Step by Step Guide to setup Elephas Cluster

1. Launcu Ec2 Instance amitsparkw1 t2.xlarge security group “launch-wizard-10”
2. Wait till its state is running
3. Connect using EC2 Instance Connect
4. commandLine Interface is up then

sudo apt update

sudo apt install software-properties-common

sudo apt install python3.8

python3.8 --version

sudo apt install pip

pip install matplotlib

pip install virtualenv

sudo apt install openjdk-8-jdk

export JAVA\_HOME=/usr/lib/jvm/java-8-openjdk-amd64

wget https://downloads.apache.org/spark/spark-3.1.2/spark-3.1.2-bin-hadoop3.2.tgz

sudo mkdir /opt/spark

sudo tar -xf spark\*.tgz -C /opt/spark --strip-component 1

sudo chmod -R 777 /opt/spark

echo "export SPARK\_HOME=/opt/spark" >> ~/.bashrc

echo "export PATH=$PATH:$SPARK\_HOME/bin:$SPARK\_HOME/sbin" >> ~/.bashrc

echo "export PYSPARK\_PYTHON=/usr/bin/python3" >> ~/.bashrc

source ~/.bashrc

sudo apt install python3-pip python3-dev

sudo -H pip3 install --upgrade pip

sudo -H pip3 install virtualenv

mkdir ~/my\_ML\_Projects

cd ~/my\_ML\_Projects

virtualenv ml\_project\_env

source ml\_project\_env/bin/activate

pip install jupyter findspark numpy matplotlib tensorflow keras pyspark pandas sklearn

pip install elephas

jupyter notebook --ip=\*

#after opening jupyter notebook open jupyter terminal

git clone <https://github.com/siyad-CT/HPE_ML_WS>

1. Launch another EC2 Instance amitsparkw2 t2.xlarge securitygroup “launch-wizard-10”
2. Repeat all above steps 2-4 for this ec2 instance
3. Launch another EC2 Instance amitsparkw3 t2.xlarge securitygroup “launch-wizard-10”
4. Repeat all above steps 2-4 for this ec2 instance

Master – amitspark

Worker1 – amitsparkw1

Worker2 – amitsparkw2

Worker3 – amitsparkw3

on master:

sudo apt install openssh-server openssh-client

cd ~/.ssh

ssh-keygen -t rsa -P ""

ls

sudo service ssh restart

cat id\_rsa.pub

ssh-rsa  ubuntu@ip-172-31-18-164

for each worker:

cd ~/.ssh

ls

nano authorisedkey

paste key

on master: (to taste connection)

ssh -i ~/.ssh/id\_rsa ubuntu@54.90.160.202

exit

ssh -i ~/.ssh/id\_rsa ubuntu@54.226.246.237

exit

ssh -i ~/.ssh/id\_rsa ubuntu@204.236.240.54

exit

on master:

cd /opt/spark/conf/

cp spark-env.sh.template spark-env.sh

nano spark-env.sh

export SPARK\_MASTER\_HOST=172.31.15.218

export JAVA\_HOME=/usr/lib/jvm/java-8-openjdk-amd64

export PYSPARK\_PYTHON=python3

cp workers.template slaves

nano slaves

172.31.28.181

172.31.20.186

172.31.29.170

cp spark-defaults.conf.template spark-defaults.conf

nano spark-defaults.conf

spark.driver.memory 5g

spark.driver.maxResultSize 2g

spark.executor.memory 5g

sh /opt/spark/sbin/start-all.sh

cd ~/my\_ML\_Projects/

source ml\_project\_env/bin/activate

jupyter notebook --ip=\*

git clone https://github.com/siyad-CT/HPE\_ML\_WS