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://upgraddetest.cyaielc9bmnf.us-east-1.rds.amazonaws.com/testdatabase \
>
> --table SRC_ATM_TRANS \
>
> --username student \
>
> --password STUDENT123 \
>
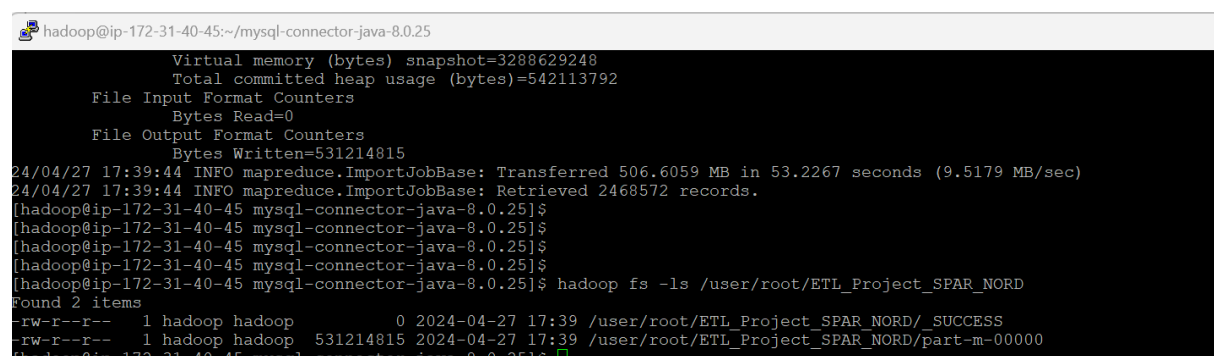> --target-dir /user/root/ETL_Project_SPAR_NORD \

This command transfers data from RDS table "SRC_ATM_TRANS" to an Hadoop.

> --connect: JDBC connection string for the RDS(MySQL) database
>
> --username: username to use to connect to the Database
>
> --password: password to use to connect to the Database
>
> --table: name of the source (MySQL) table
>
> --target-dir: Target directory where the data will be stored in HDFS
>
> --m: Number of map tasks

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

hdfs fs -ls /user/root/ETL_Project_SPAR_NORD

Screenshot of the imported data:

hadoop fs -cat /user/root/ETL_Project_SPAR_NORD/part-m-00000

<Screenshot of the imported data>