
2

1

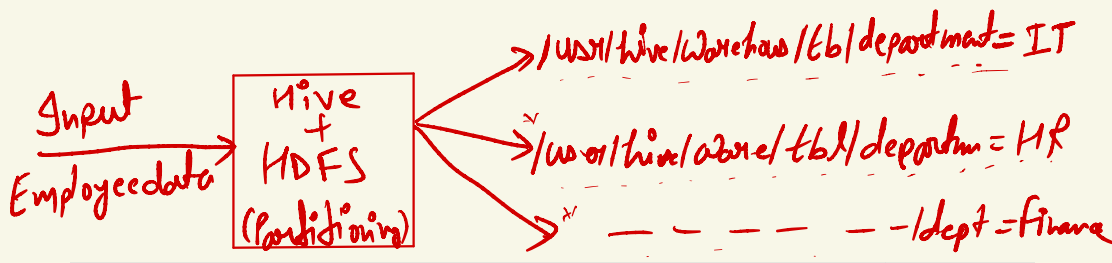


Partition in Hive

* Partitioning is a way of dividing a table into related parts based on the values of a particular column.

Employee (empname, salary, department)

Normal table data in HDFS => /usr/hive/warehouse/table-name/



Partition Column \rightarrow {Department}

\rightarrow IT
 \rightarrow HR
 \rightarrow FINANCE

Ex:

= Select * from Employee where dept='IT';

Types of Partitioning -

→ Static Partitioning

→ Dynamic Partitioning

* Static Partitioning:

→ Here value of partition column will be known to us.

→ We load the data in a specific Partition

→ Static Partitioning takes less time while loading the data.

* Dynamic Partitioning

→ Here column values will be unknown to us

→ Hive will identify the unique values from partitioned column.

→ It takes longer time while loading the data.