1. *Work Examples 7-2, 7-3, and 7-4 on CSUEB Hadoop. Type out all the commands in each step of the process and print out a screenshot of the final results in CSUEB Hadoop.*

*Hint 1: need to create a jar file including five classes: WholeFileInputFormat.class, WholeFileRecordReader.class, SmallFilesToSequenceFileConverter.class, SmallFilesToSequenceFileConverter$SequenceFileMapper.class, and JobBuilder.class*

*Note. SmallFilesToSequenceFileConverter.class is the main class. SmallFilesToSequenceFileConverter$SequenceFileMapper.class is a nested/inner class of SmallFilesToSequenceFileConverter.class.*

*JobBuilder.java can be found in Hadoop-Book-Master/common/src/main/java*

*Smallfiles folder can be found in Hadoop-Book-Master/input*

*Hint 2: -conf conf/Hadoop-localhost.xml is not needed in the hadoop jar command. So, your command will be like this: hadoop jar /home/jwu/hadoop-example.jar SmallFilesToSequenceFileConverter -D mapred.reduce.tasks=2 /home/jwu/smallfiles /home/jwu/output11*

Command for compiling Java file:

javac -classpath /home/student8/hadoop-common-2.6.1.jar:/home/student8/hadoop-mapreduce-client-core-2.6.1.jar:/home/student8/commons-cli-2.0.jar -d . SmallFilesToSequenceFileConverter.java WholeFileInputFormat.java WholeFileRecordReader.java JobBuilder.java

jar -cvf small-files-to-sequencefile-converter.jar SmallFilesToSequenceFileConverter\*.class WholeFileInputFormat.class WholeFileRecordReader.class JobBuilder.class

Command for copying a file from local disk to HDFS:

hdfs dfs -copyFromLocal smallfiles /home/student8/input1

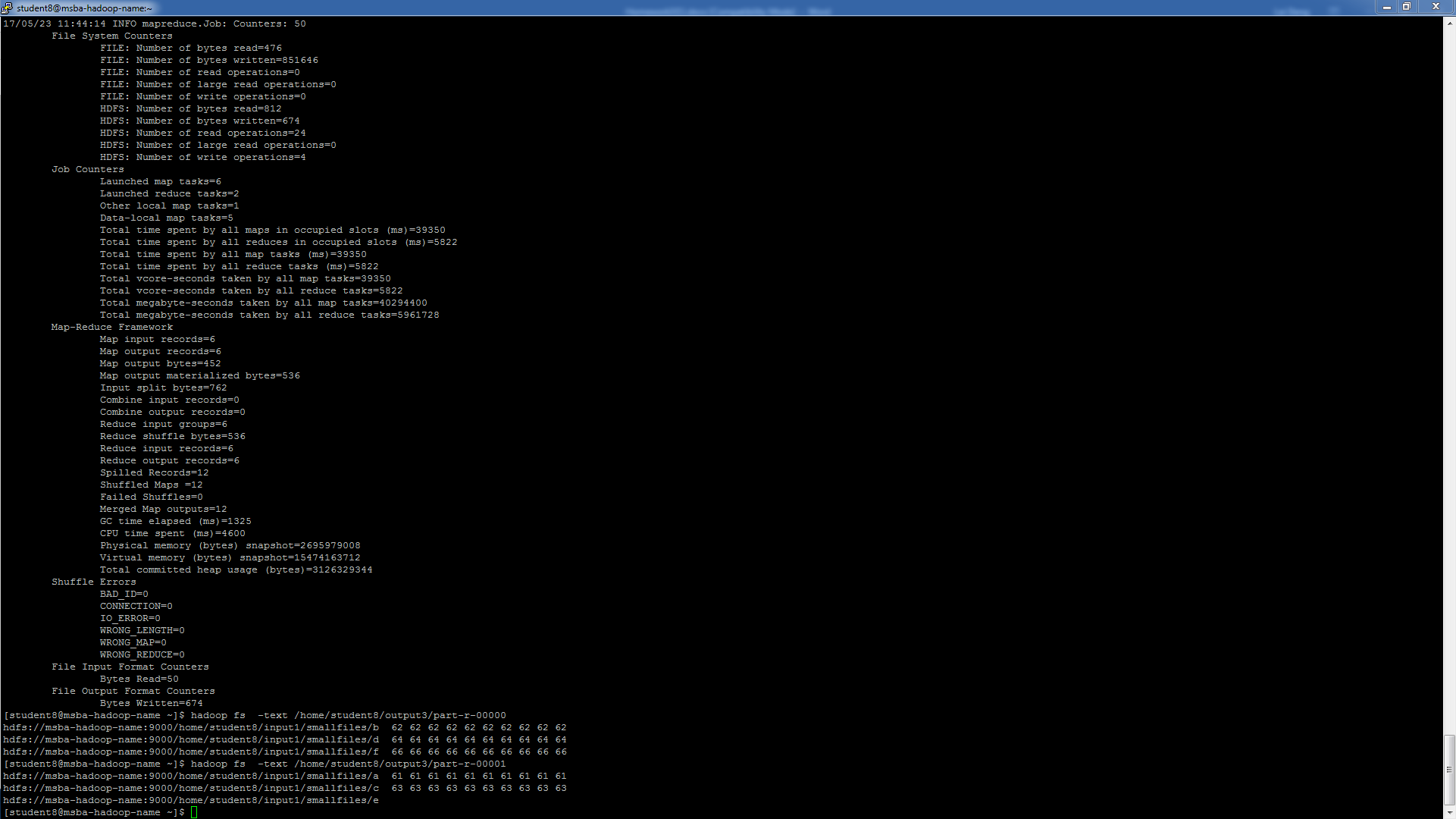
Command for running Java in Hadoop:

hadoop jar small-files-to-sequencefile-converter.jar SmallFilesToSequenceFileConverter -D mapred.reduce.tasks=2 /home/student8/input1/smallfiles /home/student8/output3

hadoop fs -text /home/student8/output3/part-r-00000

hadoop fs -text /home/student8/output3/part-r-00001

Screenshot of final results in CSUEB:



1. *Work Example 8-1 on CSUEB Hadoop. Type out all the commands in each step of the process and print out a screenshot of the final results (the counters) in CSUEB Hadoop.*

*Hint: need to create a jar file including five classes: MaxTemperatureWithCounters.class, MaxTemperatureMapperWithCounters.class, NcdcRecordParser.class, JobBuilder.class, and MaxTemperatureReducer.class.*

*Note. 1) NcdcRecordParser.java can be found in Hadoop-Book-Master/common/src/main/java, 2) you must use the data of year 1930 as input data to run the program (download data at:* [*ftp://ftp.ncdc.noaa.gov/pub/data/noaa/*](ftp://ftp.ncdc.noaa.gov/pub/data/noaa/)*), and 3) your results shall be different from those run over the complete dataset of 100 years, which are shown on page 265 in the textbook.*

Command for compiling Java file:

javac -classpath /home/student8/hadoop-common-2.6.1.jar:/home/student8/hadoop-mapreduce-client-core-2.6.1.jar:/home/student8/commons-cli-2.0.jar -d . MaxTemperatureWithCounters.java MaxTemperatureReducer.java NcdcRecordParser.java JobBuilder.java

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

Command for copying a file from local disk to HDFS:

hdfs dfs -copyFromLocal 1930 /home/student8/input1

Command for running Java in Hadoop:

hadoop jar max-temperature-withcounter.jar MaxTemperatureWithCounters /home/student8/input1/1930 /home/student8/output

Screenshot of final results in CSUEB:

