**Session 27**

**Assignment 4**

**Problem Statement:**

Run parallel jobs in hive by following the steps in the below blog link.

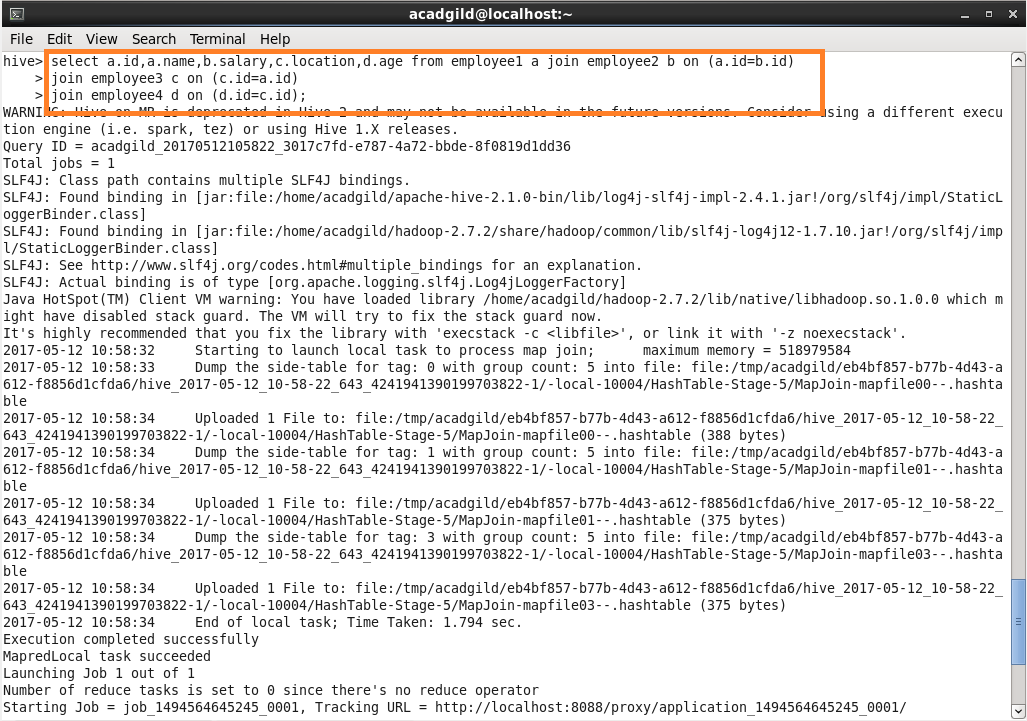
Hive can converts a query into one or more stages and to save time executes multiple jobs parallel.

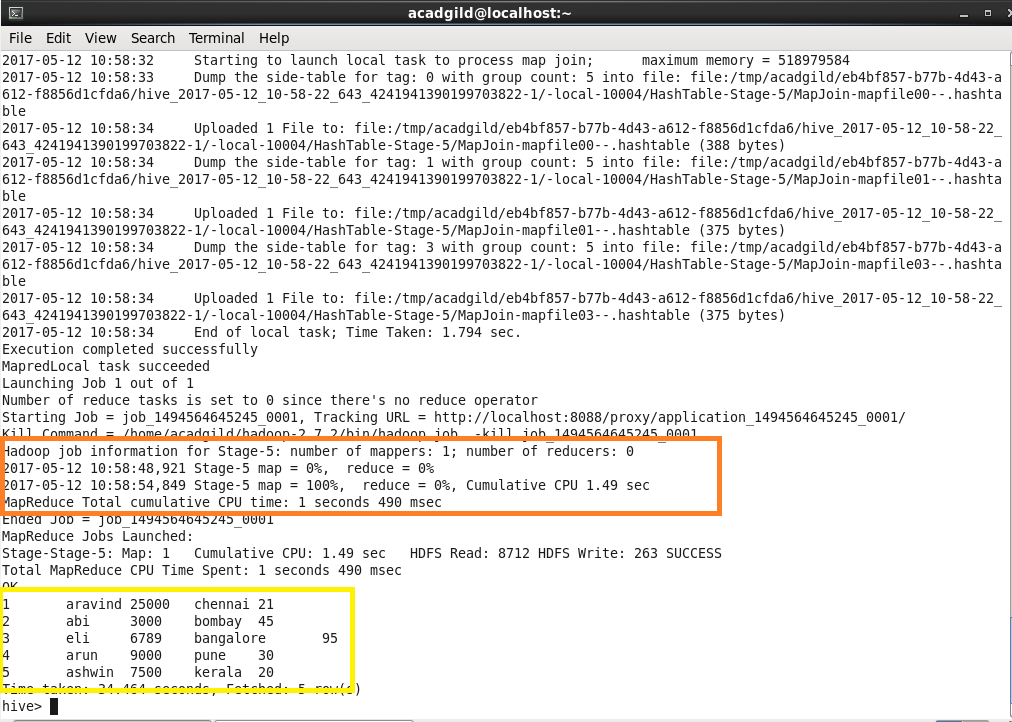
**Running Non-Parallel Jobs:**

**Step 1:** Displaying the contents of all the tables created for this assignment.



**Step 2:** Performing four joins at a time that are not parallel to each other.



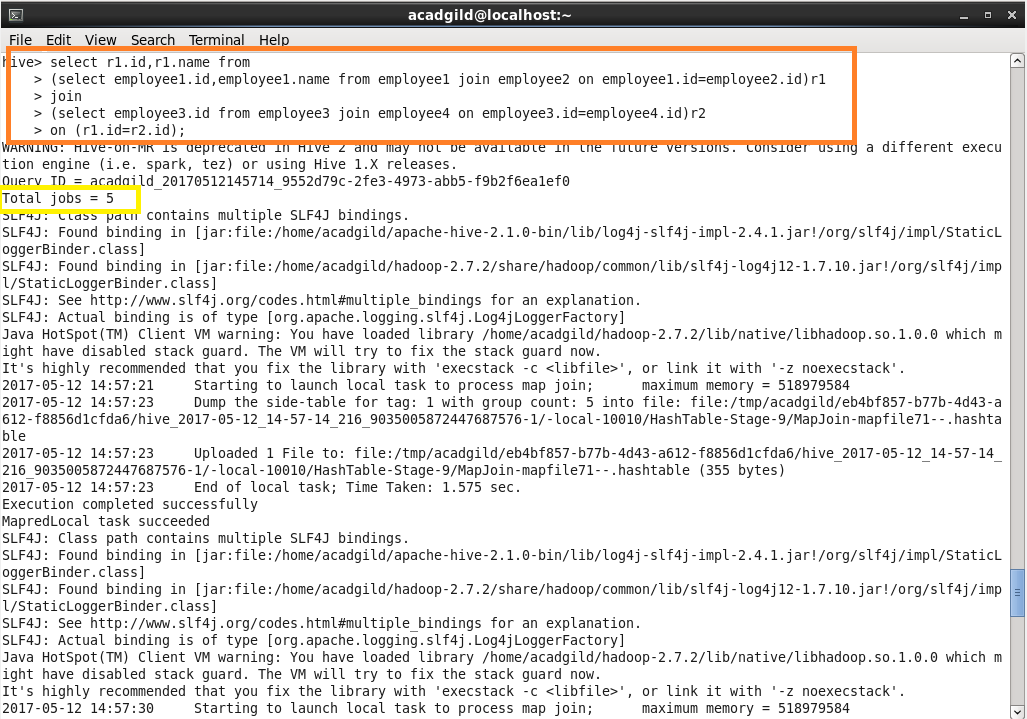


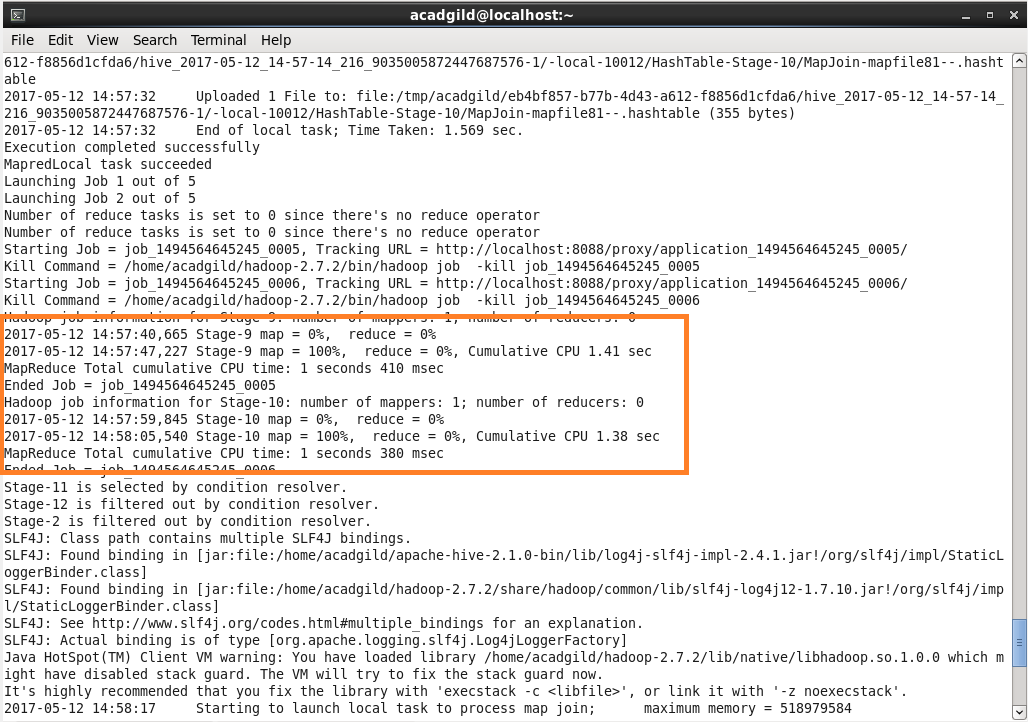
**Running Parallel Job:**

Before running parallel jobs you should set

* hive.exec.parallel=true;

**Querying:**





**Output:**

Time taken for running this four jobs is less than the actual time taken for the non-parallel job as these jobs are made parallel.

