

Assignment 10.2

Start services:

start-all.sh

mr-jobhistory-daemon.sh start historyserver

load dataset in Hadoop:

hadoop fs -put /home/acadgild/student.txt

```
drwxr-xr-x - acadgild supergroup      0 2017-09-14 19:16 hadoop
drwxr-xr-x - acadgild supergroup      0 2015-11-17 02:03 oozie-acad
drwxr-xr-x - acadgild supergroup      0 2015-11-17 02:00 share
-rw-r--r-- 1 acadgild supergroup    26204 2017-11-09 08:58 student.txt
[acadgild@localhost ~]$
```

Setting path to pig:

PIG_CLASSPATH=/home/acadgild/HADOOP/hbase-0.98.4-hadoop2/lib/hbase-server-0.98.4-hadoop2:/home/acadgild/HADOOP/hbase-0.98.4-hadoop2/lib/hbase-*.jar;

```
[acadgild@localhost ~]$ PIG_CLASSPATH=/home/hadoop/HADOOP/hbase-0.98.4-hadoop2/lib/hbase-server-0.98.4-hadoop2:/home/hadoop/HADOOP/hbase-0.98.4-hadoop2/lib/hbase-*.jar;
[acadgild@localhost ~]$
```

Hbase:

start-hbase.sh

Creating table in hbase:

create 'studentAcadgild','student_data'

```
hbase(main):001:0> create 'studentAcad','student data'
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/local/hbase/lib/slf4j-log4j12-1.6.4.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/local/hadoop-2.6.0/share/hadoop/common/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
2017-11-09 08:50:07,737 WARN [main] util.NativeCodeLoader: Unable to load native hbase library for your platform... using builtin-java classes where applicable
0 row(s) in 1.7450 seconds
```

Pig:

//loading Data

```
rawD = LOAD 'student.txt' USING PigStorage(',') AS  
(studentName:chararray,sector:chararray,DOB:chararray,qualification:chararray  
,score:int,state:chararray,randomName:chararray);
```

```
org.apache.pig.util.MapRedUtil - Total input paths to process : 1  
(StudentName,sector,DOB,qualification,,state,randomName)  
(ABROSER,goverenment,18-11-2002,MBBS,3,Pennsylvania,prattville*)  
(ALEXANDER,goverenment,20-10-2000,BSC,2,vermont,gadsden+)  
(ALEXANDER,private,20-10-2000,BE,8,arizona,decatur!)  
(ALEXANDER,goverenment,01-01-2003,BTECH,4,oregon,huntsville/)  
(AGNEW,goverenment,20-10-2000,BCOM,7,california,dothan@)  
(ATNEST,goverenment,20-10-2000,MTECH,8,arizona,decatur!)  
(BELL,goverenment,10-07-2004,BBA,9,alaska,auburn~)  
(BURR,goverenment,12-12-2001,BE,100,alabama,madison`)  
(BURD,goverenment,20-10-2000,ME,6,louisiana,hooover#)  
(BACHTEL,goverenment,28-04-2005,BE,100,alabama,madison`)  
(MULVETS,goverenment,20-10-2000,MS,8,arizona,decatur!)  
(BURTNER,goverenment,20-10-2000,BE,100,alabama,madison`)  
(BLOOM,goverenment,20-10-2000,BE,4,oregon,huntsville/)  
(BECHTEL,goverenment,20-10-2000,BE,5,Maryland,tuscaloosa$)  
(GAFF,goverenment,18-11-2002,BE,100,alabama,madison`)  
(CLARK,private,20-10-2000,BE,100,alabama,madison`)  
(COOTS,goverenment,20-10-2000,BSC,100,alabama,madison`)  
(DUNSDILL,goverenment,20-10-2000,BE,100,alabama,madison`)  
(DORONES,goverenment,20-10-2000,BTECH,100,alabama,madison`)  
(DICKEY,goverenment,20-10-2000,BBA,100,alabama,madison`)
```

//storing data to hbase table

```
STORE rawD INTO 'hbase://studentAcadgild' USING  
org.apache.pig.backend.hadoop.hbase.HBaseStorage('student_data:studentNam  
e,student_data:sector,student_data:DOB,student_data:qualification,student_data  
:score,student_data:state,student_data:randomName');
```

```

2017-11-09 10:23:42,099 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at /0.0.0.0:8032
2017-11-09 10:23:42,102 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job history server
2017-11-09 10:23:42,174 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at /0.0.0.0:8032
2017-11-09 10:23:42,179 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job history server
2017-11-09 10:23:42,239 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at /0.0.0.0:8032
2017-11-09 10:23:42,246 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job history server
2017-11-09 10:23:42,312 [main] WARN org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Unable to retrieve job to compute warning aggregation.
2017-11-09 10:23:42,312 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Success!
grunt> █

```

Hbase:

scan ('studentAcadgild')

```

SAMUEL      column=student_data:state, timestamp=1510203213260, value=prattville*
SAMUEL      column=student_data:studentName, timestamp=1510203213260, value=goverenment
SARAH       column=student_data:DOB, timestamp=1510203213261, value=BE
SARAH       column=student_data:qualification, timestamp=1510203213261, value=9
SARAH       column=student_data:score, timestamp=1510203213261, value=alaska
SARAH       column=student_data:sector, timestamp=1510203213261, value=20-10-2000
SARAH       column=student_data:state, timestamp=1510203213261, value=auburn~
SARAH       column=student_data:studentName, timestamp=1510203213261, value=goverenment
SCHNYDER    column=student_data:DOB, timestamp=1510203213194, value=MBBS
SCHNYDER    column=student_data:qualification, timestamp=1510203213194, value=100
SCHNYDER    column=student_data:score, timestamp=1510203213194, value=alabama
SCHNYDER    column=student_data:sector, timestamp=1510203213194, value=20-10-2000
SCHNYDER    column=student_data:state, timestamp=1510203213194, value=

```