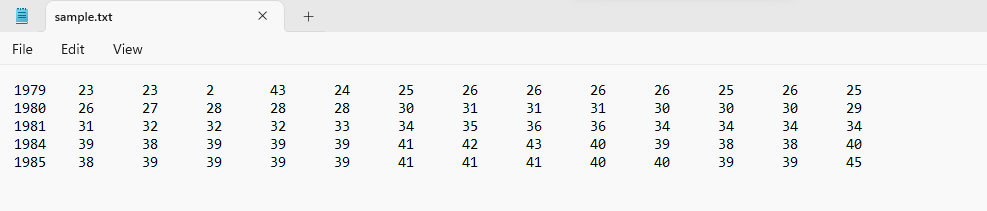
SV: Trần Thị Ngoc Trang

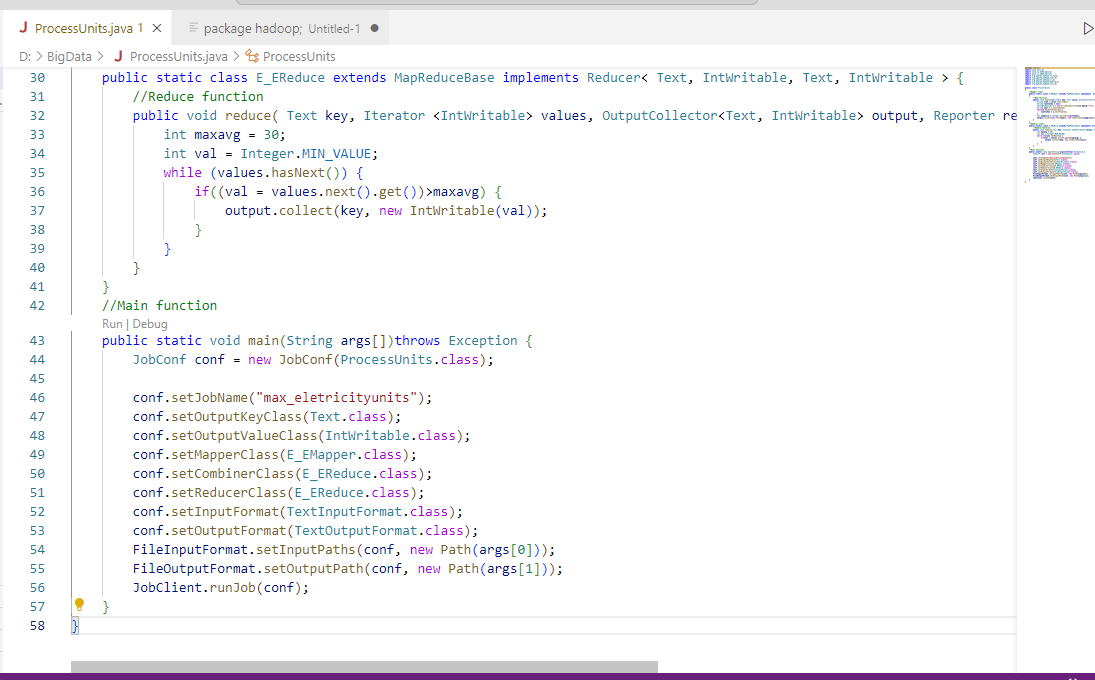
MSSV: 21133109

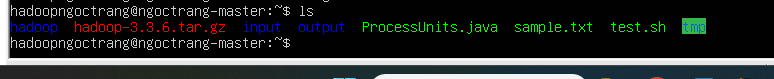
Microsoft Word - 3 Hadoop MapReduce

**Tạo tập tin sample.txt chứa dữ liệu trên**



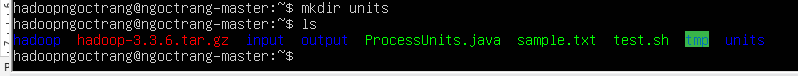
**Tạo file ProcessUnits.java chứa code sau**





1. **Tạo thư mục units chứa các file sau khi thông dịch ProcessUnits.java**

mkdir units

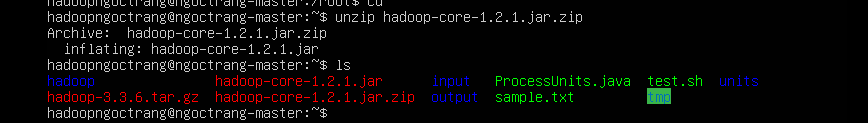


1. **Download hadoop-core-1.2.1.jar**

wget <http://www.java2s.com/Code/JarDownload/hadoop-core/hadoop-core-1.2.1.jar.zip>



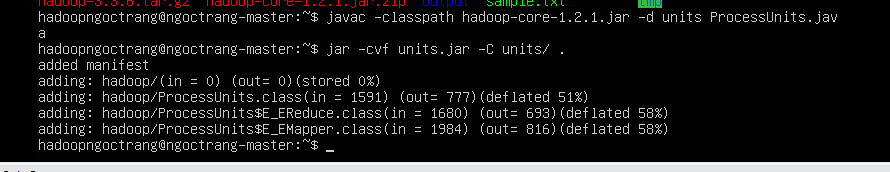
unzip hadoop-core-1.2.1.jar.zip



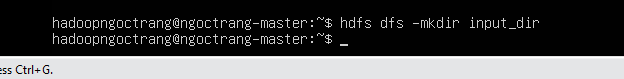
1. **Thông dịch file nguồn và tạo file jar**

javac -classpath hadoop-core-1.2.1.jar -d units ProcessUnits.java

jar -cvf units.jar -C units/ .

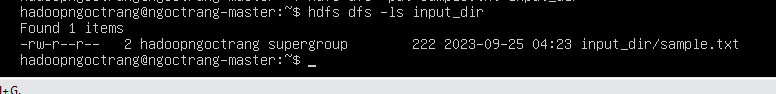


1. **Tạo thư mục input trong HDFS** hdfs dfs -mkdir input\_dir



1. **Đưa dữ liệu sample.txt vào thư mục input\_dir trong HDFS**

hdfs dfs -put sample.txt input\_dir



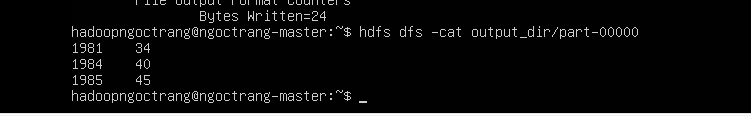
1. **Thực thi chương trình Eleunit\_max trong Hadoop**

hadoop jar units.jar hadoop.ProcessUnits input\_dir output\_dir



1. **Xuất kết quả**

hdfs dfs -cat output\_dir/part-00000



1. **Copy kết quả từ Hadoop ra thư mục bên ngoài**

hdfs dfs -get output\_dir /home/hadoopngoctrang

