

Data Ingestion from the RDS to HDFS using Sqoop

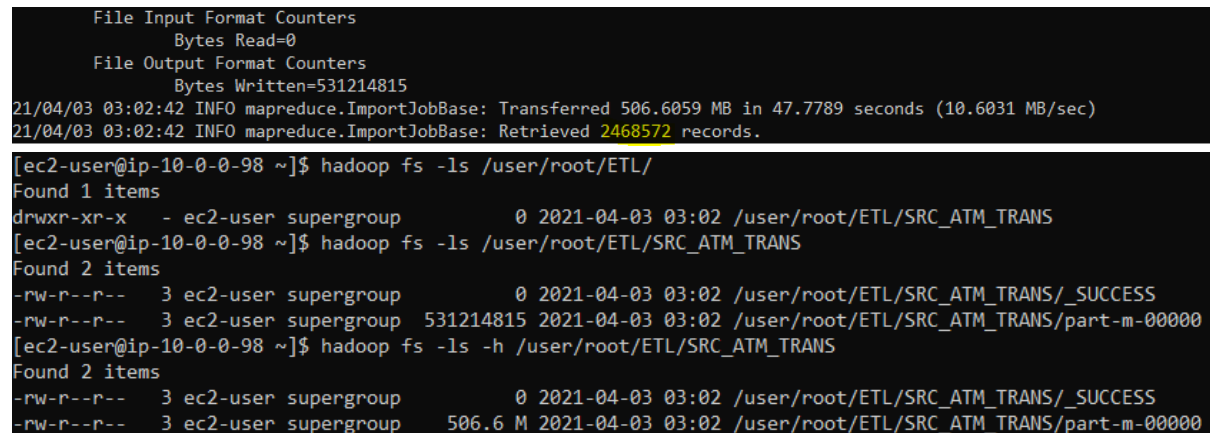
Sqoop Import command used for importing table from RDS to HDFS:

```
sqoop import --connect jdbc:mysql://xxxxxxxxxx.us-east-1.rds.amazonaws.com/testdatabase \  
--table SRC_ATM_TRANS \  
--username student \  
--password STUDENT123 \  
--target-dir /user/root/ETL/SRC_ATM_TRANS \  
-m 1
```

Command used to see the list of imported data in HDFS:

```
hadoop fs -ls /user/root/ETL/SRC_ATM_TRANS
```

Screenshot of the imported data:



```
File Input Format Counters
  Bytes Read=0
File Output Format Counters
  Bytes Written=531214815
21/04/03 03:02:42 INFO mapreduce.ImportJobBase: Transferred 506.6059 MB in 47.7789 seconds (10.6031 MB/sec)
21/04/03 03:02:42 INFO mapreduce.ImportJobBase: Retrieved 2468572 records.

[ec2-user@ip-10-0-0-98 ~]$ hadoop fs -ls /user/root/ETL/
Found 1 items
drwxr-xr-x  - ec2-user supergroup          0 2021-04-03 03:02 /user/root/ETL/SRC_ATM_TRANS
[ec2-user@ip-10-0-0-98 ~]$ hadoop fs -ls /user/root/ETL/SRC_ATM_TRANS
Found 2 items
-rw-r--r--  3 ec2-user supergroup          0 2021-04-03 03:02 /user/root/ETL/SRC_ATM_TRANS/_SUCCESS
-rw-r--r--  3 ec2-user supergroup 531214815 2021-04-03 03:02 /user/root/ETL/SRC_ATM_TRANS/part-m-00000
[ec2-user@ip-10-0-0-98 ~]$ hadoop fs -ls -h /user/root/ETL/SRC_ATM_TRANS
Found 2 items
-rw-r--r--  3 ec2-user supergroup          0 2021-04-03 03:02 /user/root/ETL/SRC_ATM_TRANS/_SUCCESS
-rw-r--r--  3 ec2-user supergroup 506.6 M 2021-04-03 03:02 /user/root/ETL/SRC_ATM_TRANS/part-m-00000
```