

EXP NO: 2 RUN A BASIC WORD COUNT MAP REDUCE PROGRAM TO UNDERSTAND MAP REDUCE PARADIGM

\$mkdir DA-Lab

\$cd DA-Lab

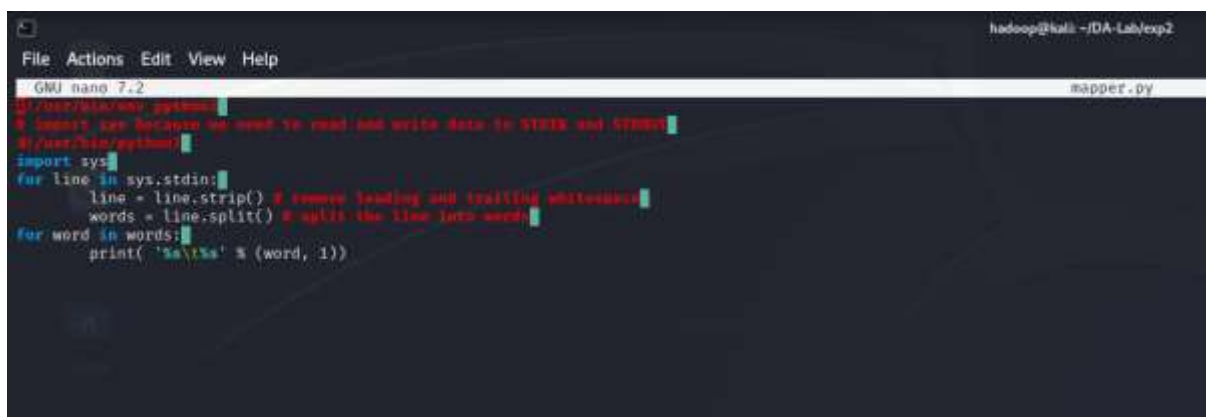
\$mkdir exp2

\$cd exp2

\$nano word_count.txt

A screenshot of a terminal window showing the nano text editor. The editor is open to a file named 'word_count.txt'. The text inside the file reads: 'Hadoop installation successful. This is a word count program executed in Hadoop.' The terminal prompt is 'hadoop@kali: ~/DA-Lab/exp2'.

\$nano mapper.py

A screenshot of a terminal window showing the nano text editor. The editor is open to a file named 'mapper.py'. The code inside is a Python script that reads from standard input, strips whitespace, splits the line into words, and prints each word followed by a tab and the number 1. The terminal prompt is 'hadoop@kali: ~/DA-Lab/exp2'.

\$nano reducer.py

A screenshot of a terminal window showing the nano text editor. The editor is open to a file named 'reducer.py'. The code inside is a Python script that reads from standard input, splits the line into words and counts, and prints the word and its total count. The terminal prompt is 'hadoop@kali: ~/DA-Lab/exp2'.

\$start-all.sh

```

hadoop@kali: ~
File Actions Edit View Help

[hadoop@kali]~$
$ start-all.sh
WARNING: Attempting to start all Apache Hadoop daemons as hadoop in 18 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 [kali]
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
2024-09-11 04:15:16.429 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Starting ResourceManager
Starting NodeManagers

```

\$ jps

```

[hadoop@kali]~$
$ jps
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
14436 NodeManager
16772 Jps
13836 SecondaryNameNode
14311 ResourceManager
13597 DataNode
13471 NameNode

[hadoop@kali]~$

```

\$hdfs dfs -mkdir /exp2**\$hdfs dfs -copyFromLocal ~/DA-Lab/exp2/word_count.txt /exp2**

```

[hadoop@kali]~/hadoop/bin$
$ ./hdfs dfs -ls /exp2
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
2024-09-11 04:05:07,484 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 1 items
drwxr-xr-x 2 hadoop supergroup 0 2024-09-11 01:05 /exp2/output
-rw-r--r-- 1 hadoop supergroup 88 2024-09-11 01:02 /exp2/word_count.txt

```

\$chmod 777 mapper.py reducer.py**\$hadoop jar \$HADOOP_STREAMING -input /exp2/word_count.txt -output /exp2/output -mapper ~/DA-Lab/exp2/mapper.py -reducer ~/DA-Lab/exp2/reducer.py**

```

kali-linux-2023a-vmware-amd64 - VMware Workstation 17 Player (Free commercial use only)
hadoop@kali: ~$
File Actions Edit View Help
hadoop@kali: ~$
hadoop@kali: ~$
hadoop@kali: ~$
[hadoop@kali]~$
$ hadoop jar $HADOOP_STREAMING -input /exp2/word_count.txt -output /exp2/output -mapper ~/DA-Lab/exp2/mapper.py -reducer ~/DA-Lab/exp2/reducer.py
2024-09-11 04:15:16.429 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [kali]
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
2024-09-11 04:15:16.429 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Starting ResourceManager
Starting NodeManagers
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
2024-09-11 04:15:16.429 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 1 items
drwxr-xr-x 2 hadoop supergroup 0 2024-09-11 01:05 /exp2/output
-rw-r--r-- 1 hadoop supergroup 88 2024-09-11 01:02 /exp2/word_count.txt

```

\$hdfs dfs -cat /exp2/output/*

```
(hadoop@kali)=[~/hadoop/bin]
$ ./hdfs dfs -cat /exp2/output/*
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
2024-09-21 00:07:24,178 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform...
Hadoop 1
This 1
a 1
count 1
executed 1
in 1
is 1
program 1
word 1
```