

BDA LAB-7

NAME : SANJAY M S

USN : 1BM18CS418

OUTPUT

```
Administrator: Command Prompt
C:\WINDOWS\system32>hadoop dfs -cat /input_dir/lab_1/
DEPRECATED: Use of this script to execute hdfs command is deprecated.
Instead use the hdfs command for it.
ravi kumar , bda lab , sample program for in Hadoop for counting the number of words in file .
C:\WINDOWS\system32>hadoop jar C:\hadoop-3.0.0\share\hadoop\mapreduce\hadoop-mapreduce-examples-3.0.0.jar wordcount /input_dir/Lab_1 /output_dir/Lab_word_count
2020-12-19 19:27:43,564 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032
2020-12-19 19:27:46,258 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadcop-yarn/staging/VINOD/.staging/job_1608385473820_0001
2020-12-19 19:27:47,897 INFO input.FileInputFormat: Total input files to process : 1
2020-12-19 19:27:48,632 INFO mapreduce.JobSubmitter: number of splits:1
2020-12-19 19:27:48,793 INFO Configuration.deprecation: yarn.resourcemanager.system-metrics-publisher.enabled is deprecated. Instead, use yarn.system-metrics-publisher.enabled
2020-12-19 19:27:49,149 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1608385473820_0001
2020-12-19 19:27:49,154 INFO mapreduce.JobSubmitter: Executing with tokens: []
2020-12-19 19:27:50,278 INFO conf.Configuration: resource-types.xml not found
2020-12-19 19:27:50,318 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.
2020-12-19 19:27:51,345 INFO impl.YarnClientImpl: Submitted application application_1608385473820_0001
2020-12-19 19:27:51,488 INFO mapreduce.Job: The url to track the job: http://DESKTOP-NNWEQSP:8088/proxy/application_1608385473820_0001/
2020-12-19 19:27:51,493 INFO mapreduce.Job: Running job: job_1608385473820_0001
2020-12-19 19:28:27,818 INFO mapreduce.Job: Job job_1608385473820_0001 running in uber mode : false
2020-12-19 19:28:27,825 INFO mapreduce.Job: map 0% reduce 0%
2020-12-19 19:28:46,742 INFO mapreduce.Job: map 100% reduce 0%
2020-12-19 19:29:07,111 INFO mapreduce.Job: map 100% reduce 100%
2020-12-19 19:29:18,254 INFO mapreduce.Job: Job job_1608385473820_0001 completed successfully
2020-12-19 19:29:18,553 INFO mapreduce.Job: Counters: 53
  File System Counters
    FILE: Number of bytes read=194
    FILE: Number of bytes written=416215
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=196
    HDFS: Number of bytes written=170
    HDFS: Number of read operations=8
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=2
  Job Counters
    Launched map tasks=1
    Launched reduce tasks=1
    Data-local map tasks=1
    Total time spent by all maps in occupied slots (ms)=15696
    Total time spent by all reduces in occupied slots (ms)=17631
    Total time spent by all map tasks (ms)=15696
    Total time spent by all reduce tasks (ms)=17631
```

bytes written=120

```
C:\WINDOWS\system32>hadoop dfs -cat /output_dir/Lab_word_count/*  
DEPRECATED: Use of this script to execute hdfs command is deprecated.  
Instead use the hdfs command for it.
```

```
2  
.  
Hadoop 1  
bda 1  
counting 1  
file 1  
for 2  
in 2  
kumar 1  
lab 1  
number 1  
of 1  
program 1  
ravi 1  
sample 1  
the 1  
words 1
```

```
C:\WINDOWS\system32>_
```