**Name - Umang Patel**

**Roll No. - 2193271**

**Assignment 6: Hive (basic commands)**

**Aim:** Implement various operations on Hive (Create, Insert, Update)

**Objective:** Study, Understand and Implement various operations on data using Hive.

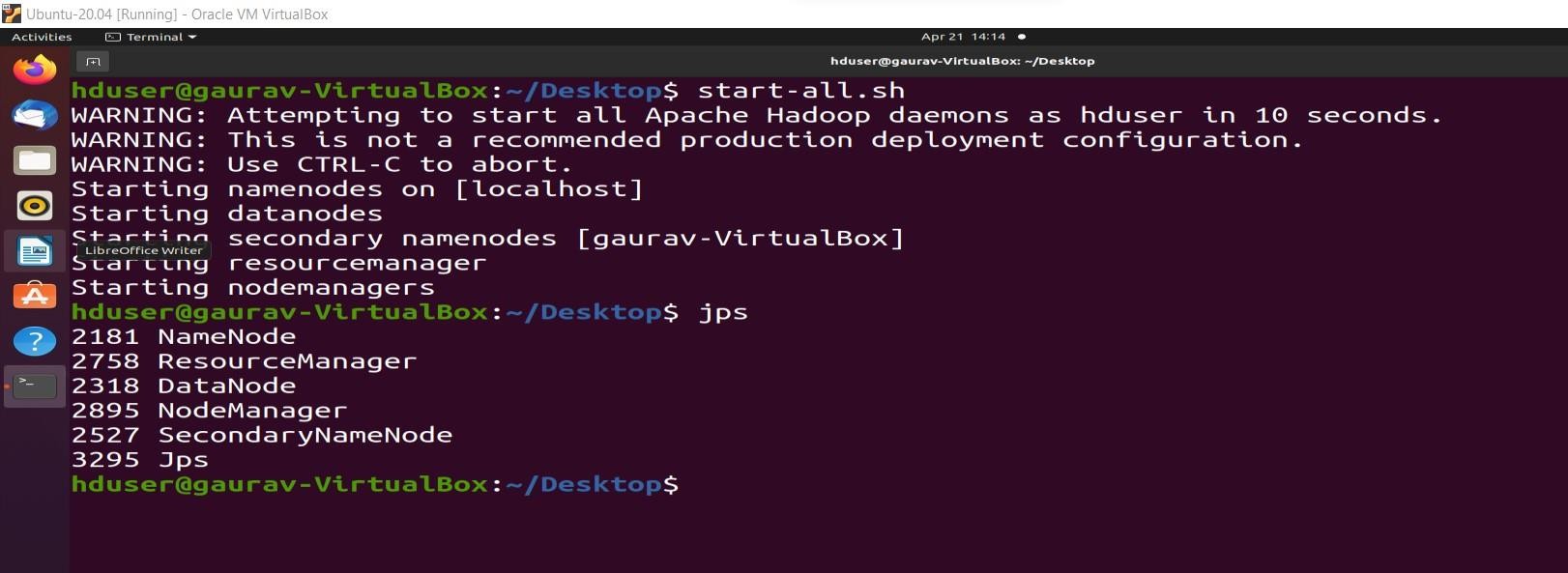
# Theory:

**Hive:**

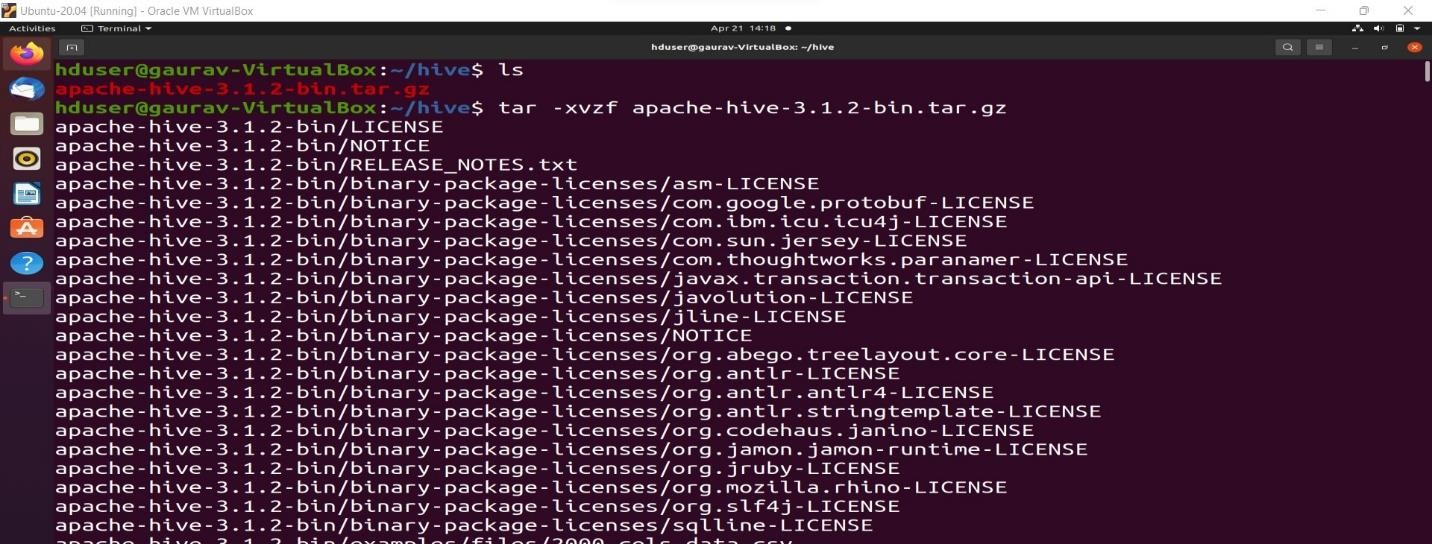
Apache Hive is a data warehouse and an ETL tool which provides an SQL-like interface between the user and the Hadoop distributed file system (HDFS) which integrates Hadoop. It is built on top of Hadoop. It is a software project that provides data query and analysis. It facilitates reading, writing and handling wide datasets that stored in distributed storage and queried by Structure Query Language (SQL) syntax. It is not built for Online Transactional Processing (OLTP) workloads. It is frequently used for data warehousing tasks like data encapsulation, Ad- hoc Queries, and analysis of huge datasets. It is designed to enhance scalability, extensibility, performance, fault-tolerance and loose-coupling with its input formats.

# Implementation:

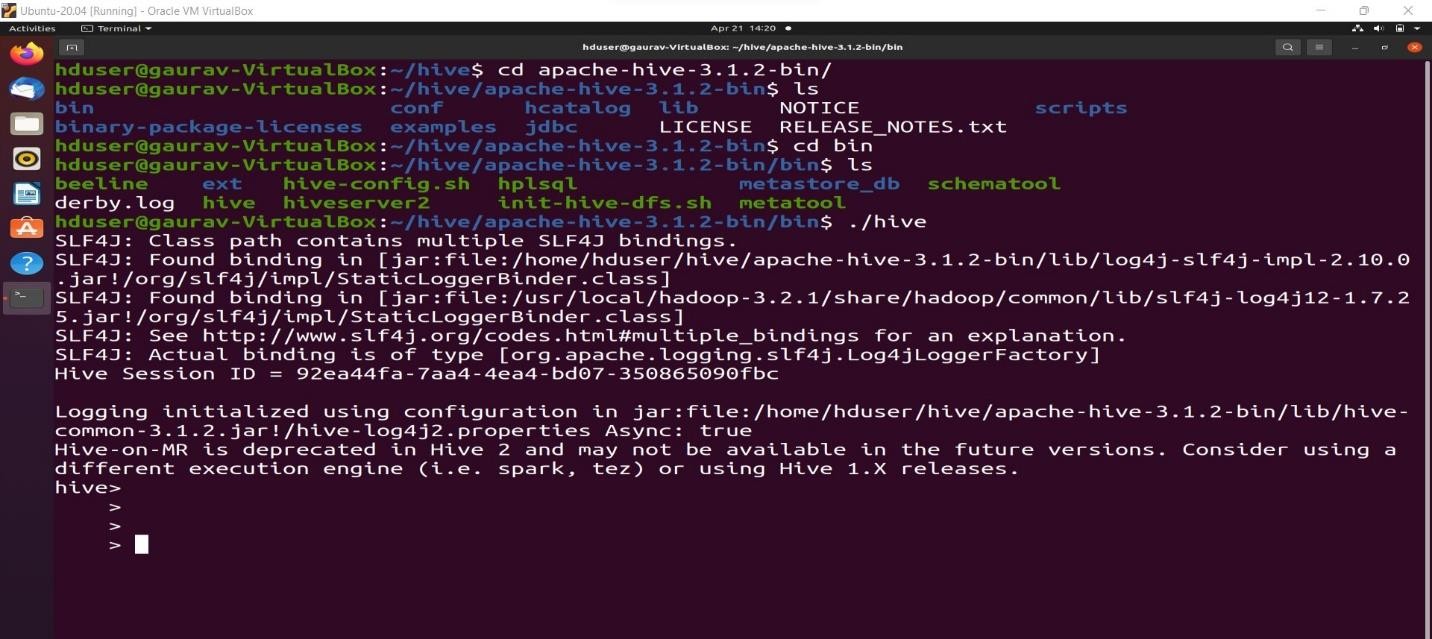
1. At the most we need to start the cluster using the start-all.sh script further check all daemons are running using the jsp command



1. Once we have the cluster running we also need to download the tar.gz file which we will be using to install the hive in our system. To install the hive we will be just extracting the data in the file tar -xzfs <tar file name>



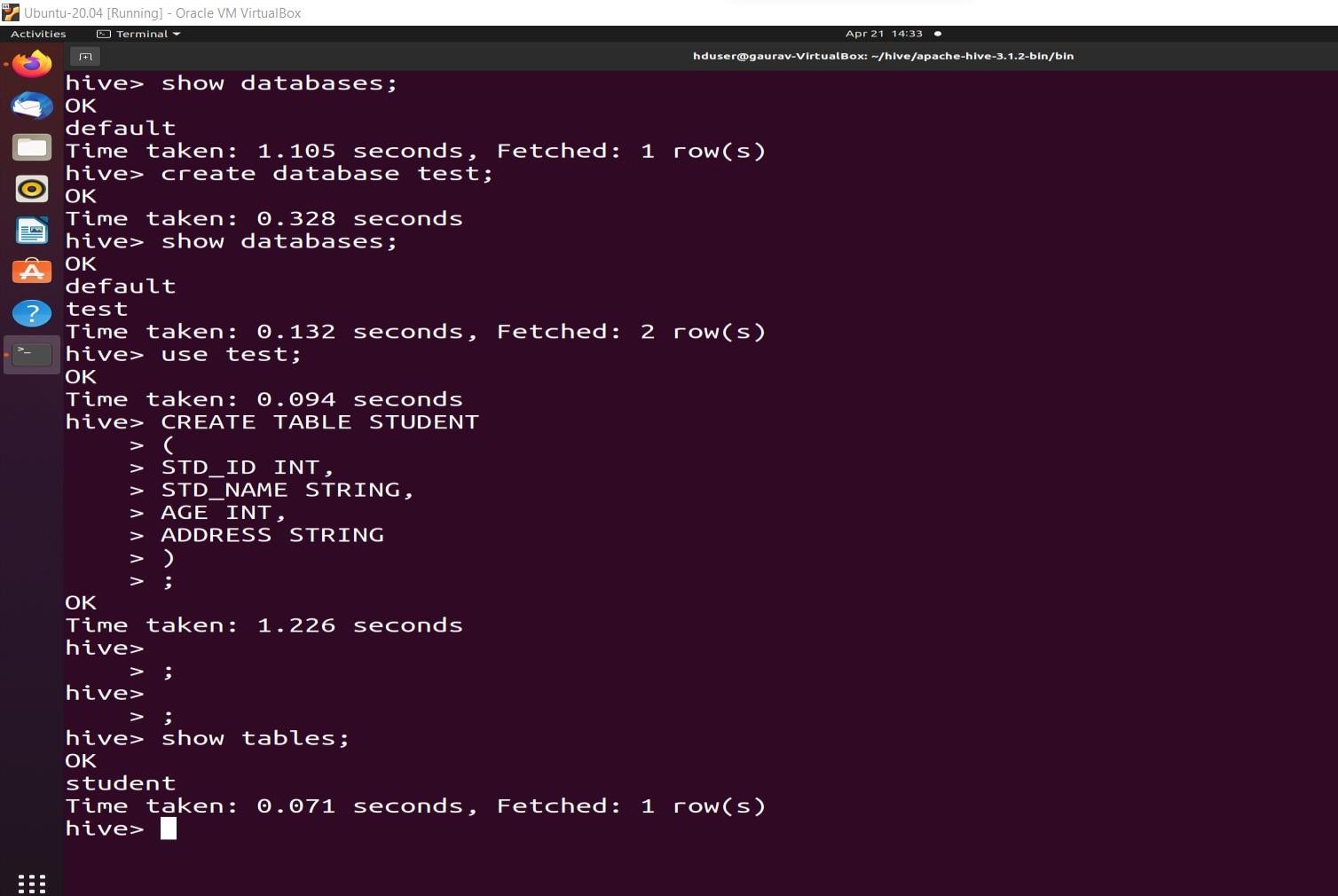
1. Now folder will be visible for Apache-hive in which we have the bin where the executable for the hive is present which will be using to get the hive shell



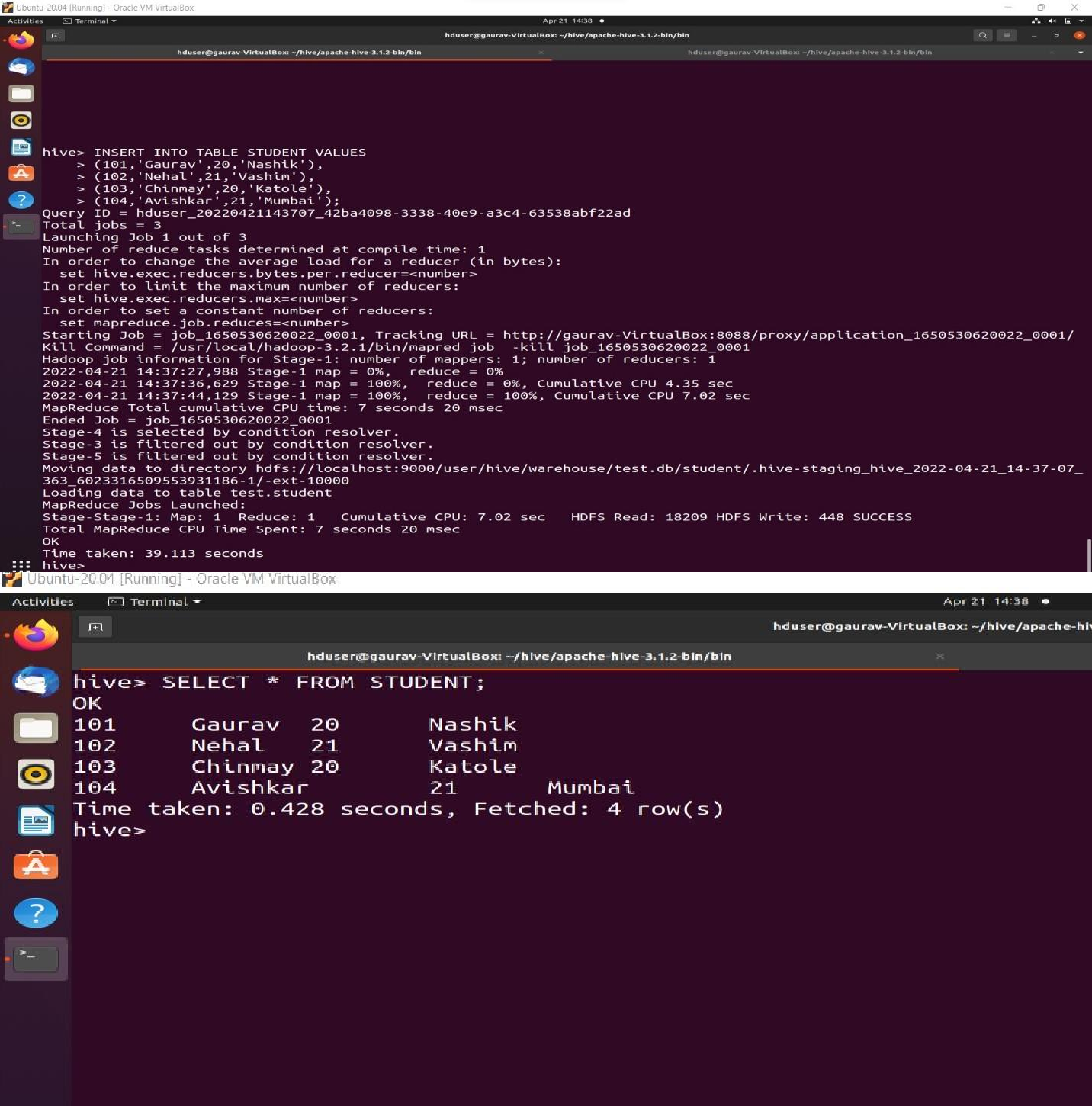
1. Once we get the hive shell ready we will be using the HQL command first we will be creating our database using the use command which will switch us to the new database and if not existed it will create and then switch.

# $create database <database\_name>

**$use <database\_name>**



1. We can ask to show the tables using the $show tables command further we can also create the table using the $create table <table\_name> (<column\_name> <type of value>, …) query.
2. We can now insert necessary values in the table using the insert values command.



**Conclusion/ outcome:** Thus we have successfully performed and understood the create, insert, . operations using the HQL in Apache hive