

WEEK-05

Word Count

```
2023-05-17 09:45:13,288 INFO mapred.Merger: Down to the last merge-pass, with 1 segments left of total size: 230 bytes
2023-05-17 09:45:13,288 INFO mapred.LocalJobRunner: 1 / 1 copied.
2023-05-17 09:45:13,361 INFO mapred.Task: Task:attempt_local1164004484_0001_r_000000_0 is done. And is in the process of committing
2023-05-17 09:45:13,363 INFO mapred.LocalJobRunner: 1 / 1 copied.
2023-05-17 09:45:13,363 INFO mapred.Task: Task:attempt_local1164004484_0001_r_000000_0 is allowed to commit now
2023-05-17 09:45:13,378 INFO output.FileOutputCommitter: Saved output of task 'attempt_local1164004484_0001_r_000000_0' to hdfs://localhost:9000/output1
2023-05-17 09:45:13,378 INFO mapred.LocalJobRunner: reduce > reduce
2023-05-17 09:45:13,379 INFO mapred.Task: Task 'attempt_local1164004484_0001_r_000000_0' done.
2023-05-17 09:45:13,379 INFO mapred.Task: Final Counters for attempt_local1164004484_0001_r_000000_0: Counters: 30
  File System Counters
    FILE: Number of bytes read=4687
    FILE: Number of bytes written=645171
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=97
    HDFS: Number of bytes written=67
    HDFS: Number of read operations=10
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=3
    HDFS: Number of bytes read erasure-coded=0
  Map-Reduce Framework
    Combine input records=0
    Combine output records=0
    Reduce input groups=10
    Reduce shuffle bytes=240
    Reduce input records=23
    Reduce output records=10
    Spilled Records=23
    Shuffled Maps =1
    Failed Shuffles=0
    Merged Map outputs=1
    GC time elapsed (ms)=0
    Total committed heap usage (bytes)=526385152
  Shuffle Errors
    BAD_ID=0
    CONNECTION=0
    IO_ERROR=0
    WRONG_LENGTH=0
    WRONG_MAP=0
    WRONG_REDUCE=0
  File Output Format Counters
    Bytes Written=67
```

```
hadoop@bmscscse-HP-Elite-Tower-600-G9-Desktop-PC: $ hadoop jar /home/hadoop/Documents/jar/WordCount.jar WCDriver /yathri/wc.txt /output1
2023-05-17 09:45:12,856 INFO impl.MetricsConfig: Loaded properties from hadoop-metrics2.properties
2023-05-17 09:45:12,890 INFO impl.MetricsSystemImpl: Scheduled Metric snapshot period at 10 second(s).
2023-05-17 09:45:12,890 INFO impl.MetricsSystemImpl: JobTracker metrics system started
2023-05-17 09:45:12,895 WARN impl.MetricsSystemImpl: JobTracker metrics system already initialized!
2023-05-17 09:45:12,949 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
2023-05-17 09:45:12,983 INFO mapred.FileInputFormat: Total input files to process : 1
2023-05-17 09:45:13,004 INFO mapreduce.JobSubmitter: number of splits:1
2023-05-17 09:45:13,056 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_local1164004484_0001
2023-05-17 09:45:13,056 INFO mapreduce.JobSubmitter: Executing with tokens: []
2023-05-17 09:45:13,106 INFO mapreduce.Job: The url to track the job: http://localhost:8080/
2023-05-17 09:45:13,107 INFO mapreduce.Job: Running job: job_local1164004484_0001
2023-05-17 09:45:13,107 INFO mapred.LocalJobRunner: OutputCommitter set in config null
2023-05-17 09:45:13,108 INFO mapred.LocalJobRunner: OutputCommitter is org.apache.hadoop.mapred.FileOutputCommitter
2023-05-17 09:45:13,110 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2023-05-17 09:45:13,110 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup failures: false
2023-05-17 09:45:13,161 INFO mapred.LocalJobRunner: Waiting for map tasks
2023-05-17 09:45:13,162 INFO mapred.LocalJobRunner: Starting task: attempt_local1164004484_0001_m_000000_0
2023-05-17 09:45:13,171 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2023-05-17 09:45:13,171 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup failures: false
2023-05-17 09:45:13,176 INFO mapred.Task: Using ResourceCalculatorProcessTree : [ ]
2023-05-17 09:45:13,178 INFO mapred.MapTask: Processing split: hdfs://localhost:9000/yathri/wc.txt:0+97
```

```
hadoop@bmscecse-HP-Elite-Tower-600-G9-Desktop-PC:~$ hdfs dfs -copyFromLocal /home/hadoop/Desktop/wc.txt /yathri
hadoop@bmscecse-HP-Elite-Tower-600-G9-Desktop-PC:~$ hadoop fs -ls /yathri
Found 2 items
drwxr-xr-x   - hadoop supergroup          0 2023-05-17 09:33 /yathri/Desktop
-rw-r--r--   1 hadoop supergroup        97 2023-05-17 09:35 /yathri/wc.txt
hadoop@bmscecse-HP-Elite-Tower-600-G9-Desktop-PC:~$ hdfs dfs -cat /yathri/wc.txt
hi how are you
hi how you are doing
hi how is your sister
how is your job
hi how is your health
```

```
hadoop@bmscecse-HP-Elite-Tower-600-G9-Desktop-PC:~$ hadoop fs -ls /output1
Found 2 items
-rw-r--r--   1 hadoop supergroup          0 2023-05-17 09:45 /output1/_SUCCESS
-rw-r--r--   1 hadoop supergroup        67 2023-05-17 09:45 /output1/part-00000
hadoop@bmscecse-HP-Elite-Tower-600-G9-Desktop-PC:~$ hadoop fs -cat /output1/part-00000
are      2
doing    1
health   1
hi        4
how       5
is        3
job       1
sister    1
you       2
your      3
hadoop@bmscecse-HP-Elite-Tower-600-G9-Desktop-PC:~$
```