146 days ago in UBUNTU AWS CASSANDRA

Setting up a Cassandra 2.1 cluster on AWS Ubuntu 14.04

Installation

Launch some new Ubuntu 14.04 instance on AWS and ssh in to each and do the following:

Install java

```
sudo add-apt-repository ppa:webupd8team/java
sudo apt-get update
sudo apt-get install oracle-java8-installer
```

Verify java is installed with:

```
java -version
```

And then let java setup environment defaults:

```
sudo apt-get install oracle-java8-set-default
```

Install Cassandra

Edit the sources file at **/etc/apt/sources.list** and add:

```
deb http://www.apache.org/dist/cassandra/debian 21x main
deb-src http://www.apache.org/dist/cassandra/debian 21x main
```

It should look like:

Now we need to add GPG keys so lets run:

```
gpg --keyserver pgp.mit.edu --recv-keys F758CE318D77295D
gpg --export --armor F758CE318D77295D | sudo apt-key add -
gpg --keyserver pgp.mit.edu --recv-keys 2B5C1B00
gpg --export --armor 2B5C1B00 | sudo apt-key add -
gpg --keyserver pgp.mit.edu --recv-keys 0353B12C
gpg --export --armor 0353B12C | sudo apt-key add -
```

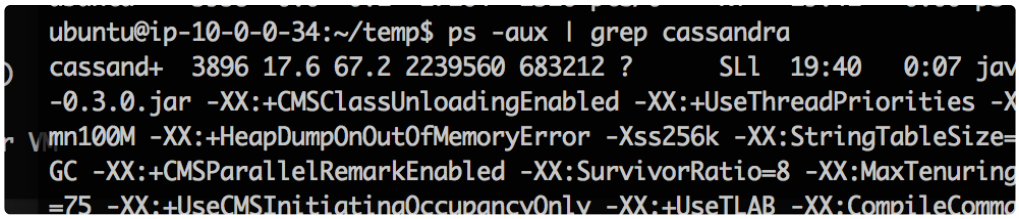
And let's install:

```
sudo apt-get update
sudo apt-get install cassandra
```

After installation Cassandra will be running and we need to kill the process in order to configure the cluster, so run:

```
ps -aux | grep cassandra
```

and you should see something like:



Notice the PID is 3896 in my example. With whatever your PID is, run:

```
sudo kill 3896
```

And let's delete the data:

```
sudo rm -rf /var/lib/cassandra/*
```

Configure the cluster

Since we're launching a three node cluster we really only need to have one seed node that the other two nodes rely on. For now, ssh into whichever server you want the seed server to be and let's edit **/etc/cassandra/cassandra.yaml**:

```
sudo nano /etc/cassandra/cassandra.yaml
```

Find the follow configuration settings:

```
cluster_name: "<name of your cluster>"
...
seed_provider:
    - class_name: org.apache.cassandra.locator.SimpleSeedProvider
      parameters:
          - seeds: "<private ip of seed server>"
...
```

```
listen_address: <private ip of current server>
...
rpc_address: <private ip of current server>
```

When completed, let's run Cassandra:

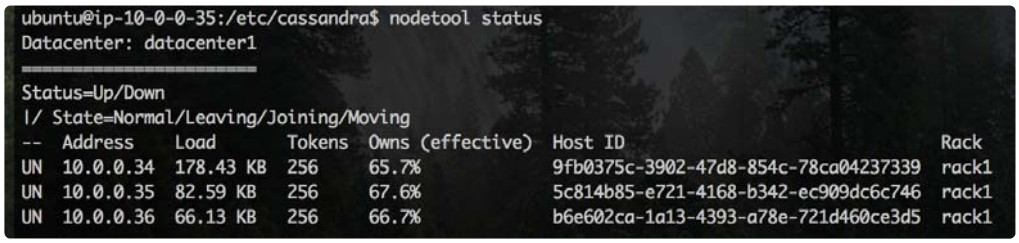
```
sudo cassandra &
```

The & is so we run Cassandra in the background.

Now, repeat these steps with the remaining nodes. When you are finished and you've launched Cassandra on all of the nodes (without errors!) you can run:

```
nodetool status
```

This outputs the status of the cluster, and should look like this:



Conclusion

That's basically it for setting up a small cluster of Cassandra nodes. You can check out the discussion on Hacker News or shoot me a tweet!

We were unable to load Disqus. If you are a moderator please see our troubleshooting guide.