hadoop@pict-OptiPlex-3020:~$ pwd

/home/hadoop

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

15884 Jps

4829 NameNode

5198 SecondaryNameNode

hadoop@pict-OptiPlex-3020:~$ hdfs dfs -mkdir /31165/input12

hadoop@pict-OptiPlex-3020:~$ hdfs dfs -put /home/hadoop/31181/a12/input\_weatherDataAverage.txt /31165/input12

hadoop@pict-OptiPlex-3020:~$ hdfs dfs -cat /31165/input12

cat: `/31165/input12': Is a directory

hadoop@pict-OptiPlex-3020:~$ hdfs dfs -cat /31165/input12/input\_weatherDataAverage.txt

51.75 1006.3 71.6

54.74 1006.3 71.6

50.59 1006.3 71.6

51.67 1006.3 71.6

65.67 1006.3 71.6

55.37 1006.3 71.6

49.26 1006.3 71.6

55.44 1006.3 71.6

64.05 1006.3 71.6

68.77 1006.3 71.6

48.93 1006.3 71.6

65.37 1006.3 71.6

69.45 1006.3 71.6

52.91 1006.3 71.6

53.69 1006.3 71.6

53.3 1006.3 71.6

66.17 1006.3 71.6

53.83 1006.3 71.6

50.54 1006.3 71.6

50.27 1006.3 71.6

59.08 1006.3 71.6

53.05 1006.3 71.6

57.97 1006.3 71.6

48.23 1006.3 71.6

47.16 1003.2 53.6

69.72 1003.2 53.6

62.71 1003.2 53.6

46.34 1003.2 53.6

53.15 1003.2 53.6

64.59 1003.2 53.6

58.26 1003.2 53.6

53.27 1003.2 53.6

43.68 1003.2 53.6

65.7 1003.2 53.6

66.27 1003.2 53.6

53.05 1003.2 53.6

68.45 1003.2 53.6

49.03 1003.2 53.6

66.59 1003.2 53.6

63.12 1003.2 53.6

49.13 1003.2 53.6

62.85 1003.2 53.6

64.67 1003.2 53.6

55.73 1003.2 53.6

56.42 1003.2 53.6

53.83 1003.2 53.6

45.14 1003.2 53.6

68.18 1003.2 53.6

48.41 1008.7 62.78

55.12 1008.7 62.78

46.48 1008.7 62.78

54.99 1008.7 62.78

50.62 1008.7 62.78

55.81 1008.7 62.78

59.28 1008.7 62.78

60.55 1008.7 62.78

52.62 1008.7 62.78

62.27 1008.7 62.78

49.9 1008.7 62.78

45.12 1008.7 62.78

62.85 1008.7 62.78

50.9 1008.7 62.78

49.73 1008.7 62.78

47.28 1008.7 62.78

46.67 1008.7 62.78

49.48 1008.7 62.78

59.53 1008.7 62.78

59.49 1008.7 62.78

52.25 1008.7 62.78

57.67 1008.7 62.78

54.04 1008.7 62.78

58.76 1008.7 62.78

48.6 1020.0 51.98

50.54 1020.0 51.98

38.55 1020.0 51.98

50.14 1020.0 51.98

34.86 1020.0 51.98

35.38 1020.0 51.98

41.52 1020.0 51.98

42.28 1020.0 51.98

45.11 1020.0 51.98

51.01 1020.0 51.98

43.6 1020.0 51.98

50.12 1020.0 51.98

55.07 1020.0 51.98

43.24 1020.0 51.98

49.84 1020.0 51.98

44.42 1020.0 51.98

58.77 1020.0 51.98

35.32 1020.0 51.98

34.83 1020.0 51.98

51.72 1020.0 51.98

52.39 1020.0 51.98

57.54 1020.0 51.98

50.37 1020.0 51.98

51.75 1020.0 51.98

hadoop@pict-OptiPlex-3020:~$

—---------------

hadoop@pict-OptiPlex-3020:~/31181/a12$ export HADOOP\_CLASSPATH=(hadoop\_classpath)

hadoop@pict-OptiPlex-3020:~/31181/a12$ echo ${HADOOP\_CLASSPATH}

hadoop\_classpath

hadoop@pict-OptiPlex-3020:~/31181/a12$ export HADOOP\_CLASSPATH=$(hadoop classpath)

hadoop@pict-OptiPlex-3020:~/31181/a12$ 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:~/31181/a12$ javac -classpath ${HADOOP\_CLASSPATH} -d 'home/hadoop/31181/a12/build

> javac -classpath ${HADOOP\_CLASSPATH} -d 'home/hadoop/31181/a12/build'

>

>

>

>

>

>

>

>

>

>

>

>

>

>

>

>

>

>

>

>

> ^C

hadoop@pict-OptiPlex-3020:~/31181/a12$ javac -classpath ${HADOOP\_CLASSPATH} -d 'home/hadoop/31181/a12/build' 'home/hadoop/31181/a12/WeatherDataAverage.java'

error: file not found: home/hadoop/31181/a12/WeatherDataAverage.java

Usage: javac <options> <source files>

use --help for a list of possible options

hadoop@pict-OptiPlex-3020:~/31181/a12$ javac -classpath ${HADOOP\_CLASSPATH} -d '/home/hadoop/31181/a12/build' '/home/hadoop/31181/a12/WeatherDataAverage.java'

hadoop@pict-OptiPlex-3020:~/31181/a12$ jar -cvf a12.jar -C build/.

Error parsing file arguments

Try `jar --help' for more information.

hadoop@pict-OptiPlex-3020:~/31181/a12$ jar -cvf a12.jar -C build/ .

added manifest

adding: WeatherDataAverage.class(in = 1705) (out= 892)(deflated 47%)

adding: WeatherDataAverage$TokenizerMapper.class(in = 2041) (out= 882)(deflated 56%)

adding: WeatherDataAverage$FloatAverageReducer.class(in = 1823) (out= 774)(deflated 57%)

hadoop@pict-OptiPlex-3020:~/31181/a12$

hadoop@pict-OptiPlex-3020:~/31181/a12$ hadoop jar /home/hadoop/31181/a12/a12.jar WeatherDataAverage /31165/input12 /31165/output12

2024-04-22 10:50:43,283 INFO impl.MetricsConfig: Loaded properties from hadoop-metrics2.properties

2024-04-22 10:50:43,361 INFO impl.MetricsSystemImpl: Scheduled Metric snapshot period at 10 second(s).

2024-04-22 10:50:43,361 INFO impl.MetricsSystemImpl: JobTracker metrics system started

2024-04-22 10:50:43,464 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 10:50:43,610 INFO input.FileInputFormat: Total input files to process : 1

2024-04-22 10:50:43,629 INFO mapreduce.JobSubmitter: number of splits:1

2024-04-22 10:50:43,761 INFO mapreduce.JobSubmitter: Submitting tokens for job: job\_local1197278168\_0001

2024-04-22 10:50:43,761 INFO mapreduce.JobSubmitter: Executing with tokens: []

2024-04-22 10:50:43,872 INFO mapreduce.Job: The url to track the job: http://localhost:8080/

2024-04-22 10:50:43,878 INFO mapreduce.Job: Running job: job\_local1197278168\_0001

2024-04-22 10:50:43,881 INFO mapred.LocalJobRunner: OutputCommitter set in config null

2024-04-22 10:50:43,908 INFO output.PathOutputCommitterFactory: No output committer factory defined, defaulting to FileOutputCommitterFactory

2024-04-22 10:50:43,910 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2

2024-04-22 10:50:43,910 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup \_temporary folders under output directory:false, ignore cleanup failures: false

2024-04-22 10:50:43,914 INFO mapred.LocalJobRunner: OutputCommitter is org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter

2024-04-22 10:50:43,976 INFO mapred.LocalJobRunner: Waiting for map tasks

2024-04-22 10:50:43,977 INFO mapred.LocalJobRunner: Starting task: attempt\_local1197278168\_0001\_m\_000000\_0

2024-04-22 10:50:44,000 INFO output.PathOutputCommitterFactory: No output committer factory defined, defaulting to FileOutputCommitterFactory

2024-04-22 10:50:44,000 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2

2024-04-22 10:50:44,000 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup \_temporary folders under output directory:false, ignore cleanup failures: false

2024-04-22 10:50:44,015 INFO mapred.Task: Using ResourceCalculatorProcessTree : [ ]

2024-04-22 10:50:44,022 INFO mapred.MapTask: Processing split: hdfs://127.0.0.1:9000/31165/input12/input\_weatherDataAverage.txt:0+1770

2024-04-22 10:50:44,069 INFO mapred.MapTask: (EQUATOR) 0 kvi 26214396(104857584)

2024-04-22 10:50:44,069 INFO mapred.MapTask: mapreduce.task.io.sort.mb: 100

2024-04-22 10:50:44,069 INFO mapred.MapTask: soft limit at 83886080

2024-04-22 10:50:44,069 INFO mapred.MapTask: bufstart = 0; bufvoid = 104857600

2024-04-22 10:50:44,069 INFO mapred.MapTask: kvstart = 26214396; length = 6553600

2024-04-22 10:50:44,073 INFO mapred.MapTask: Map output collector class = org.apache.hadoop.mapred.MapTask$MapOutputBuffer

2024-04-22 10:50:44,150 INFO mapred.LocalJobRunner:

2024-04-22 10:50:44,151 INFO mapred.MapTask: Starting flush of map output

2024-04-22 10:50:44,151 INFO mapred.MapTask: Spilling map output

2024-04-22 10:50:44,151 INFO mapred.MapTask: bufstart = 0; bufend = 4128; bufvoid = 104857600

2024-04-22 10:50:44,151 INFO mapred.MapTask: kvstart = 26214396(104857584); kvend = 26213248(104852992); length = 1149/6553600

2024-04-22 10:50:44,158 INFO mapred.MapTask: Finished spill 0

2024-04-22 10:50:44,167 INFO mapred.Task: Task:attempt\_local1197278168\_0001\_m\_000000\_0 is done. And is in the process of committing

2024-04-22 10:50:44,170 INFO mapred.LocalJobRunner: map

2024-04-22 10:50:44,170 INFO mapred.Task: Task 'attempt\_local1197278168\_0001\_m\_000000\_0' done.

2024-04-22 10:50:44,175 INFO mapred.Task: Final Counters for attempt\_local1197278168\_0001\_m\_000000\_0: Counters: 23

File System Counters

FILE: Number of bytes read=3562

FILE: Number of bytes written=648850

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=1770

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=96

Map output records=288

Map output bytes=4128

Map output materialized bytes=4710

Input split bytes=129

Combine input records=0

Spilled Records=288

Failed Shuffles=0

Merged Map outputs=0

GC time elapsed (ms)=7

Total committed heap usage (bytes)=294649856

File Input Format Counters

Bytes Read=1770

2024-04-22 10:50:44,175 INFO mapred.LocalJobRunner: Finishing task: attempt\_local1197278168\_0001\_m\_000000\_0

2024-04-22 10:50:44,176 INFO mapred.LocalJobRunner: map task executor complete.

2024-04-22 10:50:44,180 INFO mapred.LocalJobRunner: Waiting for reduce tasks

2024-04-22 10:50:44,180 INFO mapred.LocalJobRunner: Starting task: attempt\_local1197278168\_0001\_r\_000000\_0

2024-04-22 10:50:44,190 INFO output.PathOutputCommitterFactory: No output committer factory defined, defaulting to FileOutputCommitterFactory

2024-04-22 10:50:44,190 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2

2024-04-22 10:50:44,190 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup \_temporary folders under output directory:false, ignore cleanup failures: false

2024-04-22 10:50:44,191 INFO mapred.Task: Using ResourceCalculatorProcessTree : [ ]

2024-04-22 10:50:44,195 INFO mapred.ReduceTask: Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@3dcd4a6e

2024-04-22 10:50:44,197 WARN impl.MetricsSystemImpl: JobTracker metrics system already initialized!

2024-04-22 10:50:44,211 INFO reduce.MergeManagerImpl: MergerManager: memoryLimit=1443050240, maxSingleShuffleLimit=360762560, mergeThreshold=952413184, ioSortFactor=10, memToMemMergeOutputsThreshold=10

2024-04-22 10:50:44,213 INFO reduce.EventFetcher: attempt\_local1197278168\_0001\_r\_000000\_0 Thread started: EventFetcher for fetching Map Completion Events

2024-04-22 10:50:44,243 INFO reduce.LocalFetcher: localfetcher#1 about to shuffle output of map attempt\_local1197278168\_0001\_m\_000000\_0 decomp: 4706 len: 4710 to MEMORY

2024-04-22 10:50:44,251 INFO reduce.InMemoryMapOutput: Read 4706 bytes from map-output for attempt\_local1197278168\_0001\_m\_000000\_0

2024-04-22 10:50:44,256 INFO reduce.MergeManagerImpl: closeInMemoryFile -> map-output of size: 4706, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->4706

2024-04-22 10:50:44,258 INFO reduce.EventFetcher: EventFetcher is interrupted.. Returning

2024-04-22 10:50:44,259 INFO mapred.LocalJobRunner: 1 / 1 copied.

2024-04-22 10:50:44,259 INFO reduce.MergeManagerImpl: finalMerge called with 1 in-memory map-outputs and 0 on-disk map-outputs

2024-04-22 10:50:44,264 INFO mapred.Merger: Merging 1 sorted segments

2024-04-22 10:50:44,264 INFO mapred.Merger: Down to the last merge-pass, with 1 segments left of total size: 4695 bytes

2024-04-22 10:50:44,266 INFO reduce.MergeManagerImpl: Merged 1 segments, 4706 bytes to disk to satisfy reduce memory limit

2024-04-22 10:50:44,267 INFO reduce.MergeManagerImpl: Merging 1 files, 4710 bytes from disk

2024-04-22 10:50:44,267 INFO reduce.MergeManagerImpl: Merging 0 segments, 0 bytes from memory into reduce

2024-04-22 10:50:44,267 INFO mapred.Merger: Merging 1 sorted segments

2024-04-22 10:50:44,268 INFO mapred.Merger: Down to the last merge-pass, with 1 segments left of total size: 4695 bytes

2024-04-22 10:50:44,268 INFO mapred.LocalJobRunner: 1 / 1 copied.

2024-04-22 10:50:44,308 INFO Configuration.deprecation: mapred.skip.on is deprecated. Instead, use mapreduce.job.skiprecords

2024-04-22 10:50:44,399 INFO mapred.Task: Task:attempt\_local1197278168\_0001\_r\_000000\_0 is done. And is in the process of committing

2024-04-22 10:50:44,402 INFO mapred.LocalJobRunner: 1 / 1 copied.

2024-04-22 10:50:44,402 INFO mapred.Task: Task attempt\_local1197278168\_0001\_r\_000000\_0 is allowed to commit now

2024-04-22 10:50:44,434 INFO output.FileOutputCommitter: Saved output of task 'attempt\_local1197278168\_0001\_r\_000000\_0' to hdfs://127.0.0.1:9000/31165/output12

2024-04-22 10:50:44,436 INFO mapred.LocalJobRunner: reduce > reduce

2024-04-22 10:50:44,436 INFO mapred.Task: Task 'attempt\_local1197278168\_0001\_r\_000000\_0' done.

2024-04-22 10:50:44,437 INFO mapred.Task: Final Counters for attempt\_local1197278168\_0001\_r\_000000\_0: Counters: 30

File System Counters

FILE: Number of bytes read=13014

FILE: Number of bytes written=653560

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=1770

HDFS: Number of bytes written=60

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=3

Reduce shuffle bytes=4710

Reduce input records=288

Reduce output records=3

Spilled Records=288

Shuffled Maps =1

Failed Shuffles=0

Merged Map outputs=1

GC time elapsed (ms)=0

Total committed heap usage (bytes)=294649856

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=60

2024-04-22 10:50:44,441 INFO mapred.LocalJobRunner: Finishing task: attempt\_local1197278168\_0001\_r\_000000\_0

2024-04-22 10:50:44,442 INFO mapred.LocalJobRunner: reduce task executor complete.

2024-04-22 10:50:44,897 INFO mapreduce.Job: Job job\_local1197278168\_0001 running in uber mode : false

2024-04-22 10:50:44,899 INFO mapreduce.Job: map 100% reduce 100%

2024-04-22 10:50:44,902 INFO mapreduce.Job: Job job\_local1197278168\_0001 completed successfully

2024-04-22 10:50:44,911 INFO mapreduce.Job: Counters: 36

File System Counters

FILE: Number of bytes read=16576

FILE: Number of bytes written=1302410

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=3540

HDFS: Number of bytes written=60

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=96

Map output records=288

Map output bytes=4128

Map output materialized bytes=4710

Input split bytes=129

Combine input records=0

Combine output records=0

Reduce input groups=3

Reduce shuffle bytes=4710

Reduce input records=288

Reduce output records=3

Spilled Records=576

Shuffled Maps =1

Failed Shuffles=0

Merged Map outputs=1

GC time elapsed (ms)=7

Total committed heap usage (bytes)=589299712

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=1770

File Output Format Counters

Bytes Written=60

hadoop@pict-OptiPlex-3020:~/31181/a12$ hdfs dfs -cat /31165/output12/part-r-00000

DewPoint 59.99005

Temperature 53.58261

WindSpeed 1009.55023

hadoop@pict-OptiPlex-3020:~/31181/a12$