CIS 505: Software Systems

Fall 2017

Deploying your project on Amazon EC2

This document briefly summarizes how to get started with Amazon's Cloud. This document assumes that you are using the VM image; if you are not, you may need to install extra software on your computer, e.g., an ssh client.

1. Setup

1.1 Signing up for AWS

1. Go to https://aws.amazon.com/, click **Create a Free Account.** You'll be asked to enter some information, including payment info.
2. Go to https://console.aws.amazon.com/. Click on your name in the top right corner, and click "My Account". In the following screen, note down your Account ID.
3. Go to https://aws.amazon.com/education/awseducate/apply/. Click "Apply for AWS Educate for Students". Choose Student, and fill out your information. (Important: Use a Penn-affiliated email address - something that ends in `upenn.edu`!). Use the Account ID from the previous step. (Note: If you already signed up for AWS Educate for another class, you can skip this step, but keep in mind that you won't get extra free credits, so watch your resource usage carefully!)
4. Check your email, note down your verification code, and enter the code in the verification screen.
5. Accept the terms and conditions.
6. After a few days, you should get an email noting that your application was approved.
7. Make a new subdirectory `.ec2` under your home directory in the VM.
8. Go back to the AWS Management Console (https://console.aws.amazon.com/). Choose **EC2** from the "Services" dropdown list (left column, under "Compute"). Then click on **Key Pairs** in the sidebar on the left and **Create Key Pair**. Save the file as `~/.ec2/login.pem`.
9. Run the command `chmod 400 ~/.ec2/login.pem` (if the permissions on your `ssh` key are too open, `ssh` will not accept the file)

1.2 Configuring the Default Security Group

1. Go to the **AWS Management Console (**https://console.aws.amazon.com/) and sign in.
2. Choose **EC2** from the "Services" dropdown list (left column, under "Compute")
3. Choose **Security Groups** under **Networking & Security** (in the left-hand navigation bar)
4. Select the **default** security group.
5. The default permissions allow for unfirewalled access among Amazon EC2 nodes, but no access from outside. At the very least, we need to enable the HTTP protocol, which operates over TCP. Click on the **Inbound** tab, click Edit and then **Add rule**. In the drop-down select **HTTP**. The **Source** should already be filled in as '0.0.0.0/0'. Click **Add rule**.
6. If your system uses other TCP or UDP ports, be sure to open these as well.
7. Repeat, but select SSH this time. Click **Add rule**.
8. Click **Save.**

2. Elastic Compute Cloud (EC2)

2.1 Launching an EC2 Instance

Go to the **AWS Management Console (**https://console.aws.amazon.com/) and sign in.

- Choose **EC2** from the "Services" dropdown list (left column, under "Compute")
- Click on the **Launch Instance** button. Now you need to choose a type of virtual machine. Select Amazon's basic Linux machine, e.g., **Amazon Linux AMI 2016.09.0 (HVM), SSD Volume**
- Choose the type of instance you want, e.g. **Micro (t2.micro)** or **Small (t2.small)**. For better performance, choose bigger instances, but keep in mind that these are also more expensive. Click "Next: Configure Instance Details".
- Leave **Instance Details** as they are; just click **Next: Add storage**.
- On the **Add Storage** page click **Next: Add Tags**.
- On the **Add Tags** page, give a name to this instance, so that you can identify it in the Amazon Console. Click **Next: Configure Security Group**.
- On the **Configure Security Group**, select the default security group that you defined earlier.
- Click **Review and Launch**. In the review screen, choose **Launch**. Select your login keypair and **Launch Instances**. Recall that you will be billed on an hourly basis, so don't forget to turn it off later! Click **View Instances**.
- In the AWS Management Console and the **Instances** area, wait for the main **My Instances** window to indicate the instance is ready (Status 'running', with a green dot next to it). If the status is 'pending', with a yellow dot next to it, you need to wait a bit.
- Click on the instance, look at the bottom of the pane (Description), and find the **Public IP** entry. This is your instance's public IP address. Write it down.

2.2. Connecting to an EC2 Instance

You can connect to a Linux EC2 instance using `ssh` as follows:

1. Connect to your instance using the its public DNS name. For example, if the instance's IP address is `75.101.230.211`, use the following command.

   ```
   ssh -i ~/.ec2/login.pem ec2-user@75.101.230.211
   ```

   Note that you need to log in as `ec2-user`, and **not** as `cis505`, `root`, or your SEAS login! You should see a response like the following:

   ```
   The authenticity of host '75.101.230.211' can't be established.
   RSA key fingerprint is fc:8d:0c:eb:0e:a6:4a:6a:61:50:00:c4:d2:51:78:66.
   Are you sure you want to continue connecting (yes/no)?
   ```

2. Enter `yes`. You'll see a response like the following:

   ```
   Warning: Permanently added '75.101.230.211' (RSA) to the list of known hosts.
   ```

You're now logged in and can work with the instance like you would any normal server. If you need root access, e.g., for mounting or unmounting volumes, you can use `sudo`. Just remember that you are being billed while the server is alive! Log out using `exit` or `logout`.

2.3 Compiling and running your Application.

To run your code in the instance, first log out from it using exit, and then go to your local directory where you have your code. Copy your code to the instance by issuing the following command:

```
scp -r -i ~/.ec2/login.pem ./* ec2-user@<PublicDNS>:~
```

After that, login to the instance again and compile your code. Before you run your server, make sure that your frontend servers listen on port 80 (because that's the one we allowed incoming requests from in the default security group). Then, simply run your application, go to your web browser, and visit the URL http://<PublicIP>/. You should be able to see your frontend web site. If your system doesn't work as expected (e.g., certain nodes cannot connect to certain other nodes), please double-check that the correct port numbers (TCP and/or UDP) have been opened in the security group you are using.

There are some times when you will want your application to keep on running, even if you logout from the instance. To accomplish this use the nohup utility:

```
nohup ./your-binary-name-here &
exit
```

The exit command above will log you out, but if you visit your site using your web browser, you will see that your server is still running. If you need to terminate it later, you can log back in and then use the `killall` command:

```
killall your-binary-name-here
exit
```

2.4 Terminating an EC2 Instance

Please note that you will be billed for AWS instances as they are alive, so you will want to terminate them when they aren't in direct use. (If you signed up for AWS Educate, you'll have a certain number of 'free' credits, but if you run over, your personal credit card will still be charged!) Here are the Amazon instructions:

1. In the AWS Management Console (http://aws.amazon.com/console/), locate the instance in your list of instances on the **Instances** page.
2. Right-click the instance, and then click **Terminate**.
3. Click **Yes, Terminate** when prompted for confirmation.

Amazon EC2 begins terminating the instance. As soon as the instance status changes to shutting down or terminated, you stop incurring charges for that instance. AWS Dashboard has an overview of the resources you're currently using, so please be sure to check that every time you stop using AWS for some nontrivial amount of time. If you do run over your free credits and your personal credit card is charged, we cannot (!) reimburse you.