Refer link

[https://cwiki.apache.org/confluence/display/Hive/Hive+Transactions](https://cwiki.apache.org/confluence/display/Hive/Hive+Transactions#HiveTransactions-Limitations)

Set below parameters

|  |
| --- |
| - set hive.txn.manager by default it uses DummyTxnManager  set hive.txn.manager=org.apache.hadoop.hive.ql.lockmgr.DbTxnManager;  -- Also make sure these properties are set  -- hive.support.concurrency – true  -- hive.enforce.bucketing – true (Not required as of Hive 2.0)  -- hive.exec.dynamic.partition.mode – nonstrict  -- hive.compactor.initiator.on – true (for exactly one instance of the Thrift metastore service)  -- hive.compactor.worker.threads – a positive number on at least one instance of the Thrift metastore service |

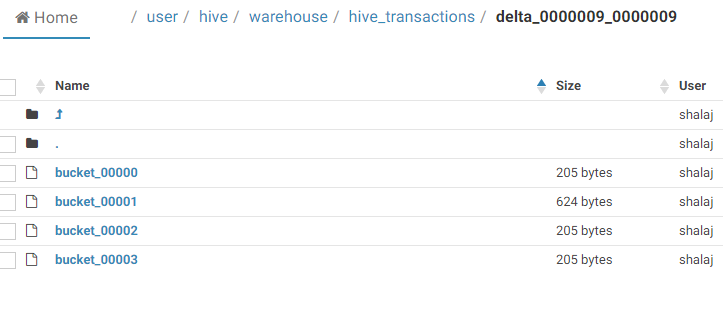
|  |
| --- |
| set hive.txn.manager=org.apache.hadoop.hive.ql.lockmgr.DbTxnManager;  set hive.support.concurrency =true;  set hive.enforce.bucketing=true;  set hive.compactor.initiator.on=true;  set hive.compactor.worker.threads=1; |

|  |
| --- |
| **-- Creating table**  **-- Make sure table is bucketed, file format is orc and**  **-- transactional is set to true under tblproperties**  create table hive\_transactions (i int, j string)  clustered by (i) into 4 buckets  stored as orc  tblproperties ('transactional'='true'); |

|  |
| --- |
| insert into table hive\_transactions values (1, 'one');  insert into table hive\_transactions values (2, 'random value'); |



Create two delta folder for two separate insert statement each folder contains 4 bucketed files



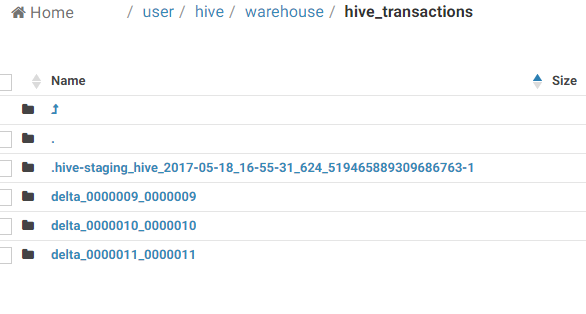
For record 1 Value persisted on bucket\_00001 only, check the file size as ORC contains some metadata that’s why all other files contains some size



For record 2 Value persisted on bucket\_00002 only, check the file size as ORC contains some metadata that’s why all other files contains some size

|  |
| --- |
| -- Updating data  update hive\_transactions set j = 'Two' where i = 2; |

Note: Cannot update bucketed column



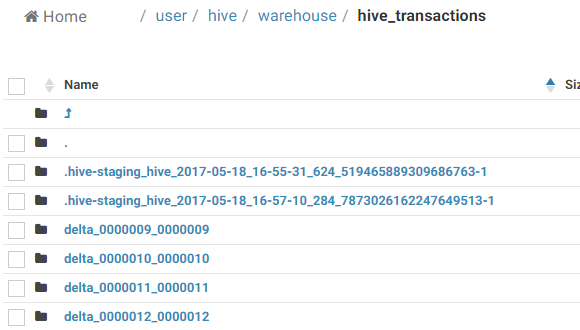
Create new delta folder **delta\_0000011\_0000011**



It is having only **bucket\_00002** file, so new folder and file created for update

|  |
| --- |
| select \* from hive\_transactions;  1 One  2 Two |

|  |
| --- |
| * Delete data   delete from hive\_transactions where i = 1; |





It will create new folder having only one bucket file that contain just one record, it is soft delete so original folder will not get deleted immediately , at the time of compaction they will get removed

|  |
| --- |
| select \* from hive\_transactions;  2 Two |