Hadoop: File Information in the Mapper

1. An InputFormat for reading a whole file as a record

Command for setting default output directory:

hdfs dfs -mkdir /home/student04/

Command for compiling Java file:

javac -classpath /home/student4/hadoop-common-2.6.1.jar:/home/student4/hadoop-mapreduce-client-core-2.6.1.jar:/home/student4/commons-cli-2.0.jar -d . WholeFileInputFormat.java WholeFileRecordReader.java SmallFilesToSequenceFileConverter.java JobBuilder.java

Command for creating Jar files:

jar -cvf small-file-sequence.jar SmallFilesToSequenceFileConverter\*.class WholeFileInputFormat.class WholeFileRecordReader.class JobBuilder.class

Command for copying a file from local disk to HDFS:

hdfs dfs -copyFromLocal smallfiles /home/student04/

Command for setting location:

export HADOOP\_CLASSPATH=/home/student4/

Run a Jar file on Hadoop:

hadoop jar small-file-sequence.jar SmallFilesToSequenceFileConverter -D mapred.reduce.tasks=2 /home/student04/smallfiles/ /home/student04/output1

Print out the output on Hadoop:

hdfs dfs -text /home/student04/output1/part-r-00000

hdfs dfs -text /home/student04/output1/part-r-00001

Screenshot of final results:

A screenshot of a social media post

Description automatically generated

1. User-Defined Java Counters: Application to run the maximum temperature job, including counting missing and malformed fields and quality codes.

Command for compiling Java file:

javac -classpath /home/student4/hadoop-common-2.6.1.jar:/home/student4/hadoop-mapreduce-client-core-2.6.1.jar:/home/student4/commons-cli-2.0.jar -d . MaxTemperatureWithCounters.java NcdcRecordParser.java JobBuilder.java MaxTemperatureReducer.java

Command for creating Jar files:

jar -cvf max-temperature.jar MaxTemperatureWithCounters\*.class NcdcRecordParser.class JobBuilder.class MaxTemperatureReducer.class

Command for copying a file from local disk to HDFS:

hdfs dfs -copyFromLocal 1930 /home/student04

Command for setting location:

export HADOOP\_CLASSPATH=/home/student4

Run a Jar file on Hadoop:

hadoop jar max-temperature.jar MaxTemperatureWithCounters

/home/student04/1930/ /home/student04/output2/

Screenshot of final results:

A screenshot of a cell phone

Description automatically generated