Go to the following link and download pig
http://mirrors.estointernet.in/apache/pig/pig-0.16.0/

To untar pig-0.16.0.tar.gz file run the following command:
$ tar xvzf pig-0.16.0.tar.gz

To create a pig folder and move pig-0.16.0 to the pig folder, execute the following command:
$ sudo mv /home/hadoop/pig-0.16.0 /usr/local/hadoop/pig

Now open the .bashrc file to edit the path and variables/settings for pig. Run the following command:
$ sudo nano .bashrc

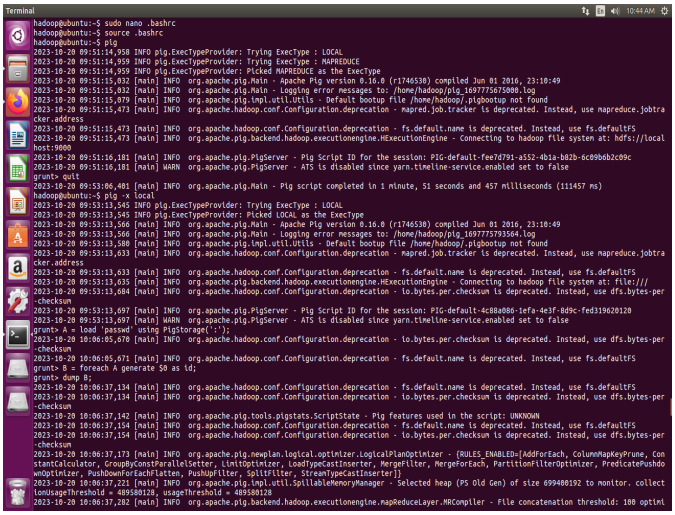
Add the below given to .bashrc file at the end and save the file.

#PIG settings
export PIG_HOME=/usr/local/hadoop/pig-0.16.0
export 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
(to exit ctrl +x)

Run the following command to make the changes effective in the .bashrc file:
$ source .bashrc

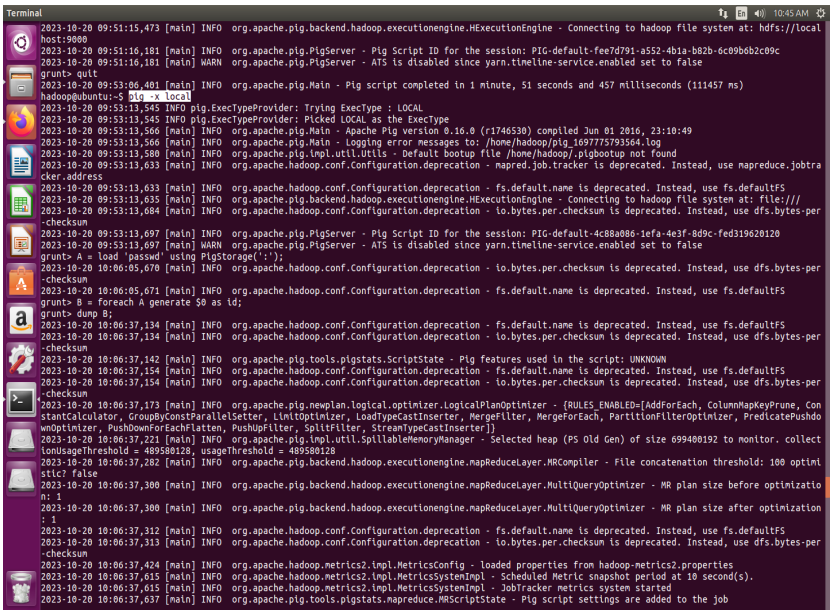
start all Hadoop daemons
cd /usr/local/hadoop/bin
start-all.sh
jps

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



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;

Quit and run pig in local mode
pig -x local

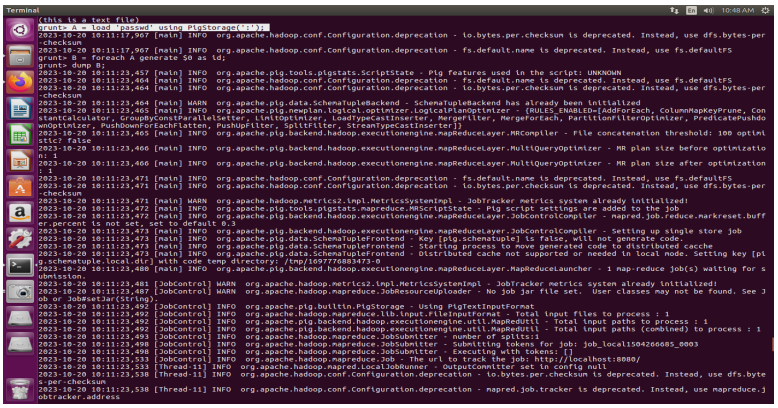


You can run Pig in interactive mode using the Grunt shell. Invoke the Grunt shell using the "pig" command and then enter your Pig Latin statements and Pig commands interactively at the command line.

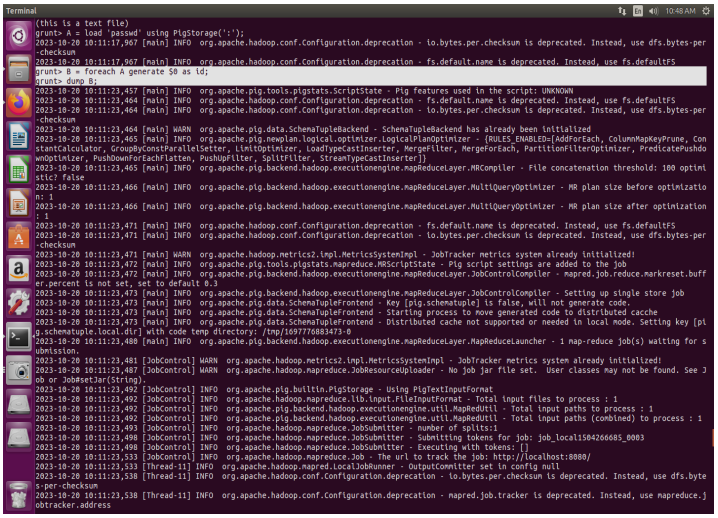
Example

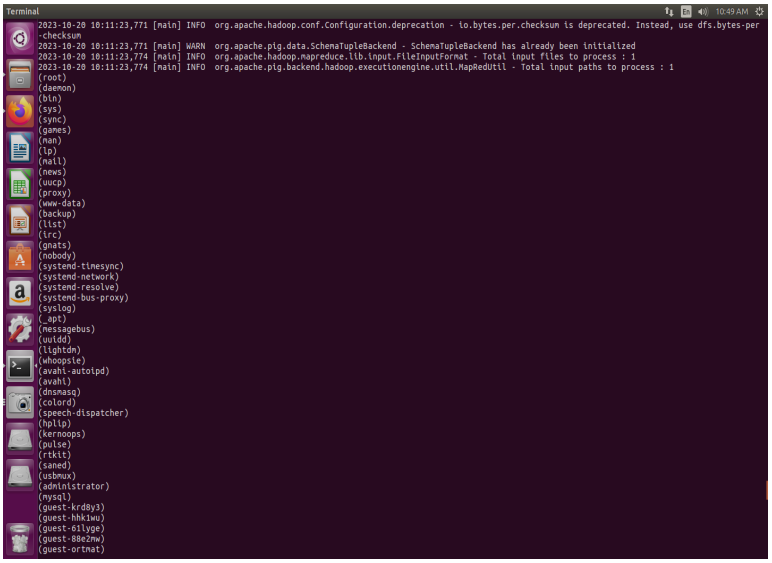
These Pig Latin statements extract all user IDs from the /etc/passwd file. First, copy the /etc/passwd file to your local working directory. Then, enter the Pig Latin statements interactively at the grunt prompt. The DUMP operator will display the results to your terminal screen.

grunt> A = load 'passwd' using PigStorage(':');



grunt> B = foreach A generate $0 as id;
grunt> dump B;





The STORE operator will write the results to a file (id.out).

/* id.pig */

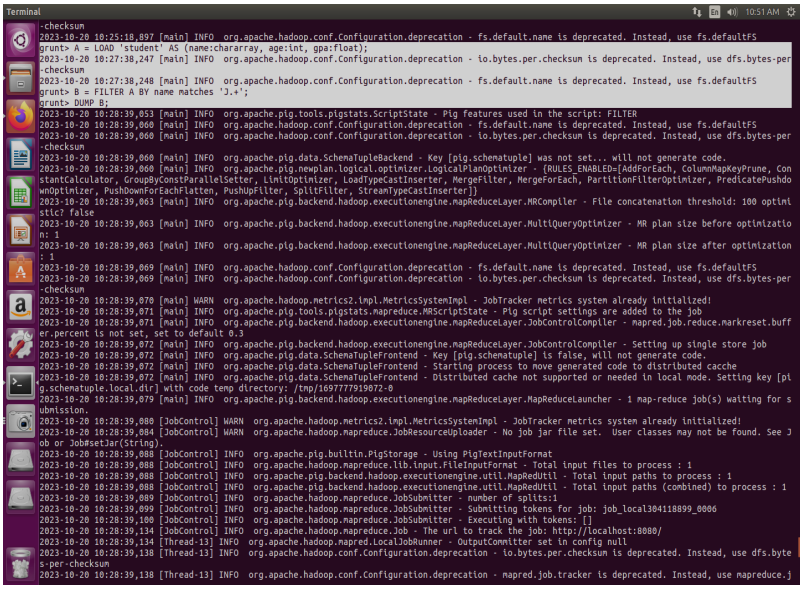
A = load 'passwd' using PigStorage(':'); -- load the passwd file
B = foreach A generate $0 as id; -- extract the user IDs
store B into '/home/hadoop/id.out'; -- write the results to a file name id.out

A = LOAD 'student' AS (name:chararray, age:int, gpa:float);

DUMP A;
(John,18,4.0F)
(Mary,19,3.7F)
(Bill,20,3.9F)
(Joe,22,3.8F)
(Jill,20,4.0F)

B = FILTER A BY name matches 'J.+';

DUMP B;
(John,18,4.0F)
(Joe,22,3.8F)
(Jill,20,4.0F)



A = LOAD 'student' AS (name:chararray, age:int, gpa:float);

B = GROUP A BY name;

C = FOREACH B GENERATE COUNT(A.age);

EXPLAIN C;

FILTER