



BIG DATA
HADOOP
&
SPARK
TRAINING

ACADGILD

ASSIGNMENT
8.3

BY :-

SAHIL
KHURANA

TRANSACTIONS IN HIVE

Transactions are provided at the row-level in Hive 0.14.

The different row-level transactions available in Hive 0.14 are as follows:

- Insert
- Delete
- Update

Creating a Table That Supports Hive Transactions After that, inserting the values in the created table.

```

ne (i.e. spark, tez) or using Hive 1.X releases.
hive> set hive.support.concurrency=true;
hive> set hive.enforce.bucketing=true;
Query returned non-zero code: 1, cause: hive configuration hive.enforce.bucketing does not exists.
hive> set hive.exec.dynamic.partition.mode=nonstrict;
hive> set hive.txn.manager = org.apache.hadoop.hive.ql.lockmgr.DbTxnManager;
hive> set hive.compactor.initiator.on=true;
hive> set hive.compactor.worker.threads = 5;
hive> create table college(clg_id int,clg_name string,clg_loc string) clustered by (clg_id) into 5 buckets stored as orc TBLP
ROPERTIES('transactional'='true');
FAILED: Execution Error, return code 1 from org.apache.hadoop.hive.ql.exec.DDLTask. AlreadyExistsException(message:Table coll
ege already exists)
hive> create table college1(clg_id int,clg_name string,clg_loc string) clustered by (clg_id) into 5 buckets stored as orc TBL
PROPERTIES('transactional'='true');
OK
Time taken: 1.202 seconds
hive> insert into table college1 values(1,'ars','shi'),(2,'kha','han'),(3,'abc','bcd'),(4,'cde','def'),(5,'efg','fgh'),(6,'ghi'
,'hij'),(7,'ijk','jkl');
FAILED: SemanticException [Error 10293]: Unable to create temp file for insert values Expression of type TOK_TABLE_OR_COL not
supported in insert/values
hive> insert into table college1 values (1,'ars','shi'),(2,'kha','han'),(3,'abc','bcd'),(4,'cde','def'),(5,'efg','fgh'),(6,'g
hi','hij'),(7,'ijk','jkl');
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execu
tion engine (i.e. spark, tez) or using Hive 1.X releases.
Query ID = acadgild_20170516231747_56a66278-a5b8-48aa-98f3-cbb8c7c18f02
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 5
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>

```

Then using 'select' keyword displaying the data of the created table.

```

Time taken: 134.864 seconds
hive> select * from college1;
OK
5      efg      fgh
6      ghi      hij
1      ars      shi
7      ijk      jkl
2      kha      han
3      abc      bcd
4      cde      def
Time taken: 0.442 seconds, Fetched: 7 row(s)

```

On again inserting the values in the table will append the data. Attached screenshot is showing Appended data using select keyword.

```

Time taken: 0.442 seconds, Fetched: 7 row(s)
hive> insert into table college1 values (1,'ars','shi'),(2,'kha','han'),(3,'abc','bcd'),(4,'cde','def'),(5,'efg','fgh'),(6,'g
hi','hij'),(7,'ijk','jkl');
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execu
tion engine (i.e. spark, tez) or using Hive 1.X releases.

```

Appended data :

```
Time taken: 113.325 seconds
hive> select * from college1;
OK
5      efg      fgh
5      efg      fgh
6      ghi      hij
1      ars      shi
6      ghi      hij
1      ars      shi
7      ijk      jkl
2      kha      han
7      ijk      jkl
2      kha      han
3      abc      bcd
3      abc      bcd
4      cde      def
4      cde      def
Time taken: 0.278 seconds, Fetched:
```

Updating the table :

From the above image, we can see that we have received an error message. This means that the Update command is not supported on the columns that are bucketed.

In this table, we have bucketed the 'clg_id' column and performing the Update operation on the same column, so we have got the error.

Then, performing the update operation.

```
FAILED: ParseException line 1:10 missing SET at 'clg_id' near '<EOF>'
hive> update college1 set clg_id=8 where clg_id=7;
FAILED: SemanticException [Error 10302]: Updating values of bucketing columns is not supported. Column clg_id.
hive> update college1 set clg_id = 8 where clg_id = 7;
FAILED: SemanticException [Error 10302]: Updating values of bucketing columns is not supported. Column clg_id.
hive> update college1 set clg_name = 'IIT' where clg_id = 6;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different
tion engine (i.e. spark, tez) or using Hive 1.X releases.
Query ID = acadgild_20170516232650_0a7e0f47-e066-4baf-9aa7-564fa69d2645
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 5
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1494954759900_0003, Tracking URL = http://localhost:8088/proxy/application_1494954759900_0003/
```

In the screenshot we can viewed that table is updated.

```
Total MapReduce CPU Time Spent: 43 seconds 580 msec
OK
Time taken: 189.609 seconds
hive> select * from college1;
OK
5      efg      fgh
5      efg      fgh
6      IIT      hij
1      ars      shi
6      IIT      hij
1      ars      shi
7      ijk      jkl
2      kha      han
7      ijk      jkl
2      kha      han
3      abc      bcd
3      abc      bcd
4      cde      def
4      cde      def
Time taken: 0.412 seconds. Fetched: 14 row(s)
```

Deleting the row from the table :

```
Time taken: 0.412 seconds. Fetched: 14 row(s)
hive> delete from college1 where clg_id=5;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future ver
sion engine (i.e. spark, tez) or using Hive 1.X releases.
```

Here, we can see that rows where clg_id=5 got deleted.

```
Total MapReduce CPU Time Spent: 39 seconds 580 msec
OK
Time taken: 185.579 seconds
hive> select * from college1;
OK
6      IIT      hij
1      ars      shi
6      IIT      hij
1      ars      shi
7      ijk      jkl
2      kha      han
7      ijk      jkl
2      kha      han
3      abc      bcd
3      abc      bcd
4      cde      def
4      cde      def
Time taken: 0.563 seconds. Fetched: 12 row(s)
```