

## Table Properties 2 – Immutable

- Hive provides table immunity from insert into command. Insert into command is used to append to a table or partition keeping the existing data intact.
- All the tables by default are mutable.

Let's see an example, Suppose we have two mutable tables, with some common and some different data.

```
hive> Select * from EMP;
OK
emp.emp_id      emp.emp_name      emp.dept_id
249972  SwatiGirhepunje  10
249973  Tanjila Pathan   10
249974  Shweta Bedmutha  20
149975  Sheela Sawant    50
249976  Rajesh Kharache  60
Time taken: 0.205 seconds, Fetched: 5 row(s)
hive> Select * from EMP_MAY;
OK
emp_may.emp_id  emp_may.emp_name      emp_may.dept_id
249972  SwatiGirhepunje  10
249977  Suvarna Katkar     10
249978  Shneha Kapgate     20
149975  Sheela Sawant      50
249976  Rajesh Kharache    60
Time taken: 0.557 seconds, Fetched: 5 row(s)
hive>
```

Now let's insert all the data from EMP\_MAY table into EMP table.

```
hive> insert into EMP select * from EMP_MAY;
Query ID = root_20200626093030_72cd82c6-2b5d-4bd8-9863-c59bd79f53d3
Total jobs = 3
Launching Job 1 out of 3
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1593160482880_0002, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1593160482880_0002/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1593160482880_0002
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 0
2020-06-26 10:59:27,625 Stage-1 map = 0%, reduce = 0%
2020-06-26 11:00:08,213 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.91 sec
MapReduce Total cumulative CPU time: 2 seconds 910 msec
Ended Job = job_1593160482880_0002
Stage-4 is selected by condition resolver.
Stage-3 is filtered out by condition resolver.
Stage-5 is filtered out by condition resolver.
Moving data to: hdfs://quickstart.cloudera:8020/user/hive/warehouse/emp/.hive-staging_hive_2020-06-26_10-58-47_512_4216166655778028332-1/-ext-10000
Loading data to table default.emp
Table default.emp stats: [numFiles=2, numRows=5, totalSize=253, rawDataSize=121]
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Cumulative CPU: 2.91 sec HDFS Read: 3459 HDFS Write: 194 SUCCESS
Total MapReduce CPU Time Spent: 2 seconds 910 msec
OK
emp_may.emp_id  emp_may.emp_name      emp_may.dept_id
Time taken: 88.078 seconds
hive> Select * from EMP;
OK
emp.emp_id      emp.emp_name      emp.dept_id
249972  SwatiGirhepunje  10
249977  Suvarna Katkar     10
249978  Shneha Kapgate     20
149975  Sheela Sawant      50
249976  Rajesh Kharache    60
249972  SwatiGirhepunje  10
249973  Tanjila Pathan     10
249974  Shweta Bedmutha    20
149975  Sheela Sawant      50
249976  Rajesh Kharache    60
Time taken: 0.261 seconds, Fetched: 10 row(s)
hive>
```

## Immutable Table

A Table can be created as immutable table, by setting its table property to True. By default this property is false.

```
create table if not exists EMP1 (emp_id int, emp_name string, dept_id int) row format delimited fields terminated by ',' lines terminated by '\n' stored as textfile tblproperties ("immutable"="true");
```

```
hive> create table if not exists EMP1 (emp_id int, emp_name string, dept_id int) row format delimited fields terminated by ',' lines terminated by '\n' stored as textfile tblproperties ("immutable"="true");
OK
Time taken: 0.452 seconds
hive>
```

- load file into EMP1 table.

```
hive> load data inpath '/user/cloudera/EMP_Data.txt' into table EMP1;
Loading data to table default.emp1
chgrp: changing ownership of 'hdfs://quickstart.cloudera:8020/user/hive/warehouse/emp1/EMP_Data.txt': Permission denied. user=root is not the owner of inode=EMP_Data.txt
chmod: changing permissions of 'hdfs://quickstart.cloudera:8020/user/hive/warehouse/emp1/EMP_Data.txt': Permission denied. user=root is not the owner of inode=EMP_Data.txt
Table default.emp1 stats: [numFiles=1, totalSize=127]
OK
Time taken: 1.069 seconds
```

```
hive> Select * from EMP1;
OK
emp1.emp_id    emp1.emp_name    emp1.dept_id
249972    SwatiGirhepunje    10
249973    Tanjila Pathan    10
249974    Shweta Bedmutha    20
149975    Sheela Sawant    50
249976    Rajesh Kharache    60
Time taken: 0.221 seconds, Fetched: 5 row(s)
hive>
```

- Verifying table properties

```
hive> show tblproperties EMP1;
OK
prpt_name      prpt_value
COLUMN_STATS_ACCURATE    true
immutable       true
numFiles        1
totalSize       127
transient_lastDdlTime    1593170910
Time taken: 0.118 seconds, Fetched: 5 row(s)
hive>
```

Immutable property allows us to load the data for the first time. After the first data insert, the successive insert into commands will fail resulting in only one set of data into table.

- Insert into command on Non-Empty Immutable Table will throw error saying, inserting into non-empty immutable table is not allowed.

```
hive> insert into EMP1 select * from EMP_MAY;
FAILED: SemanticException [Error 10256]: Inserting into a non-empty immutable table is not allowed emp1
hive>
```

- Insert Overwrite command on Non-Empty Immutable Table is not affected by immutable property

Overwrite means , clearing the table contents first and then inserting the new data.

```
hive> insert overwrite table EMP1 select * from EMP_MAY;
Query ID = root_20200626093030_72cd82c6-2b5d-4bd8-9863-c59bd79f53d3
Total jobs = 3
Launching Job 1 out of 3
Number of reduce tasks is set to 0 since there's no reduce operator
Starting Job = job_1593160482880_0003, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1593160482880_0003/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1593160482880_0003
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 0
2020-06-26 11:40:28,954 Stage-1 map = 0%, reduce = 0%
2020-06-26 11:41:30,004 Stage-1 map = 0%, reduce = 0%
2020-06-26 11:42:30,832 Stage-1 map = 0%, reduce = 0%
2020-06-26 11:43:17,607 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.89 sec
MapReduce Total cumulative CPU time: 2 seconds 890 msec
Ended Job = job_1593160482880_0003
Stage-4 is selected by condition resolver.
Stage-3 is filtered out by condition resolver.
Stage-5 is filtered out by condition resolver.
Moving data to: hdfs://quickstart.cloudera:8020/user/hive/warehouse/emp1/.hive-staging_hive_2020-06-26_11-39-23_501_869612884753087917
0-1/-ext-10000
Loading data to table default.emp1
Table default.emp1 stats: [numFiles=1, numRows=5, totalSize=126, rawDataSize=121]
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Cumulative CPU: 3.01 sec HDFS Read: 3394 HDFS Write: 195 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 10 msec
OK
emp_may.emp_id emp_may.emp_name emp_may.dept_id
Time taken: 254.744 seconds
hive>
```

```
hive> Select * from EMP1;
OK
emp1.emp_id emp1.emp_name emp1.dept_id
249972 SwatiGirhepunje 10
249977 Suvarna Katkar 10
249978 Shneha Kapgate 20
149975 Sheela Sawant 50
249976 Rajesh Kharache 60
Time taken: 0.281 seconds, Fetched: 5 row(s)
hive>
```