Performance Tuning

**Hive:**

**Tez Engine:**By Enabling Tez in hive: set hive.execution.engine=tez; in hive script

**Limit operator:** Instead of scanning all data base it will scan only limited DB

**Parallel execution:** we can enable its property in hive-site.xml

**ORC File:** Convert A table into A\_ORC & B into B\_ORC (**speed improvements**)

**Vectorization:** To improves performance of operations like scans, aggregations, filters and joins, by performing them in batches of 1024 rows at once instead of single row each time.

set hive.vectorized.execution.enabled = true;  
 set hive.vectorized.execution.reduce.enabled = true;

**Hadoop:** Swapimess, block size, heapsize, replication,