Part I:

(1) install ubuntu 16.04

Use ‘do-release-upgrade’ to upgrade ubuntu.

Chef server url: [https://packages.chef.io/files/stable/chef- server/12.15.8/ubuntu/16.04/chef-server-core\_12.15.8-1\_amd64.deb](https://packages.chef.io/files/stable/chef-%20server/12.15.8/ubuntu/16.04/chef-server-core_12.15.8-1_amd64.deb)

Change Vagrantfile:

Vagrant.configure("2") do |config|

config.vm.box = "ubuntu/xenial64"

end

(2) install java8

manually download java8 from browser

sudo mkdir /usr/lib/jvm

cd /usr/lib/jvm

sudo tar -xvzf ~/Downloads/jdk-8u212-linux-x64.tar.gz

follow this link for the following steps:

<https://www.vultr.com/docs/how-to-manually-install-java-8-on-ubuntu-16-04>

Java 8 installation completed.

(3) install Cassandra

Download Cassandra from browser

echo "deb http://www.apache.org/dist/cassandra/debian 311x main" | sudo tee -a /etc/apt/sources.list.d/cassandra.sources.list

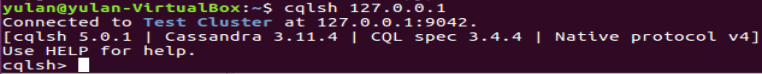
curl https://www.apache.org/dist/cassandra/KEYS | sudo apt-key add -

sudo apt-get update

sudo apt-get install cassandra

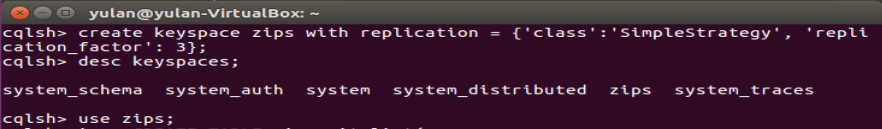
sudo service cassandra start

nodetool status

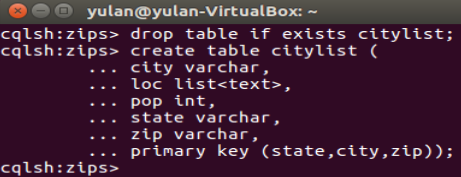


Part II:

1.



2.





3.

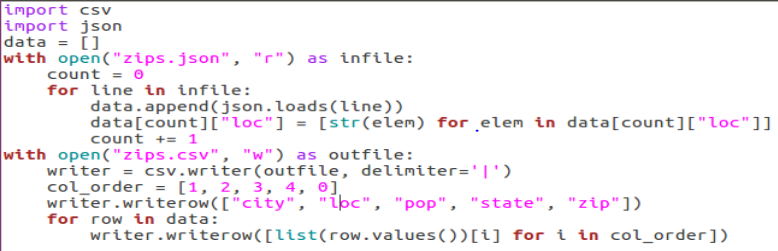
Install python3.6 to use pandas

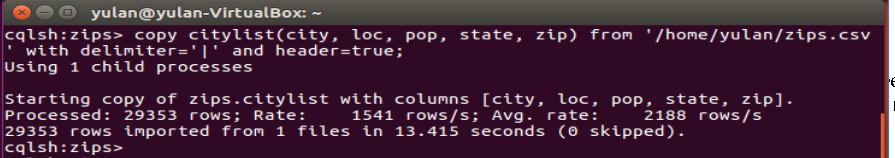
sudo add-apt-repository ppa:jonathonf/python-3.6

sudo apt-get update

sudo apt-get install python3.6

create zips.py file like this under /home/yulan.





This python file helps to convert json file to csv file. The json file given is a list of json objects. I need to first make an array of them. I store the array as list ‘data’ in the python file. After Json.loads, a json object is stored as a dictionary in the list ‘data’. For the values of the key ‘loc’: I need to change it into a list of strings. If I don’t change it, after conversion, I will get a string whose contect is a list. I used python to do that. Col\_order specifies the col order because I want to change the col order of the given json file into the order of the cols in table. Using values in col\_order as index, I can import values into cassandra table in the order of columns: “city”, “loc”, “pop”, “state”, “zip”. Then, data is imported successfully into Cassandra table.

4.

