Assignment 10.1 ( Oozie and Sqoop)

Step1: Create job.properties

Create job.properties with content below:

nameNode=hdfs://localhost:8020

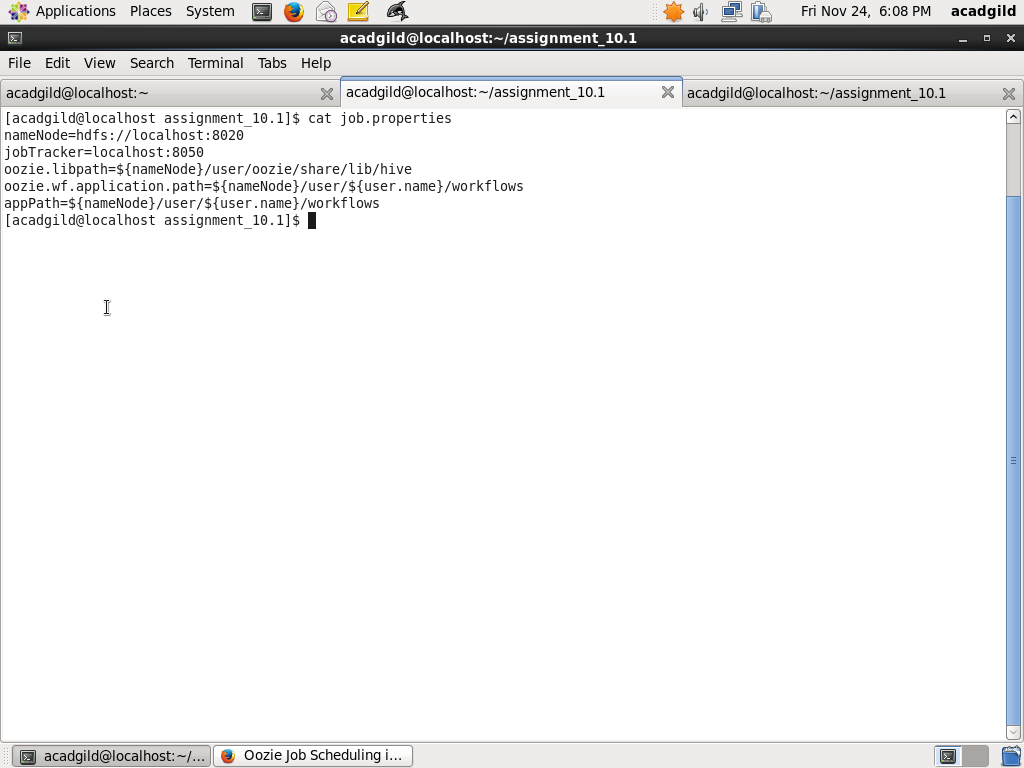
jobTracker=localhost:8050

oozie.libpath=${nameNode}/user/oozie/share/lib/hive

oozie.wf.application.path=${nameNode}/user/${user.name}/workflows

appPath=${nameNode}/user/${user.name}/workflows

Screenshot is as below:



Step2: Create workflow.xml

Create workflow.xml. It contains all the details of files, scripts,  required to schedule and run Oozie job

<workflow-app name="DemoOozie" xmlns="uri:oozie:workflow:0.1">

<start to="demo-hive"/>

<action name="demo-hive">

<hive xmlns="uri:oozie:hive-action:0.2">

<job-tracker>${jobTracker}</job-tracker>

<name-node>${nameNode}</name-node>

<job-xml>${appPath}/hive-site.xml</job-xml>

<configuration>

<property>

<name>oozie.hive.defaults</name>

<value>${appPath}/hive-site.xml</value>

</property>

<property>

<name>hadoop.proxyuser.acadgild.hosts</name>

<value>\*</value>

</property>

<property>

<name>hadoop.proxyuser.acadgild.groups</name>

<value>\*</value>

</property>

</configuration>

<script>create\_table.hql</script>

</hive>

<ok to="end"/>

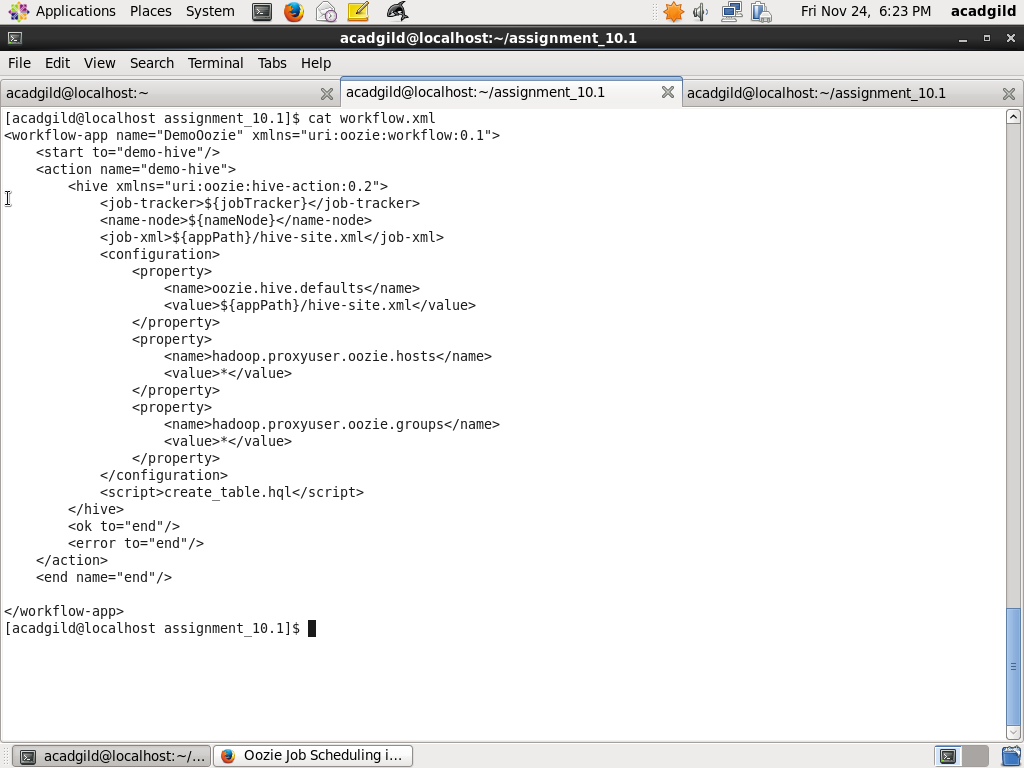
<error to="end"/>

</action>

<end name="end"/>

</workflow-app>

Screenshot is as below:



Step3 : Write script create\_table.hql

Create hive script file which will be used by Oozie workflow as a action

use default;

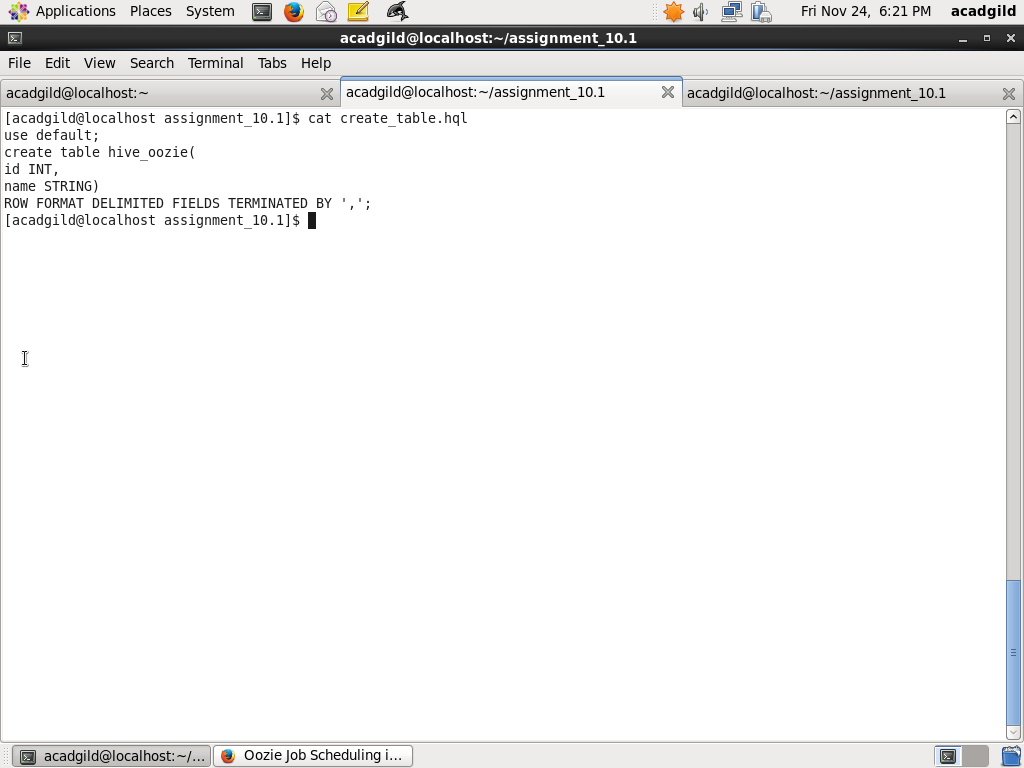
create table hive\_oozie(

id INT,

name STRING)

ROW FORMAT DELIMITED FIELDS TERMINATED BY ',';

The screenshot with query and output is as below:



Step4: Copy files to HDFS

In HDFS create a directory /user/oozie/workflows and copy the files workflow.xml, create\_table.hql, hive-site.xml to the directory using commands below:

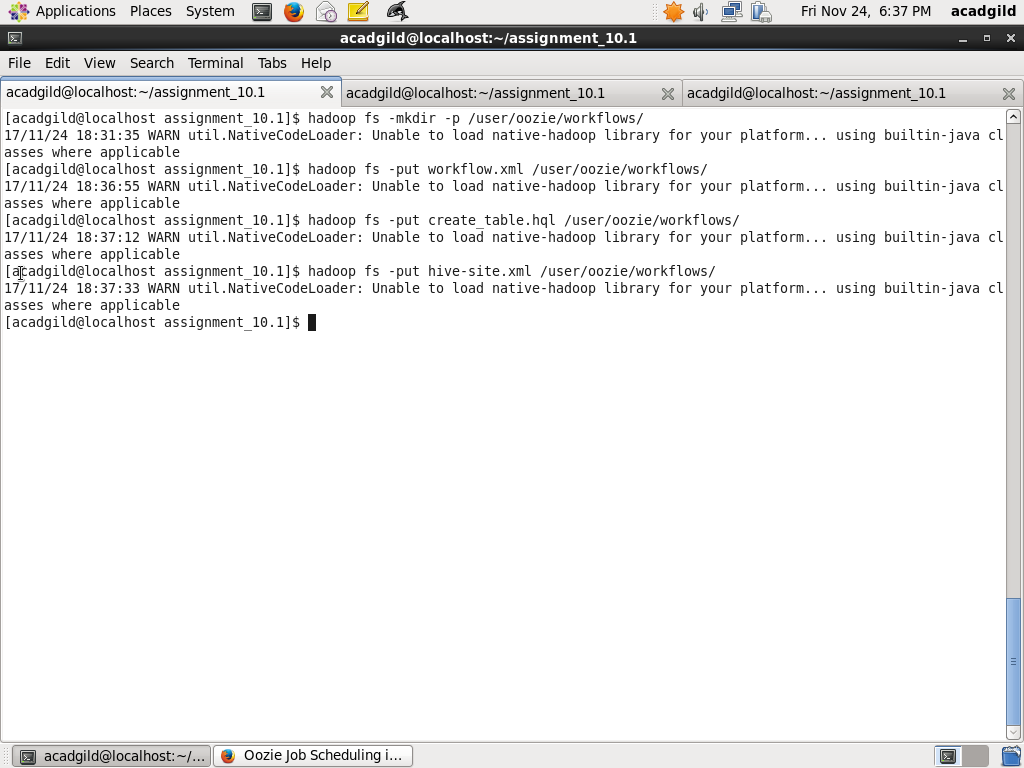
hadoop fs –mkdir –p /user/oozier/workflows

hadoop fs –put workflow.xml /user/oozier/workflows

hadoop fs –put create\_table.hql /user/oozier/workflows

hadoop fs –put hive-site.xml /user/oozier/workflows

Screenshot is as below



Step5: Start oozie

Start oozie using command below:

sudo /usr/localoozie-4.1.0/bin/oozie job -oozie http://127.0.0.1:11000/oozie -config job.properties –run

However, I am facing Connection error on acadgild VM. Raising a support ticket for this issue

Screenshot is as below:

