

## Data Ingestion from the RDS to HDFS using Sqoop

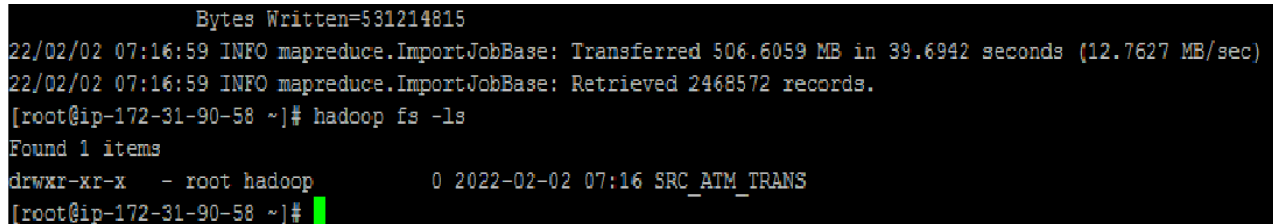
### Sqoop Import command used for importing table from RDS to HDFS:

```
sqoop import \  
--connect jdbc:mysql://upgraddetest.cyaiehc9bmnf.us-east-  
1.rds.amazonaws.com/testdatabase \  
--table SRC_ATM_TRANS \  
--username student --password STUDENT123 \  
-m 1
```

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

```
hadoop fs -ls
```

### Screenshot of the imported data:



```
Bytes Written=531214815  
22/02/02 07:16:59 INFO mapreduce.ImportJobBase: Transferred 506.6059 MB in 39.6942 seconds (12.7627 MB/sec)  
22/02/02 07:16:59 INFO mapreduce.ImportJobBase: Retrieved 2468572 records.  
[root@ip-172-31-90-58 ~]# hadoop fs -ls  
Found 1 items  
drwxr-xr-x  - root hadoop          0 2022-02-02 07:16 SRC_ATM_TRANS  
[root@ip-172-31-90-58 ~]#
```

### Explanation:

- After running EMR cluster , install MYSQL.
- After installing MYSQL, write the import sqoop command to import from RDS to HDFS as written above in the root user.
- After that just do hadoop fs -ls, which will show the loaded data like found in the above.