**Hands on Activity**

**Scenario: We will create sqoop job. We will attach sqoop job and sqoop metastore together in oozie workflow.**

**Dataset: We will use movielense dataset**

1. Run your sqoop meta store .To do so rum below command on terminal

sqoop-metastore

1. Create sqoop job to do so open new terminal or Putty then run following script on terminal/ putty

sqoop job

--meta-connect jdbc:hsqldb:hsql://localhost:16000/sqoop

--create movieJob

-- import

--connect jdbc:mysql://localhost/movielens

--driver com.mysql.jdbc.Driver

--query

'SELECT \* FROM movies WHERE $CONDITIONS'

--fields-terminated-by '\t'

--incremental append

--check-column id

--last-value '1000'

--target-dir /user/maria\_dev/metadata\_demo/new

-m1

1. Below command print out all the list of the job created under metastore

sqoop job --list -meta-connect jdbc:hsqldb:hsql://localhost:16000/sqoop

1. Create workflow.xml file and put following code in the workflow.xml file as mention below (You can give any name to .xml file)

<?xml version="1.0" encoding="UTF-8"?>

<workflow-app xmlns="uri:oozie:workflow:0.2" name="movie-wf">

<start to="sqoopmovies\_node"/>

<action name="sqoopmovies\_node">

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

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

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

<configuration>

<property>

<name>mapred.job.queue.name</name>

<value>${queueName}</value>

</property>

</configuration>

<command>job --meta-connect jdbc:hsqldb:hsql://localhost:16000/sqoop --exec movieJob</command>

<archive>/user/maria\_dev/lib/java-json.jar#java-json.jar</archive>

</sqoop>

<ok to="end"/>

<error to="fail"/>

</action>

<kill name="fail">

<message> Workflow failed, error message[${wf:errorMessage(wf:lastErrorNode())}]</message>

</kill>

<end name="end"/>

</workflow-app>

1. Create job.properties file and add the following code in job.properties file

nameNode=hdfs://sandbox.hortonworks.com:8020

jobTracker=http://sandbox.hortonworks.com:8050

queueName=default

oozie.use.system.libpath=true

oozie.wf.application.path=${nameNode}/user/maria\_dev/sqoop\_movie/wk.xml

1. Upload workflow.xml file under this path "user/maria\_dev/ sqoop\_movie "
2. Upload job.properties file inside VMware local directory e.g Documents/foldername/job.properties
3. Execute oozie job as mention below

oozie job -oozie http://localhost:11000/oozie -config .Documents/job.properties -run