**PDrive location**

[\\uskihfil4\Public\KIS\Asorey\ProjectsByTigerTeams\MAP](file:///\\uskihfil4\Public\KIS\Asorey\ProjectsByTigerTeams\MAP)

\\uskihfil4\public\KIS\Asorey\ProjectsByTigerTeams\MAP\SEARCH\Iteration 3\03 Development\SywInteractions\Hadoop (**Cleansing data JAR FILES**)

**Hdfs**

-bash-4.1$ hls /user/smoham3/tdexp

/user/smoham3/tdexp/Inferred\_Internal.out

**Morphline**

sudo -u mapadmin nohup /appl/solr/map/utils/mr\_update\_collection.sh -c sample-syw-profilegold-06-dv01 \

-i hdfs://heahadn01.hadoop.searshc.com:8020/user/smoham3/sample\_collection \

-o hdfs://heahadn01.hadoop.searshc.com:8020/user/smoham3/sample\_collection\_index \

-m /home/auto/smoham3/morphline/sample\_06/gold\_new.conf \

-s /home/auto/smoham3/morphline/sample\_06/schema\_new.xml > /home/auto/smoham3/morphline/sample\_06/Morphline6.log 2>&1 &

**PSCP**

pscp smoham3@151.149.131.21:/home/auto/smoham3/SampleHadoopJob/Inferred/Inferred\_Internal.pig C:\Users\smoham3\Desktop\

pscp C:\Users\smoham3\Downloads\SYW\_AggregatedCollection.pig smoham3@151.149.131.19:/home/auto/smoham3/

pscp C:\Users\smoham3\Desktop\Trip\_Count.jar smoham3@151.149.131.19:/home/auto/smoham3/scripts

**kill job**

Hadoop job –list

hadoop job -kill job\_1420170492480\_53610

**unix terminal kill**

**jobs** ->to check running jobs only started in that shell

**ps aux|grep smoham3** 🡪 all jobs started by this use

**kill -9 PID** (2nd column of the above command)

**fexp command**

**hdfs comand**

hrm -r /user/smoham3/cleanse/CleanData.jar

**To see hidden delimiter**

cat Inferred\_Internal.out | less -U

**To generate SQL script:**

fexp\_gen\_run\_v2.sh -t L2\_SelfServAnlysEngineDev\_S.SywCtlg -P -o ~/SYW/SywCtlg

./fexp\_gen\_run\_v2.sh -**P** -t L2\_MAPSrcViews\_S.StrFmtCdLu **–o ~/SYW/**StrFmtCdLu.fexp

fexp\_gen\_run\_v2.sh –P -t L2\_MAPSRCVIEWS\_S.SywrSegTagMstrCrm –o ~/SYW/ SywrSegTagMstrCrm.fexp

fexp\_gen\_run\_v2.sh -t L2\_SelfServAnlysEngineDev\_S.Vi\_segment\_mem\_TY\_Ly\_V3 -P -o ~/SYW/sam\_segments

/appl/scripts/TDextractor\_v2.sh -q ~/SampleHadoopJob/Inferred\_query -i /user/smoham3/tdexp -u smoham3 -p password -s tdadhoc.intra.searshc.com

fexp< Rd\_Kmt\_Soar\_Bu

fexp< rd\_srs\_soar\_bu

**Webserver PID**

sudo -u mapadmin netstat -plntu | grep 8081

sudo -u mapadmin ps -ef | grep -i carte

cd /logs/map/carte/

cd /appl/map/scripts/

**Hadoop file size**

**du**

Usage: hdfs dfs -du [-s] [-h] URI [URI ...]

Displays sizes of files and directories contained in the given directory or the length of a file in case its just a file.

Options:

* The -s option will result in an aggregate summary of file lengths being displayed, rather than the individual files.
* The -h option will format file sizes in a "human-readable" fashion (e.g 64.0m instead of 67108864)

Example:

* hdfs dfs -du /user/hadoop/dir1 /user/hadoop/file1 hdfs://nn.example.com/user/hadoop/dir1

Exit Code: Returns 0 on success and -1 on error.

**Hadoop file delete without trashing(fast delete)**

sudo -u mapadmin hadoop fs -rm -r -skipTrash /work/syw\_made/Inferred\_internal