hadoop@pict-OptiPlex-5070:~/Desktop/final\_weather$ mkdir build

hadoop@pict-OptiPlex-5070:~/Desktop/final\_weather$ javac -classpath /home/hadoop/hadoop-3.3.5/share/hadoop/common/hadoop-common-3.3.5.jar:/home/hadoop/hadoop-3.3.5/share/hadoop/mapreduce/hadoop-mapreduce-client-core-3.3.5.jar -d ./build /home/hadoop/Desktop/final\_weather/WeatherDataAverage.java

hadoop@pict-OptiPlex-5070:~/Desktop/final\_weather$ jar -cvf weatherData.jar -C build/ .

added manifest

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

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

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

hadoop@pict-OptiPlex-5070:~/Desktop/final\_weather$ start-dfs.sh

Starting namenodes on [localhost]

Starting datanodes

Starting secondary namenodes [pict-OptiPlex-5070]

hadoop@pict-OptiPlex-5070:~/Desktop/final\_weather$ start-yarn.sh

Starting resourcemanager

Starting nodemanagers

hadoop@pict-OptiPlex-5070:~/Desktop/final\_weather$ hdfs dfs -mkdir /user39

hadoop@pict-OptiPlex-5070:~/Desktop/final\_weather$ hdfs dfs -mkdir /user39/input

hadoop@pict-OptiPlex-5070:~/Desktop/final\_weather$ hdfs dfs -put weather\_input.txt /user39/input

hadoop@pict-OptiPlex-5070:~/Desktop/final\_weather$ hdfs dfs -ls /user39/input

Found 1 items

-rw-r--r-- 1 hadoop supergroup 1770 2024-04-16 12:15 /user39/input/weather\_input.txt

hadoop@pict-OptiPlex-5070:~/Desktop/final\_weather$ hadoop jar './weatherData.jar' WeatherDataAverage /user39/input /user39/output

2024-04-16 12:16:51,116 INFO client.DefaultNoHARMFailoverProxyProvider: Connecting to ResourceManager at /127.0.0.1:8032

2024-04-16 12:16:51,295 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.

2024-04-16 12:16:51,333 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/hadoop/.staging/job\_1713249890056\_0001

2024-04-16 12:16:51,588 INFO input.FileInputFormat: Total input files to process : 1

2024-04-16 12:16:51,694 INFO mapreduce.JobSubmitter: number of splits:1

2024-04-16 12:16:52,253 INFO mapreduce.JobSubmitter: Submitting tokens for job: job\_1713249890056\_0001

2024-04-16 12:16:52,254 INFO mapreduce.JobSubmitter: Executing with tokens: []

2024-04-16 12:16:52,376 INFO conf.Configuration: resource-types.xml not found

2024-04-16 12:16:52,376 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.

2024-04-16 12:16:52,510 INFO impl.YarnClientImpl: Submitted application application\_1713249890056\_0001

2024-04-16 12:16:52,541 INFO mapreduce.Job: The url to track the job: http://pict-OptiPlex-5070:8088/proxy/application\_1713249890056\_0001/

2024-04-16 12:16:52,542 INFO mapreduce.Job: Running job: job\_1713249890056\_0001

2024-04-16 12:16:57,598 INFO mapreduce.Job: Job job\_1713249890056\_0001 running in uber mode : false

2024-04-16 12:16:57,600 INFO mapreduce.Job: map 0% reduce 0%

2024-04-16 12:17:00,653 INFO mapreduce.Job: map 100% reduce 0%

2024-04-16 12:17:04,687 INFO mapreduce.Job: map 100% reduce 100%

2024-04-16 12:17:05,708 INFO mapreduce.Job: Job job\_1713249890056\_0001 completed successfully

2024-04-16 12:17:05,754 INFO mapreduce.Job: Counters: 54

File System Counters

FILE: Number of bytes read=4710

FILE: Number of bytes written=562097

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=1887

HDFS: Number of bytes written=60

HDFS: Number of read operations=8

HDFS: Number of large read operations=0

HDFS: Number of write operations=2

HDFS: Number of bytes read erasure-coded=0

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)=1371

Total time spent by all reduces in occupied slots (ms)=1517

Total time spent by all map tasks (ms)=1371

Total time spent by all reduce tasks (ms)=1517

Total vcore-milliseconds taken by all map tasks=1371

Total vcore-milliseconds taken by all reduce tasks=1517

Total megabyte-milliseconds taken by all map tasks=1403904

Total megabyte-milliseconds taken by all reduce tasks=1553408

Map-Reduce Framework

Map input records=96

Map output records=288

Map output bytes=4128

Map output materialized bytes=4710

Input split bytes=117

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)=29

CPU time spent (ms)=970

Physical memory (bytes) snapshot=485785600

Virtual memory (bytes) snapshot=5591339008

Total committed heap usage (bytes)=312475648

Peak Map Physical memory (bytes)=280535040

Peak Map Virtual memory (bytes)=2795143168

Peak Reduce Physical memory (bytes)=205250560

Peak Reduce Virtual memory (bytes)=2796195840

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-5070:~/Desktop/final\_weather$ hdfs dfs -ls /user39/output

Found 2 items

-rw-r--r-- 1 hadoop supergroup 0 2024-04-16 12:17 /user39/output/\_SUCCESS

-rw-r--r-- 1 hadoop supergroup 60 2024-04-16 12:17 /user39/output/part-r-00000

hadoop@pict-OptiPlex-5070:~/Desktop/final\_weather$ hdfs dfs -cat /user39/output/part-r-00000

DewPoint 59.99005

Temperature 53.58261

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

Java Code for WeatherDataAverage

import java.io.\*;

import org.apache.hadoop.conf.\*;

import org.apache.hadoop.fs.\*;

import org.apache.hadoop.io.\*;

import org.apache.hadoop.mapreduce.\*;

import org.apache.hadoop.mapreduce.lib.input.\*;

import org.apache.hadoop.mapreduce.lib.output.\*;

public class WeatherDataAverage {

public static class TokenizerMapper extends Mapper<Object, Text, Text, FloatWritable> {

private Text category = new Text();

private FloatWritable temperature = new FloatWritable();

private FloatWritable windSpeed = new FloatWritable();

private FloatWritable dewPoint = new FloatWritable();

public void map(Object key, Text value, Context context) throws IOException, InterruptedException {

String[] cols = value.toString().split(" ");

float temp = Float.parseFloat(cols[0]);

float wind = Float.parseFloat(cols[1]);

float dew = Float.parseFloat(cols[2]);

category.set("Temperature");

temperature.set(temp);

context.write(category, temperature);

category.set("WindSpeed");

windSpeed.set(wind);

context.write(category, windSpeed);

category.set("DewPoint");

dewPoint.set(dew);

context.write(category, dewPoint);

}

}

public static class FloatAverageReducer extends Reducer<Text, FloatWritable, Text, FloatWritable> {

private FloatWritable result = new FloatWritable();

public void reduce(Text key, Iterable<FloatWritable> values, Context context)

throws IOException, InterruptedException {

float sum = 0;

int count = 0;

for (FloatWritable a : values){sum += a.get();count++;}

float avg = sum / count;

result.set(avg);

context.write(key, result);

}

}

public static void main(String[] args) throws Exception {

Configuration conf = new Configuration();

Job job = Job.getInstance(conf, "weather data average");

job.setJarByClass(WeatherDataAverage.class);

job.setMapperClass(TokenizerMapper.class);

job.setReducerClass(FloatAverageReducer.class);

job.setMapOutputKeyClass(Text.class);

job.setMapOutputValueClass(FloatWritable.class);

job.setOutputKeyClass(Text.class);

job.setOutputValueClass(FloatWritable.class);

job.setInputFormatClass(TextInputFormat.class);

job.setOutputFormatClass(TextOutputFormat.class);

TextInputFormat.addInputPath(job, new Path(args[0]));

TextOutputFormat.setOutputPath(job, new Path(args[1]));

System.exit(job.waitForCompletion(true) ? 0 : 1);

}

}