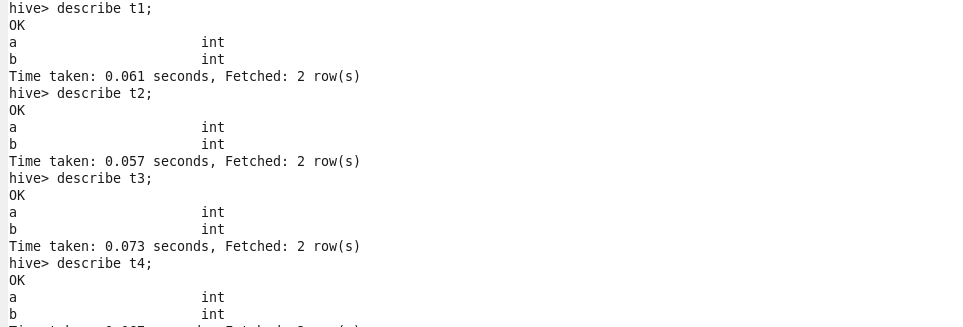
**Assignment 27.4**

**Problem Statement:**

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

<https://acadgild.com/blog/hive-parallel-jobs/>

**Input:**





**Output:**

* **Hive Parallel Jobs:**

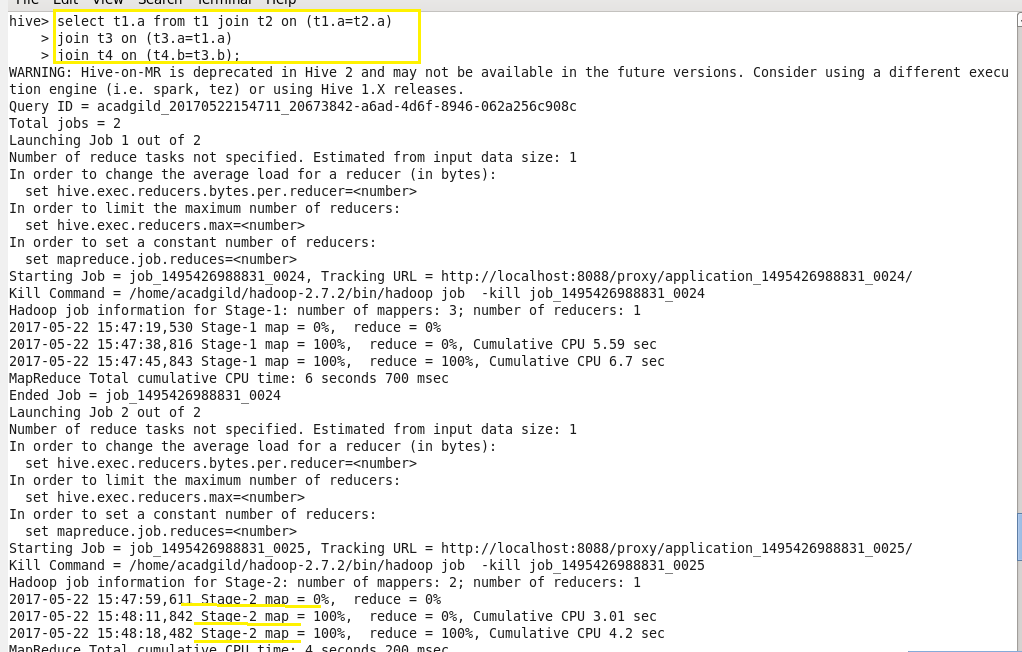
Hive sometimes can take lot of time to complete a job.

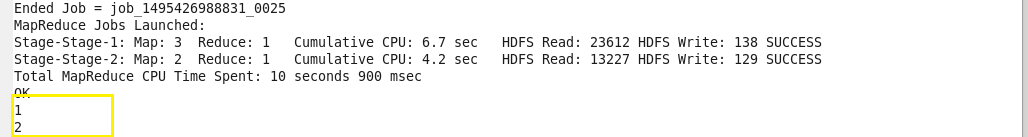
The jobs may have many different stages to get completed. By default, Hive executes these stages one at a time.

Different stages may include a Map stage, Reduce stage, a sampling stage, a merge stage, a limit stage, or other possible tasks Hive needs to do.

A particular job may consist of some stages that are not dependent on each other and could be executed in parallel, possibly allowing the overall job to complete more quickly.

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





The configuration file named hive-site.xml shall be created in hive/conf/ directory, where we need to change the properties for parallel execution and override this property.

We can do enable parallel execution of job stages by setting hive.exec.parallel to true.

