**MYSQL DATA BASE SIDE**

**HOW TO CONNECT TO MYSQL DB**

mysql -u root –p

NOTE: Here password is “root”

**TO SEE THE AVAILABLE DATABASES**

mysql> show databases;

**TO SEE AVAILABLE TABLES**

mysql> show tables;

**TO CREATE A TABLE IN MYSQL**

**mysql> use gopaldb;**

**mysql> create table emp(empid int primary key,ename varchar(50),esal int);**

Query OK, 0 rows affected (0.00 sec)

**TO INSERT THE DATA INTO TABLE**

**mysql> insert into emp values(100,'EMP1',12000);**

**Query OK, 1 row affected (0.00 sec)**

**NOTE: insert multiple records in the same way like above**

**mysql> create table empdetails(id int,name varchar(50),sal int,address varchar(50));**

Query OK, 0 rows affected (0.00 sec)

**OTHER TABLES CREATION & DATA INSERTION**

**mysql> insert into empdetails VALUES (100,'Raju',12000,'Hyd'),(101,'Ravi',14000,'Pune'),(102,'Ramya',15000,'Chn'),(103,'Ranga',16000,'Ban');**

**mysql> create table depttab(deptid int primary key,dname varchar(50),dloc varchar(60));**

**Query OK, 0 rows affected (0.00 sec)**

**mysql> insert into depttab VALUES (100,'HRD','HYD'),(101,'TRANSPORT','BAN'),(102,'R&D','CHN'),(103,'ADMIN','HYD'),(104,'SALES','PUNE'),(105,'FINANCE','MUM'),(106,'MARKETING','HYD');**

**mysql> create table student(sid int primary key,name varchar(50),age int,marks int);**

**Query OK, 0 rows affected (0.00 sec)**

**mysql> insert into student VALUES (1,'ABHIRAM',21,80),(2,'BHAVYA',23,90),(3,'CHAITU',28,67),(4,'ESWAR',22,91);**

**SQOOP COMMANDS**

**TO CHECK MYSQL “DATABASES” , “TABLES” & “TABLES DATA”**

sqoop **list-databases** --connect jdbc:mysql://localhost:3306/ --username root --password root;

sqoop **list-tables** --connect jdbc:mysql://localhost:3306/gopaldb --username root --password root;

sqoop **eval** --connect jdbc:mysql://localhost:3306/gopaldb --username root --password root --query "select \* from emp limit 3";

sqoop **eval** --connect jdbc:mysql://localhost:3306/gopaldb --username root --password root --query "select \* from emp where esal > 25000";

sqoop **eval** **--options-file** /home/gopalkrishna/PRAC/SQOOP/ConnectionDetails.txt --query "show tables";

**DATA INGETION ON HDFS**

**TO IMPORT THE DATA FROM “RDBMS” to “HDFS”**

sqoop **import** --connect jdbc:mysql://localhost:3306/gopaldb --username root –password root --table emp;

sqoop import --connect jdbc:mysql://localhost:3306/gopaldb --username root --password root --table emp --target-dir=/Import -m 1;

sqoop import --connect jdbc:mysql://localhost:3306/gopaldb --username root --password root --table emp --target-dir=/Import1 -m 2;

**TO IMPORT ALL TABLES DATA FROM RDBMS TO DEFAULT PATH OF HDFS**

sqoop **import-all-tables** --options-file /home/gopalkrishna/PRAC/SQOOP/ConnectionDetails.txt -m 1;

**TO IMPORT ALL TABLES DATA FROM RDBMS TO HIVE WAREHOUSE PATH**

sqoop **import-all-tables** --options-file /home/gopalkrishna/PRAC/SQOOP/ConnectionDetails.txt -m 1 --warehouse-dir=/user/hive/warehouse/sqoopdata.db;

**TO IMPORT SELECTED RANGE OF TABLES DATA [--exclude-tables]**

sqoop **import-all-tables** --options-file /home/gopalkrishna/PRAC/SQOOP/ConnectionDetails.txt -m 1 **--exclude-tables** depttab,emp,student;

**TO IMPORT RANGE OF DATA FROM A PARTICULAR TABLE [--boundary-query]**

sqoop import --connect jdbc:mysql://localhost:3306/gopaldb --username root --password root --table emp --target-dir=/ImportBoundaryDirNew -m 1 --boundary-query "select 100,105 from emp";

**IMPORTING THE DATA FROM RDBMS TO HDFS IN DIFFRENT FORMS**

sqoop **import** --options-file /home/gopalkrishna/PRAC/SQOOP/ConnectionDetails.txt -m 1 --table emp **--as-textfile** --target-dir=/IMPORTTEXTFILE;

sqoop **import** --options-file /home/gopalkrishna/PRAC/SQOOP/ConnectionDetails.txt -m 1 --table emp **--as-sequencefile** --target-dir=/IMPORT\_SEQFILE;

sqoop **import** --options-file /home/gopalkrishna/PRAC/SQOOP/ConnectionDetails.txt -m 1 --table emp **--as-avrodatafile** --target-dir=/IMPORT\_AVROFILE;

**IMPORT THE DATA WITH SELECTED COLUMNS DATA ONLY**

sqoop import --options-file /home/gopalkrishna/PRAC/SQOOP/ConnectionDetails.txt -m 1 --table emp **--columns** 'ename,esal' --target-dir=/ImportCols;

**IMPORT THE DATA IN A SPECIFIED FORMAT [other than Default - CSV]**

sqoop **import** --options-file /home/gopalkrishna/PRAC/SQOOP/ConnectionDetails.txt --table emp --target-dir=/ImportFormat **--fields-terminated-by** '\t' -m 1;

**IMPORT THE DATA WITH SELECTED CRITERIA [--where]**

sqoop import --options-file /home/gopalkrishna/PRAC/SQOOP/ConnectionDetails.txt --table emp --target-dir=/ImportCond --fields-terminated-by '|' **--where** 'esal > 22000' -m 1;

**IMPORT THE DATA WITH “SPLIT-BY” OPTION**

sqoop import --options-file /home/gopalkrishna/PRAC/SQOOP/ConnectionDetails.txt --table emp --target-dir=/ImportSplit --fields-terminated-by '|' --where 'esal > 22000' **--split-by** empid -m 1;

**IMPORTING THE DATA WITH “COMPRESSION TECHNIQUES”**

sqoop import --options-file /home/gopalkrishna/PRAC/SQOOP/ConnectionDetails.txt --table emp --**compression-codec** **GzipCodec** -m 1 --target-dir=/ImportGZIP;

sqoop import --options-file /home/gopalkrishna/PRAC/SQOOP/ConnectionDetails.txt --table emp --**compression-codec** **BZip2Codec** -m 1 --target-dir=/ImportBZIP

**IMPORT WITH “–query” OPTION [\$CONDITIONS]**

sqoop **import** --connect jdbc:mysql://localhost:3306/gopaldb --username root --password root **--query** "select \* from emp **WHERE \$CONDITIONS**" --target-dir=/ImportQuery1 -m 1;

sqoop import --connect jdbc:mysql://localhost:3306/gopaldb --username root --password root --query "select e.empid,ename,esal,deptid,dname,dloc from emp e JOIN dept d ON (e.empid = d.empid) AND \$CONDITIONS" -m 1 --target-dir=/INNERJOIN26;

sqoop import --connect jdbc:mysql://localhost:3306/gopaldb --username root --password root --query "select e.empid,ename,esal,deptid,dname,dloc from emp e LEFT OUTER JOIN dept d ON (e.empid = d.empid) AND \$CONDITIONS" -m 1 --target-dir=/LEFTJOIN26;

sqoop import --connect jdbc:mysql://localhost:3306/gopaldb --username root --password root --query "select e.empid,ename,esal,deptid,dname,dloc from emp e RIGHT OUTER JOIN dept d ON (e.empid = d.empid) AND \$CONDITIONS" -m 1 --target-dir=/RIGHTJOIN26;

**SQOOP JOB**

sqoop **job** **--create gopaljob** -- import --connect jdbc:mysql://localhost/gopaldb --username root --password root --table emp -m 1 --target-dir /AppedJob **--append;**

sqoop **job** **--create gopaljobnew** -- import --connect jdbc:mysql://localhost/gopaldb --username root --password root --table emp -m 1 --fields-terminated-by '\t' --lines-terminated-by '\n' --target-dir /AppedJobNew **--append**;

**SQOOP JOB Options**

**To see the available jobs**

sqoop job --list;

**To See the Details about a Job**

sqoop job --show gopaljob;

**To Execute a Job**

sqoop job --exec gopaljob;

**To Delete a Job**

Sqoop job --delete gopaljob

**SQOOP INCREMENTAL LOADING**

sqoop **job** --create incrjob -- import --connect jdbc:mysql://localhost/gopaldb --username root --password root --table emp **--incremental append --check-column empid --last-value 0** --target-dir /IncrJob -m 1;

**To Execute the above Job:** sqoop job --exec incrjob;

**SQOOP EXPORT**

sqoop **export** --connect jdbc:mysql://localhost/gopaldb --username root --password root --table exporttab --export-dir /IncrJob/part-m-\*;

sqoop **export** --connect jdbc:mysql://localhost/gopaldb --username root --password root --table exporttab --export-dir /IncrJob/part-r-00001 --fields-terminated-by '\t';

sqoop **export** --connect jdbc:mysql://localhost/gopaldb --username root --password root --table exporttab --export-dir /IncrJob/part-r-00001 --fields-terminated-by '\t' **--outdir java\_code**

**sqoop export --options-file ConnectionDetails.txt -m 1 --table empnew --export-dir /ExportPath/initialFile.log --fields-terminated-by '#';**

**For MORE PERFOMANCE 🡪 USE –direct**

**sqoop export --options-file ConnectionDetails.txt -m 1 --table empnew --export-dir /ExportPath/initialFile.log --fields-terminated-by '#' --direct;**

**update-key**

**sqoop export --options-file ConnectionDetails.txt -m 1 --table empnew --export-dir /ExportPath/updatedFile.log --update-mode allowinsert --update-key id --fields-terminated-by '#';**

**SQOOP – INTEGRATION WITH – HIVE**

sqoop import-all-tables --connect jdbc:mysql://localhost/gopaldb --username root --password root --hive-import --hive-overwrite --create-hive-table --warehouse-dir=/user/hive/warehouse/sqoopImport1.db --compression-codec BZip2Codec --num-mappers 1 --outdir java\_files77

**NOTE:** **Once above command is successfully completed, we can access all these tables in hive default database**.

sqoop import --connect "jdbc:mysql://localhost:3306/gopaldb" --username=root --password=root --table student --fields-terminated-by '|' --lines-terminated-by '\n' **--hive-import --hive-overwrite --hive-table studetails** --outdir java\_codefiles2;

sqoop import --connect jdbc:mysql://localhost:3306/gopaldb --username=root --password=root --table student --fields-terminated-by '|' --lines-terminated-by '\n' **--hive-import --create-hive-table --hive-table studetailsnew** --outdir java\_codefiles;

**SQOOP – INTEGRATION WITH – HBASE**

sqoop import --connect jdbc:mysql://localhost/gopaldb --username root --password root --table emp --hbase-create-table --hbase-table empnew --column-family colfam1 --hbase-row-key empid --outdir java\_files -m 1;

**NOTE: Go to hbase shell and in the tables list , we can see “emp” table and if we use “scan emp” command-> we can see all the data of RDBMS emp table data over.**