Exercise : 03 210701134

Implement a MapReduce program to process a weather dataset

Aim:

To Implement a MapReduce program to process a weather dataset

Procedure:

1. Open the terminal and start Hadoop using start-all.sh command
2. Open the browser and go to the URL localhost:9870.
3. In the terminal using the command hadoop fs -mkdir /user create a directory called user.
4. Upload the sample_weather.txt file to hdfs using the command

   hadoop fs -put sample_weather.txt /user.

   Then perform the mapreduce operation using the command

   hadoop jar /path/to/hadoop-streaming.jar \

   -files /path/to/mapper.py, /path/to/reducer.py \ -input /path/to/input \

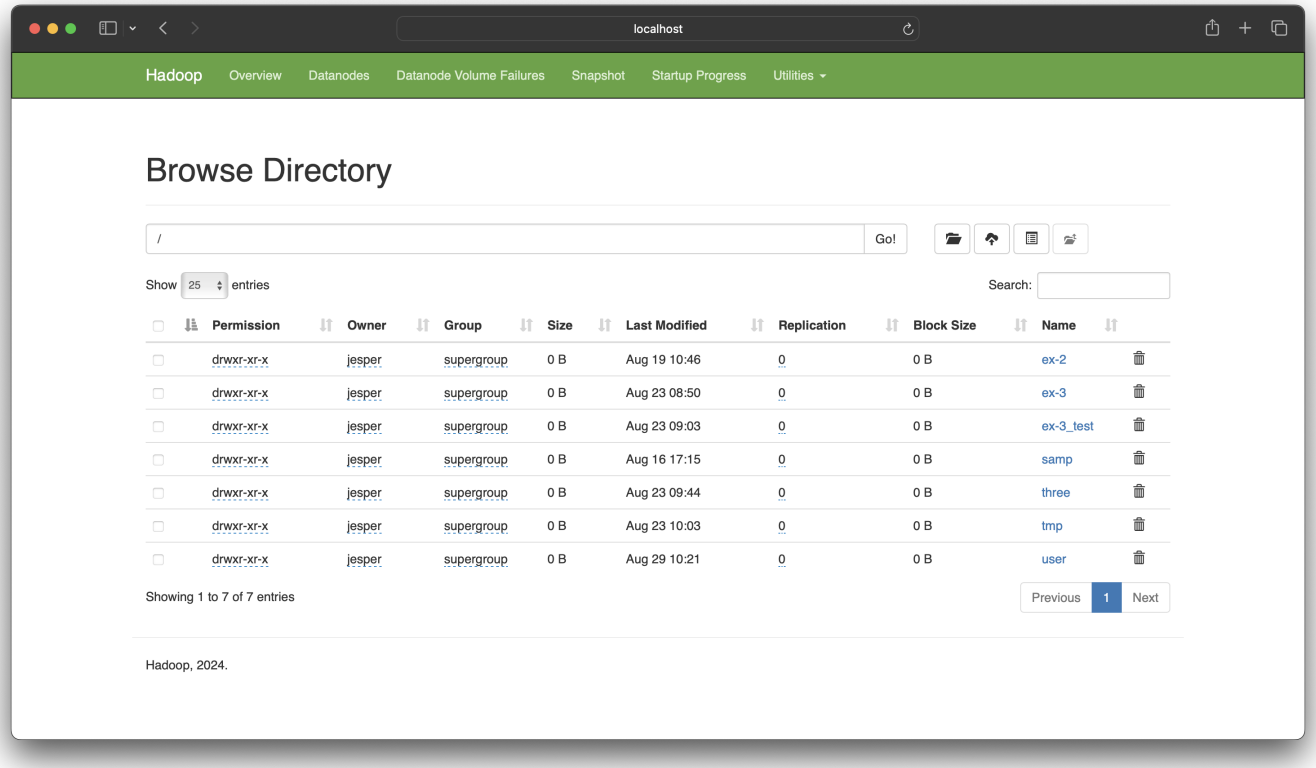
   -output /path/to/output \

   -mapper mapper.py \

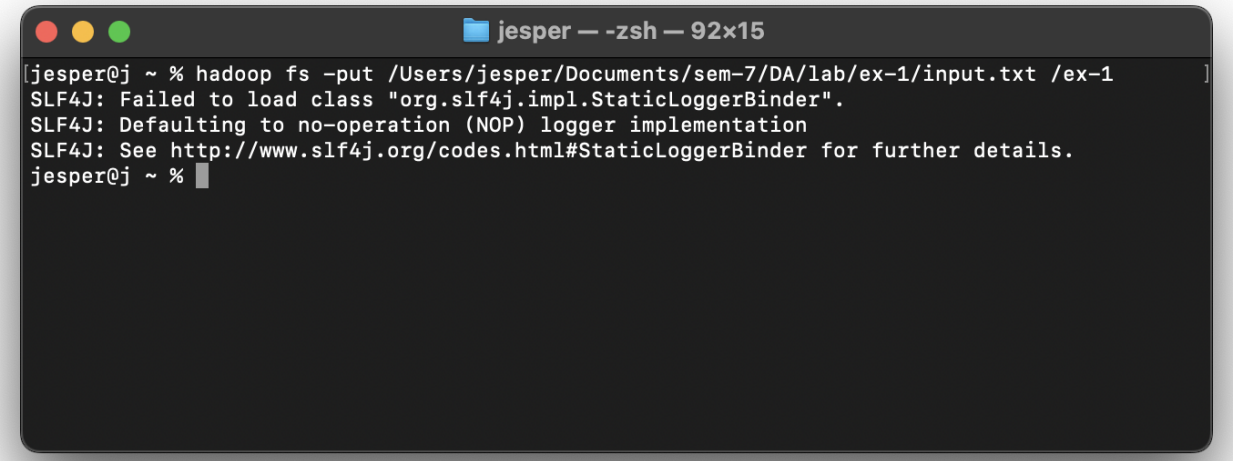
   -reducer reducer.py
5. Check the output using the command hadoop fs -cat /user/output/part-00000.

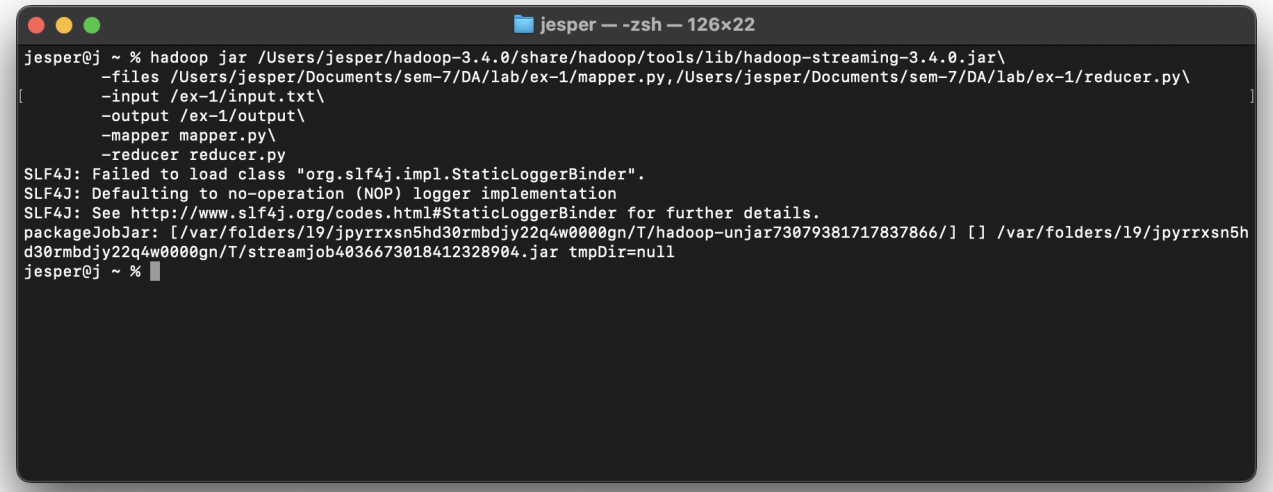
Output:

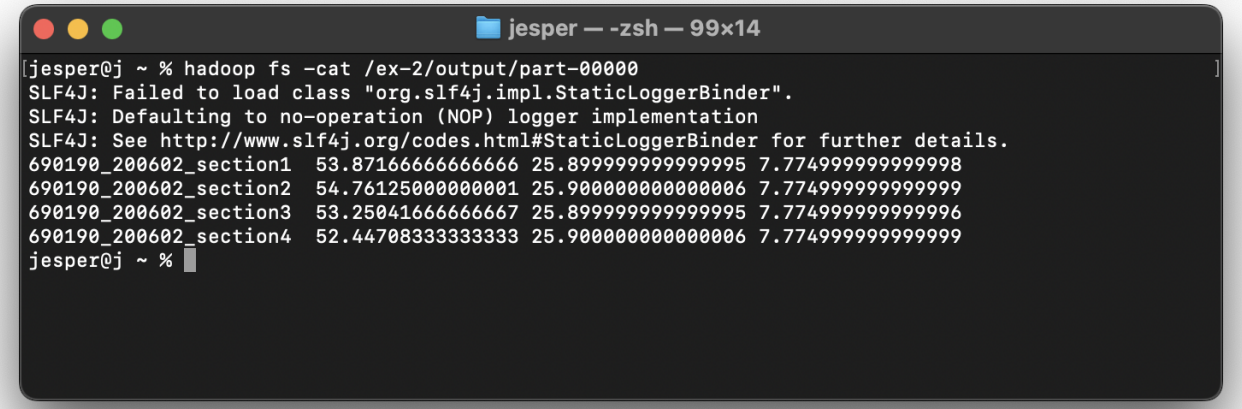
```
jesper — -zsh — 78×15
[jesper@j ~ % start-all.sh                                                    ]
WARNING: Attempting to start all Apache Hadoop daemons as jesper in 10 seconds
.
WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [0.0.0.0]
Starting datanodes
Starting secondary namenodes [j.local]
SLF4J: Failed to load class "org.slf4j.impl.StaticLoggerBinder".
SLF4J: Defaulting to no-operation (NOP) logger implementation
SLF4J: See http://www.slf4j.org/codes.html#StaticLoggerBinder for further deta
ils.
Starting resourcemanager
Starting nodemanagers
jesper@j ~ %
```




```
jesper — -zsh — 78×15
[jesper@j ~ % hadoop fs -mkdir /ex-1                                          ]
SLF4J: Failed to load class "org.slf4j.impl.StaticLoggerBinder".
SLF4J: Defaulting to no-operation (NOP) logger implementation
SLF4J: See http://www.slf4j.org/codes.html#StaticLoggerBinder for further deta
ils.
jesper@j ~ %
```







Result:

Thus the Installation, Configuration and run Hadoop and HDFS is successfully executed.