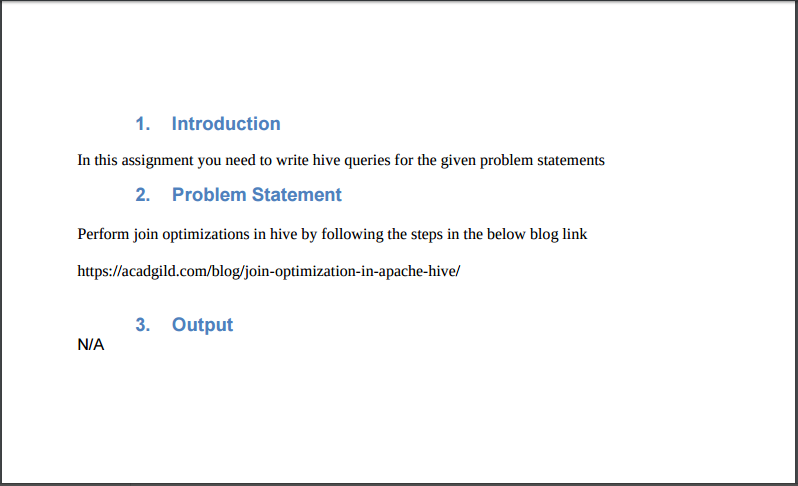
Assignment-27.5:



Join is one of the most important feature of Hive Join Optimization means the use of joins in such a manner to ensure efficient querying.

Join optimization could be done by

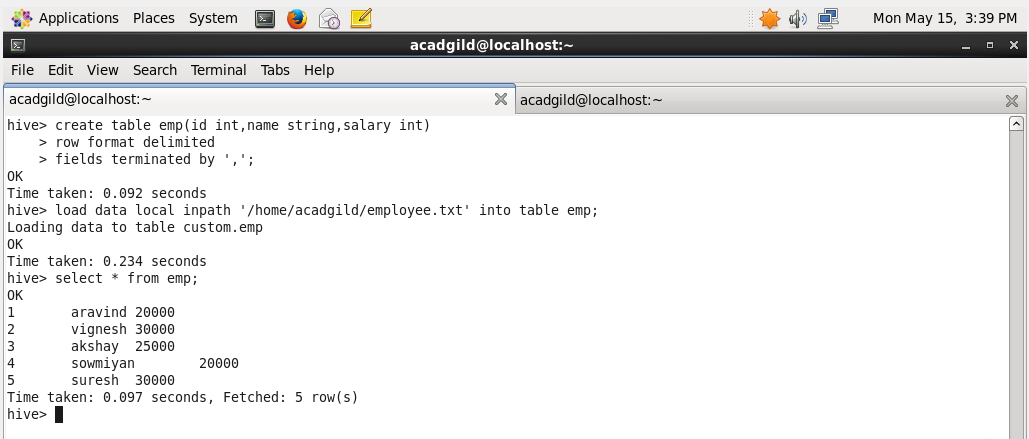
**1. Join table ordering (Largest table last).**

**2. Map Side Join**

**3. Bucket map join**

**4. Sorted Merge Bucket Map Join**

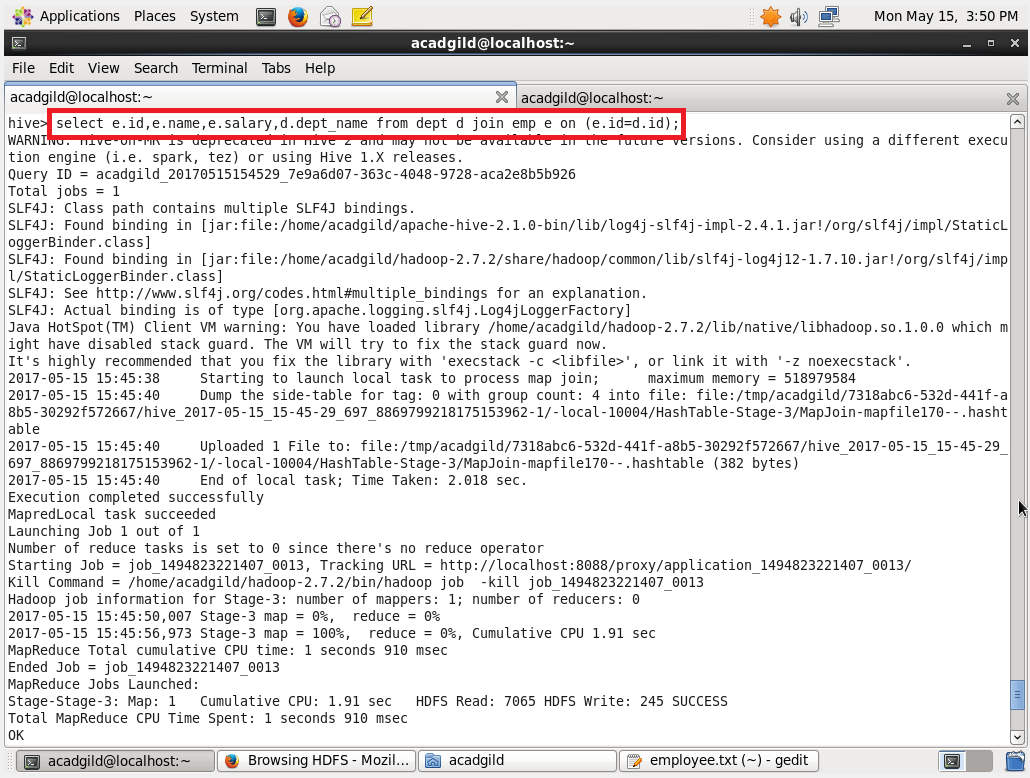
I am creating two tables named employee and department.

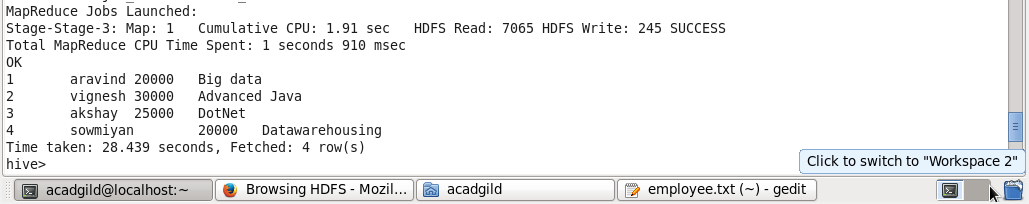




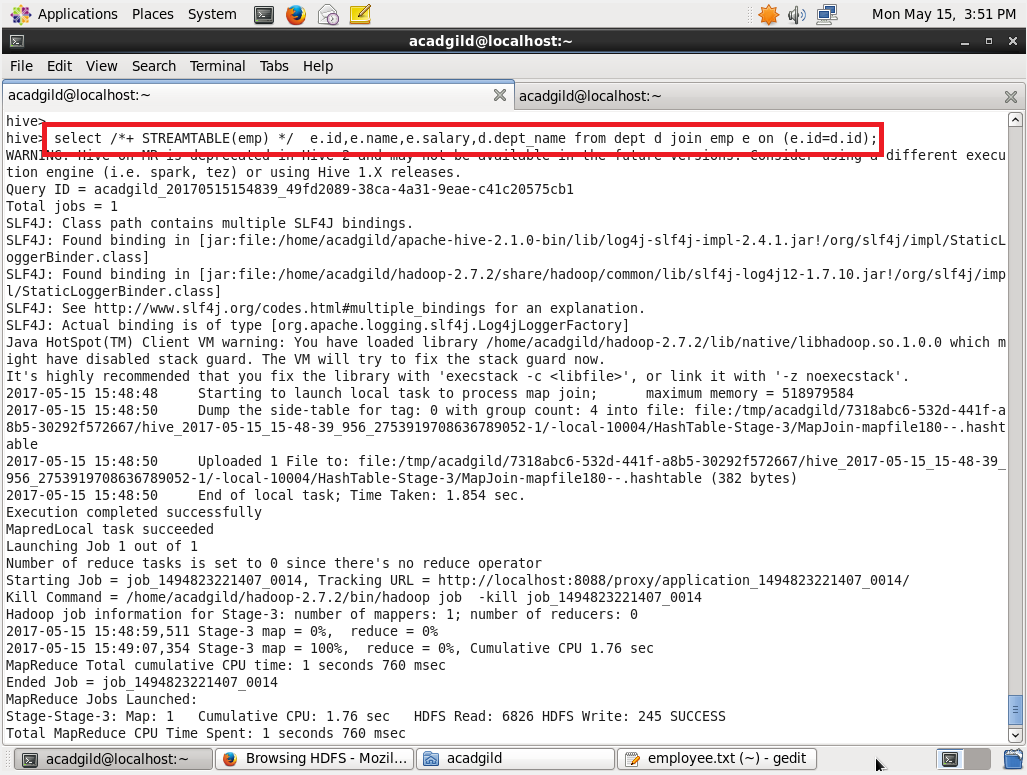
**Join table ordering (Largest table last).**

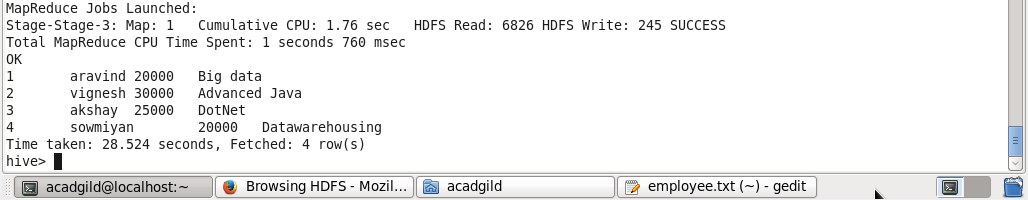
**In our case the large table is employee table so it is joined as a last table**





**Instead we can also explicitly stream the table using stream table command**



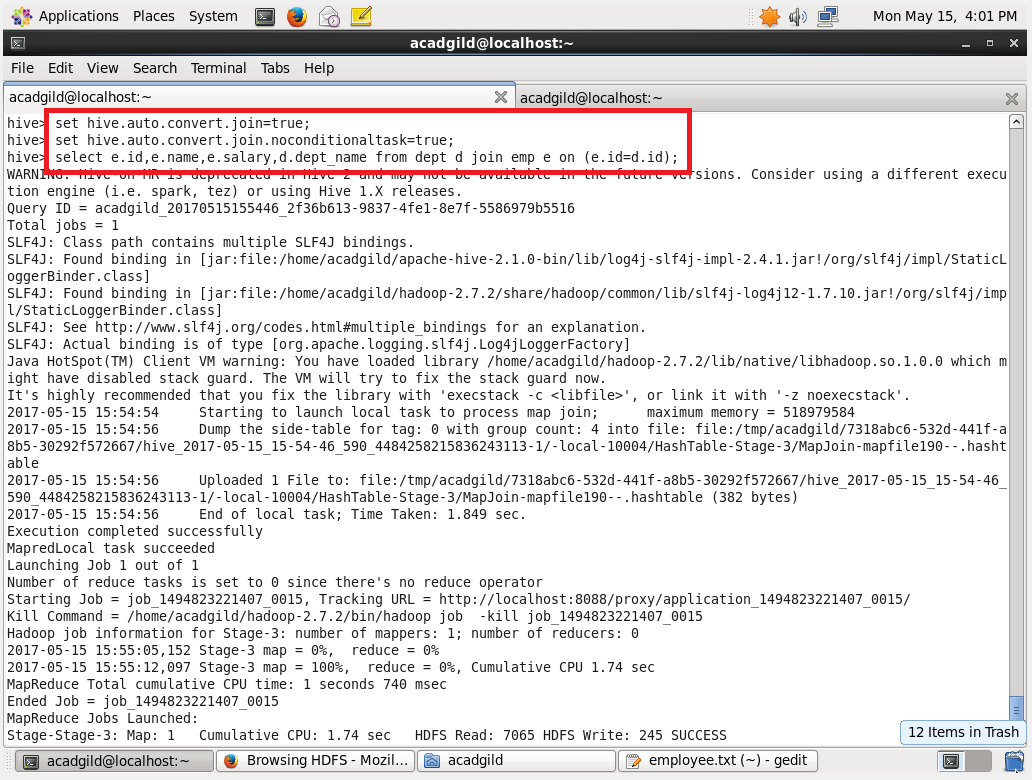


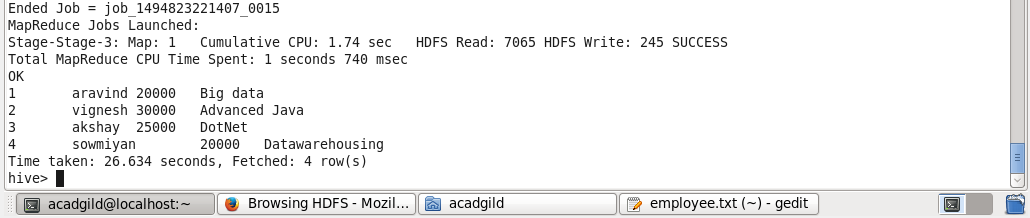
**Map Side Join:**

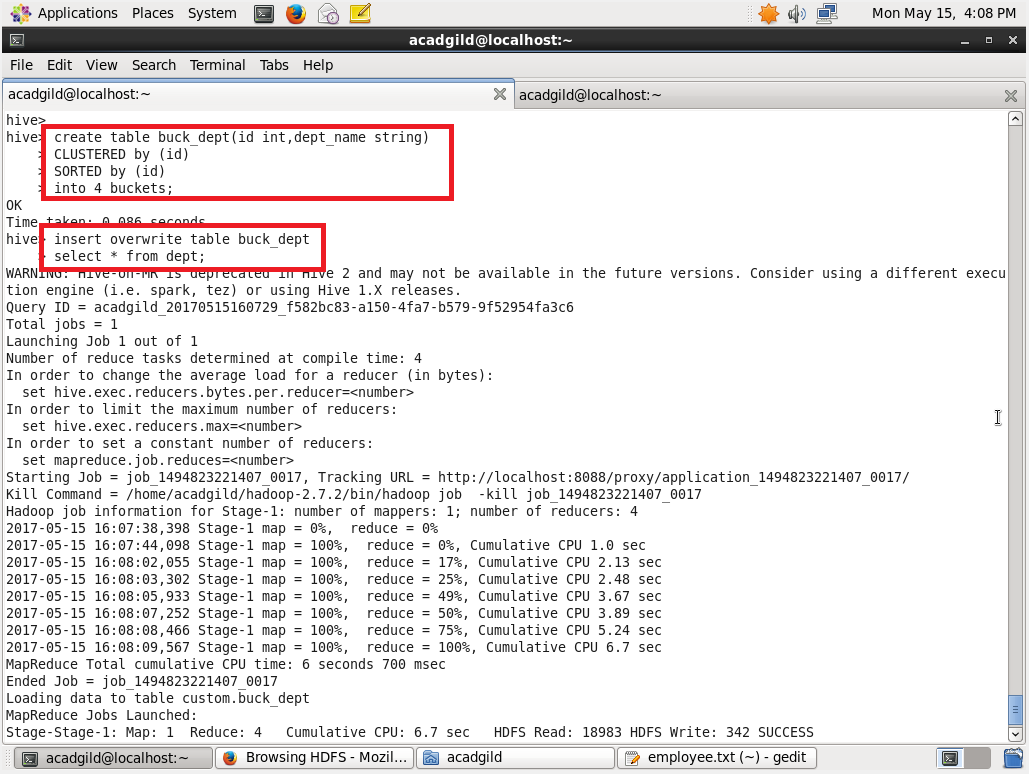
**In order to do map side join**

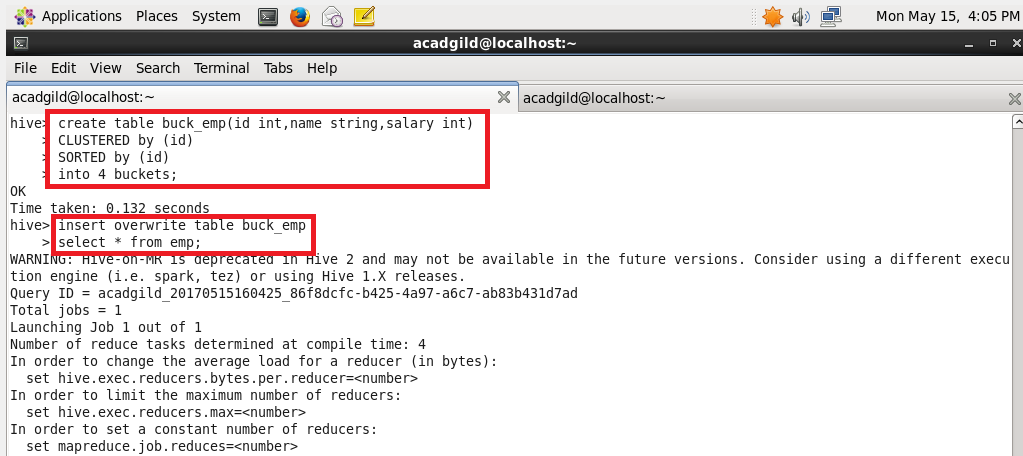
**Set hive.auto.convert=true;**

**Set hive.nonconditional.task=true;**

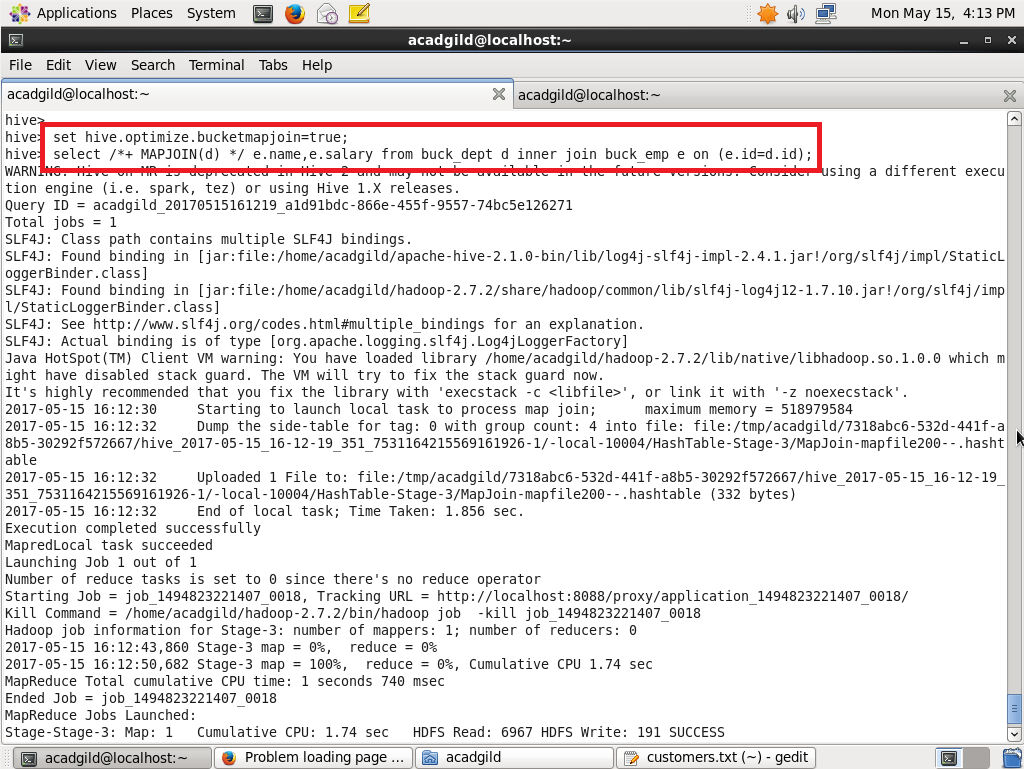


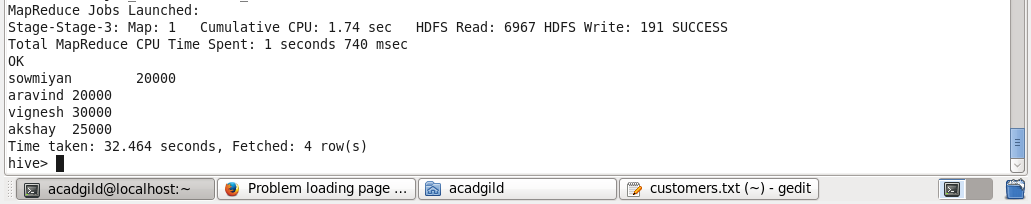


**Bucket Map Join:**

Creating the bucket table for both the employee and department. 

**Then in order to perform bucket map join set hive.optimize.bucketmapjoin=true and perform map join**





**Sort-Merge-Bucket-Map Join**

