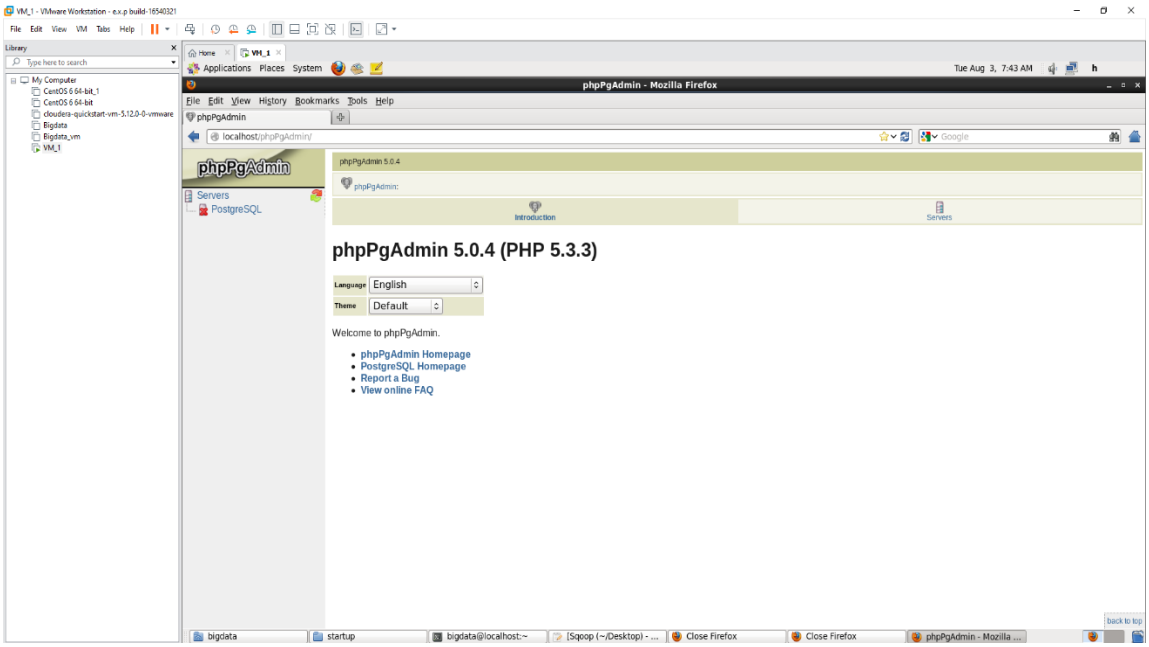
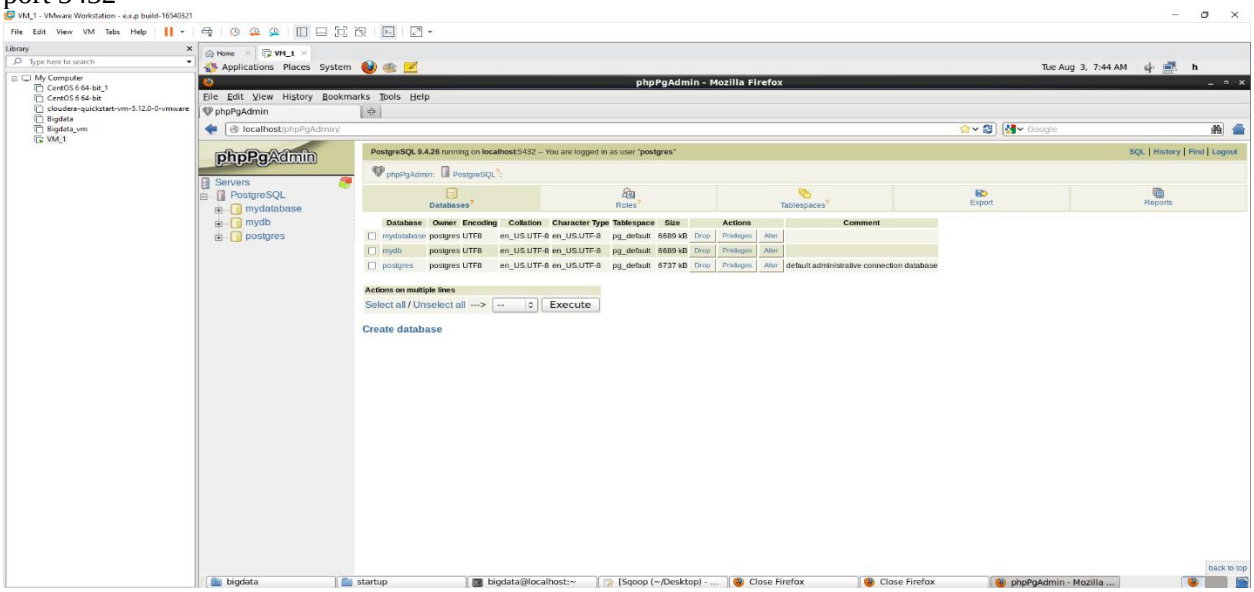
Sqoop_Lab

Start your Hadoop HDFS & YARN
Makesure you have sqoop installed
Inside the sqoop lib check for postgres JDBC

#postgres access
#GUI
http://localhost/phpPgAdmin/



Select postgres
username postgres
pass :post
username bob
pass :bob
port 5432



#CLI

username postgres
pass :post
(base) [bigdata@localhost ~]$ su root
Password: 123456
[root@localhost bigdata]# su postgres
bash-4.1$ psql
Password: post
psql (9.4.26)
Type "help" for help.
postgres=#

#listing databases in PostgreSQL

postgres=# \l+

List of data

bases

```
   Name    | Owner   | Encoding |   Collate   |    Ctype    |   Access privile
ges   | Size   | Tablespace |              Description
-----------+----------+----------+-------------+-------------+-----------------
------+---------+------------+-------------------------------------------
 mydatabase | postgres | UTF8     | en_US.UTF-8 | en_US.UTF-8 | =Tc/postgres
    +| 6689 kB | pg_default |
          |          |          |             |             | postgres=CTc/pos
tgres+|          |            |
          |          |          |             |             | mm=CTc/postgres
     |          |            |
 mydb       | postgres | UTF8     | en_US.UTF-8 | en_US.UTF-8 | =Tc/postgres
    +| 6689 kB | pg_default |
          |          |          |             |             | postgres=CTc/pos
tgres+|          |            |
          |          |          |             |             | bob=CTc/postgres
     |          |            |
 postgres   | postgres | UTF8     | en_US.UTF-8 | en_US.UTF-8 |
     | 6852 kB | pg_default | default administrative connection database
 template0  | postgres | UTF8     | en_US.UTF-8 | en_US.UTF-8 | =c/postgres
    +| 6681 kB | pg_default | unmodifiable empty database
          |          |          |             |             | postgres=CTc/pos
tgres |          |            |
 template1  | postgres | UTF8     | en_US.UTF-8 | en_US.UTF-8 | =c/postgres
    +| 6804 kB | pg_default | default template for new databases
          |          |          |             |             | postgres=CTc/pos
tgres |          |            |
(5 rows)
```

#listing tables in PostgreSQL

- \dt or \dt+

```
postgres=# \dt
        List of relations
 Schema |  Name   | Type  |  Owner
--------+---------+-------+----------
 public | company | table | postgres
 public | weather | table | postgres
```

##To get out of psql, type:

```
 mydb=> \q

postgres=# SELECT version();
       version

-----------------------------------------------------------------------------------------------------------
 PostgreSQL 9.4.26 on x86_64-unknown-linux-gnu, compiled by gcc (GCC) 4.4.7 2012
0313 (Red Hat 4.4.7-23), 64-bit
(1 row)

postgres=# SELECT current_date;
```

#CREATE TABLE

```
-----------------------------------------------------------------------------------------------------------
CREATE TABLE COMPANY1(
   ID INT PRIMARY KEY     NOT NULL,

   NAME          TEXT    NOT NULL,
   AGE           INT     NOT NULL,
   ADDRESS       CHAR(50),
   SALARY        REAL
);
postgres=# CREATE TABLE COMPANY1(
postgres(#    ID INT PRIMARY KEY     NOT NULL,
postgres(#    NAME          TEXT    NOT NULL,
postgres(#    AGE           INT     NOT NULL,
postgres(#    ADDRESS       CHAR(50),
postgres(#    SALARY        REAL
postgres(# );
CREATE TABLE

INSERT INTO COMPANY1  VALUES (2, 'Ahmed', 25, 'Giza', '  20000.00');
INSERT INTO COMPANY1  VALUES (3, 'Lobna', 23, 'Giza', 20000.00 );
INSERT INTO COMPANY1   VALUES (1, 'Marwa', 32, 'Cairo', 30000.00);
INSERT INTO COMPANY1   VALUES (4, 'Amal', 32, 'Cairo', 30000.00);

select * FROM COMPANY1;

postgres=# INSERT INTO COMPANY1  VALUES (2, 'Ahmed', 25, 'Giza', '  20000.00');
INSERT 0 1
postgres=# INSERT INTO COMPANY1  VALUES (3, 'Lobna', 23, 'Giza', 20000.00 );
INSERT 0 1
```

```
postgres=# INSERT INTO COMPANY1   VALUES (1, 'Marwa', 32, 'Cairo', 30000.00);
INSERT 0 1
postgres=# INSERT INTO COMPANY1   VALUES (4, 'Amal', 32, 'Cairo', 30000.00);
INSERT 0 1
postgres=# select * FROM COMPANY1;
 id | name  | age |                    address
         | salary
----+-------+-----+-----------------------------------------
----------+--------
  2 | Ahmed |  25 | Giza
         |  20000
  3 | Lobna |  23 | Giza
         |  20000
  1 | Marwa |  32 | Cairo
         |  30000
  4 | Amal  |  32 | Cairo
         |  30000
(4 rows)
```

```
CREATE TABLE weather2 (
city varchar(80),
temp_lo int, -- low temperature
temp_hi int, -- high temperature
date date
);
postgres=# CREATE TABLE weather2 (
postgres(# city varchar(80),
postgres(# temp_lo int, -- low temperature
postgres(# temp_hi int, -- high temperature
postgres(# date date
postgres(# );
CREATE TABLE
INSERT  INTO weather2 VALUES (’ Giza' ,’20’,'40', '2020-8-10');
INSERT  INTO  weather2 VALUES (’ Cairo' ,’19’,'38', '2020-8-11');
INSERT  INTO  weather2 VALUES (’ Alexanderia' ,’15’,'30', '2020-8-11');
INSERT  INTO  weather2 VALUES (’ Sohag' ,’22’,'43', '2020-8-11');
INSERT  INTO  weather2 VALUES (’Asuit' ,’23’,'42', '2020-8-11');

SELECT city, temp_lo, temp_hi, date FROM weather2;
select * FROM weather2;

DROP TABLE weather2;
postgres=# \q
bash-4.1$ exit
exit
[root@localhost bigdata]# exit
exit
(base) [bigdata@localhost ~]$
```

#Sqoop
tar -xzvf sqoop-1.4.7.bin__hadoop-2.6.0.tar.gz

#check the JDBC dirver for PostgreSQL

cd /home/bigdata/sqoop-1.4.7.bin__hadoop-2.6.0/lib

#import table from DB to directory on HDFS

cd /home/bigdata/sqoop-1.4.7.bin__hadoop-2.6.0/bin

./sqoop import --connect 'jdbc:postgresql://127.0.0.1:5432/postgres' --username 'postgres' --password post --table 'company1' --target-dir '/sqp1' -m 1

##check HDFS
(base) [bigdata@localhost ~]$ hdfs dfs -cat /sqp1/part-m-00000
2,Ahmed,25,Giza ,20000.0
3,Lobna,23,Giza ,20000.0
1,Marwa,32,Cairo ,30000.0
4,Amal,32,Cairo ,30000.0