LAPORAN TUGAS HADOOP
KOMPUTASI TERSEBAR DAN PARALEL
"Hadoop Single Node Cluster"

Disusun Oleh:

Farhan Dwicahyo - 24060120140099

Luthfi Arya Manggala - 24060120140150

PROGRAM STUDI S-1 INFORMATIKA

FAKULTAS SAINS DAN MATEMATIKA

UNIVERSITAS DIPONEGORO

SEMARANG

2022

Link Demo Video:
https://drive.google.com/drive/folders/18wjvMsR2xGn55j3lj2HGyEZIQEw9HlIW?usp=sharing

Pembuatan Cluster Hadoop

Menginstall Java dan Pembuatan User Hadoop

```
hadoop@jay-VirtualBox:~$ sudo apt install default-jdk default-jre -y
hadoop@jay-VirtualBox:~$ java -version
hadoop@jay-VirtualBox:~$ sudo adduser hadoop
hadoop@jay-VirtualBox:~$ sudo usermod -aG sudo hadoop
hadoop@jay-VirtualBox:~$ sudo su - hadoop
hadoop@jay-VirtualBox:~$ apt install openssh-server openssh-client -y
hadoop@jay-VirtualBox:~$ ssh-keygen -t rsa
hadoop@jay-VirtualBox:~$ sudo cat ~/.ssh/id_rsa.pub >>
~/.ssh/authorized_keys
hadoop@jay-VirtualBox:~$ sudo chmod 640
~/.ssh/authorized_keys
hadoop@jay-VirtualBox:~$ ssh localhost
```

Menginstall Apache Hadoop

```
hadoop@jay-VirtualBox:~$ sudo su - hadoop
hadoop@jay-VirtualBox:~$ wget
https://downloads.apache.org/hadoop/common/hadoop-3.3.1/hadoop-3.3.1.tar.gz
hadoop@jay-VirtualBox:~$ tar -xvzf hadoop-3.3.1.tar.gz
hadoop@jay-VirtualBox:~$ sudo mv hadoop-3.3.1/usr/local/hadoop
hadoop@jay-VirtualBox:~$ sudo mkdir /usr/local/hadoop/logs
hadoop@jay-VirtualBox:~$ sudo chown -R hadoop:hadoop/usr/local/hadoop
```

Konfigurasi Hadoop

```
hadoop@jay-VirtualBox:~$ sudo nano ~/.bashrc export HADOOP_HOME=/usr/local/hadoop
export HADOOP_INSTALL=$HADOOP_HOME export HADOOP_MAPRED_HOME=$HADOOP_HOME export
HADOOP_COMMON_HOME=$HADOOP_HOME export HADOOP_HDFS_HOME=$HADOOP_HOME
hadoop@jay-VirtualBox:~$ sudo nano ~/.bashrc
```

```
//masukkan kode berikut
export YARN_HOME=$HADOOP_HOME
export HADOOP_COMMON_LIB_NATIVE_DIR=$HADOOP_HOME/lib/native export
PATH=$PATH:$HADOOP_HOME/sbin:$HADOOP_HOME/bin
export HADOOP_OPTS="Djava.library.path=$HADOOP_HOME/lib/native"
hadoop@jay-VirtualBox:~$ source ~/.bashrc
```

Konfigurasi Java Environment Variables

```
hadoop@jay-VirtualBox:~$ which javac
hadoop@jay-VirtualBox:~$ readlink -f /usr/bin/javac
hadoop@jay-VirtualBox:~$ sudo nano $HADOOP_HOME/etc/hadoop/hadoop-env.sh
export JAVA_HOME=/usr/lib/jvm/java-11-openjdk-amd64 export HADOOP_CLASSPATH+="
$HADOOP_HOME/lib/*.jar"
hadoop@jay-VirtualBox:~$ cd /usr/local/hadoop/lib hadoop@jay-VirtualBox:~$ sudo
wget
https://jcenter.bintray.com/javax/activation/javax.activation-api/1
.2.0/javax.activation-api-1.2.0.jar
hadoop@jay-VirtualBox:~$ hadoop version
hadoop@jay-VirtualBox:~$ sudo nano
hadoop@jay-VirtualBox:~$HADOOP_HOME/etc/hadoop/core-site.x ml
//masukkan kode berikut
<configuration> <property>
<name>fs.default.name</name>
<value>hdfs://0.0.0.0:9000</value>
<description>The default file system URI</description> </property>
</configuration>
hadoop@jay-VirtualBox:~$ sudo mkdir -p /home/hadoop/hdfs/{namenode,datanode}
hadoop@jay-VirtualBox:~$ sudo chown -R hadoop:hadoop /home/hadoop/hdfs
hadoop@jay-VirtualBox:~$ sudo nano $HADOOP_HOME/etc/hadoop/hdfs-site.xml
//masukkan kode berikut
<configuration> <property>
<name>dfs.replication</name>
<value>1</value>
</property>

<property>
<name>dfs.name.dir</name>
<value>file:///home/hadoop/hdfs/namenode</value>
</property>
<property>
<name>dfs.data.dir</name>
<value>file:///home/hadoop/hdfs/datanode</value> </property> </configuration>
hadoop@jay-VirtualBox:~$ sudo nano $HADOOP_HOME/etc/hadoop/mapred-site.xml
```

```
//masukkan kode berikut
<property>
<name>mapreduce.framework.name</name>
<value>yarn</value>
</property>
<property>
<name>yarn.app.mapreduce.am.env</name>
<value>HADOOP_MAPRED_HOME=${HADOOP_HOME}</value>
</property>
<property>
<name>mapreduce.map.env</name>
<value>HADOOP_MAPRED_HOME=${HADOOP_HOME}</value>
</property>
<property>
<name>mapreduce.reduce.env</name>
<value>HADOOP_MAPRED_HOME=${HADOOP_HOME}</value>
</property>
hadoop@jay-VirtualBox:~$ sudo nano $HADOOP_HOME/etc/hadoop/yarn-site.xml
<configuration> <property>
<name>yarn.nodemanager.aux-services</name>
<value>mapreduce_shuffle</value>
</property> </configuration>
hadoop@jay-VirtualBox:~$ sudo su - hadoop
hadoop@jay-VirtualBox:~$ hdfs namenode -format
```

Memulai Apache Hadoop

```
hadoop@jay-VirtualBox:~$ start-dfs.sh
hadoop@jay-VirtualBox:~$ start-yarn.sh
hadoop@jay-VirtualBox:~$ jps
```

Dataset dan Program

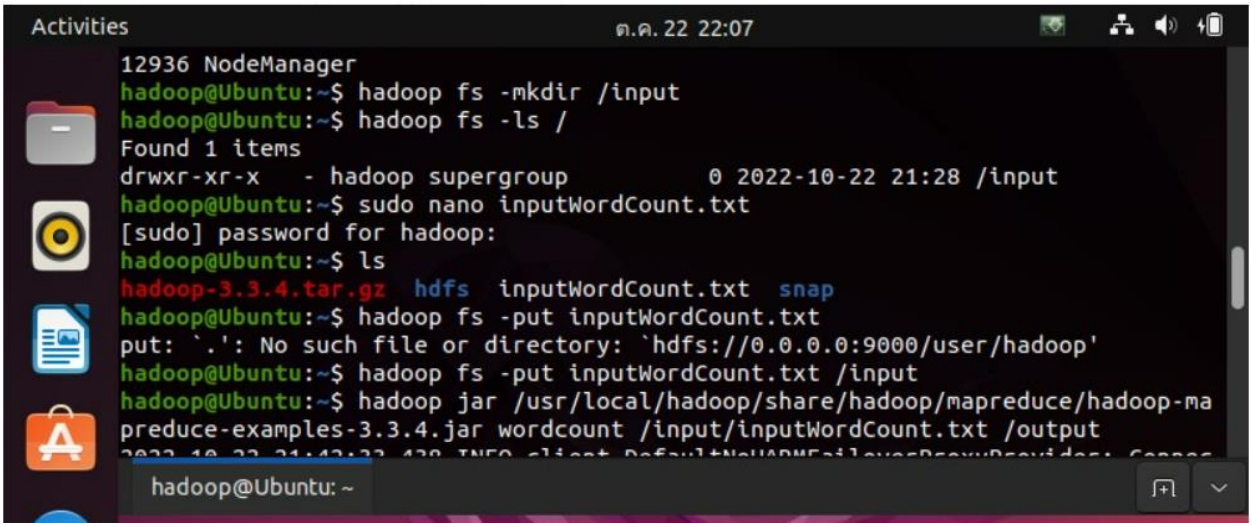
Data Set

Cardamom seed, dried, ground
Cinnamon, dried, ground
Cloves, dried, ground
Coriander seed, dried, ground
Cumin (cummin) seed, dried, ground

Curry powder
Fenugreek seed, dried
Ginger, dried, ground
Mustard powder
Nutmeg, dried, ground
Oregano, dried
Paprika, dry powder
Pepper, black, ground
Rosemary, dried
Sage, dried
Thyme, dried, ground
Turmeric, dried, ground
Salt substitute, potassium chloride
Salt, table, iodised
Salt, table, non-iodised
Stock, dry powder or cube
Taco seasoning mix, chilli-based
Baking powder, dry powder
Baking soda (bicarbonate), dry powder
Cream of tartar, dry powder
Gelatine, all types
Gluten, from wheat (vital wheat gluten)
Starch, potato
Vanilla, artificial or imitation
Vanilla bean extract
Yeast, dry powder
Beer, high alcohol (5% v/v & above)
Beer, full strength (alcohol 4-4.9% v/v)
Beer, full strength (alcohol 4-4.9% v/v), carbohydrate modified
Beer, mid-strength (alcohol 3-3.9% v/v)
Beer, light (alcohol 1-2.9% v/v)
Alcoholic beverage, spirit, approximately 40% v/v, all (Brandy, Gin, Rum, Vodka and Whisky)
Alcoholic beverage, spirit, approximately 30% v/v, all (Brandy, Gin, Rum, Vodka and Whisky), cooked
Cider, apple (alcohol approximately 4-5% v/v)
Wine, red, cabernet sauvignon
Wine, red, merlot
Wine, red, pinot noir
Wine, red, shiraz
Wine, red
Wine, red, cooked
Wine, red, sparkling

Wine, rose
Wine, white, sauvignon blanc
Wine, white, semillon
Wine, white, chardonnay
Wine, white, riesling
Wine, white
Wine, white, cooked
Wine, white, sparkling
Wine, white, sweet dessert style
Wine, fortified, port
Wine, fortified, sherry, dry style (approximately 1% sugars)
Wine, fortified, sherry, sweet style (approximately 11% sugars)
Cocoa powder
Beverage base, chocolate flavour, added vitamins A, B1, B2, C & D, Ca & Fe (Milo)
Beverage, chocolate flavour, from Milo powder, with regular fat cows milk
Beverage base, drinking chocolate, unfortified
Beverage, chocolate flavour, from drinking chocolate, with regular fat cows milk
Coffee, instant, dry powder or granules
Coffee, black, from instant coffee powder
Coffee, instant, dry powder or granules, decaffeinated
Coffee, black, from instant coffee powder, decaffeinated
Coffee, espresso, from ground coffee beans
Coffee, flat white/latte/cappuccino, from ground coffee beans, with regular fat cows milk
Coffee, long black, from ground coffee beans
Cordial base, 25% citrus fruit juice, regular
Cordial, 25% citrus fruit juice, regular, recommended dilution
Cordial base, 40% citrus fruit juice, regular
Cordial, 40% citrus fruit juice, regular, recommended dilution
Cordial base, blackcurrant juice, regular
Cordial, blackcurrant juice, regular, recommended dilution
Fruit drink, apple juice
Fruit drink, cranberry
Juice, apple, commercial, added vitamin C
Juice, apple & blackcurrant, commercial
Juice, lemon
Juice, lime
Juice, orange, commercial
Juice, orange & mango, commercial
Mineral water, natural, unflavoured
Mineral water, citrus flavoured
Soft drink, cola flavour
Soft drink, cola flavour, decaffeinated

Soft drink, cola flavour, intense sweetened or diet
Soft drink, cola flavour, intense sweetened or diet, decaffeinated
Soft drink, energy drink, Red Bull
Soft drink, energy drink, V
Soft drink, fruit flavours
Soft drink, tonic water
Soft drink, tonic water, intense sweetened or diet
Tea, green, plain, without milk
Tea, regular, black, brewed from leaf or teabags, without milk
Tea, decaffeinated, black, brewed from leaf or teabags, without milk
Water, bottled, still
Water, tap
Water, soda



1. hadoop fs -mkdir /input

   Untuk membuat file directory yang bernama /input

2. hadoop fs -ls /

   Untuk mengetahui apakah directory yang sebelumnya telah dibuat sudah berhasil dibuat atau belum

3. sudo nano inputWordCount.txt

   Untuk membuat file txt yang berisikan dataset yang ingin dihitung katanya

4. ls

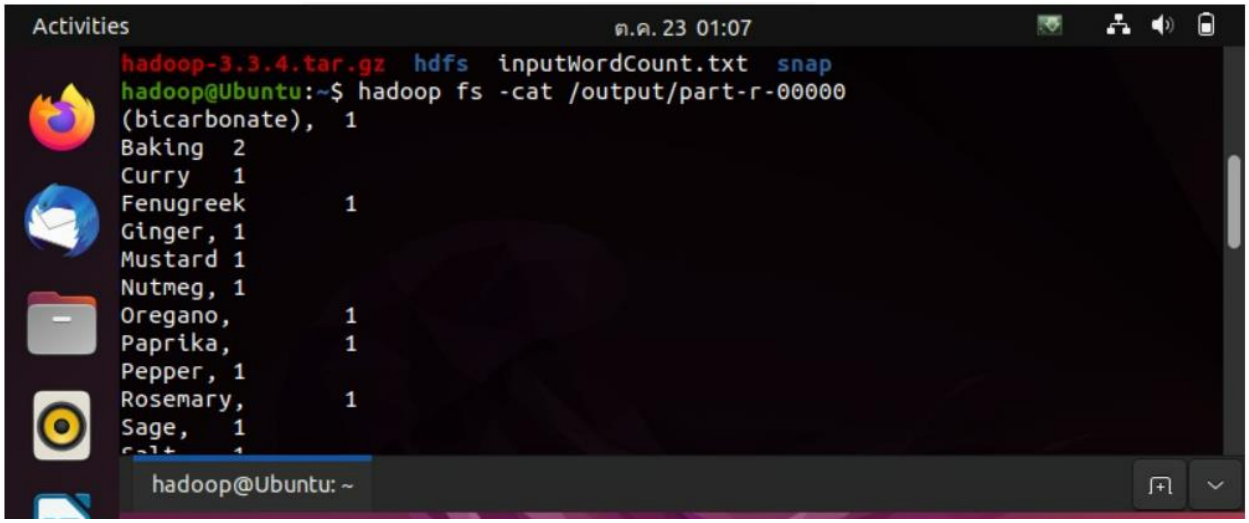
   Untuk mengetahui apakah file txt yang sebelumnya sudah dibuat itu berhasil atau belum

5. hadoop fs -put inputWordCount.txt /input

   Untuk memasukkan file txt yang sudah di buat ke dalam directory /input

6. .hadoop jar /usr/local/hadoop/mapreduce/hadoop-mapreduce-example-3.3.4.jar wordcount /input/inputWordCount.txt /output,

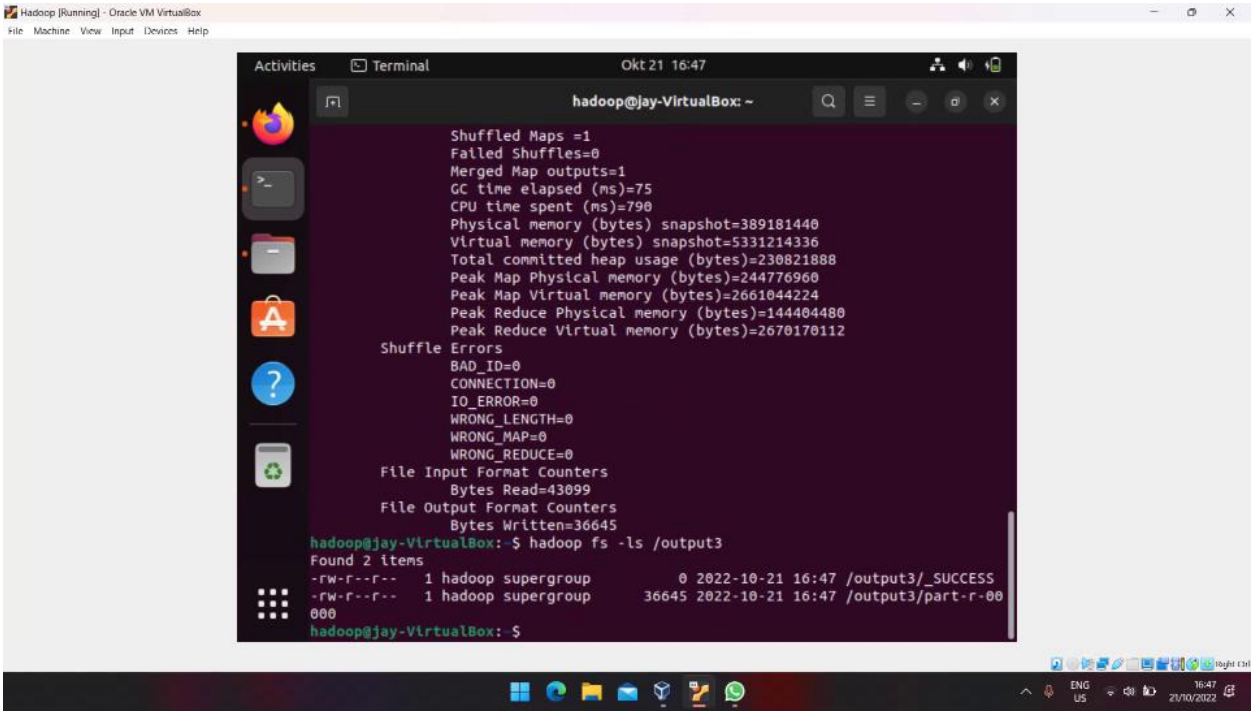
   untuk mengcompile file txt yang sudah berada di dalam directory /input dan membuat directory /output yang dimana file yang sudah di compile akan berada di dalam directory tersebut.



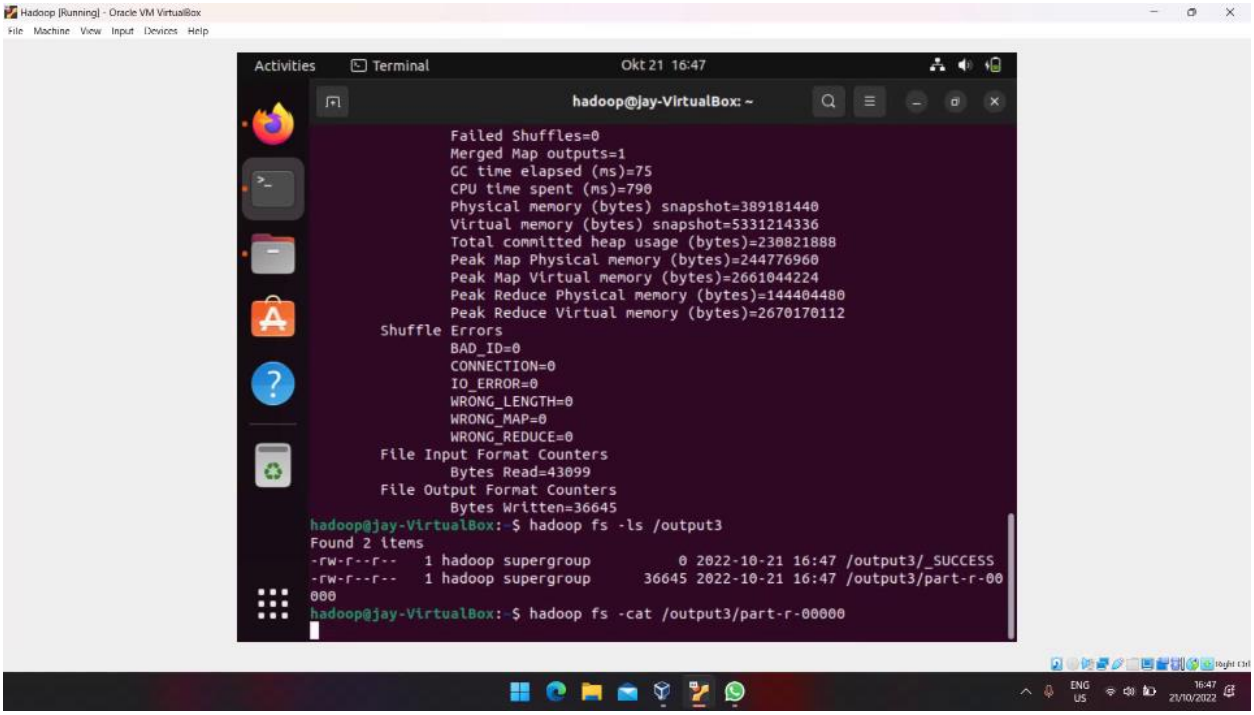
7. hadoop fs -cat /output/part-r-00000

   Untuk memunculkan hasil wordcount yang telah di compile

   Kemudian membuat buat folder input3 dengan perintah mkdir. Setelah itu file dataset .txt yang dibuat pindah kedalam folder input dengan perintah -put. Kemudian jalankan java nya dengan command jar, sesuai argument yang sudah dimasukkan, file result akan masuk ke folder output3



Setelah dijalankan file java nya, dapat dilihat pada folder output3, sudah ada file outputnya



Dengan perintah cat, akan munculkan hasil compile tadi sehingga hasilnya akan seperti dibawah ini

1. hasil dari wordcount (`hadoop fs -cat`).

Hasil Eksekusi Program

```
(bicarbonate),  1
Baking 2
Curry   1
Fenugreek       1
Ginger,         1
Mustard         1
Nutmeg,         1
Oregano,        1
Paprika,        1
Pepper,         1
Rosemary,       1
Sage,   1
Salt    1
Salt,   2
Stock,  1
Taco    1
Thyme,          1
Turmeric,       1
black,  1
chilli-based    1
chloride        1
cube    1
dried   4
dried,  4
dry     4
ground 5
iodised 1
mix,    1
non-iodised     1
or      1
potassium       1
powder6
powder,         1
seasoning       1
seed,   1
soda    1
substitute,     1
```

table, 2