

Implement word count/frequency programs using MapReduce

AIM:

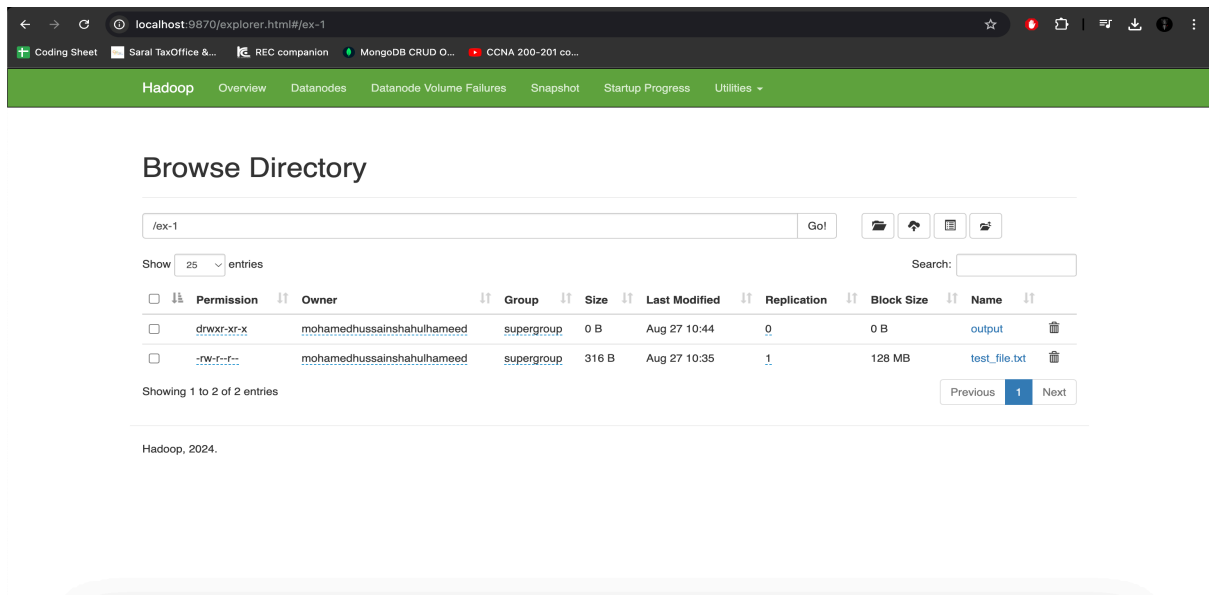
To implement word count/frequency programs using MapReduce

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 input.txt file to hdfs using the command `hadoop fs -put input.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:

```
mohamedhussainshahulhameed@Mohameds-Laptop ~ % start-all.sh
WARNING: Attempting to start all Apache Hadoop daemons as mohamedhussainshahulhameed in 10 seconds.
WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [Mohameds-Laptop.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 details.
Starting resourcemanager
Starting nodemanagers
```



```
mohamedhussainshahulhameed@Mohameds-Laptop ~ % hadoop fs -put /Users/mohamedhussainshahulhameed/Desktop/DA/ex1/test_file.txt /ex1
```

```
mohamedhussainshahulhameed@Mohameds-Laptop ~ % hadoop jar //Users/mohamedhussainshahulhameed/hadoop-3.4.0/share/hadoop/tools/lib/hadoop-streaming-3.4.0.jar \
-files /Users/mohamedhussainshahulhameed/Desktop/DA/ex1/mapper.py, /Users/mohamedhussainshahulhameed/Desktop/DA/ex1/reducer.py \
-input /ex1/test_file.txt \
-output /ex1/output \
-mapper mapper.py \
-reducer reducer.py
```

```
mohamedhussainshahulhameed@Mohameds-Laptop ~ % hadoop fs -cat /ex-1/output/part-00000
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 details.
, 1
-files 1
-input 1
-mapper 1
-output 1
-put 1
-reducer 1
/Users/mohamedhussainshahulhameed/Desktop/DA/ex1/test_file.txt 1
/path/to/hadoop-streaming.jar 1
/path/to/input 1
/path/to/mapper.py,/path/to/reducer.py 1
/path/to/output 1
/user 1
Hadoop 1
\ 5
fs 1
hadoop 1
hello 2
hi 5
jar 1
mapper.py 1
reducer.py 1
there 1
welcome 1
world 1
mohamedhussainshahulhameed@Mohameds-Laptop ~ %
```

RESULT:

Thus the above program to Implement word count/frequency programs using MapReduce has been implemented successfully.