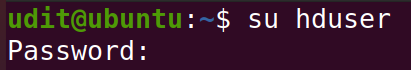
**PRACTICAL NO: 03**

**Aim:** Implement word count / frequency programs using MapReduce

**Code:**

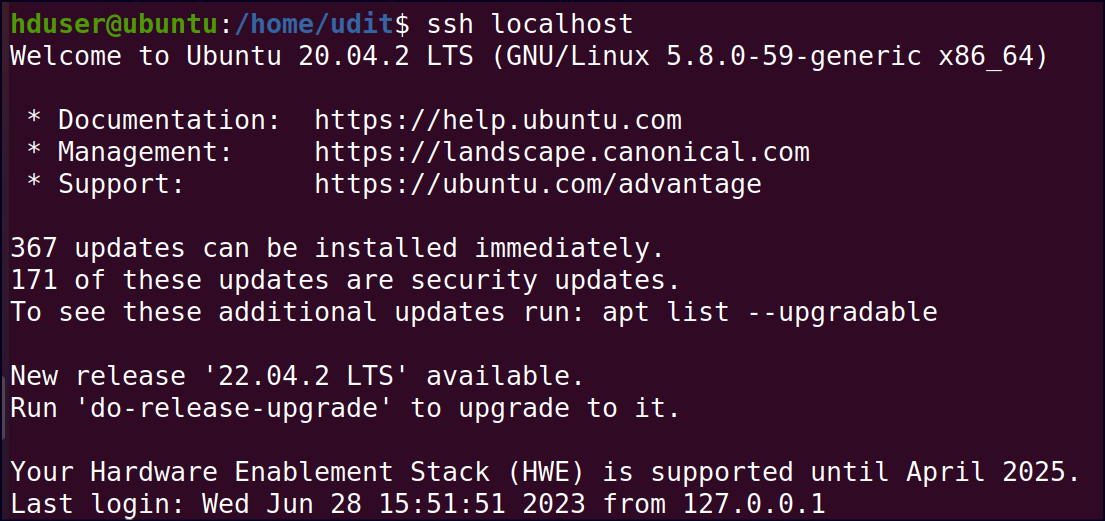
*#Login to hadoop user if not already logged in*

su hduser

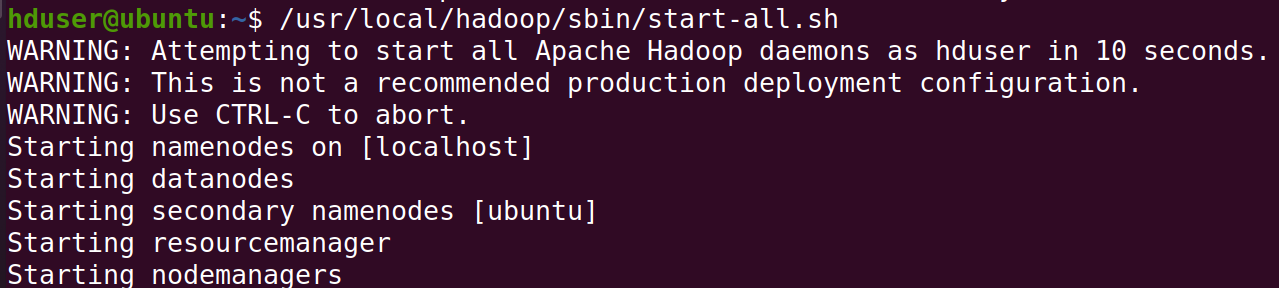


*# Start hadoop services*

ssh localhost

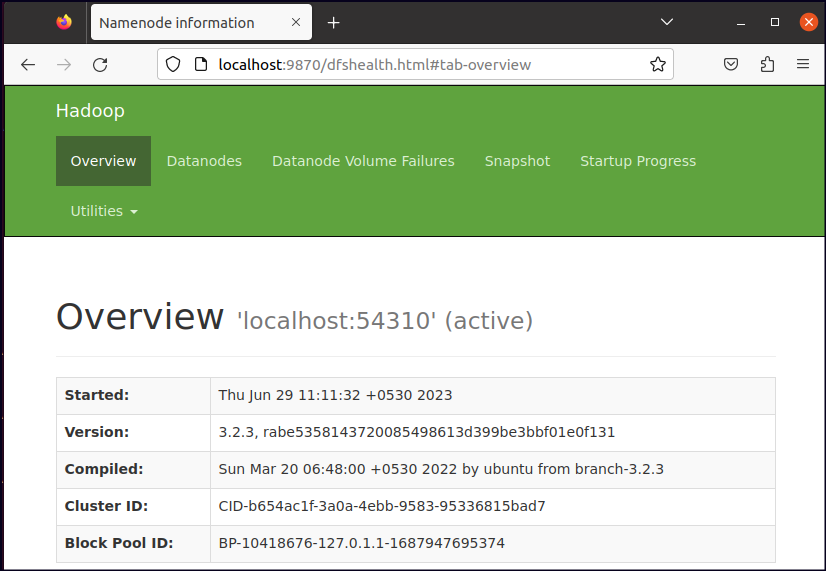


/usr/local/hadoop/sbin/start-all.sh



*# Start localhost*

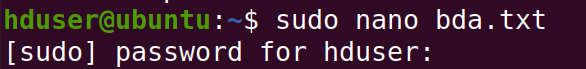
<http://localhost:9870>



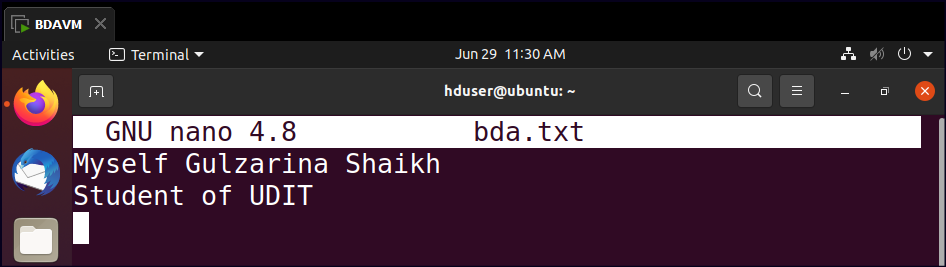
*# 2. Create a text file with some words at local file system*

*# (Necessary to include same and repeated words)*

sudo nano bda.txt



*# Write 2/3 Sentences of your own make sure it must have same and repeated words.*



sudo chown hduser:hadoop /home/hduser/bda.txt



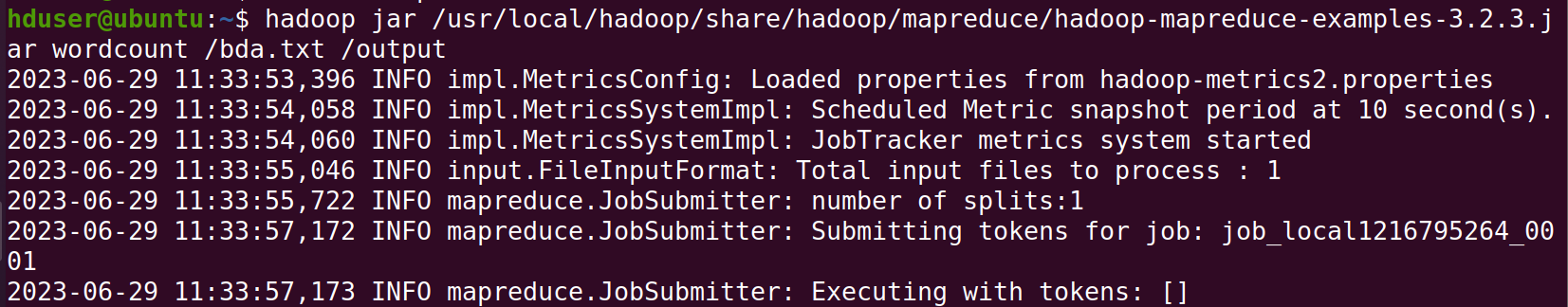
*# 3. Move bda.txt file to HDFS*

hdfs dfs -put /home/hduser/bda.txt /



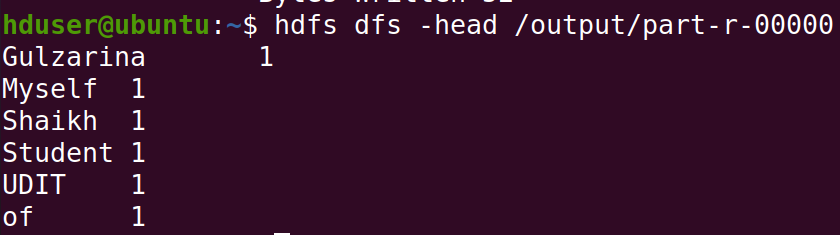
# 4. *Running* MapReduce for wordcount file bda.txt

hadoop jar /usr/local/hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-3.2.3.jar wordcount /bda.txt /output



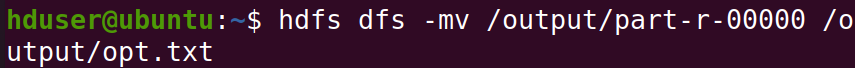
*# 5 Check output at default output location*

hdfs dfs -head /output/part-r-00000



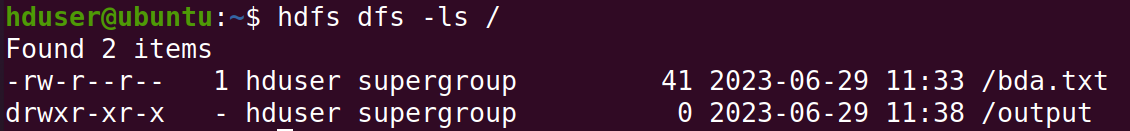
*#6. To get output in a .txt file in HDFS*

hdfs dfs -mv /output/part-r-00000 /output/opt.txt



*# 7. To check the file system in HDFS*

hdfs dfs -ls /

****

**Conclusion:** The program on implementation of word count / frequency using MapReduce has been demonstrated successfully.

# #login to hadoop user if not already logged in (su hduser OR su – hduser)

* su hduser

# #starting Hadoop services

* ssh localhost
* /usr/local/hadoop/sbin/start-all.sh  
    
  OR

ssh localhost

/usr/local/hadoop/sbin/start-dfs.sh

/usr/local/hadoop/sbin/start-yarn.sh

# #start localhost to Check Hadoop Web UI (optional)

* http://localhost:9870

# # Create a text file with some words at local file system

# # (Necessary to include same and repeated words)

* sudo nano bda.txt
* sudo chown hduser:hadoop /home/hduser/bda.txt

# # Move bda.txt file to HDFS

* hdfs dfs -put /home/hduser/bda.txt /

# # 4. Running MapReduce for wordcount file bda.txt

* hadoop jar /usr/local/hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-3.2.3.jar wordcount /bda.txt /output

✅ This is correct, if you are using Hadoop 3.2.3.

* **If the /output directory already exists, the job will fail.**  
  To prevent this, delete it before running the job: hdfs dfs -rm -r /output

# # Check output at default output location

* hdfs dfs -head /output/part-r-00000

OR

* + - hdfs dfs -cat /output/part-r-00000 | head

# #To get output in a .txt file in HDFS

* hdfs dfs -mv /output/part-r-00000 /output/opt.txt

# # To check the file system in HDFS

* hdfs dfs -ls /