[cloudera@quickstart saiclass]$ cat task\_25

mysql> create database custserdb;

mysql> use custserdb;

mysql> create table cust\_job(custid int,firstname varchar(22),createdt date);

mysql> show tables;

+---------------------+

| Tables\_in\_custserdb |

+---------------------+

| cust\_job |

+---------------------+

mysql> insert into cust\_job values(1,'Arun','2015-09-20');

mysql> insert into cust\_job values(2,'sirini','2015-09-21');

mysql> insert into cust\_job values(3,'vasu','2015-09-22');

mysql> select \* from cust\_job;

+--------+-----------+------------+

| custid | firstname | createdt |

+--------+-----------+------------+

| 1 | Arun | 2015-09-20 |

| 2 | sirini | 2015-09-21 |

| 3 | vasu | 2015-09-22 |

+--------+-----------+------------+

|==============================================|

| Edge node create the job, validate, execute |

|==============================================|

[cloudera@quickstart saiclass]$ sqoop job --create incjoob -- import --connect jdbc:mysql://localhost/custserdb --username root --password cloudera --table cust\_job -m 1 --target-dir /user/cloudera/job\_out --incremental append --check-column custid --last-value 0

[cloudera@quickstart saiclass]$ sqoop job --list

[cloudera@quickstart saiclass]$ sqoop job --exec incjoob

-------------------------------------

--> give password cloudera and enter

-------------------------------------

[cloudera@quickstart saiclass]$ sqoop job -list

Available jobs:

incjob

incjoob

[cloudera@quickstart saiclass]$ hadoop fs -ls /user/cloudera/job\_out

Found 1 items

-rw-r--r-- 1 cloudera supergroup 56 2021-12-26 09:04 /user/cloudera/job\_out/part-m-00000

[cloudera@quickstart saiclass]$ hadoop fs -cat /user/cloudera/job\_out/part-m-00000

1,Arun,2015-09-20

2,sirini,2015-09-21

3,vasu,2015-09-22

[cloudera@quickstart]$ cd

[cloudera@quickstart]$ cd .sqoop

[cloudera@quickstart .sqoop]$ ll

total 16

-rw-rw-r-- 1 cloudera cloudera 419 Dec 26 09:06 metastore.db.properties

-rw-rw-r-- 1 cloudera cloudera 11743 Dec 26 09:06 metastore.db.script

[cloudera@quickstart .sqoop]$ cat metastore.db.script | grep 'last.value'

INSERT INTO SQOOP\_SESSIONS VALUES('incjob','incremental.last.value','0','SqoopOptions')

INSERT INTO SQOOP\_SESSIONS VALUES('incjoob','incremental.last.value','3','SqoopOptions')

mysql> insert into cust\_job values(4,'momohad','2015-09-23');

mysql> insert into cust\_job values(5,'ravi','2015-09-24');

mysql> select \* from cust\_job;

+--------+-----------+------------+

| custid | firstname | createdt |

+--------+-----------+------------+

| 1 | Arun | 2015-09-20 |

| 2 | sirini | 2015-09-21 |

| 3 | vasu | 2015-09-22 |

| 4 | momohad | 2015-09-23 |

| 5 | ravi | 2015-09-24 |

+--------+-----------+------------+

-------------------------------------------------------

Execute the job again

-------------------------------------------------------

[cloudera@quickstart saiclass]$ sqoop job -list

Available jobs:

incjoob

[cloudera@quickstart saiclass]$ sqoop job --exec incjoob

[cloudera@quickstart saiclass]$ cat ../.sqoop/metastore.db.script | grep 'last.value'

INSERT INTO SQOOP\_SESSIONS VALUES('incjob','incremental.last.value','0','SqoopOptions')

INSERT INTO SQOOP\_SESSIONS VALUES('incjoob','incremental.last.value','5','SqoopOptions')

[cloudera@quickstart saiclass]$ hadoop fs -ls /user/cloudera/job\_out

Found 2 items

-rw-r--r-- 1 cloudera supergroup 56 2021-12-26 09:04 /user/cloudera/job\_out/part-m-00000

-rw-r--r-- 1 cloudera supergroup 39 2021-12-26 09:15 /user/cloudera/job\_out/part-m-00001

[cloudera@quickstart saiclass]$ hadoop fs -cat /user/cloudera/job\_out/part-m-00000

1,Arun,2015-09-20

2,sirini,2015-09-21

3,vasu,2015-09-22

[cloudera@quickstart saiclass]$ hadoop fs -cat /user/cloudera/job\_out/part-m-00001

4,momohad,2015-09-23

5,ravi,2015-09-24

[cloudera@quickstart saiclass]$ sqoop import --connect jdbc:mysql://localhost/custserdb --username root --password-file file:////home/cloudera/saiclass/pwfile --table cust\_job -m 1 --delete-target-dir --target-dir /user/cloudera/job\_out\_asparquet --as-parquetfile

[cloudera@quickstart saiclass]$ sqoop import --connect jdbc:mysql://localhost/custserdb --username root --password-file file:hdfs/cloudera/saiclass/pwfile --table cust\_job -m 1 --delete-target-dir --target-dir /user/cloudera/job\_out\_asparquet --as-parquetfile

[cloudera@quickstart saiclass]$ hadoop fs -cat /user/cloudera/job\_out\_asparquet/\*

cat: `/user/cloudera/job\_out\_asparquet/.metadata': Is a directory

cat: `/user/cloudera/job\_out\_asparquet/.signals': Is a directory

PAR144,

$­

(fh,

vasuArun3\­ArunsiriniXvasumomohadravi\X,

&O]O.D­]O8qO&OLcust\_job%custid

% firstname%%createdt

<custid

vv<&~ firstname

&~<vasuArun&createdt

&<&O]O

parquet.avro.schema{"type":"record","name":"cust\_job","doc":"Sqoop import of cust\_job","fields":[{"name":"custid","type":["null","int"],"default":null,"columnName":"custid","sqlType":"4"},{"name":"firstname","type":["null","string"],"default":null,"columnName":"firstname","sqlType":"12"},{"name":"createdt","type":["null","long"],"default":null,"columnName":"createdt","sqlType":"91"}],"tableName":"cust\_job"}9parquet-mr version 1.5.0-cdh5.12.0 (build ${buildNumber})PAR1[cloudera@quickstart saiclass]$

[cloudera@quickstart saiclass]$ sqoop import --connect jdbc:mysql://localhost/custserdb --username root --password-file file:hdfs/cloudera/saiclass/pwfile --table cust\_job -m 1 --delete-target-dir --target-dir /user/cloudera/job\_out\_asparquet --as-parquetfile