Hive

Hadoop Cluster

Hive

→ Warehouse
→ Metastore

HDFS

→ Warehouse
/ hive

Sharding

Sharding

MySQL

→ memory
→ cpu

Metastore

Transactional

Externally

MySQL

HDFS

/user/ hive/ warehouse

table + data

Internal Table data

Internal Table

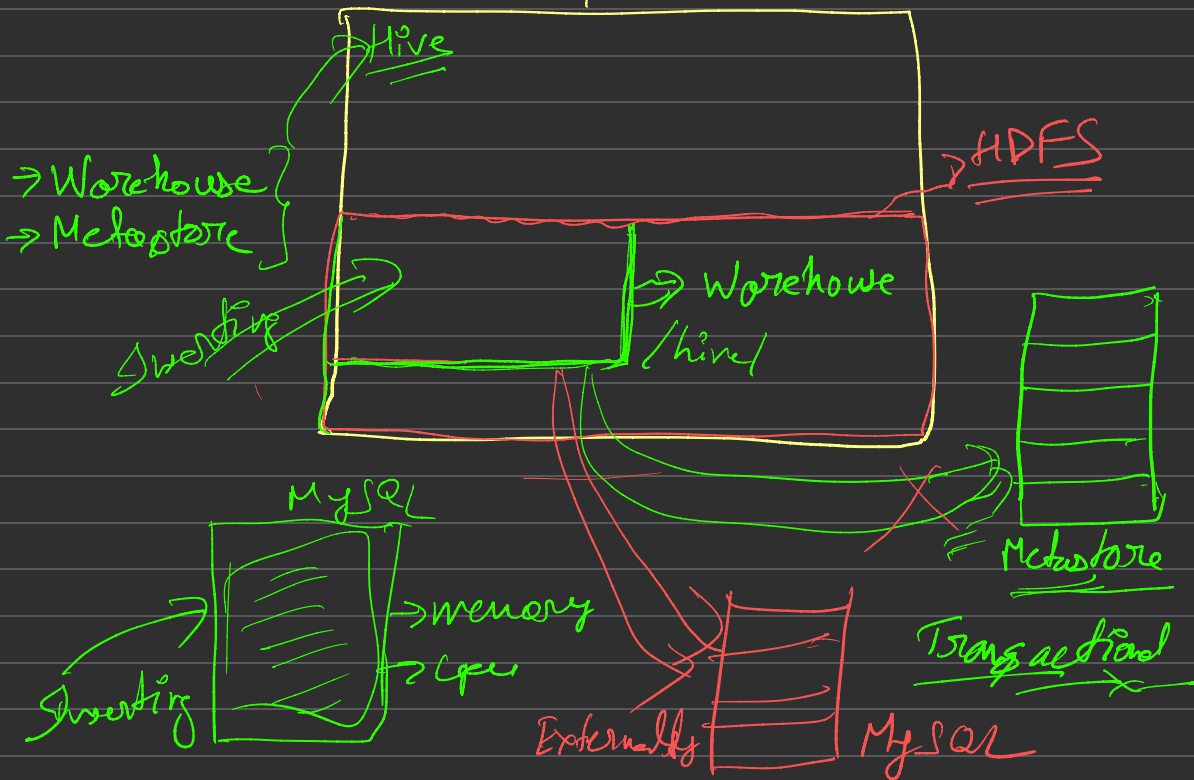external table

Raw data

client

Metastore

external data

table + metadata

20-Dec

/tmp/ input-dat / deport_data_1.csv (100)
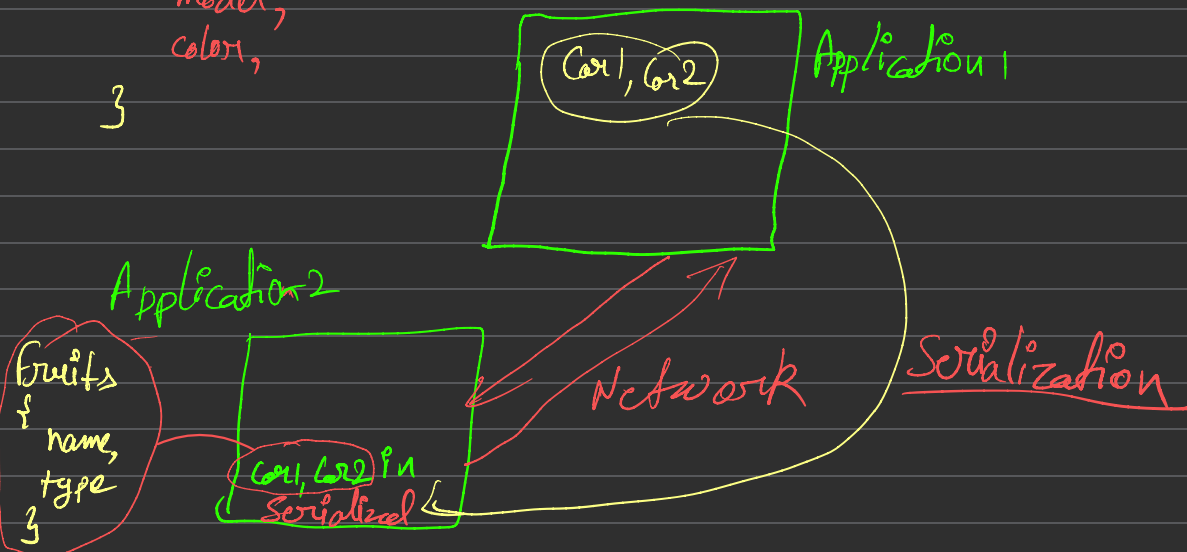
21-Dec → /tmp/ input-data / de port_data-2.csv (200)

19-Dec you created external table

20-De when i run the query → How many record?

100 record

21-Dec when i run query → How many record?

300 record

Car
{
   name,
   model,
   color,
}

Car1 = Car ('Greta', 'Hyundai', 'White')
Car2 = Car ('XUV700', 'Mahendra', 'Silver')

Car1, Car2   Application 1

Application 2

fruits
{
   name,
   type
}

Car1, Car2 in
Serialized

Network   Serialization

Table1
operates
on
CSV

Table 2
operates
on
ORC
data

New-data

emp-name
country
designation
time in compgy

Insert into Table_2 as { Select * from table_1 where country = 'INDIA';

name , Address

Amit , Lucknow

Rahul , "x+, Lucknow, U.P"

Nitin , " Abc, "Lucknow", U.P "