Spark:

1)Setting up single node cluster for Spark:

Configure hadoop single node in i3.large instance and Configure spark

Install scala

* apt-get install scala

Download spark

Wget <https://archive.apache.org/dist/spark/spark-1.6.1/spark-1.6.1-bin-hadoop2.6.tgz>

Untar it.

tar -zxvf spark-1.6.1-bin-hadoop2.6.tgz.

mv spark-1.6.1-bin-hadoop2.6 /spark

Configure the spark path in the .bashrc

Export PATH=$PATH: /root/spark/bin.

You can get into the shell using command

Spark-shell

ii)Setting 8 node cluster with.’ spark

Export AWS Access key and the Secret key

Launch the cluster using the following command

./spark-ec2 -k <key> -I <Pem file> -s <no of instances> --spot price <spot price> launch <Cluster name>

Execution

Spark-shell -I <scala file name>”.