**EX : 04 REG.NO:210701307 Create UDF in PIG**

**Step-by-step installation of Apache Pig on Hadoop cluster on Ubuntu**

**Pre-requisite:**

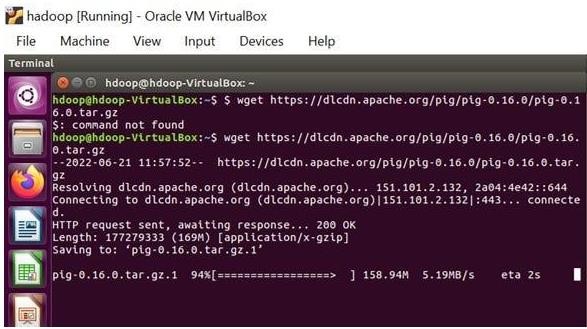
· Ubuntu 16.04 or higher version running (I have installed Ubuntu on Oracle VM (Virtual

Machine) VirtualBox),

· Run Hadoop on ubuntu (I have installed Hadoop 3.2.1 on Ubuntu 16.04). You may refer to my blog “How to install Hadoop installation” click here for Hadoop installation).

**Pig installation steps**

**Step 1:** Login into Ubuntu



**Step 2:** Go to https://pig.apache.org/releases.html and copy the path of the latest version of pig that you want to install. Run the following comment to download Apache Pig in Ubuntu:

$ wge[t https://dlcdn.apache.org/pig/pig-0.16.0/pig-0.16.0.tar.gz](https://dlcdn.apache.org/pig/pig-0.16.0/pig-0.16.0.tar.gz)



**Step 3:** To untar pig-0.16.0.tar.gz file run the following command:

$ tar xvzf pig-0.16.0.tar.gz

**Step 4:** To create a pig folder and move pig-0.16.0 to the pig folder, execute the following command:

$ sudo mv /home/hdoop/pig-0.16.0 /home/hdoop/pig

**Step 5:** Now open the .bashrc file to edit the path and variables/settings for pig. Run the following command:

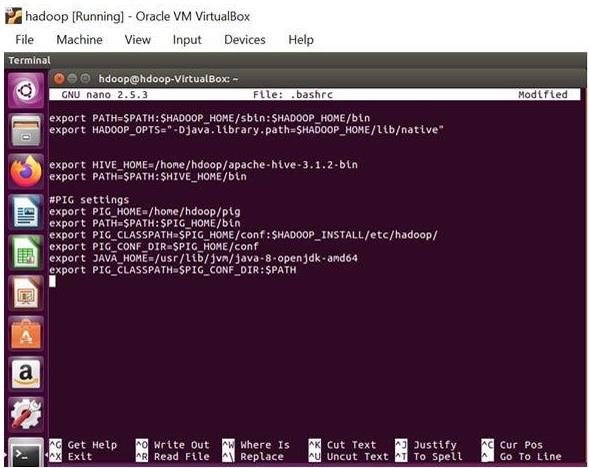
$ sudo nano .bashrc

Add the below given to .bashrc file at the end and save the file.

#PIG settingsexport PIG\_HOME=/home/hdoop/pigexport

PATH=$PATH:$PIG\_HOME/binexport

PIG\_CLASSPATH=$PIG\_HOME/conf:$HADOOP\_INSTALL/etc/hadoop/export PIG\_CONF\_DIR=$PIG\_HOME/confexport JAVA\_HOME=/usr/lib/jvm/java-8-openjdk amd64export PIG\_CLASSPATH=$PIG\_CONF\_DIR:$PATH#PIG setting ends



**Step 6:** Run the following command to make the changes effective in the .bashrc file:

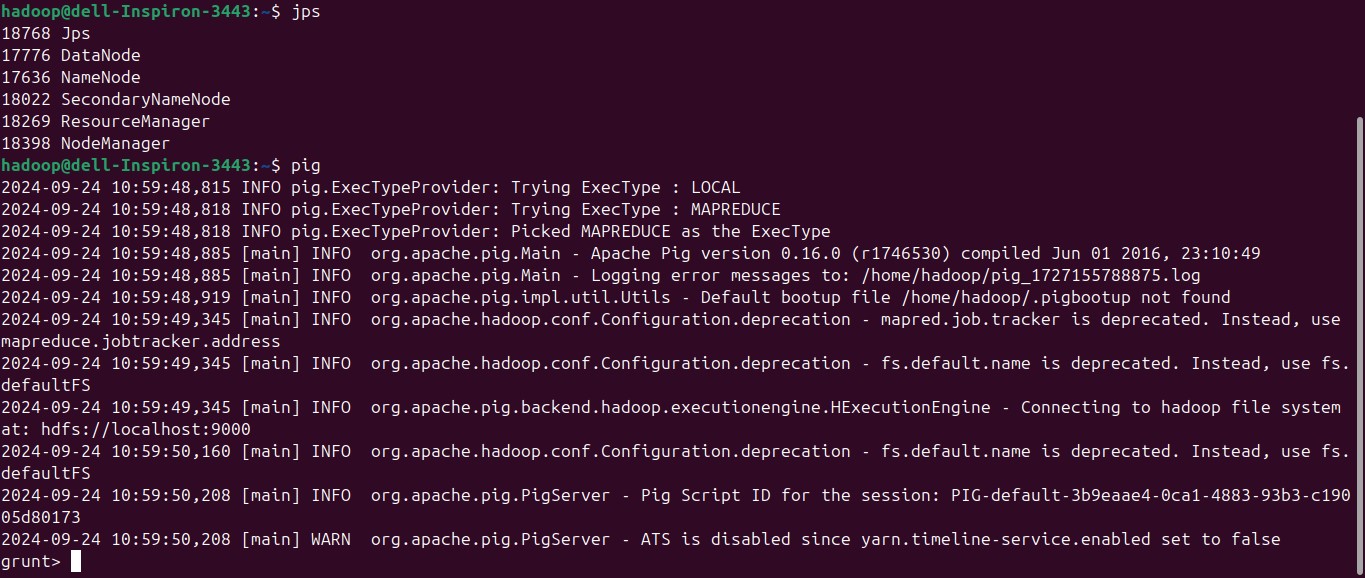
$ source .bashrc

**Step 7:** To start all Hadoop daemons, navigate to the hadoop-3.2.1/sbin folder and run the following commands:

$ ./start-dfs.sh$ ./start-yarn$ jps



**Step 8:** Now you can launch pig by executing the following command: $ pig



**Step 9:** Now you are in pig and can perform your desired tasks on pig. You can come out of the pig by the quit command:

> quit;