**Session 28**

**Assignment 1**

**Problem Statement:**

**Part 1:**

What are ACID properties and Explain Transactions in Hive?

**A.C.I.D:**

An acronym for Atomicity, Consistency Isolation, Durability and is a concept that Database Professionals generally look for when evaluating databases and application architectures. For a reliable database all these four attributes should be achieved.

**Properties of ACID:**

* **Atomicity** is an all-or-none proposition, that is, a transaction should complete successfully or else it should fail completely but it should not be left partially.
* **Consistency** guarantees that a transaction never leaves your database in a half-finished state by ensuring that any transaction will bring the database from one valid state to another state.
* **Isolation** states that every transaction should be independent of each other i.e. one transaction should not affect another by keeping transactions separated from each other until they’re finished.
* **Durability** states that if a transaction is completed, it should be preserved in the database even if the machine state is lost or a system failure might occur.

These ACID properties are essential for a transaction and every transaction should ensure that these properties are met.

**Transactions in Hive:**

Transactions in Hive are introduced in Hive 0.13, but they only partially fulfill the ACID properties like atomicity, consistency, durability, at the partition level. Here, Isolation can be provided by turning on one of the locking mechanisms available with zookeeper or in memory.

But in Hive 0.14, new API’s have been added to completely fulfill the ACID properties while performing any transaction.

Transactions are provided at the row-level in Hive 0.14. The different row-level transactions available in Hive 0.14 are as follows:

1. Insert
2. Delete
3. Update

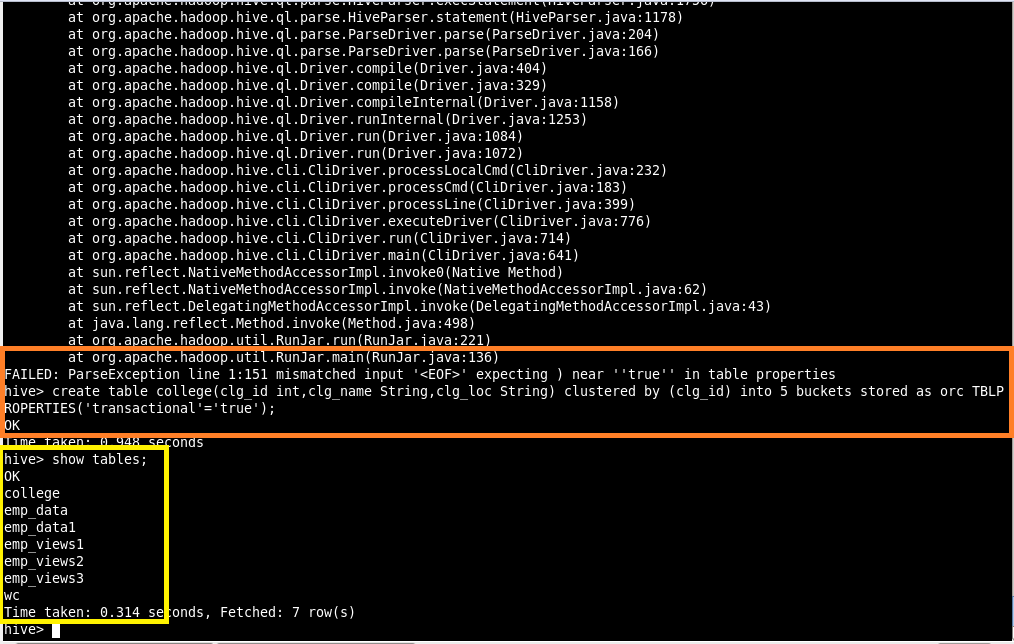
There are numerous limitations with the present transactions available in Hive 0.14. ORC is the file format supported by Hive transaction. It is now essential to have ORC file format for performing transactions in Hive. The table needs to be bucketed in order to support transactions.

**Part 2:**

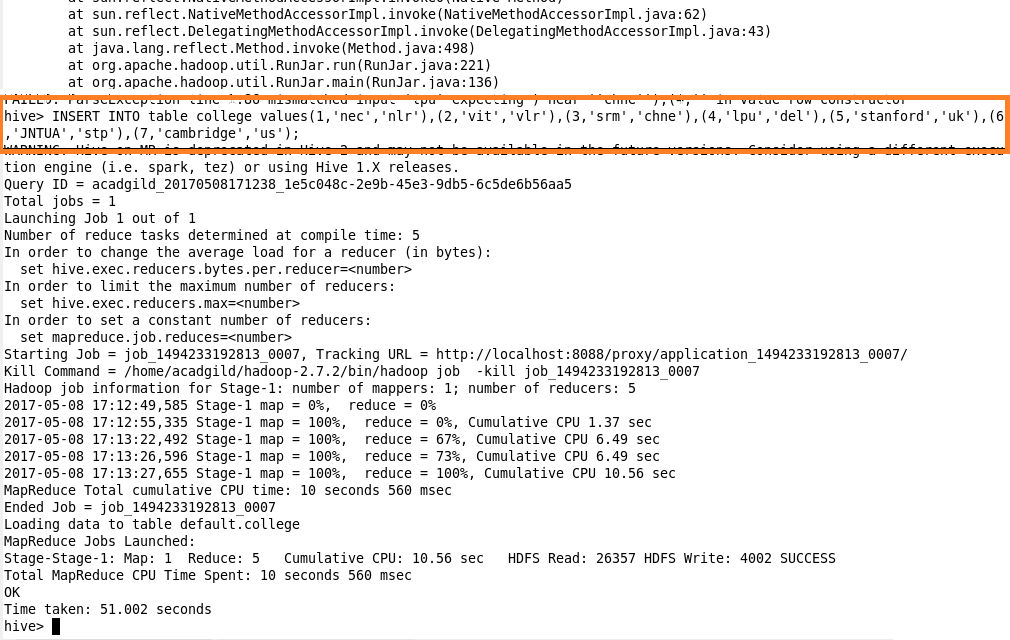
Refer our below blog and perform Hive Transaction Operations.

* https://acadgild.com/blog/transactions-in-hive/

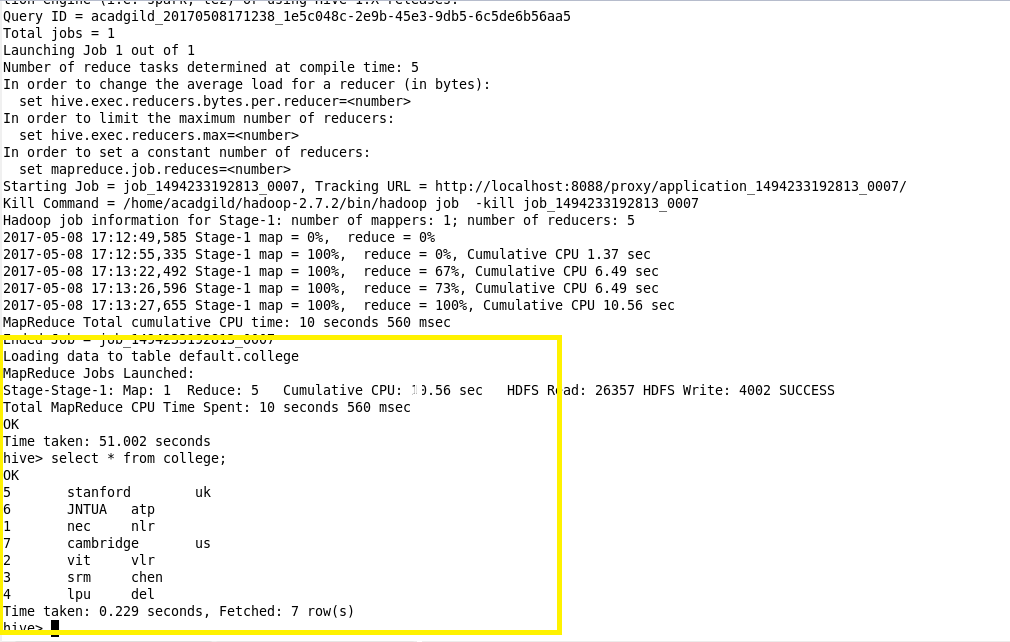
**Step 1:** Creating a table.



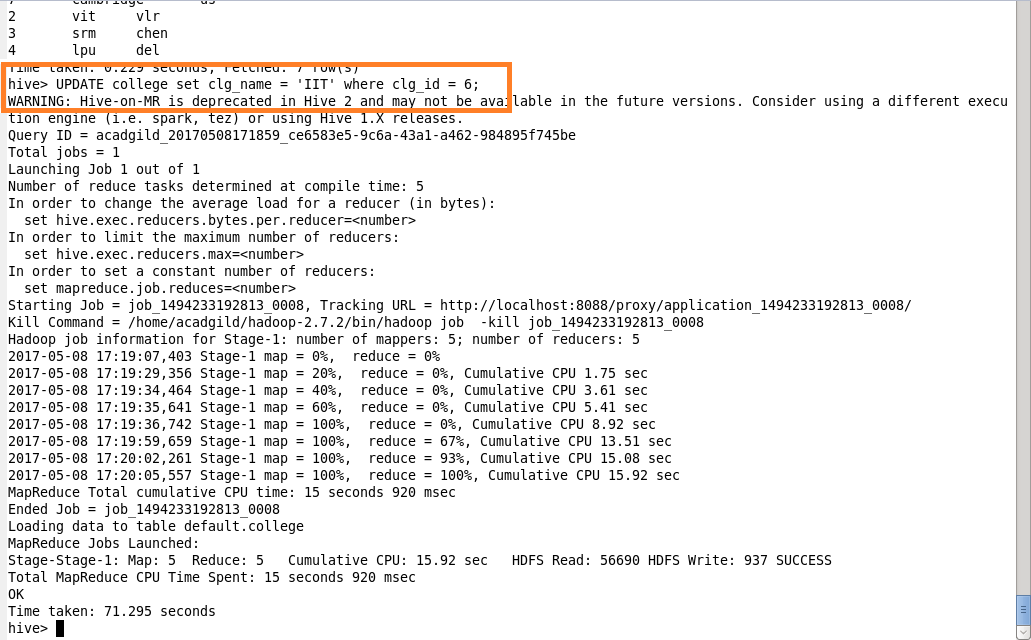
**Step 2:** Inserting values into the table created.



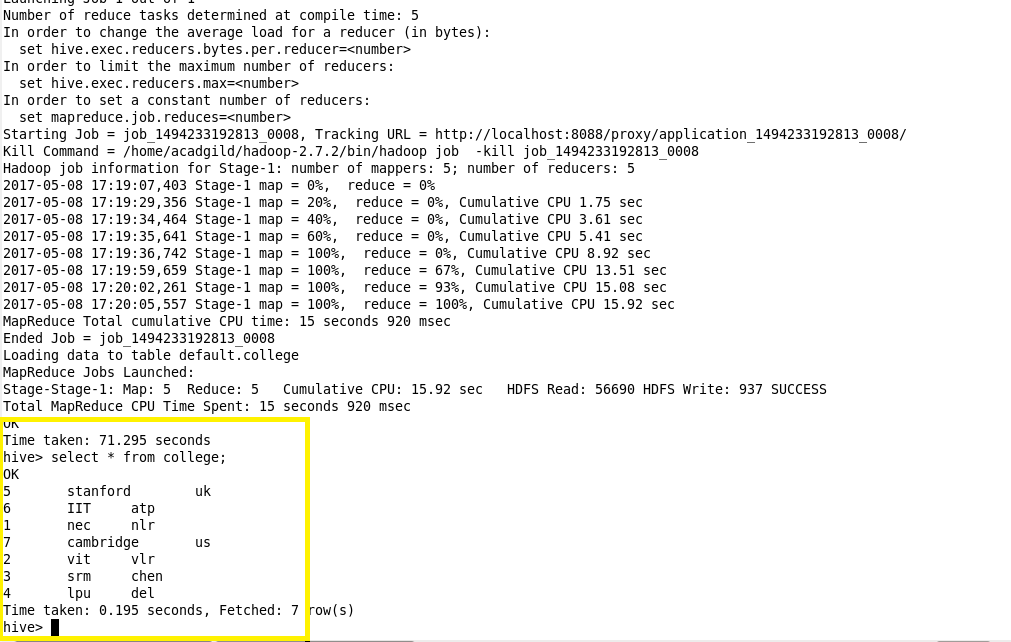
**Step 3:** Displaying the data stored in the table.



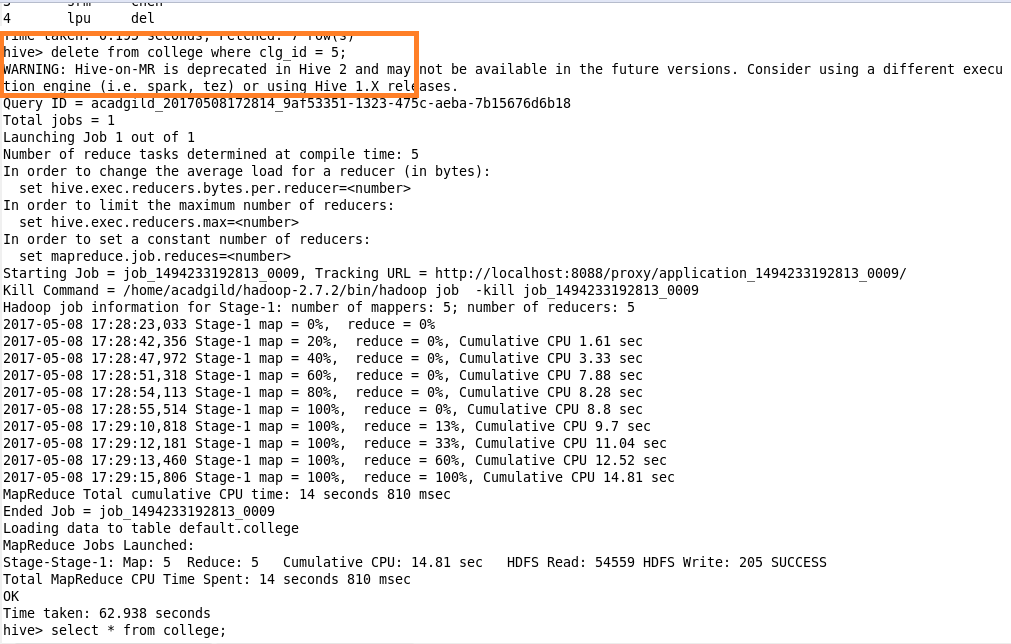
**Step 4:** Updating the data stored in the table where clg\_id is ‘6’ as clg\_name=’IIT’.



**Step 5:** Displaying the updated data from the table.



**Step 6:** Deleting an entry data from the table where clg\_id is 5.



**Step 7:** Displaying the edited data from the table.

