**Steps \_**

**1 mod****ify query condition**

rownum <= 1000

2. Update “Config.xml” file

example: "home/kfang/Desktop/rec\_data/runLog/" for log folder

3. Create Network Connection For Oracle

1. cd /usr/lib/oracle/11.2/client64 (or client32)

2. sudo mkdir network

3. sudo mkdir network/admin

4. sudo cp installTools/tnsnames.ora /usr/lib/oracle/11.2/client64/network/admin/

4. Create A Symbolic Link For Python (since it's usually in “bin” directory)

sudo ln -s /usr/bin/python /usr/local/bin/python

5. Start Hadoop

/home/kfang/hadoop-1.0.4/bin => "start-all.sh"

note: may need to wait for a while

http://localhost:50070/dfshealth.jsp

http://localhost:50030/jobtracker.jsp

6. Run The Engine

python processingManager.py -fcr

(if crashed after Hadoop part,can keep those files and -fcr -skip\_fcr\_export\_phase )

7. Clean For Next Run

delete "runtime" folder