# Practical No: 7

**Hadoop Installation:**

**Date: 17/04/2023**

## Aim: Implement word count / frequency programs using MapReduce. Solution:

C:\hadoop-3.3.0\sbin>start-dfs.cmd C:\hadoop-3.3.0\sbin>start-yarn.cmd

Open a command prompt as administrator and run the following command to create an input and output folder on the Hadoop file system, to which we will be moving the sample.txt file for our analysis.

C:\hadoop-3.3.0\bin>cd\

C:\>hadoop dfsadmin -safemode leave

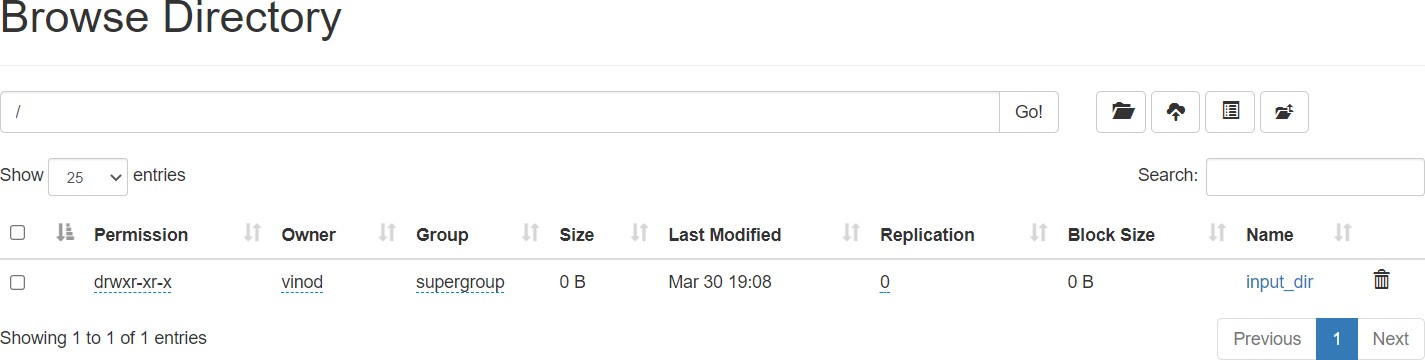
DEPRECATED: Use of this script to execute hdfs command is deprecated. Instead use the hdfs command for it.

Safe mode is OFF

C:\>hadoop fs -mkdir /input\_dir

Check it by giving the following URL at browser http://localhost:9870

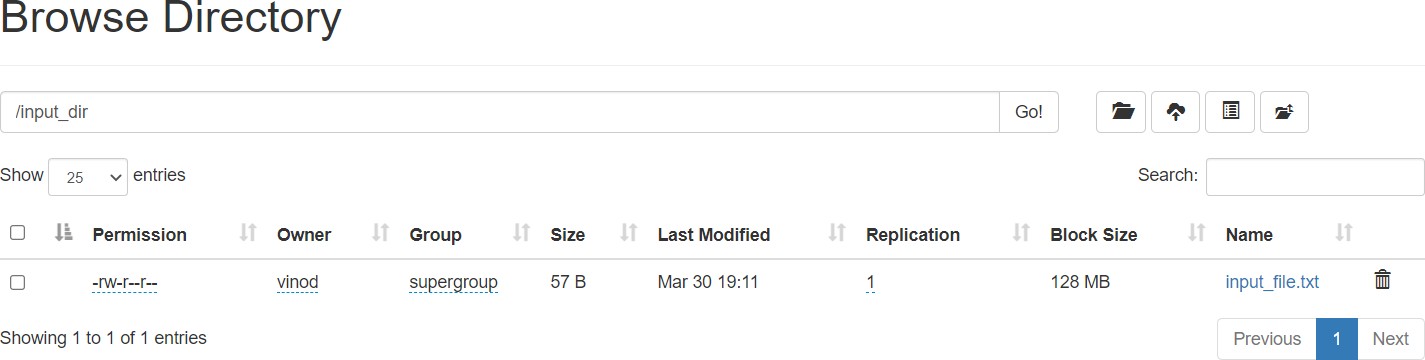
Utilities -> browse the file system



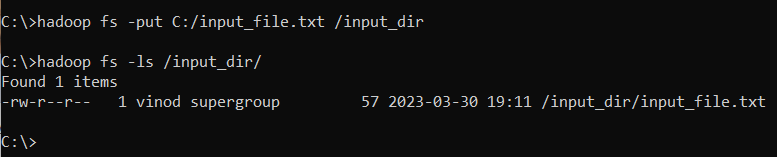
Copy the input text file named input\_file.txt in the input directory (input\_dir) of HDFS.

Make a file in c:\input\_file.txt and write following content in it. Hadoop Window version is easy compared to Ubuntu version Now apply the following command at c:\>

C:\> hadoop fs -put C:/input\_file.txt /input\_dir

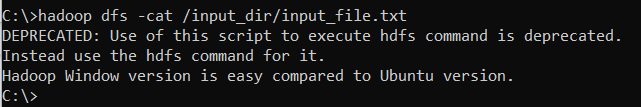


Verify input\_file.txt available in HDFS input directory (input\_dir). C:\>Hadoop fs -ls /input\_dir/



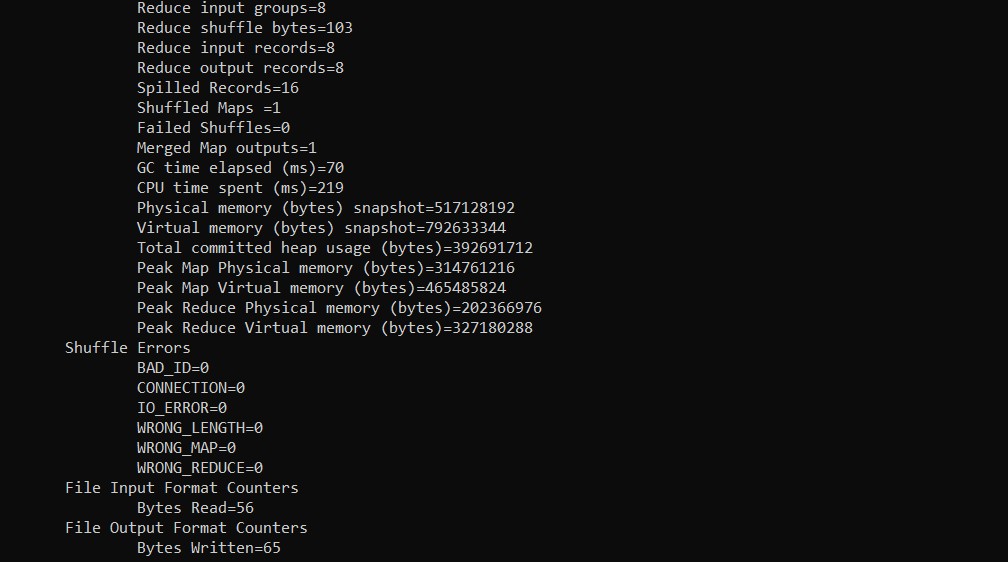
Verify content of the copied file C:\>hadoop dfs -cat /input\_dir/input\_file.txt

You can see the file content displayed on the CMD.



Run MapReduceClient.jar and also provide input and out directories.

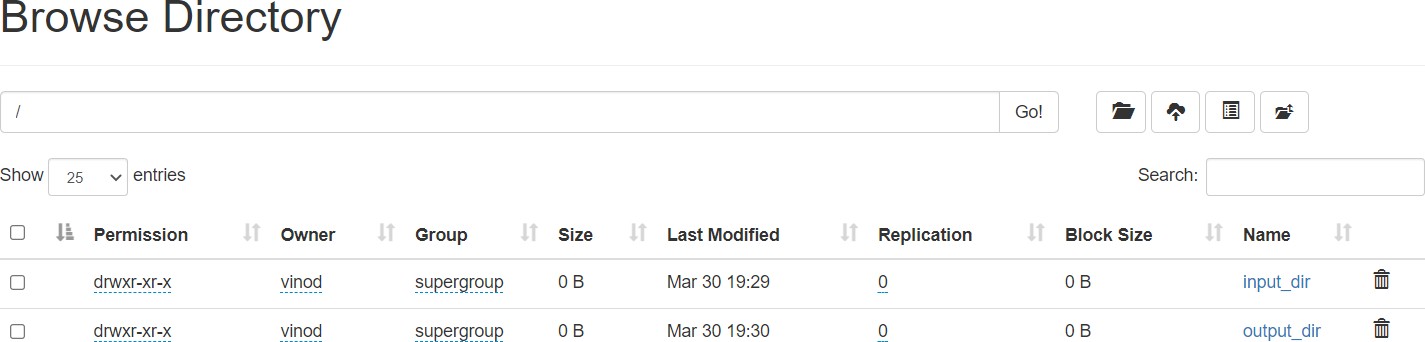
C:\>hadoop jar C:/hadoop-3.3.0/share/hadoop/mapreduce/hadoop-mapreduce-examples- 3.3.0.jar wordcount /input\_dir /output\_dir



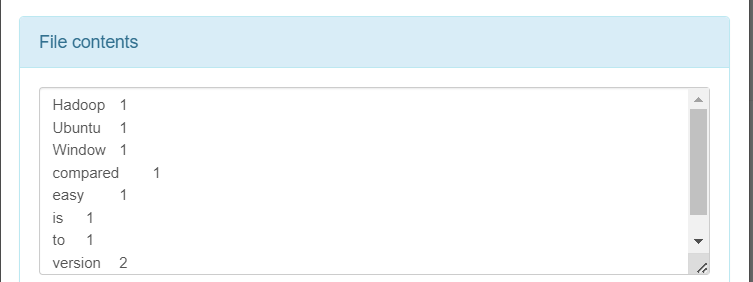
In case, there is some error in executing then copy the file MapReduceClient.jar in C:\ and run the program with the jar file using existing MapReduceClient.jar file as:

*C:\> hadoop jar C:/MapReduceClient.jar wordcount /input\_dir /output\_dir*

Now, check the output\_dir on browser as follows:

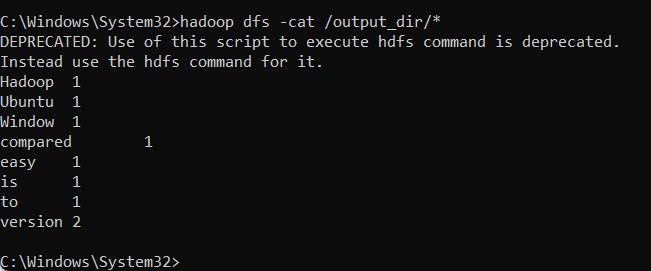


Click on output\_dir → part-r-00000 → Head the file (first 32 K) and check the file contentas the output.

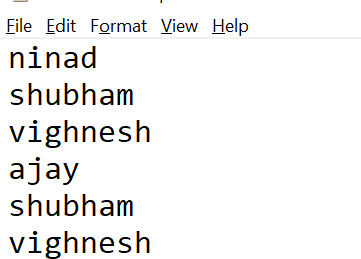


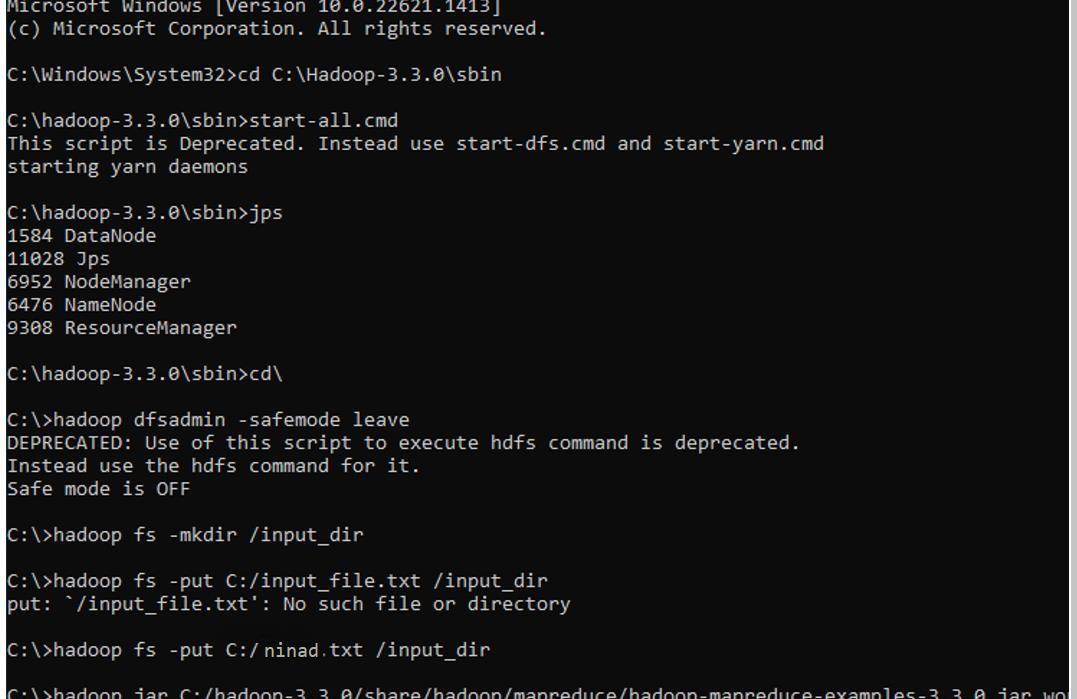
Alternatively, you may type the following command on CMD window as: C:\> hadoop dfs -cat /output\_dir/\*

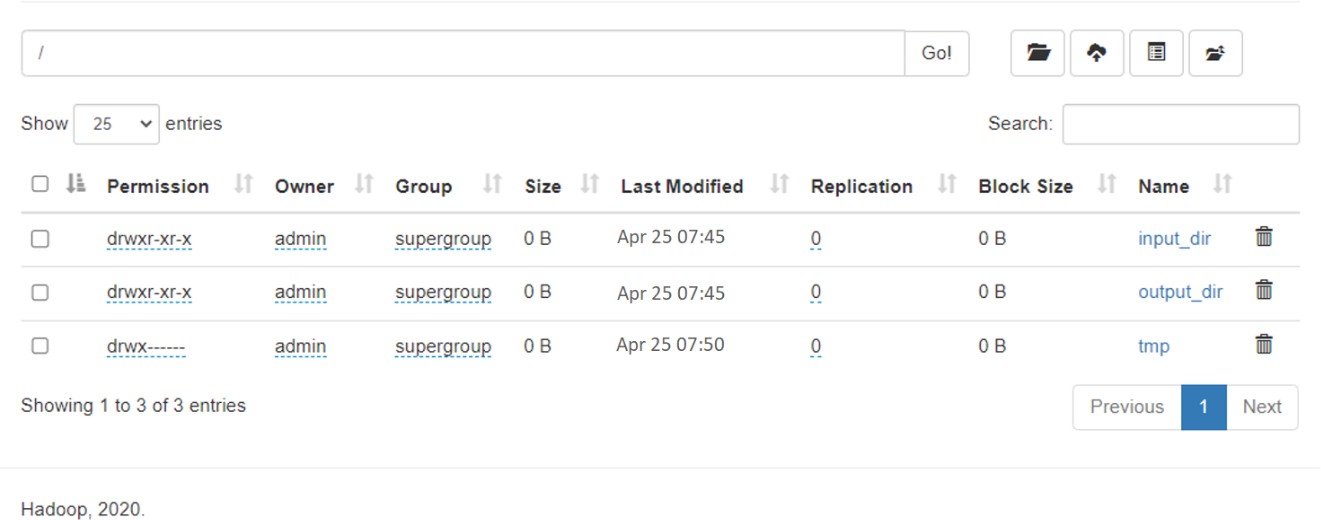
You can get the following output



## Output:







**Output:**

