Ex No 2

Run a basic Word Count Map Reduce program to understand Map Reduce Paradigm.

AIM:

To run a basic Word Count MapReduce program using Hadoop.

PROCEDURE:

Step 1: Start the Hadoop cluster

1. Open Terminal in administrative mode:
    - Open a terminal window.
    - Run Hadoop's startup scripts to start the cluster:

```
cd /usr/local/Cellar/hadoop/3.4.0/libexec/sbin
./start-dfs.sh
./start-yarn.sh
```

2. Verify that all nodes are up by running:

```
jps
```

Step 2: Create an input directory in HDFS

Create an HDFS directory where you will place the input file for the MapReduce job. You can name it "input_dir":

```
hadoop fs -mkdir /input_dir
```

Step 3: Copy the input text file to the input directory

Prepare your input file (named `input_file.txt`), or create a sample text file on your local system:

```
echo "Hadoop is a distributed computing framework" >
~/input_file.txt
```

Copy the input file to HDFS:

```
hadoop fs -put ~/input_file.txt /input_dir
```

Step 4: Verify if the file is copied to HDFS
List files in the input directory:

```
hadoop fs -ls /input_dir
```

Check the content of the copied file:

```
hadoop fs -cat /input_dir/input_file.txt
```

Step 5: Run the MapReduce Word Count job

1. Run the MapReduce job:
    - Use the built-in WordCount example that comes with Hadoop.
    - Run the following command, specifying the input directory (`/input_dir`) and an output directory (`/output_dir`):

```
hadoop jar
/usr/local/Cellar/hadoop/3.4.0/libexec/share/hadoop/mapreduce/ha
doop-mapreduce-examples-3.4.0.jar wordcount /input_dir
/output_dir
```

Step 6: Verify the output generated
Check the content of the output directory:

```
hadoop fs -ls /output_dir
```

View the content of the output file:

```
hadoop fs -cat /output_dir/part-r-00000
```

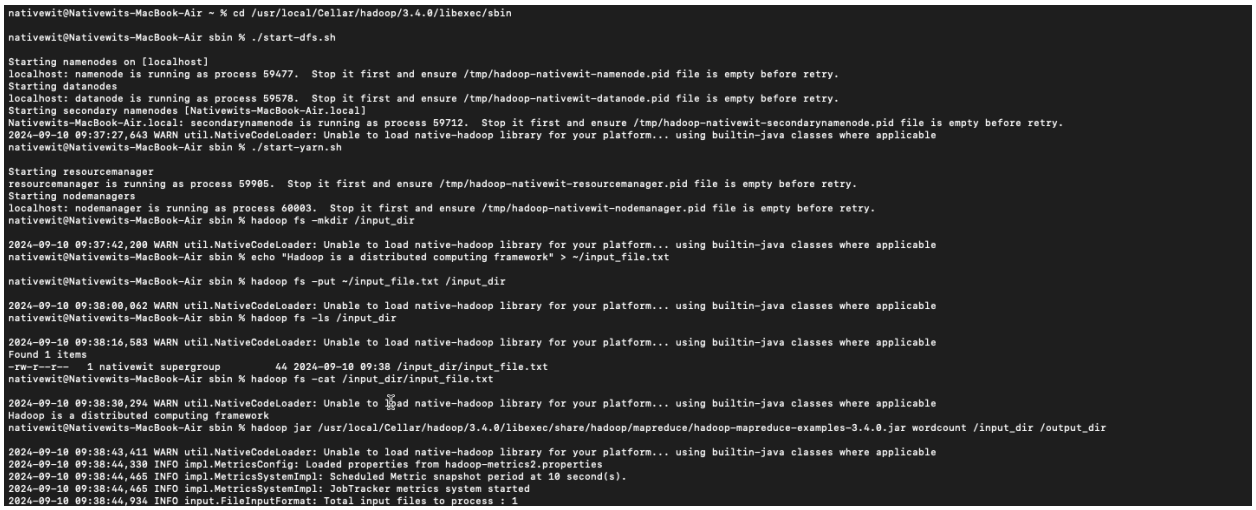
Step 7: Useful Hadoop Commands
To delete a file from HDFS directory:

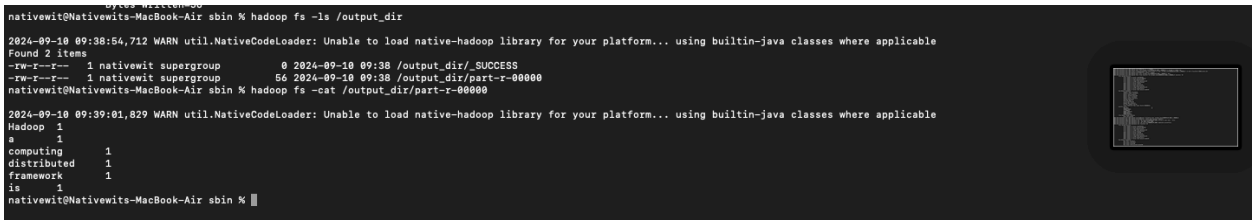
```
hadoop fs -rm -r /input_dir/input_file.txt
```

To delete a directory from HDFS directory:

```
hadoop fs -rm -r /input_dir
```

Output:





RESULT:

Thus, the program for basic Word Count Map Reduce has been executed successfully.