Lab 01 Specification – Setting up a Compute Cloud using EC2
Due Monday, 29 Jan 2017
50 points

Lab Goals

- How to create Amazon EC2 instances in the Cloud?
- What needs to be done in order to setup the communication between those nodes in the Cluster?

Assignment Details

Firstly creating individual machines in the Cloud by using the GUI provided by Amazon EC2. In this lab, one would find out there is a need for configuring security firewall connectivity in order to open necessary ports on the machine. Secondly, how to setup the communication between those machines and understand what are some basics behind creating any cluster architecture.

Provisioner

Provision 4 nodes (Ubuntu 14, with default volume configurations) manually using the Amazon EC2 interface. There are two options for becoming an user in Amazon EC2 environment. You could go in the direct route by sign up for new account with a credit card (with free trial period). You could also sign up as an AWS Educate user and then become a registered user.

After registration, you can use the Amazon EC2 console and then set up the IAM rule similar to what we discussed in class to launch instances.

Nodes provisioned form a small cluster with the following cluster configuration:
Set 1 node as the Master node
Set 3 nodes as the Slave nodes

Accessing Nodes

EC2 Nodes can be accessed in two ways:

1. How to ssh with a pem file?
   - sudo chmod 600 ”PATH TO PEM FILE”
   - ssh -i ”PATH TO PEM FILE” ubuntu@PUBLIC-DNS-NAME PUBLIC-DNS-NAME can be copied from your EC2 console
   - or else ssh -i ”PATH TO PEM FILE” ubuntu@PUBLIC-IP PUBLIC-IP can be copied from your EC2 console
2. How to ssh without a pem file?
   - sudo su -

- passwd ubuntu – set the password for ubuntu user
- nano /etc/ssh/sshd_config – change PasswordAuthentication no to yes
- sudo service ssh restart

3. Okay now you can login to the EC2 node from any machine without using pem file

Software Installation

So, let us install Java on all the nodes.

1. sudo apt-get update
2. sudo apt-get install default-jre
3. sudo apt-get install default-jdk
4. Test java –version to see if java is installed correctly.
5. Configure JAVA_HOME, so that other applications can know where Java is installed The command

   ```
   sudo update-alternatives --config java
   ```

   will tell where Java is installed.
   –The path will be for example:
   /usr/lib/jvm/java-7-openjdk-amd64/
6. Add the JAVA_HOME to the file below:
   sudo nano /etc/environment
   JAVA_HOME=”YOURPATH”
   For example: JAVA_HOME=”/usr/lib/jvm/java-7-openjdk-amd64/”
7. source /etc/environment
8. Okay now, if you enter the command

   ```
   echo $JAVA\_HOME
   ```

   it will show where Java is installed.

Machine Report Creation

Write a program in Java that generates the machine report IP address, Disk space, and Memory of the respective nodes.
One requirement of the program is that the resulting text file Node-IPAddress.txt should be saved in /home/ubuntu/cluster-info/ directory.

A Sample report is shown below:

IP Address: 172.31.19.18
Disk Space: Total: 52428795904 bytes; Free: 33677811712 bytes
Memory: Total: 128516096 bytes; Free: 127173744 bytes
The java program is then run on all 4 machines to generate Node-IPAddress.txt on all nodes.

Each of the slave node will send their machine reports to the master node. So that the entire cluster information is organized and kept in Master node inside the cluster-info directory. Hint: How can we move the data in linux terminal from one machine to another?

Heart Beat Checker

Write a bash script in the master machine that will check if a slave machine is active or not? Note: You can use ping command to check if a machine is running. **Important:** Make sure the All Traffic configuration is set to enabled in your instance IAM role for source from anywhere.

A sample code is below:

```
#!/bin/bash
RESULT="64"
PING=$(ping $1 -c 1 | grep 64 | awk '{print $1}')
if [ "$RESULT" != "$PING" ]
then
   echo "ping failed"
else
   echo "ping successful, do ssh here"
fi
```

Submission Details

There is no submission needed for this lab. On or before the due date, you are responsible to demonstrate the correct working of your code by doing a detailed demo to the Instructor. The points for the lab will be based on the evaluations done on your demo and then evaluating the correct working of your code. In addition, the feedback on your lab will be given by the instructor after your demo.