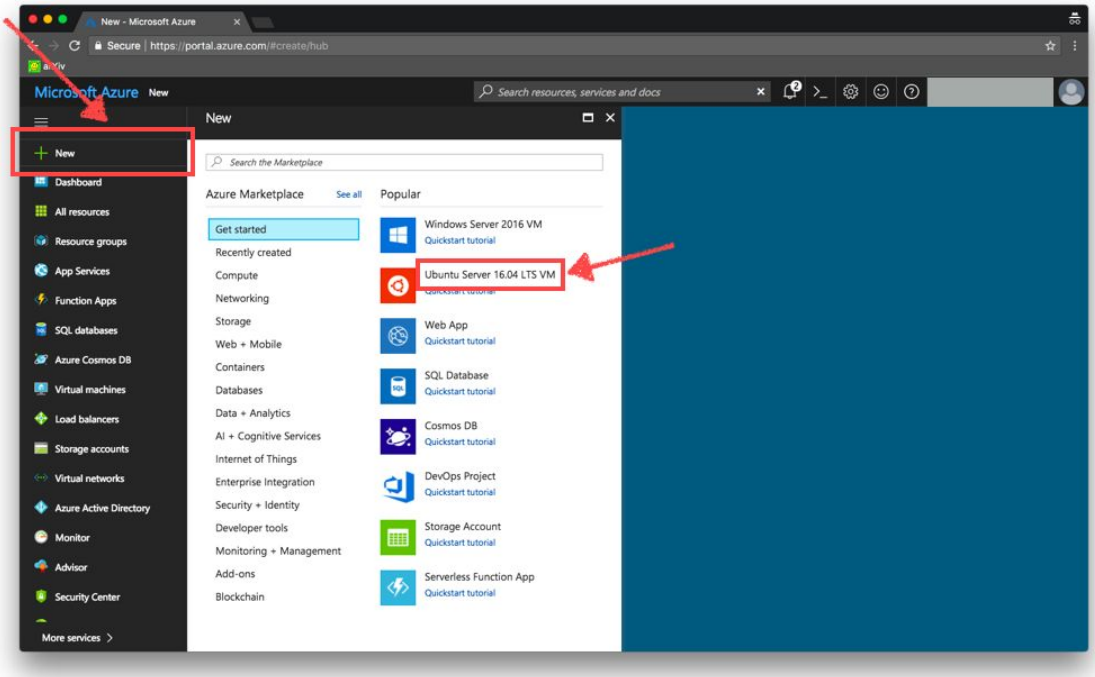


8. You've created a VM! Continue to Using Azure.

Using your own image (45 min)

This walkthrough provides an example for the general steps for creating your own image.

1. Click the `+` in the left sidebar menu and click `Ubuntu Server 16.04 LTS VM`. You might need to type this term into the Search bar if it does not appear initially.



2. Follow steps 3-8 of the Using a predefined image walkthrough.
3. Follow steps 1-3 of the Managing a VM walkthrough.
4. Follow steps 1-2 of the Connecting to a VM walkthrough.
5. Now you will finish configuration of the VM by manually installing the required software. ssh into the VM, download and unzip the setup scripts from the course website, and cd into the resulting directory.

```
$ wget https://web.stanford.edu/class/cs224n/azure/azure-setup.zip
$ sudo apt-get -y install unzip
$ unzip azure-setup.zip
$ cd azure-setup
$ ls
gpu-setup-part1.sh  gpu-setup-part2.sh  README
```

6. Run `gpu-setup-part1.sh`. This will install some libraries, fetch and install NVIDIA drivers, and trigger a reboot. (The command will take some time to run.)

```
$ ./gpu-setup-part1.sh
```

7. Once your VM has finished restarting, ssh into the VM again. cd into the setup directory and run `gpu-setup-part2.sh`. This script installs tensorflow-gpu, CUDA, and cuDNN.

```
$ cd azure-setup
$ ./gpu-setup-part2.sh
```

8. Source `~/.bashrc`.

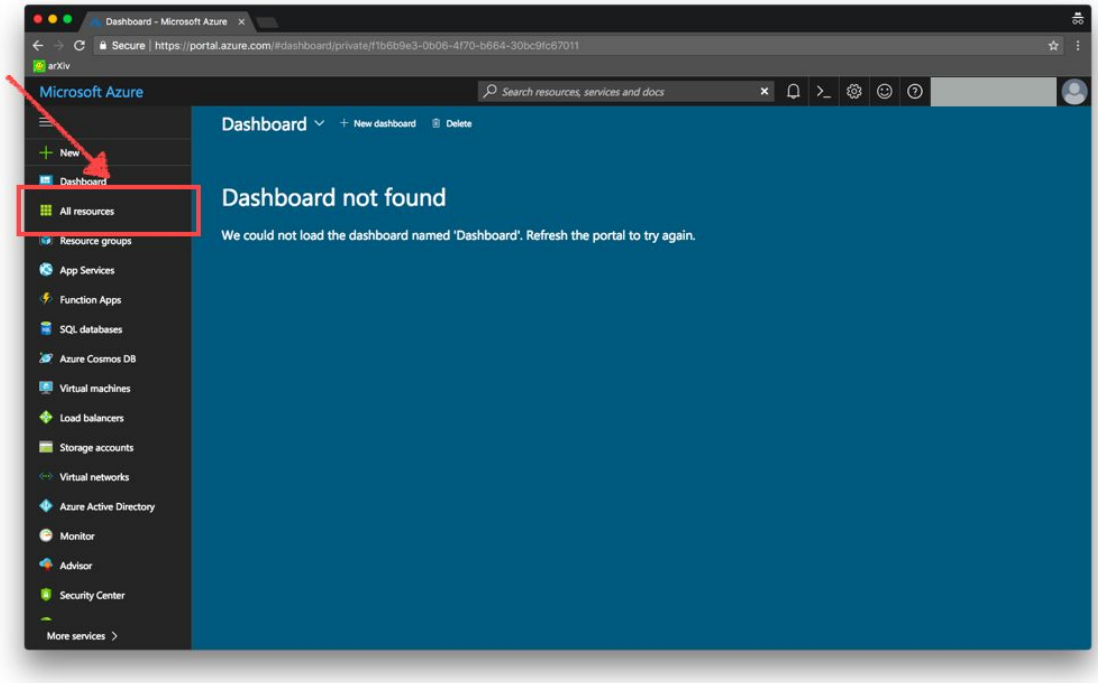
```
$ source ~/.bashrc
```

9. Follow step 3 of the Connecting to a VM walkthrough.

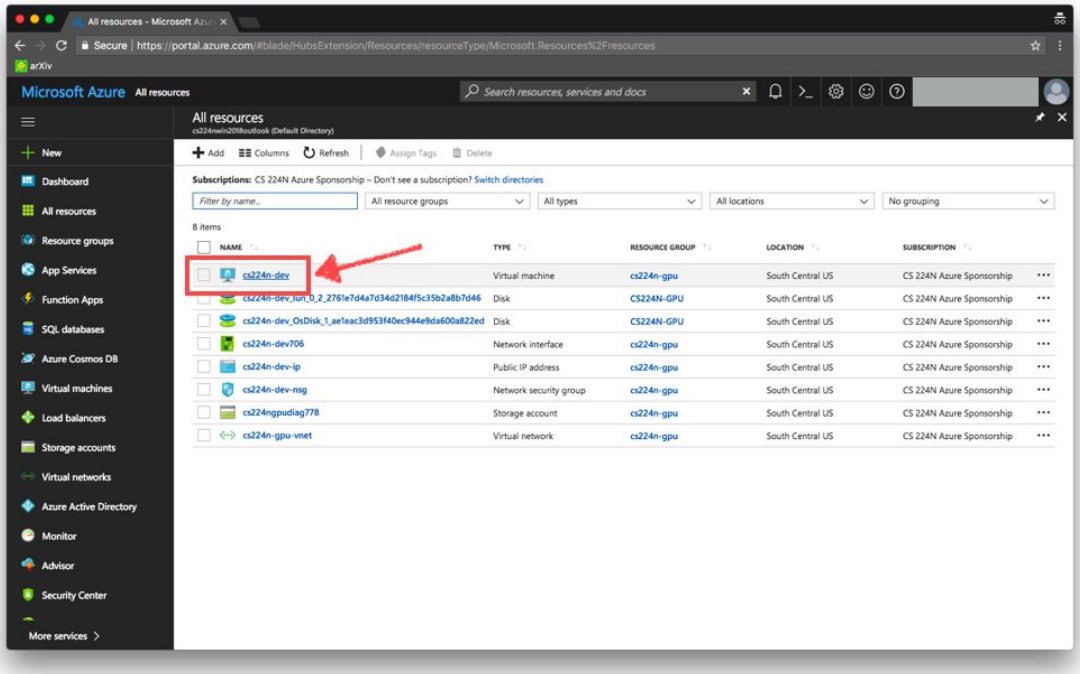
Using Azure

Managing a VM

1. Click the `All resources` in the left sidebar menu.



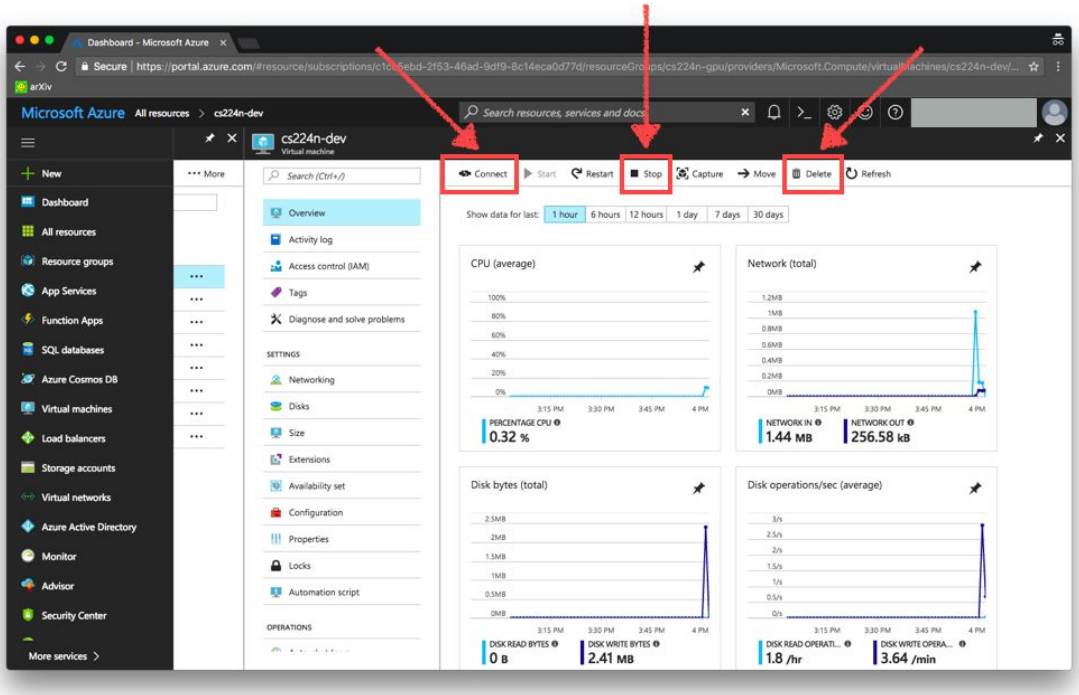
2. Click the name of your VM. You might need to wait up to 10 minutes after creating the VM for it to appear on this menu.



3. There are a few important options. Click `Connect` for an ssh command to connect to your instance. Click `Start`/`Stop` to start or stop the instance. If you want to delete the instance, click `Delete`.

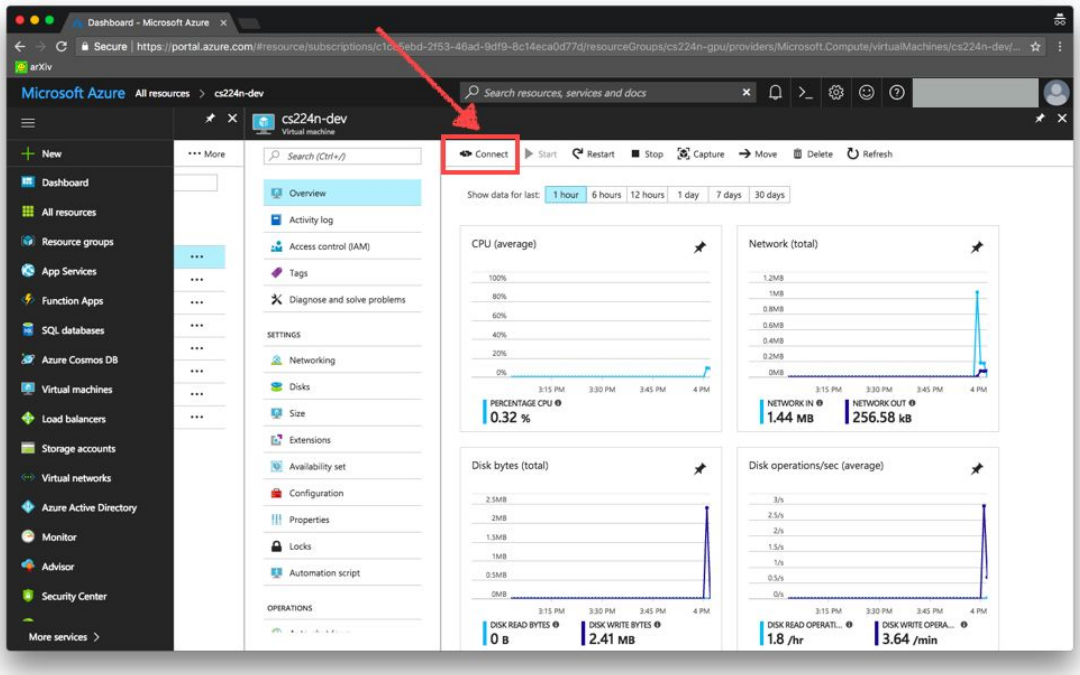
   Note that if your instance is stopped but not deleted, it will still accrue charge for storage. If your instanced is not stopped or deleted, it will accrue charge for storage and VM.

Again, ***do not leave your machine running when you are not using it****.*

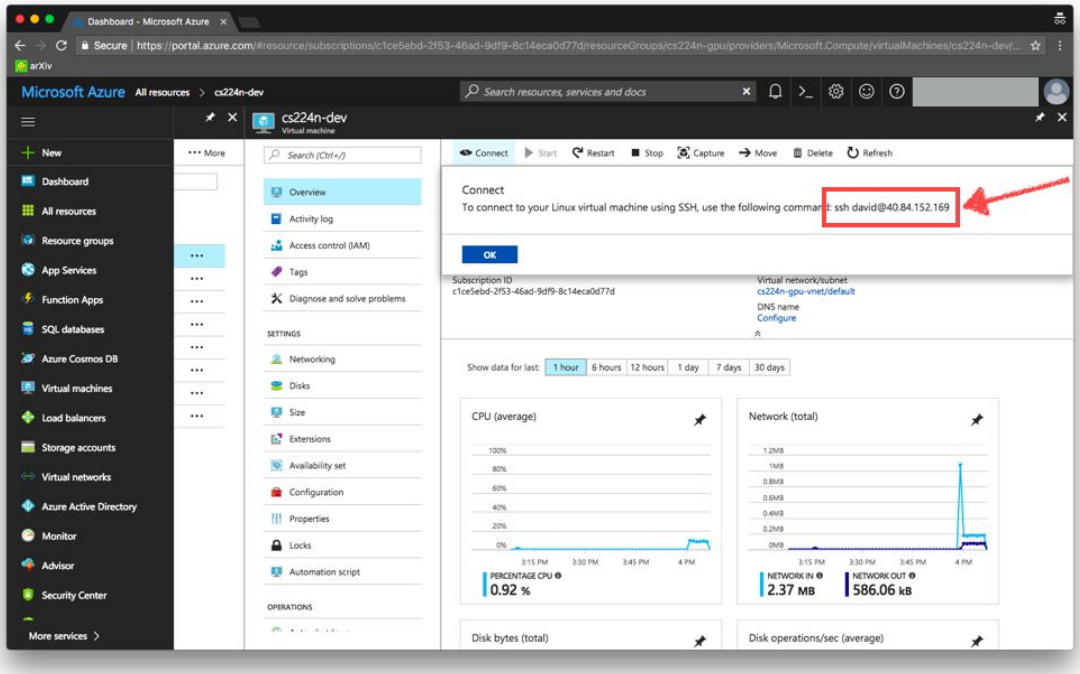


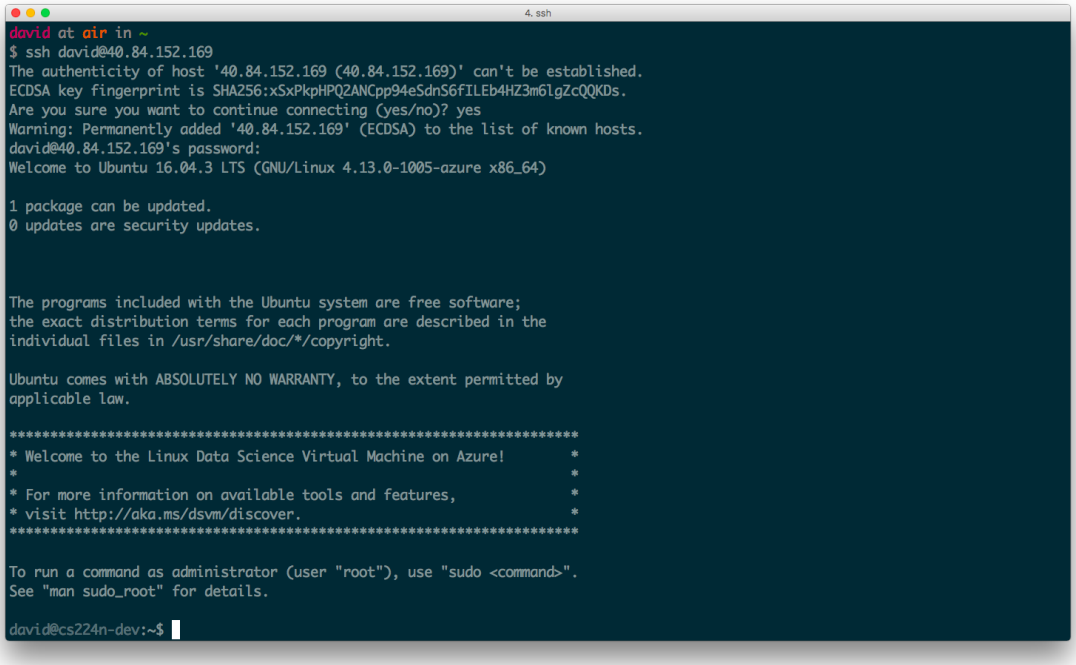
Connecting to a VM

1. Click `Connect` from the previous menu.



2. Copy and paste the command into your terminal.



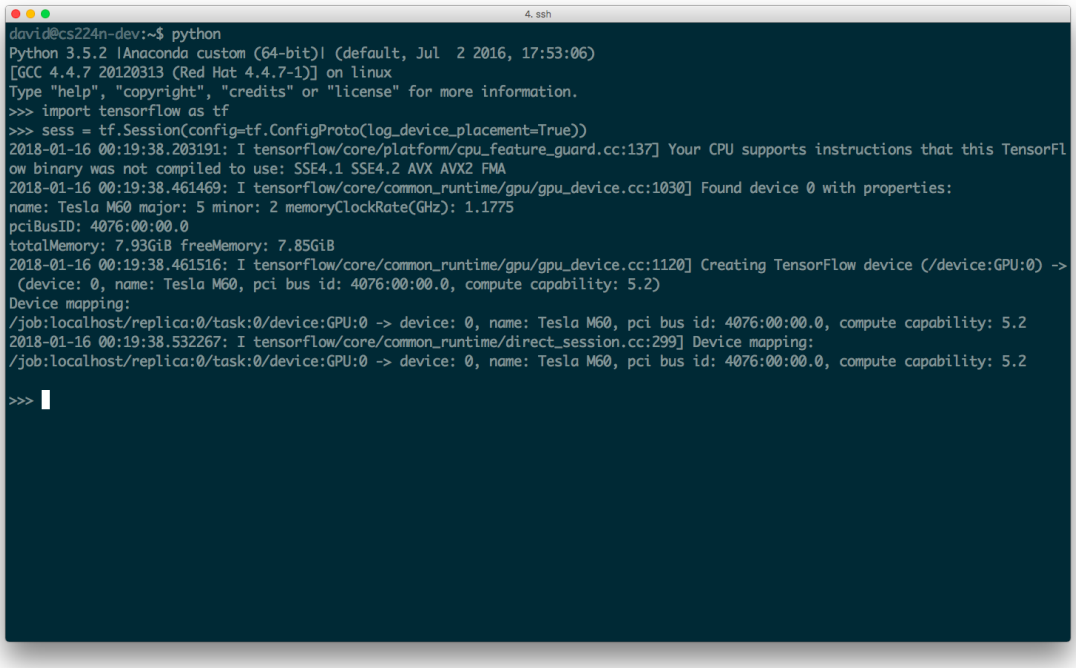


3. Check that Tensorflow can access the GPUs by opening Python and typing the following:

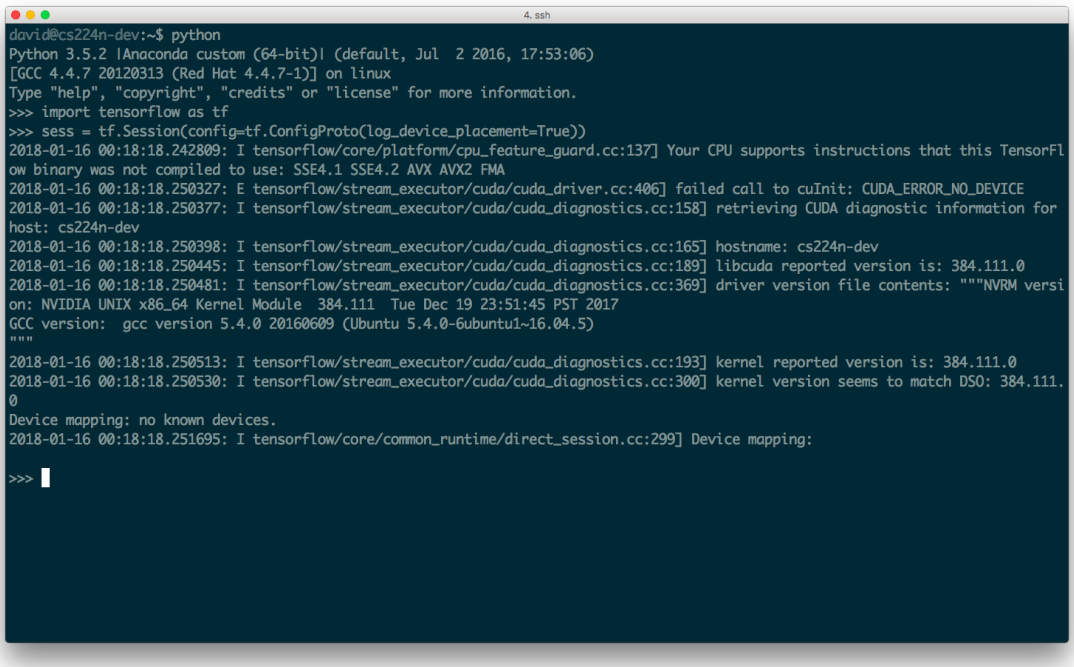
```
import tensorflow as tf
```

```
sess = tf.Session(config=tf.ConfigProto(log_device_placement=True))
```

You should see something like this:



If instead it says `Device mapping: no known devices`, then post on Piazza for additional assistance.



FAQs

How do I check my remaining balance?

Go to https://www.microsoftazuresponsorships.com/manage. Note that Azure bills at midnight every business day, so this figure usually reflects your credit as of the last billing time.

What happens when I exceed my credit?

Azure bills at midnight every business day. So if, at midnight on a business day, you have exceeded your credit, Azure will turn off your VM. Note that at this point, your account may be in debt. If/when this happens to you, let us know on Piazza and tag your post with the "azure" tag.

Can I add a personal credit card to the account?

Sure, though we do not recommend it. If you exhaust the funds from your CS 224N subscription, your personal credit card will be charged without warning.

Can I select more powerful instances?

Though we recommend the NV6, you are free to use any of the instances. Just keep in mind that you have a budget!