7/1/2025

Data Storage Solutions

Dr. Muhammad Iqbal

Leandro G. Damasceno ID 2019299

2019299@student.cct.ie

<https://github.com/LeandroGD2017/RepeatCA-Muhammad>

Contents

[1. 1](#_Toc203328802)

[2. 2](#_Toc203328803)

[3. 3](#_Toc203328804)

[4. 4](#_Toc203328805)

[5. 4](#_Toc203328806)

[6. 5](#_Toc203328807)

[7. 5](#_Toc203328808)

[8. 6](#_Toc203328809)

[9. 6](#_Toc203328810)

[10. 7](#_Toc203328811)

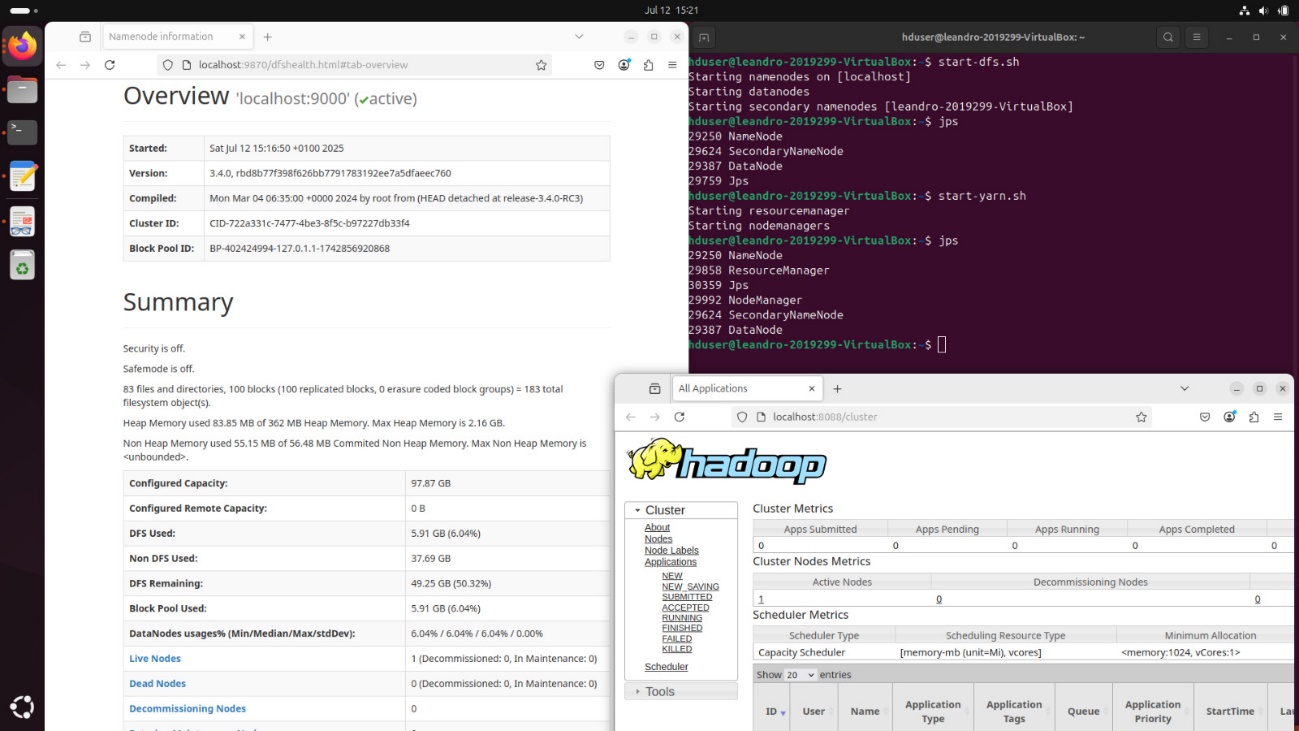
[11. 7](#_Toc203328812)

[12. 8](#_Toc203328813)

[Reference 9](#_Toc203328814)

# 1.

Hadoop Installed



# 2.

hadoop jar /usr/local/hadoop-3.4.0/share/hadoop/toos/lib/hadoop-streaming-3.4.0.jar \

-input /user/hduser/ev\_project/Electric-Vehicle.csv \

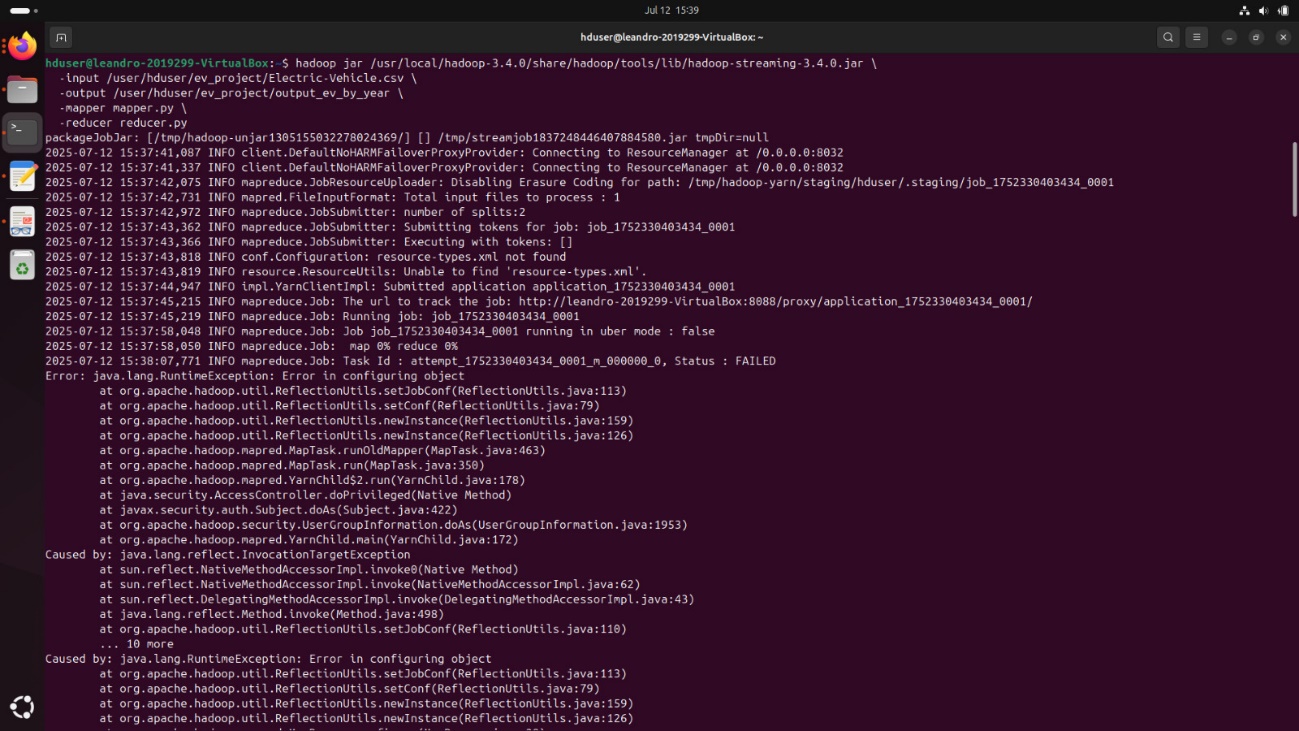
-output /user/hduser/ev\_project/output\_ev\_by\_year \

-mapper mapper.py \

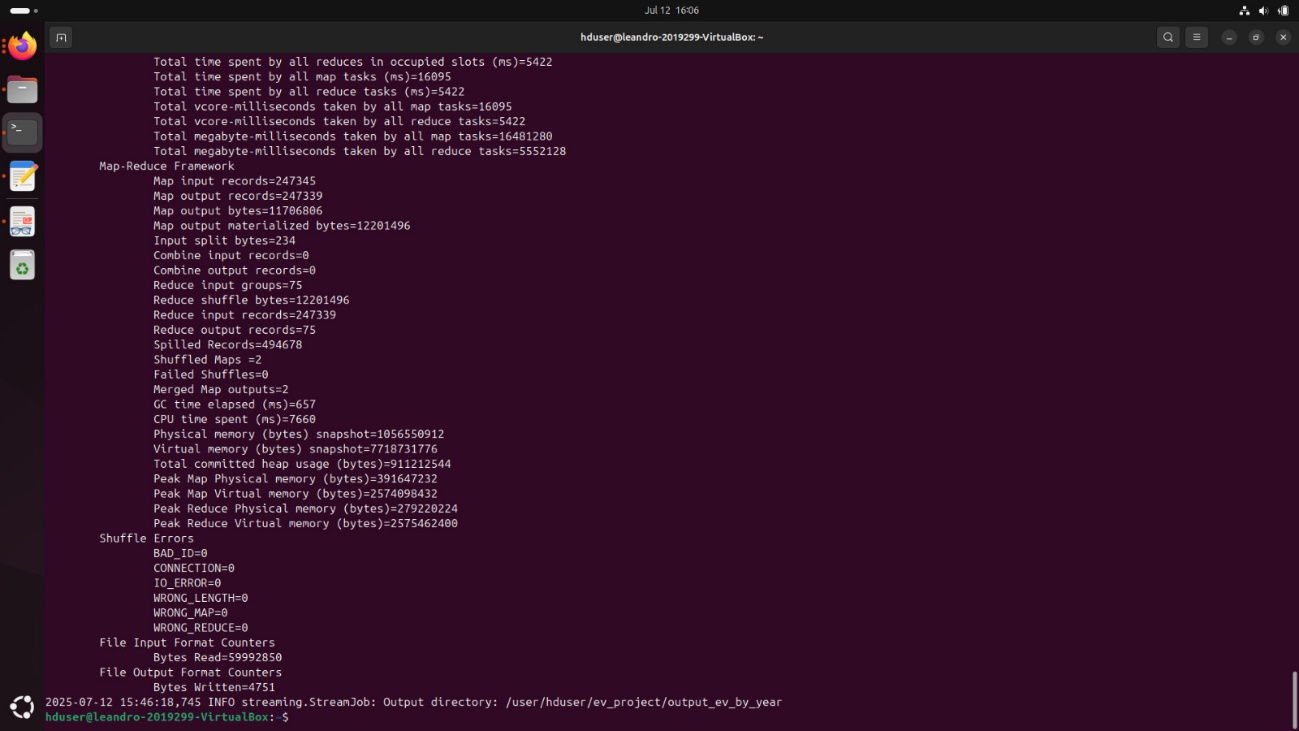
-reducer reducer.py \

-file mapper.py \

-file reducer.py

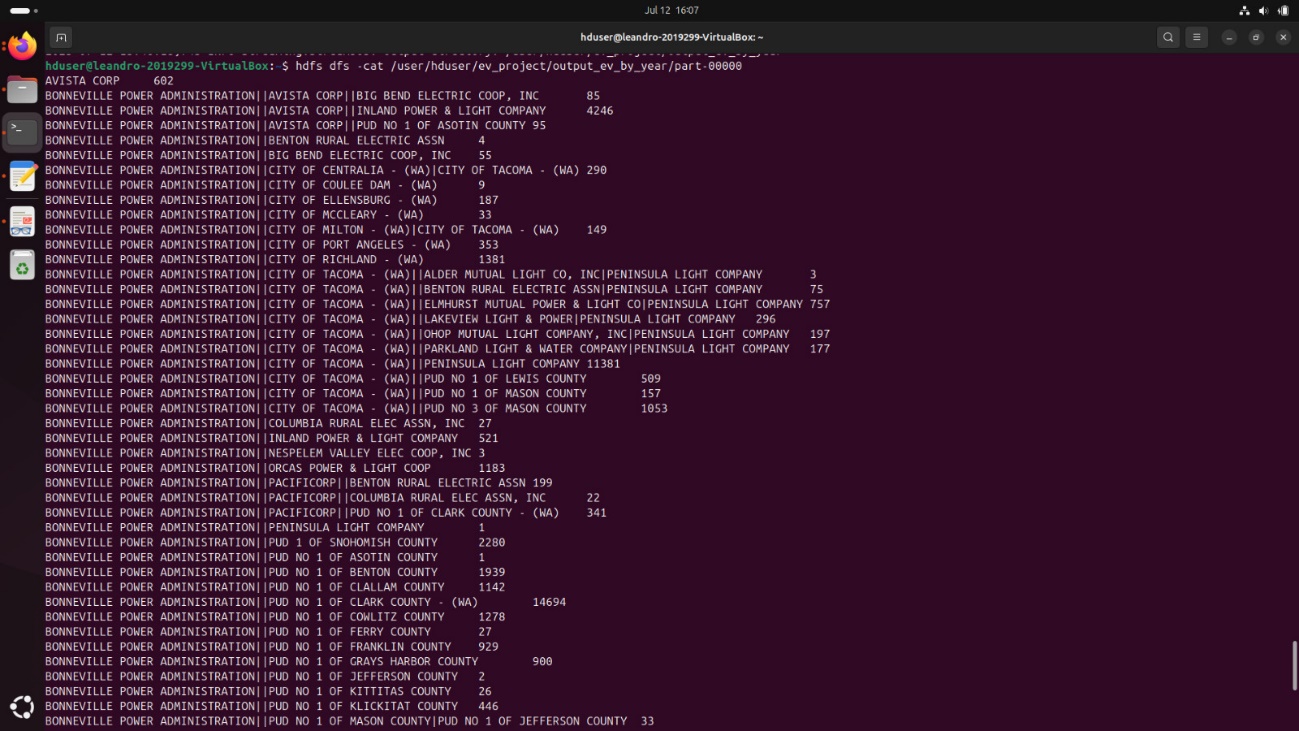


# 3.



# 4.

hdfs dfs -cat /user/hduser/ev\_project/output\_ev\_by\_year/part-00000



# 5.

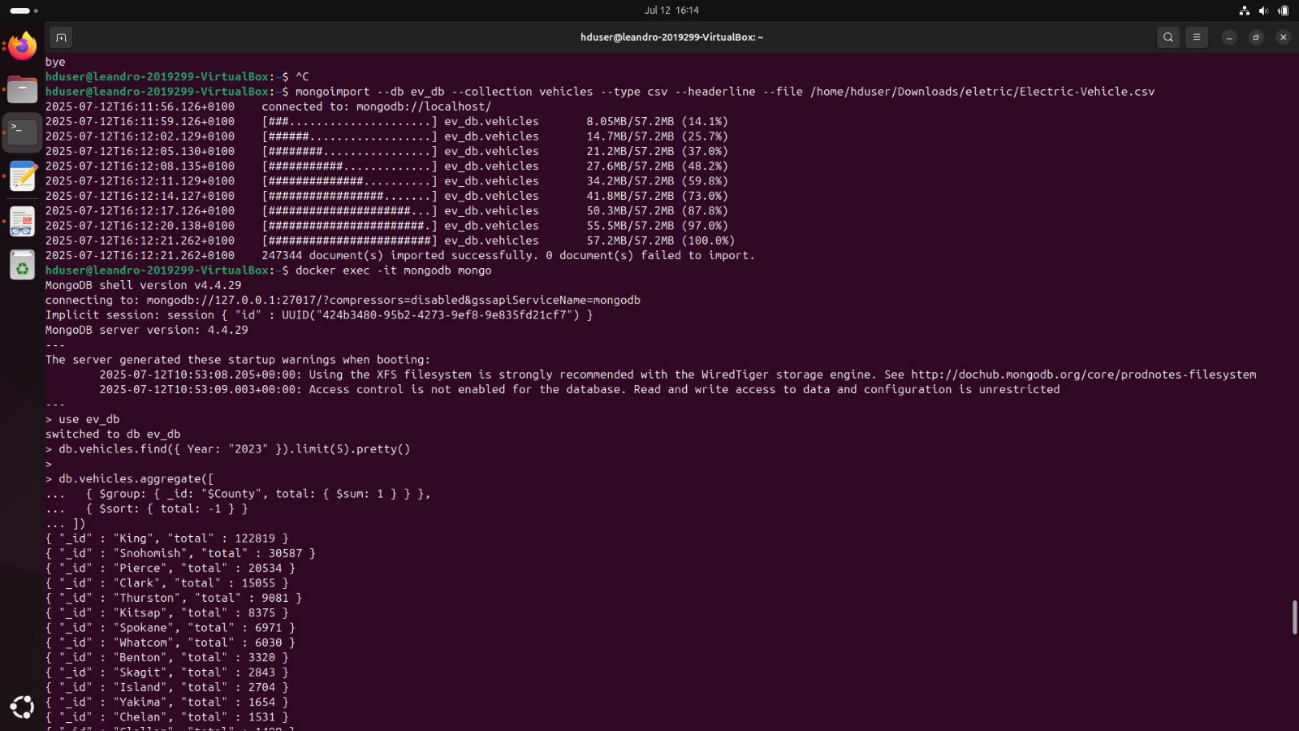


I had, a lot of problems with my Ubuntu version 24.4.2 and Mongodb refused worked with it using VM, I had to use docker.

# 6.

mongoimport --db ev\_db --collection vehicles --type csv --headerline --file

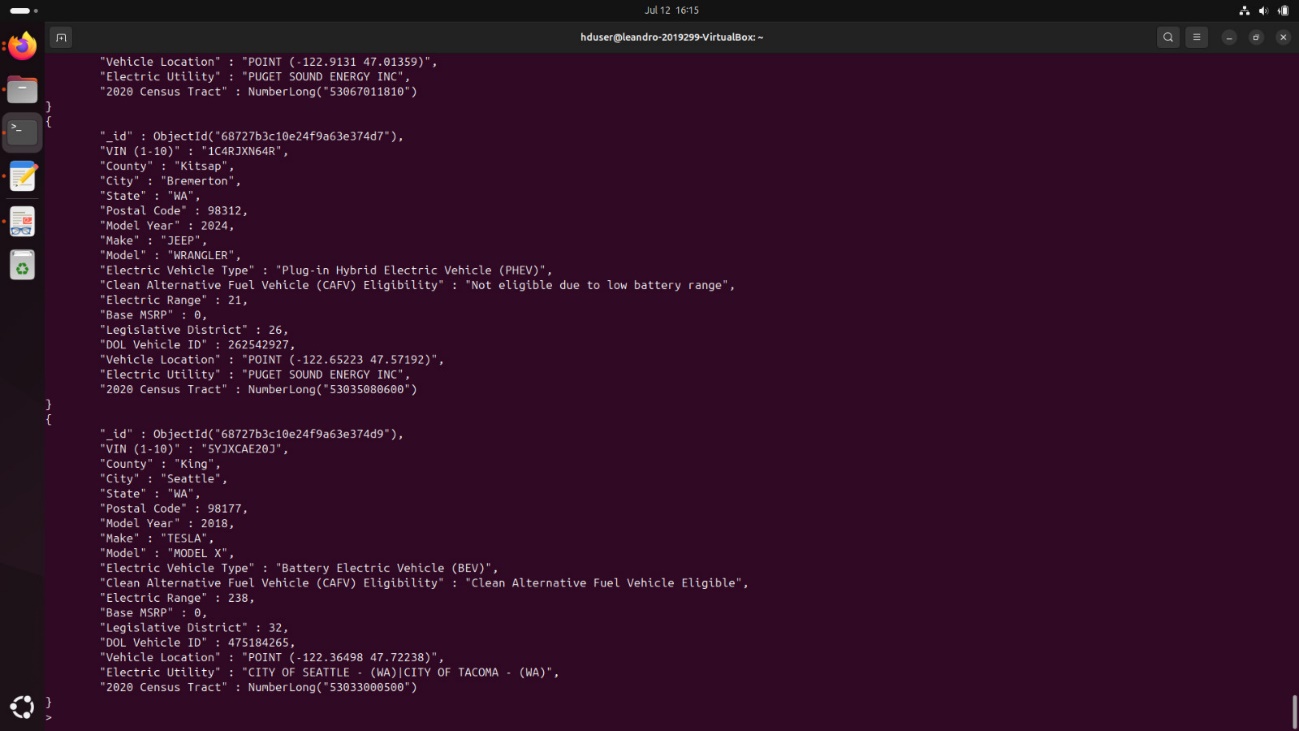
/home/hduser/Downloads/eletric/Electric-Vehicles.csv



Docker exec -it mongodb mongo

Use ev\_db

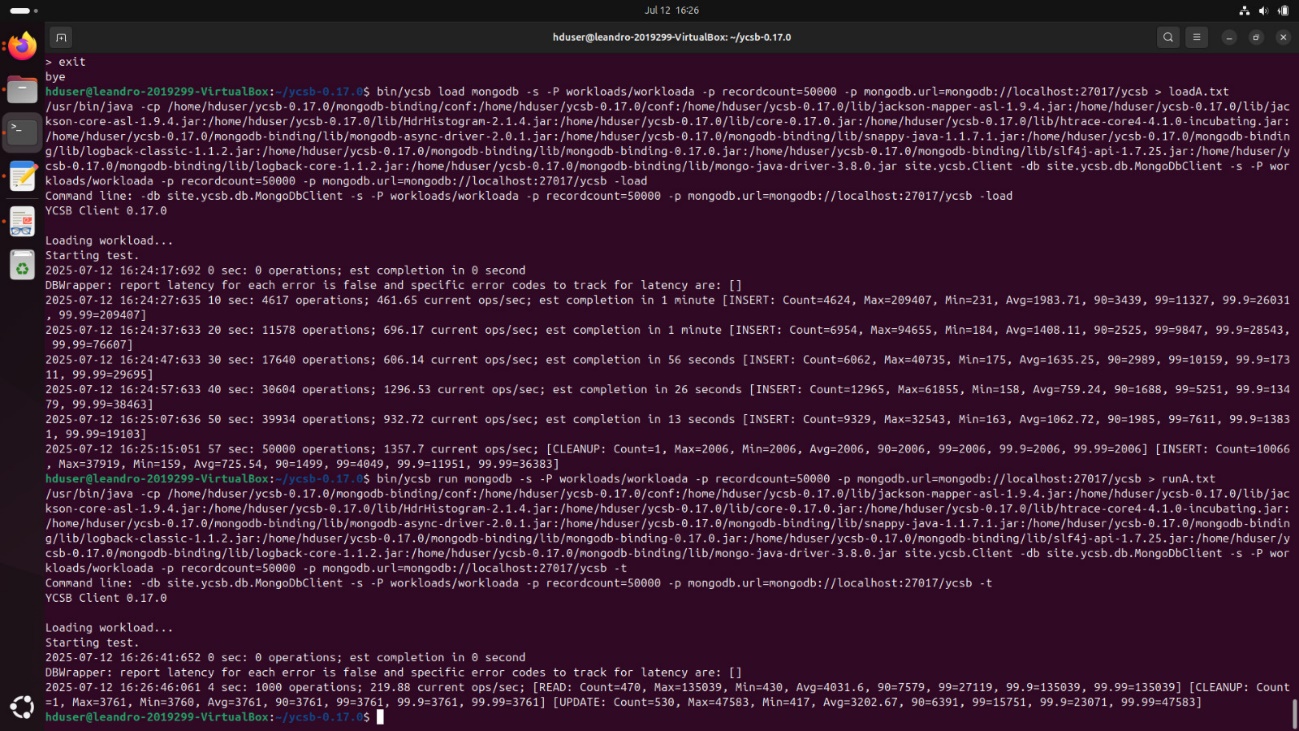
# 7.

db.vehicles.find({ Year: “2023” }).limit(5).pretty()

# 8.

bin/ycsb load mongodb -s -P workloads/workloada -p recordcount=50000 -p

mongodb.url=mongodb://localhost:2707/ycsb > loadA.txt

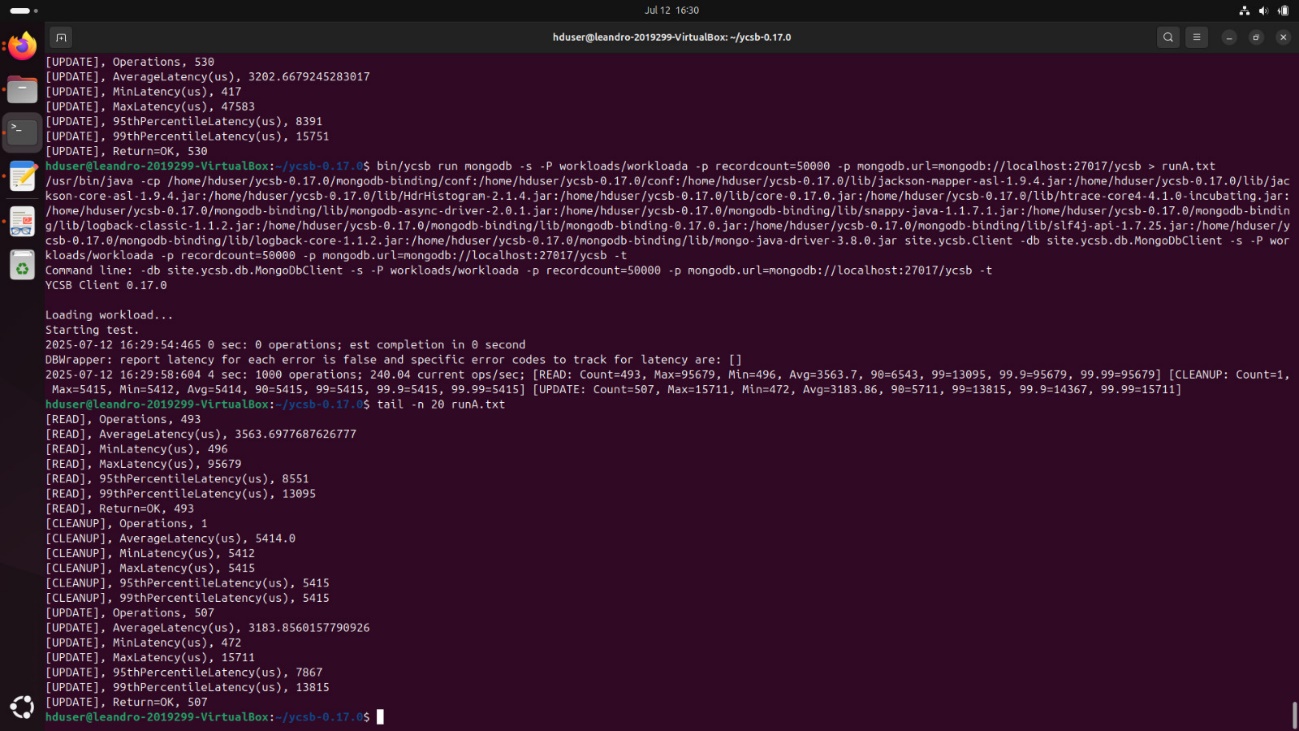


# 9.

bin/ycsb run mongodb -s -P workloads/workloada -p recordcount=50000 -p

mongodb.url=mongodb://localhost:2707/ycsb > runA.txt

tail -n 20 runA.txt

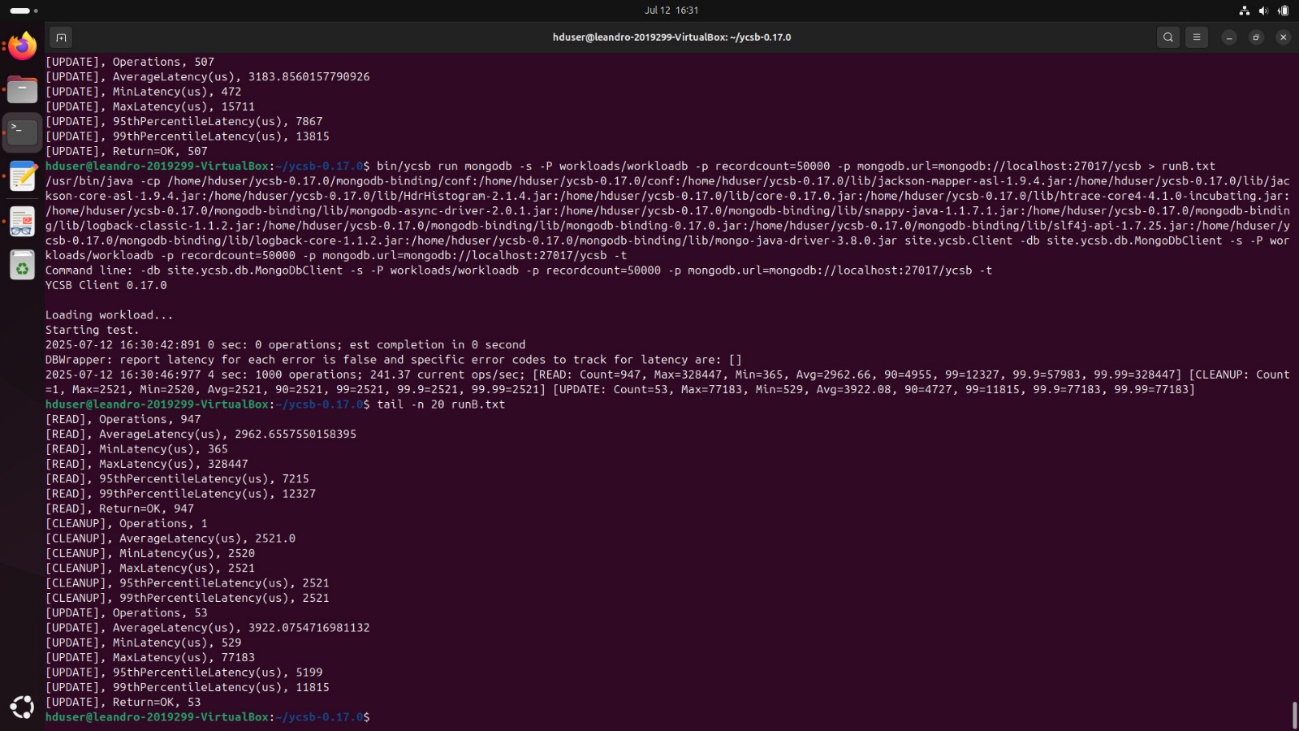


# 10.

bin/ycsb run mongodb -s -P workloads/workloadb -p recordcount=50000 -p

mongodb.url=mongodb://localhost:2707/ycsb > runB.txt

tail -n 20 runB.txt

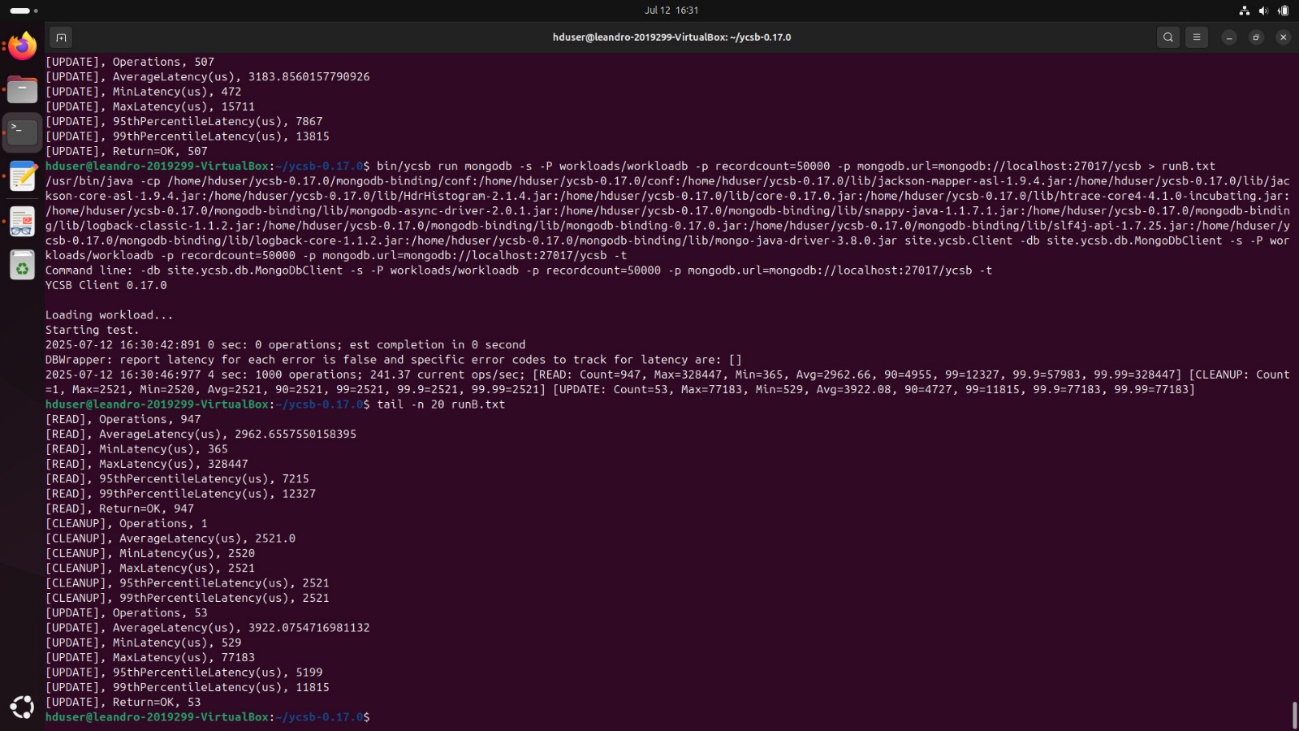


# 11.

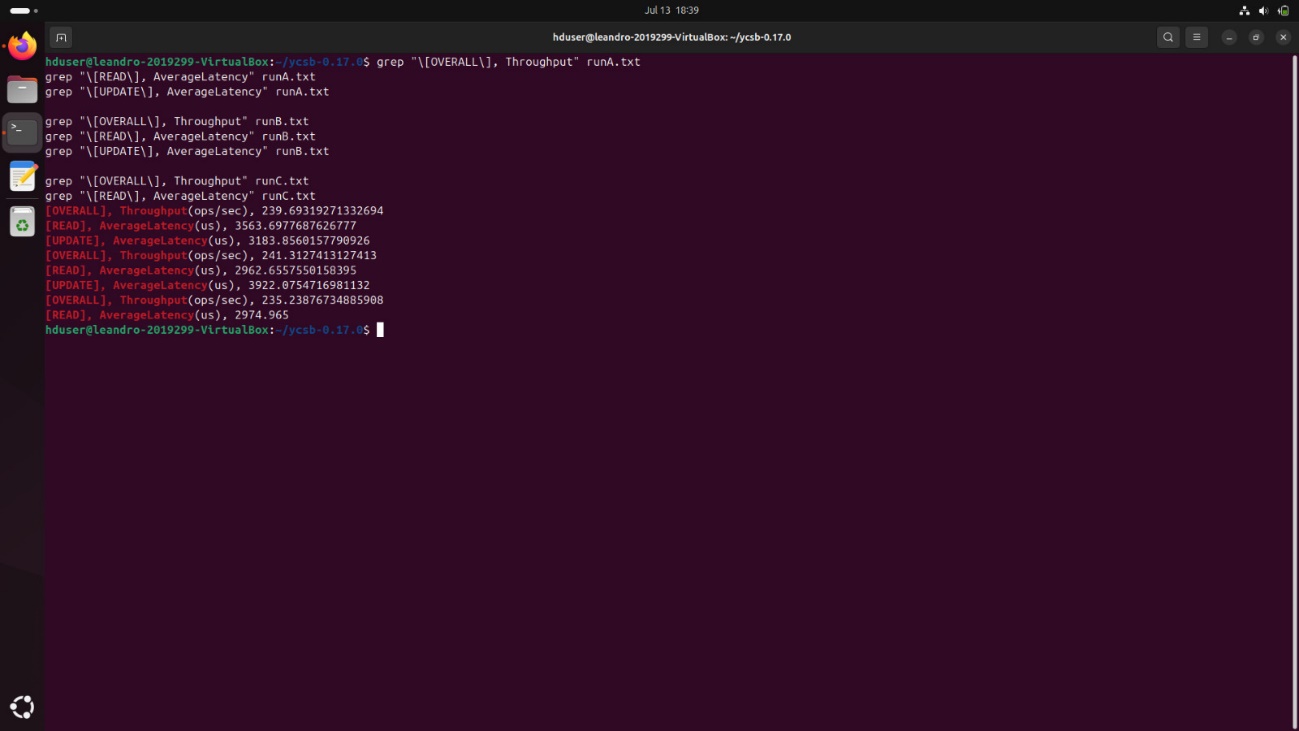
bin/ycsb run mongodb -s -P workloads/workloadc -p recordcount=50000 -p

mongodb.url=mongodb://localhost:2707/ycsb > runC.txt

tail -n 20 runC.txt



# 12.



# Reference

<https://catalog.data.gov/dataset/electric-vehicle-population-data>