[root@ip-172-31-49-44 /]# cat \*.sh

#! /bin/bash

cd $HOME

umount /data

echo "using drive " $1

echo "WARNING!! This will format the drive at" $1

read -rsp $'Press any key to continue or control-C to quit...\n' -n1 key

#make a new ext4 filesystem

mkfs.ext4 $1

#mount the new filesystem under /data

mount -t ext4 $1 /data

chmod a+rwx /data

#format the hadoop namenode

sudo -u hdfs hdfs namenode -format

#start hdfs

for x in `cd /etc/init.d ; ls hadoop-hdfs-\*` ; do sudo service $x restart ; done

#make the hadoop directories

/usr/lib/hadoop/libexec/init-hdfs.sh

sudo -u hdfs hdfs dfs -mkdir /user/w205

sudo -u hdfs hdfs dfs -chown w205 /user/w205

#start YARN services

service hadoop-yarn-resourcemanager restart

service hadoop-yarn-nodemanager restart

service hadoop-mapreduce-historyserver restart

#set up directories for postgres

mkdir /data/pgsql

mkdir /data/pgsql/data

mkdir /data/pgsql/logs

chown -R postgres /data/pgsql

sudo -u postgres initdb -D /data/pgsql/data

#setup pg\_hba.conf

sudo -u postgres echo "host all all 0.0.0.0 0.0.0.0 md5" >> /data/pgsql/data/pg\_hba.conf

#setup postgresql.conf

sudo -u postgres echo "listen\_addresses = '\*'" >> /data/pgsql/data/postgresql.conf

sudo -u postgres echo "standard\_conforming\_strings = off" >> /data/pgsql/data/postgresql.conf

#make start postgres file

cd /data

cat > /data/start\_postgres.sh <<EOF

#! /bin/bash

sudo -u postgres pg\_ctl -D /data/pgsql/data -l /data/pgsql/logs/pgsql.log start

EOF

chmod +x /data/start\_postgres.sh

#make a stop postgres file

cat > /data/stop\_postgres.sh <<EOF

#! /bin/bash

sudo -u postgres pg\_ctl -D /data/pgsql/data -l /data/pgsql/logs/pgsql.log stop

EOF

chmod +x /data/stop\_postgres.sh

#start postgres

/data/start\_postgres.sh

sleep 5

#write setup script for hive metastore

cat > /data/setup\_hive\_for\_postgres.sql <<EOF

CREATE USER hiveuser WITH PASSWORD 'hive';

CREATE DATABASE metastore;

\c metastore

\i /usr/lib/hive/scripts/metastore/upgrade/postgres/hive-schema-1.1.0.postgres.sql

\i /usr/lib/hive/scripts/metastore/upgrade/postgres/hive-txn-schema-0.13.0.postgres.sql

\c metastore

\pset tuples\_only on

\o /tmp/grant-privs

SELECT 'GRANT SELECT,INSERT,UPDATE,DELETE ON "' || schemaname || '". "' ||tablename ||'" TO hiveuser ;'

FROM pg\_tables

WHERE tableowner = CURRENT\_USER and schemaname = 'public';

\o

\pset tuples\_only off

\i /tmp/grant-privs

\q

EOF

#run the metastore creation sql

sudo -u postgres psql -f /data/setup\_hive\_for\_postgres.sql

#make the new hive configuration directory

sudo -u hadoop mkdir -p /data/hadoop/hive/conf

#setup the hive-site file

cat > /data/hadoop/hive/conf/hive-site.xml <<EOF

<?xml version="1.0"?>

<!--

Licensed to the Apache Software Foundation (ASF) under one or more

contributor license agreements. See the NOTICE file distributed with

this work for additional information regarding copyright ownership.

The ASF licenses this file to You under the Apache License, Version 2.0

(the "License"); you may not use this file except in compliance with

the License. You may obtain a copy of the License at

http://www.apache.org/licenses/LICENSE-2.0

Unless required by applicable law or agreed to in writing, software

distributed under the License is distributed on an "AS IS" BASIS,

WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.

See the License for the specific language governing permissions and

limitations under the License.

-->

<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>

<configuration>

<!-- Hive Configuration can either be stored in this file or in the hadoop configuration files -->

<!-- that are implied by Hadoop setup variables. -->

<!-- Aside from Hadoop setup variables - this file is provided as a convenience so that Hive -->

<!-- users do not have to edit hadoop configuration files (that may be managed as a centralized -->

<!-- resource). -->

<!-- Hive Execution Parameters -->

<property>

<name>javax.jdo.option.ConnectionURL</name>

<value>jdbc:postgresql://localhost:5432/metastore</value>

</property>

<property>

<name>javax.jdo.option.ConnectionDriverName</name>

<value>org.postgresql.Driver</value>

</property>

<property>

<name>javax.jdo.option.ConnectionUserName</name>

<value>hiveuser</value>

</property>

<property>

<name>javax.jdo.option.ConnectionPassword</name>

<value>hive</value>

</property>

<property>

<name>datanucleus.autoCreateSchema</name>

<value>false</value>

</property>

<!-- <property>

<name>hive.metastore.uris</name>

<value>thrift://localhost:9083</value>

<description>IP address (or fully-qualified domain name) and port of the metastore host</description>

</property>

-->

<property>

<name>hive.metastore.schema.verification</name>

<value>true</value>

</property>

</configuration>

EOF

#setup zeppelin

cat > setup\_zeppelin.sh <<EOF

mkdir /data/w205

chown w205 /data/w205

sudo -u w205 wget -O /data/apache-maven-3.3.3-bin.tar.gz http://www.trieuvan.com/apache/maven/maven-3/3.3.3/binaries/apache-maven-3.3.3-bin.tar.gz

cd /data/ && sudo -u w205 tar xvzf /data/apache-maven-3.3.3-bin.tar.gz

sudo -u w205 git clone https://github.com/apache/incubator-zeppelin.git /data/zeppelin

cd /data/zeppelin

/data/apache-maven-3.3.3/bin/mvn -Pspark-1.5 -Dhadoop.version=2.6.0 -DskipTests -Phadoop-2.6 clean package

cp conf/zeppelin-env.sh.template conf/zeppelin-env.sh

cp /etc/hadoop/conf/\*.xml conf/

cp /data/hadoop/hive/conf/hive-site.xml conf/

echo 'export ZEPPELIN\_MEM="-Xmx2048m"' >> conf/zeppelin-env.sh

echo 'export SPARK\_HOME=/home/w205/spark15' >> conf/zeppelin-env.sh

EOF

chmod +x setup\_zeppelin.sh

[root@ip-172-31-49-44 /]#