A11 :

hadoop@pict-OptiPlex-3020:~$ start-dfs.sh

Starting namenodes on [localhost]

localhost: namenode is running as process 4829. Stop it first and ensure /tmp/hadoop-hadoop-namenode.pid file is empty before retry.

Starting datanodes

localhost: datanode is running as process 4981. Stop it first and ensure /tmp/hadoop-hadoop-datanode.pid file is empty before retry.

Starting secondary namenodes [pict-OptiPlex-3020]

pict-OptiPlex-3020: secondarynamenode is running as process 5198. Stop it first and ensure /tmp/hadoop-hadoop-secondarynamenode.pid file is empty before retry.

hadoop@pict-OptiPlex-3020:~$ start-yarn.sh

Starting resourcemanager

resourcemanager is running as process 5468. Stop it first and ensure /tmp/hadoop-hadoop-resourcemanager.pid file is empty before retry.

Starting nodemanagers

localhost: nodemanager is running as process 5618. Stop it first and ensure /tmp/hadoop-hadoop-nodemanager.pid file is empty before retry.

hadoop@pict-OptiPlex-3020:~$ jps

5618 NodeManager

4981 DataNode

5468 ResourceManager

4829 NameNode

5198 SecondaryNameNode

14015 Jps

hadoop@pict-OptiPlex-3020:~$ hdfs dfs -mkdir /31165

hadoop@pict-OptiPlex-3020:~$ hdfs dfs -ls /31165

hadoop@pict-OptiPlex-3020:~$ hdfs dfs -mkdir /31165/input

hadoop@pict-OptiPlex-3020:~$ hdfs dfs -put /home/hadoop/31181/test.txt /31165/input

hadoop@pict-OptiPlex-3020:~$ hdfs dfs -ls /31165/input

Found 1 items

-rw-r--r-- 3 hadoop supergroup 58 2024-04-22 10:22 /31165/input/test.txt

hadoop@pict-OptiPlex-3020:~$ hdfs dfs -cat /31165/input

cat: `/31165/input': Is a directory

hadoop@pict-OptiPlex-3020:~$ hdfs dfs -cat /31165/input/test.txt

Hello world.

Vedant lives in Mumbai.

Vivek lives in Pune.

hadoop@pict-OptiPlex-3020:~$ export HADOOP\_CLASSPATH=$(hadoop classpath)

hadoop@pict-OptiPlex-3020:~$ echo $HADOOP\_CLASSPATH

/home/hadoop/hadoop-3.3.5/etc/hadoop:/home/hadoop/hadoop-3.3.5/share/hadoop/common/lib/\*:/home/hadoop/hadoop-3.3.5/share/hadoop/common/\*:/home/hadoop/hadoop-3.3.5/share/hadoop/hdfs:/home/hadoop/hadoop-3.3.5/share/hadoop/hdfs/lib/\*:/home/hadoop/hadoop-3.3.5/share/hadoop/hdfs/\*:/home/hadoop/hadoop-3.3.5/share/hadoop/mapreduce/\*:/home/hadoop/hadoop-3.3.5/share/hadoop/yarn:/home/hadoop/hadoop-3.3.5/share/hadoop/yarn/lib/\*:/home/hadoop/hadoop-3.3.5/share/hadoop/yarn/\*

hadoop@pict-OptiPlex-3020:~$ javac -classpath ${HADOOP\_CLASSPATH} -d '/home/hadoop/31181/build' 'home/hadoop/31181/WordCount.java'

error: file not found: home/hadoop/31181/WordCount.java

Usage: javac <options> <source files>

use --help for a list of possible options

hadoop@pict-OptiPlex-3020:~$ javac -classpath ${HADOOP\_CLASSPATH} -d '/home/hadoop/31181/build' '/home/hadoop/31181/WordCount.java'

hadoop@pict-OptiPlex-3020:~$

—-------------------------------------------

hadoop@pict-OptiPlex-3020:~/31181$ jar -cvf a11.jar -C build/ .

added manifest

adding: WordCount.class(in = 1511) (out= 825)(deflated 45%)

adding: WordCount$IntSumReducer.class(in = 1755) (out= 749)(deflated 57%)

adding: WordCount$TokenizerMapper.class(in = 1752) (out= 764)(deflated 56%)

hadoop@pict-OptiPlex-3020:~/31181$ hadoop jar '/home/hadoop/31181/a11.jar' WordCount /31165/input /31165/output

2024-04-22 10:33:15,757 INFO impl.MetricsConfig: Loaded properties from hadoop-metrics2.properties

2024-04-22 10:33:15,834 INFO impl.MetricsSystemImpl: Scheduled Metric snapshot period at 10 second(s).

2024-04-22 10:33:15,834 INFO impl.MetricsSystemImpl: JobTracker metrics system started

2024-04-22 10:33:15,960 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.

2024-04-22 10:33:16,080 INFO input.FileInputFormat: Total input files to process : 1

2024-04-22 10:33:16,098 INFO mapreduce.JobSubmitter: number of splits:1

2024-04-22 10:33:16,255 INFO mapreduce.JobSubmitter: Submitting tokens for job: job\_local369299969\_0001

2024-04-22 10:33:16,255 INFO mapreduce.JobSubmitter: Executing with tokens: []

2024-04-22 10:33:16,362 INFO mapreduce.Job: The url to track the job: http://localhost:8080/

2024-04-22 10:33:16,364 INFO mapreduce.Job: Running job: job\_local369299969\_0001

2024-04-22 10:33:16,365 INFO mapred.LocalJobRunner: OutputCommitter set in config null

2024-04-22 10:33:16,375 INFO output.PathOutputCommitterFactory: No output committer factory defined, defaulting to FileOutputCommitterFactory

2024-04-22 10:33:16,379 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2

2024-04-22 10:33:16,379 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup \_temporary folders under output directory:false, ignore cleanup failures: false

2024-04-22 10:33:16,383 INFO mapred.LocalJobRunner: OutputCommitter is org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter

2024-04-22 10:33:16,450 INFO mapred.LocalJobRunner: Waiting for map tasks

2024-04-22 10:33:16,451 INFO mapred.LocalJobRunner: Starting task: attempt\_local369299969\_0001\_m\_000000\_0

2024-04-22 10:33:16,478 INFO output.PathOutputCommitterFactory: No output committer factory defined, defaulting to FileOutputCommitterFactory

2024-04-22 10:33:16,478 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2

2024-04-22 10:33:16,478 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup \_temporary folders under output directory:false, ignore cleanup failures: false

2024-04-22 10:33:16,492 INFO mapred.Task: Using ResourceCalculatorProcessTree : [ ]

2024-04-22 10:33:16,496 INFO mapred.MapTask: Processing split: hdfs://127.0.0.1:9000/31165/input/test.txt:0+58

2024-04-22 10:33:16,550 INFO mapred.MapTask: (EQUATOR) 0 kvi 26214396(104857584)

2024-04-22 10:33:16,551 INFO mapred.MapTask: mapreduce.task.io.sort.mb: 100

2024-04-22 10:33:16,551 INFO mapred.MapTask: soft limit at 83886080

2024-04-22 10:33:16,551 INFO mapred.MapTask: bufstart = 0; bufvoid = 104857600

2024-04-22 10:33:16,552 INFO mapred.MapTask: kvstart = 26214396; length = 6553600

2024-04-22 10:33:16,568 INFO mapred.MapTask: Map output collector class = org.apache.hadoop.mapred.MapTask$MapOutputBuffer

2024-04-22 10:33:16,648 INFO mapred.LocalJobRunner:

2024-04-22 10:33:16,650 INFO mapred.MapTask: Starting flush of map output

2024-04-22 10:33:16,650 INFO mapred.MapTask: Spilling map output

2024-04-22 10:33:16,650 INFO mapred.MapTask: bufstart = 0; bufend = 98; bufvoid = 104857600

2024-04-22 10:33:16,650 INFO mapred.MapTask: kvstart = 26214396(104857584); kvend = 26214360(104857440); length = 37/6553600

2024-04-22 10:33:16,659 INFO mapred.MapTask: Finished spill 0

2024-04-22 10:33:16,669 INFO mapred.Task: Task:attempt\_local369299969\_0001\_m\_000000\_0 is done. And is in the process of committing

2024-04-22 10:33:16,672 INFO mapred.LocalJobRunner: map

2024-04-22 10:33:16,672 INFO mapred.Task: Task 'attempt\_local369299969\_0001\_m\_000000\_0' done.

2024-04-22 10:33:16,678 INFO mapred.Task: Final Counters for attempt\_local369299969\_0001\_m\_000000\_0: Counters: 24

File System Counters

FILE: Number of bytes read=3262

FILE: Number of bytes written=639352

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=58

HDFS: Number of bytes written=0

HDFS: Number of read operations=5

HDFS: Number of large read operations=0

HDFS: Number of write operations=1

HDFS: Number of bytes read erasure-coded=0

Map-Reduce Framework

Map input records=3

Map output records=10

Map output bytes=98

Map output materialized bytes=103

Input split bytes=107

Combine input records=10

Combine output records=8

Spilled Records=8

Failed Shuffles=0

Merged Map outputs=0

GC time elapsed (ms)=7

Total committed heap usage (bytes)=201326592

File Input Format Counters

Bytes Read=58

2024-04-22 10:33:16,678 INFO mapred.LocalJobRunner: Finishing task: attempt\_local369299969\_0001\_m\_000000\_0

2024-04-22 10:33:16,679 INFO mapred.LocalJobRunner: map task executor complete.

2024-04-22 10:33:16,682 INFO mapred.LocalJobRunner: Waiting for reduce tasks

2024-04-22 10:33:16,682 INFO mapred.LocalJobRunner: Starting task: attempt\_local369299969\_0001\_r\_000000\_0

2024-04-22 10:33:16,691 INFO output.PathOutputCommitterFactory: No output committer factory defined, defaulting to FileOutputCommitterFactory

2024-04-22 10:33:16,691 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2

2024-04-22 10:33:16,691 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup \_temporary folders under output directory:false, ignore cleanup failures: false

2024-04-22 10:33:16,692 INFO mapred.Task: Using ResourceCalculatorProcessTree : [ ]

2024-04-22 10:33:16,697 INFO mapred.ReduceTask: Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@2e6dbceb

2024-04-22 10:33:16,699 WARN impl.MetricsSystemImpl: JobTracker metrics system already initialized!

2024-04-22 10:33:16,715 INFO reduce.MergeManagerImpl: MergerManager: memoryLimit=1443050240, maxSingleShuffleLimit=360762560, mergeThreshold=952413184, ioSortFactor=10, memToMemMergeOutputsThreshold=10

2024-04-22 10:33:16,718 INFO reduce.EventFetcher: attempt\_local369299969\_0001\_r\_000000\_0 Thread started: EventFetcher for fetching Map Completion Events

2024-04-22 10:33:16,744 INFO reduce.LocalFetcher: localfetcher#1 about to shuffle output of map attempt\_local369299969\_0001\_m\_000000\_0 decomp: 99 len: 103 to MEMORY

2024-04-22 10:33:16,747 INFO reduce.InMemoryMapOutput: Read 99 bytes from map-output for attempt\_local369299969\_0001\_m\_000000\_0

2024-04-22 10:33:16,749 INFO reduce.MergeManagerImpl: closeInMemoryFile -> map-output of size: 99, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->99

2024-04-22 10:33:16,750 INFO reduce.EventFetcher: EventFetcher is interrupted.. Returning

2024-04-22 10:33:16,751 INFO mapred.LocalJobRunner: 1 / 1 copied.

2024-04-22 10:33:16,751 INFO reduce.MergeManagerImpl: finalMerge called with 1 in-memory map-outputs and 0 on-disk map-outputs

2024-04-22 10:33:16,756 INFO mapred.Merger: Merging 1 sorted segments

2024-04-22 10:33:16,756 INFO mapred.Merger: Down to the last merge-pass, with 1 segments left of total size: 91 bytes

2024-04-22 10:33:16,758 INFO reduce.MergeManagerImpl: Merged 1 segments, 99 bytes to disk to satisfy reduce memory limit

2024-04-22 10:33:16,758 INFO reduce.MergeManagerImpl: Merging 1 files, 103 bytes from disk

2024-04-22 10:33:16,759 INFO reduce.MergeManagerImpl: Merging 0 segments, 0 bytes from memory into reduce

2024-04-22 10:33:16,759 INFO mapred.Merger: Merging 1 sorted segments

2024-04-22 10:33:16,759 INFO mapred.Merger: Down to the last merge-pass, with 1 segments left of total size: 91 bytes

2024-04-22 10:33:16,760 INFO mapred.LocalJobRunner: 1 / 1 copied.

2024-04-22 10:33:16,795 INFO Configuration.deprecation: mapred.skip.on is deprecated. Instead, use mapreduce.job.skiprecords

2024-04-22 10:33:16,878 INFO mapred.Task: Task:attempt\_local369299969\_0001\_r\_000000\_0 is done. And is in the process of committing

2024-04-22 10:33:16,880 INFO mapred.LocalJobRunner: 1 / 1 copied.

2024-04-22 10:33:16,880 INFO mapred.Task: Task attempt\_local369299969\_0001\_r\_000000\_0 is allowed to commit now

2024-04-22 10:33:16,894 INFO output.FileOutputCommitter: Saved output of task 'attempt\_local369299969\_0001\_r\_000000\_0' to hdfs://127.0.0.1:9000/31165/output

2024-04-22 10:33:16,895 INFO mapred.LocalJobRunner: reduce > reduce

2024-04-22 10:33:16,895 INFO mapred.Task: Task 'attempt\_local369299969\_0001\_r\_000000\_0' done.

2024-04-22 10:33:16,895 INFO mapred.Task: Final Counters for attempt\_local369299969\_0001\_r\_000000\_0: Counters: 30

File System Counters

FILE: Number of bytes read=3500

FILE: Number of bytes written=639455

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=58

HDFS: Number of bytes written=65

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=8

Reduce shuffle bytes=103

Reduce input records=8

Reduce output records=8

Spilled Records=8

Shuffled Maps =1

Failed Shuffles=0

Merged Map outputs=1

GC time elapsed (ms)=0

Total committed heap usage (bytes)=201326592

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=65

2024-04-22 10:33:16,895 INFO mapred.LocalJobRunner: Finishing task: attempt\_local369299969\_0001\_r\_000000\_0

2024-04-22 10:33:16,896 INFO mapred.LocalJobRunner: reduce task executor complete.

2024-04-22 10:33:17,369 INFO mapreduce.Job: Job job\_local369299969\_0001 running in uber mode : false

2024-04-22 10:33:17,371 INFO mapreduce.Job: map 100% reduce 100%

2024-04-22 10:33:17,373 INFO mapreduce.Job: Job job\_local369299969\_0001 completed successfully

2024-04-22 10:33:17,385 INFO mapreduce.Job: Counters: 36

File System Counters

FILE: Number of bytes read=6762

FILE: Number of bytes written=1278807

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=116

HDFS: Number of bytes written=65

HDFS: Number of read operations=15

HDFS: Number of large read operations=0

HDFS: Number of write operations=4

HDFS: Number of bytes read erasure-coded=0

Map-Reduce Framework

Map input records=3

Map output records=10

Map output bytes=98

Map output materialized bytes=103

Input split bytes=107

Combine input records=10

Combine output records=8

Reduce input groups=8

Reduce shuffle bytes=103

Reduce input records=8

Reduce output records=8

Spilled Records=16

Shuffled Maps =1

Failed Shuffles=0

Merged Map outputs=1

GC time elapsed (ms)=7

Total committed heap usage (bytes)=402653184

Shuffle Errors

BAD\_ID=0

CONNECTION=0

IO\_ERROR=0

WRONG\_LENGTH=0

WRONG\_MAP=0

WRONG\_REDUCE=0

File Input Format Counters

Bytes Read=58

File Output Format Counters

Bytes Written=65

hadoop@pict-OptiPlex-3020:~/31181$ hdfs dfs -ls /31165/output

Found 2 items

-rw-r--r-- 3 hadoop supergroup 0 2024-04-22 10:33 /31165/output/\_SUCCESS

-rw-r--r-- 3 hadoop supergroup 65 2024-04-22 10:33 /31165/output/part-r-00000

hadoop@pict-OptiPlex-3020:~/31181$ hdfs dfs -cat /31165/output/part-r-00000

Hello 1

Mumbai. 1

Pune. 1

Vedant 1

Vivek 1

in 2

lives 2

world. 1

hadoop@pict-OptiPlex-3020:~/31181$