

EXP NO: 3 MAP REDUCE PROGRAM TO PROCESS A WEATHER DATASET

\$cd DA-Lab

\$mkdir exp3

\$cd exp3

\$nano dataset.txt

\$nano mapper.py

\$nano reducer.py

```
[hadoop@fedora da]$ cd exp3
[hadoop@fedora exp3]$ nano mapper.py
[hadoop@fedora exp3]$ nano reducer.py
[hadoop@fedora exp3]$ chmod 777 mapper.py reducer.py
[hadoop@fedora exp3]$ hadoop fs -mkdir -p /weatherdata
[hadoop@fedora exp3]$ hadoop fs -copyFromLocal /home/hadoop/Downloads/dataset.txt /weatherdata
[hadoop@fedora exp3]$ hadoop jar -hadoop /usr/hadoop/hadoop-3.3.6-jar -input /weatherdata/dataset.txt -output /weatherdata/new_output -mapper -jd/da/exp3/mapper.py -reducer -da/exp3/reducer.py
2024-09-11 14:13:47,831 INFO Impl.MetricConfig: loaded properties from hadoop-metrics2.properties
2024-09-11 14:13:48,881 INFO Impl.MetricsSystemImpl: Scheduled Metric snapshot period at 10 second(s).
2024-09-11 14:13:48,882 INFO Impl.MetricsSystemImpl: JobTracker metrics system started
2024-09-11 14:13:48,123 WARN Impl.MetricsSystemImpl: JobTracker metrics system already initialized!
2024-09-11 14:13:48,794 INFO mapred.FileInputFormat: Total input files to process : 1
2024-09-11 14:13:48,852 INFO mapreduce.JobSubmitter: number of splits:1
2024-09-11 14:13:49,495 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_local961705618_0001
2024-09-11 14:13:49,495 INFO mapreduce.JobSubmitter: Executing with tokens: []
2024-09-11 14:13:49,765 INFO mapreduce.Job: The url to track the job: http://localhost:8880/
2024-09-11 14:13:49,771 INFO mapreduce.Job: Running job: job_local961705618_0001
2024-09-11 14:13:49,772 INFO mapred.LocalJobRunner: OutputCommitter set in config null
2024-09-11 14:13:49,780 INFO mapred.LocalJobRunner: OutputCommitter is org.apache.hadoop.mapred.FileOutputCommitter
2024-09-11 14:13:49,793 INFO output.FileOutputCommitter: File OutputCommitter Algorithm version is 2
2024-09-11 14:13:49,793 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup_temporary folders under output directory:false, ignore cleanup failures: false
2024-09-11 14:13:50,181 INFO mapred.LocalJobRunner: Waiting for map tasks
2024-09-11 14:13:50,489 INFO mapred.LocalJobRunner: Starting task: attempt_local961705618_0001_0_000008_0
2024-09-11 14:13:50,155 INFO output.FileOutputCommitter: File OutputCommitter Algorithm version is 2
2024-09-11 14:13:50,155 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup_temporary folders under output directory:false, ignore cleanup failures: false
2024-09-11 14:13:50,193 INFO mapred.Task: Using ResourceCalculatorProcessTree: {}
2024-09-11 14:13:50,214 INFO mapred.MapTask: Processing split: hdfs://localhost:9000/weatherdata/dataset.txt:1=9568
2024-09-11 14:13:50,280 INFO mapred.MapTask: numReduceTasks: 1
2024-09-11 14:13:50,343 INFO mapred.MapTask: (EQ=ATQ) = but 26214396 (104857504)
2024-09-11 14:13:50,361 INFO mapred.MapTask: mapreduce.task.io.sort.mb: 100
2024-09-11 14:13:50,362 INFO mapred.MapTask: soft limit at 83866880
2024-09-11 14:13:50,362 INFO mapred.MapTask: bufsize = 0, buffered = 104857600
2024-09-11 14:13:50,362 INFO mapred.MapTask: kvsyst = 26214396; length = 5535968
2024-09-11 14:13:50,372 INFO mapred.MapTask: Map output collector class = org.apache.hadoop.mapred.MapTaskMapOutputBuffer
2024-09-11 14:13:50,389 INFO streaming.PipeMapRed: PipeMapRed exec: [/home/hadoop/da/exp3/mapper.py]
2024-09-11 14:13:50,419 INFO Configuration.deprecation: mapred.work.output.dir is deprecated. Instead, use mapreduce.task.output.dir
2024-09-11 14:13:50,414 INFO Configuration.deprecation: mapred.local.dir is deprecated. Instead, use mapreduce.cluster.local.dir
2024-09-11 14:13:50,415 INFO Configuration.deprecation: map.input.file is deprecated. Instead, use mapreduce.map.input.file
2024-09-11 14:13:50,416 INFO Configuration.deprecation: map.input.length is deprecated. Instead, use mapreduce.map.input.length
2024-09-11 14:13:50,418 INFO Configuration.deprecation: mapred.job.id is deprecated. Instead, use mapreduce.job.id
2024-09-11 14:13:50,419 INFO Configuration.deprecation: mapred.task.partition is deprecated. Instead, use mapreduce.task.partition
2024-09-11 14:13:50,424 INFO Configuration.deprecation: map.input.start is deprecated. Instead, use mapreduce.map.input.start
2024-09-11 14:13:50,425 INFO Configuration.deprecation: mapred.task.id is deprecated. Instead, use mapreduce.task.attempt.id
2024-09-11 14:13:50,425 INFO Configuration.deprecation: mapred.task.id is deprecated. Instead, use mapreduce.task.attempt.id
2024-09-11 14:13:50,427 INFO Configuration.deprecation: mapred.tip.id is deprecated. Instead, use mapreduce.task.id
2024-09-11 14:13:50,428 INFO Configuration.deprecation: mapred.slp.on is deprecated. Instead, use mapreduce.job.slprecords
2024-09-11 14:13:50,433 INFO Configuration.deprecation: user.name is deprecated. Instead, use mapreduce.job.user.name
2024-09-11 14:13:50,497 INFO streaming.PipeMapRed: R/W=1/0/0 tInNA [rec/s] outNA [rec/s]
2024-09-11 14:13:50,497 INFO streaming.PipeMapRed: R/W=1/0/0 tInNA [rec/s] outNA [rec/s]
2024-09-11 14:13:50,592 INFO streaming.PipeMapRed: R/W=1/0/0 tInNA [rec/s] outNA [rec/s]
2024-09-11 14:13:50,727 INFO streaming.PipeMapRed: Records R/W=365/1
2024-09-11 14:13:50,738 INFO streaming.PipeMapRed: MRErrorThread done
2024-09-11 14:13:50,739 INFO streaming.PipeMapRed: mapRedFinished
2024-09-11 14:13:50,749 INFO mapred.LocalJobRunner:
2024-09-11 14:13:50,749 INFO mapred.MapTask: Starting flush of map output
```

```
hadoop@fedora:~$ find /usr/lib/jvm/java-11-openjdk/bin/java -Dprog.jar -Djava.library.path=/home/hadoop/lib/native -Xmx1000m -Dpig.log.dir=/home/hadoop/pig/logs -Dpig.log.file=pig.log -Dpig.home.dir=/home/hadoop/pig -Dyarn.log.dir=/
```

```
storage/avro/avro_test_files/test_dir/  
pig-0.16.0/contrib/piggybank/java/src/test/java/org/apache/pig/piggybank/test/storage/avro_avro_test_files/test_dir1/test_subdir/  
pig-0.16.0/contrib/piggybank/java/src/test/java/org/apache/pig/piggybank/test/storage/avro_avro_test_files/test_dir/  
piggybank/java/src/test/java/org/apache/pig/piggybank/test/storage/avro_avro_test_files/test_merge_schemas_default/  
pig-0.16.0/contrib/piggybank/java/src/test/java/org/apache/pig/piggybank/test/storage/avro_avro_test_files/test_primitive_types/  
pig-0.16.0/contrib/piggybank/java/src/test/java/org/apache/pig/piggybank/test/storage/avro_avro_test_files/test_user_defined_load_schema/  
pig-0.16.0/docs/  
pig-0.16.0/docs/api/  
pig-0.16.0/docs/api/org/  
pig-0.16.0/docs/api/org/apache/  
pig-0.16.0/docs/api/org/apache/hadoop/  
pig-0.16.0/docs/api/org/apache/hadoop/hive/  
pig-0.16.0/docs/api/org/apache/hadoop/hive/serde2/  
pig-0.16.0/docs/api/org/apache/hadoop/hive/serde2/objectinspector/  
pig-0.16.0/docs/api/org/apache/hadoop/hive/serde2/objectinspector/primitive/  
pig-0.16.0/docs/api/org/apache/hadoop/hive/serde2/objectinspector/primitive/class-use/  
pig-0.16.0/docs/api/org/apache/pig/  
pig-0.16.0/docs/api/org/apache/pig/backend/  
pig-0.16.0/docs/api/org/apache/pig/backend/class-use/  
pig-0.16.0/docs/api/org/apache/pig/backend/dataloader/  
pig-0.16.0/docs/api/org/apache/pig/backend/dataloader/class-use/  
pig-0.16.0/docs/api/org/apache/pig/backend/exeuctionengine/  
pig-0.16.0/docs/api/org/apache/pig/backend/exeuctionengine/class-use/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/accumulo/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/accumulo/class-use/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/class-use/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/dataloader/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/dataloader/class-use/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/exeuctionengine/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/exeuctionengine/class-use/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/exeuctionengine/fetch/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/exeuctionengine/fetch/class-use/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/exeuctionengine/mapreducelayer/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/exeuctionengine/mapreducelayer/class-use/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/exeuctionengine/mapreducelayer/partitioners/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/exeuctionengine/mapreducelayer/partitioners/class-use/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/exeuctionengine/mapreducelayer/plans/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/exeuctionengine/mapreducelayer/plans/class-use/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/exeuctionengine/optimize/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/exeuctionengine/optimize/class-use/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/exeuctionengine/physicallayer/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/exeuctionengine/physicalLayer/class-use/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/exeuctionengine/physicalLayer/expressionoperators/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/exeuctionengine/physicalLayer/expressionOperators/class-use/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/exeuctionengine/physicalLayer/expressionOperators/regx/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/exeuctionengine/physicalLayer/expressionOperators/regx/class-use/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/exeuctionengine/physicalLayer/plans/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/exeuctionengine/physicalLayer/plans/class-use/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/exeuctionengine/physicalLayer/relationOperators/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/exeuctionengine/physicalLayer/relationOperators/class-use/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/exeuctionengine/physicalLayer/util/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/exeuctionengine/physicalLayer/util/class-use/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/exeuctionengine/tet/  
pig-0.16.0/docs/api/org/apache/pig/backend/hadoop/exeuctionengine/tet/class-use/
```

```
$start-all.sh
```

```
$ jps
```

```
$hdfs dfs -mkdir /exp3
```

```
$hdfs dfs -copyFromLocal ~/DA-Lab/exp3/dataset.txt /exp2
```

```
$chmod 777 mapper.py reducer.py
```

```
$hadoop jar $HADOOP_STREAMING -input /exp3/dataset.txt -output /exp3/output -mapper  
~/DA-Lab/exp3/mapper.py -reducer ~/DA-Lab/exp3/reducer.py
```

```
hadoop@fedora:~$ ./src/lib/hym/java-11-openjdk/bin/java -Dproc.jar -Djava.library.path=/home/hadoop/hadoop/lib/native -Xmx1000m -Dpig.log.dir=/home/hadoop/pig/logs -Dpig.log.file=pig.log -Dpig.home=/home/hadoop/pig -Dyarn.log.dir=...  
pig-0.16.0/test/unit-tests  
pig-0.16.0/tutorial/build.xml  
pig-0.16.0/tutorial/build/classes/org/apache/pig/tutorial/ExtractHour.class  
pig-0.16.0/tutorial/build/classes/org/apache/pig/tutorial/NgramGenerator.class  
pig-0.16.0/tutorial/build/classes/org/apache/pig/tutorial/NonURLDetector.class  
pig-0.16.0/tutorial/build/classes/org/apache/pig/tutorial/ScoreGenerator.class  
pig-0.16.0/tutorial/build/classes/org/apache/pig/tutorial/Tolower.class  
pig-0.16.0/tutorial/build/classes/org/apache/pig/tutorial/tutorialTest.class  
pig-0.16.0/tutorial/build/classes/org/apache/pig/tutorial/tutorialUtil.class  
pig-0.16.0/tutorial/build/output/pigtmp/excite-small.log  
pig-0.16.0/tutorial/build/output/pigtmp/excite-log.b22  
pig-0.16.0/tutorial/build/output/pigtmp/script1-hadoop.pig  
pig-0.16.0/tutorial/build/output/pigtmp/script1-local.pig  
pig-0.16.0/tutorial/build/output/pigtmp/script2-hadoop.pig  
pig-0.16.0/tutorial/build/output/pigtmp/script2-local.pig  
pig-0.16.0/tutorial/build/output/pigtmp/tutorial.jar  
pig-0.16.0/tutorial/build/pigtutorial.tar  
pig-0.16.0/tutorial/data/excite-small.log  
pig-0.16.0/tutorial/data/excite-log.b22  
pig-0.16.0/tutorial/pigtutorial.tar.gz  
pig-0.16.0/tutorial/scripts/script1-hadoop.pig  
pig-0.16.0/tutorial/scripts/script1-local.pig  
pig-0.16.0/tutorial/scripts/script2-hadoop.pig  
pig-0.16.0/tutorial/scripts/script2-local.pig  
pig-0.16.0/tutorial/src/org/apache/pig/tutorial/ExtractHour.java  
pig-0.16.0/tutorial/src/org/apache/pig/tutorial/NgramGenerator.java  
pig-0.16.0/tutorial/src/org/apache/pig/tutorial/NonURLDetector.java  
pig-0.16.0/tutorial/src/org/apache/pig/tutorial/ScoreGenerator.java  
pig-0.16.0/tutorial/src/org/apache/pig/tutorial/Tolower.java  
pig-0.16.0/tutorial/src/org/apache/pig/tutorial/tutorialTest.java  
pig-0.16.0/tutorial/src/org/apache/pig/tutorial/tutorialUtil.java  
pig-0.16.0/bin/pig  
pig-0.16.0/bin/pig.cmd  
pig-0.16.0/bin/pig.py  
[hadoop@fedora ~]$ sudo mv /home/hadoop/pig-0.16.0 /home/hadoop/pig  
(sudo) password for hadoop:  
[hadoop@fedora ~]$ sudo nano .bashrc  
[hadoop@fedora ~]$ sudo nano .bashrc  
[hadoop@fedora ~]$ source .bashrc  
[hadoop@fedora ~]$ ./start-dfs.sh  
[hadoop@fedora ~]$ ./start-yarns jps  
bash: ./start-dfs.sh: No such file or directory  
[hadoop@fedora ~]$ pig  
2024-09-11 14:24:17,530 INFO pig.ExecTypeProvider: Trying ExecType : LOCAL  
2024-09-11 14:24:17,533 INFO pig.ExecTypeProvider: Picked MAPREDUCE as the ExecType  
2024-09-11 14:24:17,534 INFO pig.ExecTypeProvider: Picked MAPREDUCE as the ExecType  
2024-09-11 14:24:17,604 (main) INFO org.apache.pig.Main - Apache Pig version 0.16.0 (r1746530) compiled Jun 01 2016, 23:10:49  
2024-09-11 14:24:17,605 (main) INFO org.apache.pig.Main - Logging error messages to: /home/hadoop/pig_1726044857669.log  
2024-09-11 14:24:17,925 (main) INFO org.apache.pig.impl.util.Utils - Default bootstrap file /home/hadoop/.pigbootstrap not found  
2024-09-11 14:24:18,799 (main) INFO org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address  
2024-09-11 14:24:18,799 (main) INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS  
2024-09-11 14:24:18,799 (main) INFO org.apache.pig.backend.hadoop.executionengine.MExecutionEngine - Connecting to hadoop file system at: hdfs://localhost:9000  
2024-09-11 14:24:20,183 (main) INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS  
2024-09-11 14:24:20,228 (main) INFO org.apache.pig.PigServer - Pig script ID for the session: PIG-default-fe6d793-ea34-4f62-bedb-45665ab7e22  
2024-09-11 14:24:20,229 (main) WARN org.apache.pig.PigServer - ATS is disabled since yarn.timeline-service.enabled set to false  
grunt> quit
```

```
hadoop@fedora:~$  
pig-0.16.0/tutorial/build/classes/org/apache/pig/tutorial/ExtractHour.class  
pig-0.16.0/tutorial/build/classes/org/apache/pig/tutorial/NgramGenerator.class  
pig-0.16.0/tutorial/build/classes/org/apache/pig/tutorial/NonURLDetector.class  
pig-0.16.0/tutorial/build/classes/org/apache/pig/tutorial/ScoreGenerator.class  
pig-0.16.0/tutorial/build/classes/org/apache/pig/tutorial/Tolower.class  
pig-0.16.0/tutorial/build/classes/org/apache/pig/tutorial/tutorialTest.class  
pig-0.16.0/tutorial/build/classes/org/apache/pig/tutorial/tutorialUtil.class  
pig-0.16.0/tutorial/build/output/pigtmp/excite-small.log  
pig-0.16.0/tutorial/build/output/pigtmp/excite-log.b22  
pig-0.16.0/tutorial/build/output/pigtmp/script1-hadoop.pig  
pig-0.16.0/tutorial/build/output/pigtmp/script1-local.pig  
pig-0.16.0/tutorial/build/output/pigtmp/script2-hadoop.pig  
pig-0.16.0/tutorial/build/output/pigtmp/script2-local.pig  
pig-0.16.0/tutorial/build/output/pigtmp/tutorial.jar  
pig-0.16.0/tutorial/build/pigtutorial.tar  
pig-0.16.0/tutorial/data/excite-small.log  
pig-0.16.0/tutorial/data/excite-log.b22  
pig-0.16.0/tutorial/pigtutorial.tar.gz  
pig-0.16.0/tutorial/scripts/script1-hadoop.pig  
pig-0.16.0/tutorial/scripts/script1-local.pig  
pig-0.16.0/tutorial/scripts/script2-hadoop.pig  
pig-0.16.0/tutorial/scripts/script2-local.pig  
pig-0.16.0/tutorial/src/org/apache/pig/tutorial/ExtractHour.java  
pig-0.16.0/tutorial/src/org/apache/pig/tutorial/NgramGenerator.java  
pig-0.16.0/tutorial/src/org/apache/pig/tutorial/NonURLDetector.java  
pig-0.16.0/tutorial/src/org/apache/pig/tutorial/ScoreGenerator.java  
pig-0.16.0/tutorial/src/org/apache/pig/tutorial/Tolower.java  
pig-0.16.0/tutorial/src/org/apache/pig/tutorial/tutorialTest.java  
pig-0.16.0/tutorial/src/org/apache/pig/tutorial/tutorialUtil.java  
pig-0.16.0/bin/pig  
pig-0.16.0/bin/pig.cmd  
pig-0.16.0/bin/pig.py  
[hadoop@fedora ~]$ sudo mv /home/hadoop/pig-0.16.0 /home/hadoop/pig  
(sudo) password for hadoop:  
[hadoop@fedora ~]$ sudo nano .bashrc  
[hadoop@fedora ~]$ sudo nano .bashrc  
[hadoop@fedora ~]$ source .bashrc  
[hadoop@fedora ~]$ ./start-dfs.sh  
[hadoop@fedora ~]$ ./start-yarns jps  
bash: ./start-dfs.sh: No such file or directory  
[hadoop@fedora ~]$ pig  
2024-09-11 14:24:17,530 INFO pig.ExecTypeProvider: Trying ExecType : LOCAL  
2024-09-11 14:24:17,533 INFO pig.ExecTypeProvider: Picked MAPREDUCE as the ExecType  
2024-09-11 14:24:17,534 INFO pig.ExecTypeProvider: Picked MAPREDUCE as the ExecType  
2024-09-11 14:24:17,604 (main) INFO org.apache.pig.Main - Apache Pig version 0.16.0 (r1746530) compiled Jun 01 2016, 23:10:49  
2024-09-11 14:24:17,605 (main) INFO org.apache.pig.Main - Logging error messages to: /home/hadoop/pig_1726044857669.log  
2024-09-11 14:24:17,925 (main) INFO org.apache.pig.impl.util.Utils - Default bootstrap file /home/hadoop/.pigbootstrap not found  
2024-09-11 14:24:18,799 (main) INFO org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address  
2024-09-11 14:24:18,799 (main) INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS  
2024-09-11 14:24:18,799 (main) INFO org.apache.pig.backend.hadoop.executionengine.MExecutionEngine - Connecting to hadoop file system at: hdfs://localhost:9000  
2024-09-11 14:24:20,183 (main) INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS  
2024-09-11 14:24:20,228 (main) INFO org.apache.pig.PigServer - Pig script ID for the session: PIG-default-fe6d793-ea34-4f62-bedb-45665ab7e22  
2024-09-11 14:24:20,229 (main) WARN org.apache.pig.PigServer - ATS is disabled since yarn.timeline-service.enabled set to false  
grunt> quit  
2024-09-11 14:25:27,771 (main) INFO org.apache.pig.Main - Pig script completed in 1 minute, 10 seconds and 347 milliseconds (70347 ms)  
[hadoop@fedora ~]$
```

\$hdfs dfs -cat /exp3/output/*

```
hadoop@fedora:~/da/exp3
FILE: Number of bytes read=290172
FILE: Number of bytes written=1582686
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=159136
HDFS: Number of bytes written=96
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=365
  Map output records=365
  Map output bytes=2916
  Map output materialized bytes=3652
  Input split bytes=97
  Combine input records=0
  Combine output records=0
  Reduce input groups=12
  Reduce shuffle bytes=3652
  Reduce input records=365
  Reduce output records=12
  Spilled Records=730
  Shuffled Maps =1
  Failed Shuffles=0
  Merged Map outputs=1
  GC time elapsed (ms)=20
  Total committed heap usage (bytes)=574619648
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=79558
File Output Format Counters
  Bytes Written=96
2024-09-11 14:13:52.431 INFO streaming.StreamJob: Outout directory: /weatherdata/new_output
hadoop@fedora exp3$ hdfs dfs -cat /weatherdata/new_output/*
91 26.5
92 26.6
93 29.1
94 30.8
95 31.1
96 33.6
97 38.5
98 40.2
99 36.5
10 36.9
11 27.6
12 25.9
hadoop@fedora exp3$
```