Importing data into Hive

sqoop import --connect jdbc:mysql://localhost/fleetdb \

--username root --password root \

--table driver \

-m 3 --split-by driverId \

--target-dir '/user/fleetdb/' --delete-target-dir \

--hive-import \

--create-hive-table \

--hive-table fleetdb.driver;

[hduser@Inceptez ~]$ hadoop fs -ls /user/hive/warehouse

Found 10 items

drwxr-xr-x - hduser supergroup 0 2019-06-20 07:55 /user/hive/warehouse/\_tmp.txnrecord\_es

drwxr-xr-x - hduser supergroup 0 2019-06-24 10:17 /user/hive/warehouse/\_tmp.txnrecord\_es\_id

drwxr-xr-x - hduser supergroup 0 2019-03-30 04:05 /user/hive/warehouse/custdb.db

drwxr-xr-x - hduser supergroup 0 2019-03-26 23:16 /user/hive/warehouse/customerall

drwxr-xr-x - hduser supergroup 0 2019-06-25 23:37 /user/hive/warehouse/fleetdb.db

drwxr-xr-x - hduser supergroup 0 2019-03-29 20:49 /user/hive/warehouse/retail.db

drwxrwxr-x - hduser supergroup 0 2019-03-24 21:13 /user/hive/warehouse/retail1.db

drwxr-xr-x - hduser supergroup 0 2019-06-20 07:47 /user/hive/warehouse/txnrecord

drwxr-xr-x - hduser supergroup 0 2019-06-20 07:50 /user/hive/warehouse/txnrecord\_es

drwxr-xr-x - hduser supergroup 0 2019-06-24 10:17 /user/hive/warehouse/txnrecord\_es\_id

Even though we give target-dir, the data is imported into the database folder created in the warehouse folder.

Lets see inside the table folder.

[hduser@Inceptez ~]$ hadoop fs -ls /user/hive/warehouse/fleetdb.db/driver

Found 3 items

-rw-r--r-- 1 hduser supergroup 636 2019-06-25 23:37 /user/hive/warehouse/fleetdb.db/driver/part-m-00000

-rw-r--r-- 1 hduser supergroup 652 2019-06-25 23:37 /user/hive/warehouse/fleetdb.db/driver/part-m-00001

-rw-r--r-- 1 hduser supergroup 617 2019-06-25 23:37 /user/hive/warehouse/fleetdb.db/driver/part-m-00002

When we again run the same command, the data is appended to the same folder “driver” but as new files.

[hduser@Inceptez ~]$ hadoop fs -ls /user/hive/warehouse/fleetdb.db/driver

Found 6 items

-rw-r--r-- 1 hduser supergroup 636 2019-06-25 23:37 /user/hive/warehouse/fleetdb.db/driver/part-m-00000

-rw-r--r-- 1 hduser hadoop 636 2019-06-25 23:52 /user/hive/warehouse/fleetdb.db/driver/part-m-00000\_copy\_1

-rw-r--r-- 1 hduser supergroup 652 2019-06-25 23:37 /user/hive/warehouse/fleetdb.db/driver/part-m-00001

-rw-r--r-- 1 hduser hadoop 652 2019-06-25 23:52 /user/hive/warehouse/fleetdb.db/driver/part-m-00001\_copy\_1

-rw-r--r-- 1 hduser supergroup 617 2019-06-25 23:37 /user/hive/warehouse/fleetdb.db/driver/part-m-00002

-rw-r--r-- 1 hduser hadoop 617 2019-06-25 23:52 /user/hive/warehouse/fleetdb.db/driver/part-m-00002\_copy\_1

**Below are the major commands used in the –hive-import..**

| **Sqoop Command Option** | **Description** |
| --- | --- |
| --hive-home <directory> | Overrides $HIVE\_HOME. |
| --hive-import | Imports tables into Hive using Hive's default delimiters if none are explicitly set. |
| --hive-overwrite | Overwrites existing data in the Hive table. |
| --create-hive-table | Creates a hive table during the operation. If this option is set and the Hive table already exists, the job will fail. Set to false by default. |
| --hive-table <table\_name> | Specifies the table name to use when importing data into Hive. |
| --hive-drop-import-delims | Drops the delimiters \n, \r, and \01 from string fields when importing data into Hive. |
| --hive-delims-replacement | Replaces the delimiters \n, \r, and \01 from strings fields with a user-defined string when importing data into Hive. |
| --hive-partition-key | Specifies the name of the Hive field on which a sharded database is partitioned. |
| --hive-partition-value <value> | A string value that specifies the partition key for data imported into Hive. |
| --map-column-hive <map> | Overrides the default mapping from SQL type to Hive type for configured columns. |

Importing data into HBase

sqoop import --connect jdbc:mysql://localhost/fleetdb --username root --password root \

--table driver -m 3 --split-by driverId \

--hbase-table fleetdb \

--column-family driver \

--hbase-row-key driverId ;