

## 1. MapReduce MaxTemperature example

```

Linux nyu-dataproc-m 5.10.0-0.bpo.12-amd64 #1 SMP Debian 5.10.103-1~bpo10+1 (2022-03-08) x86_64

Last login: Thu Sep 21 21:55:19 2023 from 35.235.244.33
pi2018_nyu_edu@nyu-dataproc-m:~$ javac -classpath 'yarn classpath' -d . MaxTemperatureMapper.java
pi2018_nyu_edu@nyu-dataproc-m:~$ javac -classpath 'yarn classpath' -d . MaxTemperatureReducer.java
pi2018_nyu_edu@nyu-dataproc-m:~$ javac -classpath 'yarn classpath':. -d . MaxTemperature.java
Note: MaxTemperature.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
pi2018_nyu_edu@nyu-dataproc-m:~$ jar -cvf maxTemp.jar *.class
added manifest
adding: MaxTemperature.class(in = 1452) (out= 821)(deflated 43%)
adding: MaxTemperatureMapper.class(in = 1876) (out= 802)(deflated 57%)
adding: MaxTemperatureReducer.class(in = 1687) (out= 714)(deflated 57%)
pi2018_nyu_edu@nyu-dataproc-m:~$ ls
HW1                                MaxTemperatureMapper.java          MaxTemperatureReducer.java          maxTemp.jar
MaxTemperature.class               MaxTemperatureMapper.class          MaxTemperatureReducer.class          input.txt
pi2018_nyu_edu@nyu-dataproc-m:~$

```

```

Linux nyu-dataproc-m 5.10.0-0.bpo.12-amd64 #1 SMP Debian 5.10.103-1-bpo10+1 (2022-03-08) x86_64

Last login: Fri Sep 22 16:55:13 2023 from 35.235.244.33
pi2018_nyu_edu@nyu-dataproc-m:~$ ls
HW1 MaxTemperature.java MaxTemperatureMapper.java MaxTemperatureReducer.java maxTemp.jar
MaxTemperature.class MaxTemperatureMapper.class MaxTemperatureReducer.class input.txt
pi2018_nyu_edu@nyu-dataproc-m:~$ hdfs dfs -ls
Found 4 items
-rw-r--r-- 1 pi2018_nyu_edu pi2018_nyu_edu 1092 2023-09-21 16:41 MaxTemperatureMapper.java
drwxrwxr-x+ - pi2018_nyu_edu pi2018_nyu_edu 0 2023-09-14 15:42 dirToShareAccess
drwxr-xr-x - pi2018_nyu_edu pi2018_nyu_edu 0 2023-09-21 16:32 hw
-rw-r--r-- 1 pi2018_nyu_edu pi2018_nyu_edu 279 2023-09-21 16:40 input.txt
pi2018_nyu_edu@nyu-dataproc-m:~$ hadoop jar maxTemp.jar MaxTemperatureMapper input.txt new_output
2023-09-22 16:58:05,208 INFO client.RMPProxy: Connecting to ResourceManager at nyu-dataproc-m/192.168.1.38:8032
2023-09-22 16:58:05,389 INFO client.AHSPProxy: Connecting to Application History server at nyu-dataproc-m/192.168.1.38:10200
2023-09-22 16:58:05,548 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
2023-09-22 16:58:05,627 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/pi2018_nyu_edu/.staging/job_1691775874963_2074
2023-09-22 16:58:05,919 INFO Input.FileInputFormat: Total input files to process : 1
2023-09-22 16:58:05,989 INFO mapreduce.JobSubmitter: number of splits:1
2023-09-22 16:58:06,144 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1691775874963_2074
2023-09-22 16:58:06,146 INFO mapreduce.JobSubmitter: Executing with tokens: []
2023-09-22 16:58:06,315 INFO conf.Configuration: resource-types.xml not found
2023-09-22 16:58:06,315 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.
2023-09-22 16:58:06,538 INFO impl.YarnClientImpl: Submitted application application_1691775874963_2074
2023-09-22 16:58:06,572 INFO mapreduce.Job: The url to track the job: http://nyu-dataproc-m:8088/proxy/application_1691775874963_2074/
2023-09-22 16:58:06,573 INFO mapreduce.Job: Running job: job_1691775874963_2074
2023-09-22 16:58:13,662 INFO mapreduce.Job: Job job_1691775874963_2074 running in uber mode : false
2023-09-22 16:58:13,664 INFO mapreduce.Job: map 0% reduce 0%
2023-09-22 16:58:19,729 INFO mapreduce.Job: map 100% reduce 0%
2023-09-22 16:58:24,759 INFO mapreduce.Job: map 100% reduce 100%
2023-09-22 16:58:25,775 INFO mapreduce.Job: Job job_1691775874963_2074 completed successfully
2023-09-22 16:58:25,867 INFO mapreduce.Job: Counters: 54
File System Counters
FILE: Number of bytes read=61
FILE: Number of bytes written=492155
FILE: Number of read operations=0
FILE: Number of large read operations=0

```

```
Linux nyu-edu@nyu-dataproc-m 5.10.0-0.bpo.12-amd64 #1 SMP Debian 5.10.103-1-bpo10+1 (2022-03-08) x86_64

Last login: Thu Sep 21 16:40:00 2023 from 35.235.244.32
pi2018_nyu_edu@nyu-dataproc-m:~$ ls
HW1                               MaxTemperature.java              MaxTemperatureMapper.java        MaxTemperatureReducer.java      maxTemp.jar
MaxTemperature.class              MaxTemperatureMapper.class        MaxTemperatureReducer.class       input.txt
pi2018_nyu_edu@nyu-dataproc-m:~$ hdfs dfs -ls
Found 6 items
drwxr-xr-x   - pi2018_nyu_edu pi2018_nyu_edu           0 2023-09-21 16:32 -mkdir
-rw-r--r--   1 pi2018_nyu_edu pi2018_nyu_edu          1092 2023-09-21 16:41 MaxTemperatureMapper.java
drwxrwxr-x+  - pi2018_nyu_edu pi2018_nyu_edu           0 2023-09-14 15:42 dirToShareAccess
drwxr-xr-x   - pi2018_nyu_edu pi2018_nyu_edu           0 2023-09-21 16:32 hw
-rw-r--r--   1 pi2018_nyu_edu pi2018_nyu_edu          279 2023-09-21 16:40 input.txt
drwxr-xr-x   - pi2018_nyu_edu pi2018_nyu_edu           0 2023-09-21 16:46 new_output
pi2018_nyu_edu@nyu-dataproc-m:~$ hdfs dfs -ls new_output
Found 2 items
-rw-r--r--   1 pi2018_nyu_edu pi2018_nyu_edu           0 2023-09-21 16:46 new_output/_SUCCESS
-rw-r--r--   1 pi2018_nyu_edu pi2018_nyu_edu          17 2023-09-21 16:46 new_output/part-r-00000
pi2018_nyu_edu@nyu-dataproc-m:~$ hdfs dfs -cat new_output/part-r-00000
1949      111
1950       22
pi2018_nyu_edu@nyu-dataproc-m:~$
```