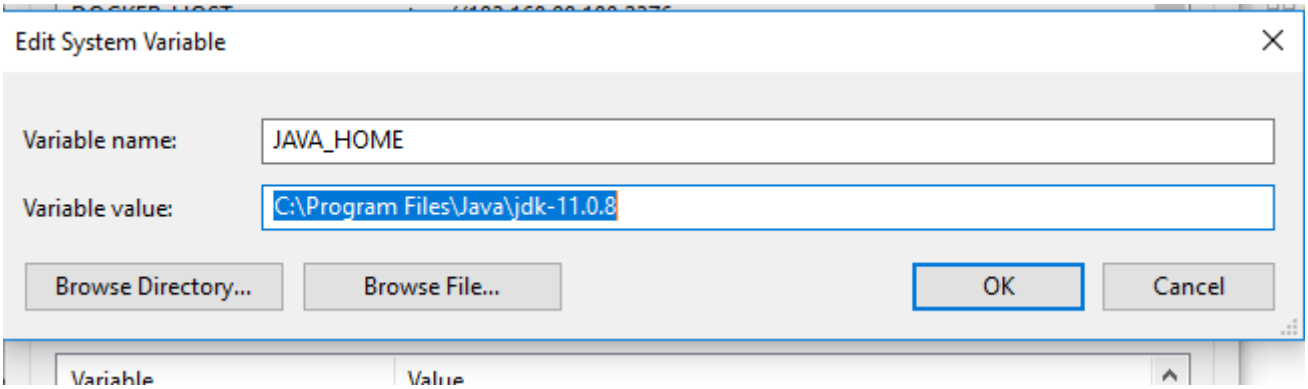
Instalasi dan Konfigurasi Java

Download Java 1.8 dari Oracle Website dan instalasikan ke folder lokal

```
https://www.oracle.com/java/technologies/javase/javase-jdk8-downloads.html
```

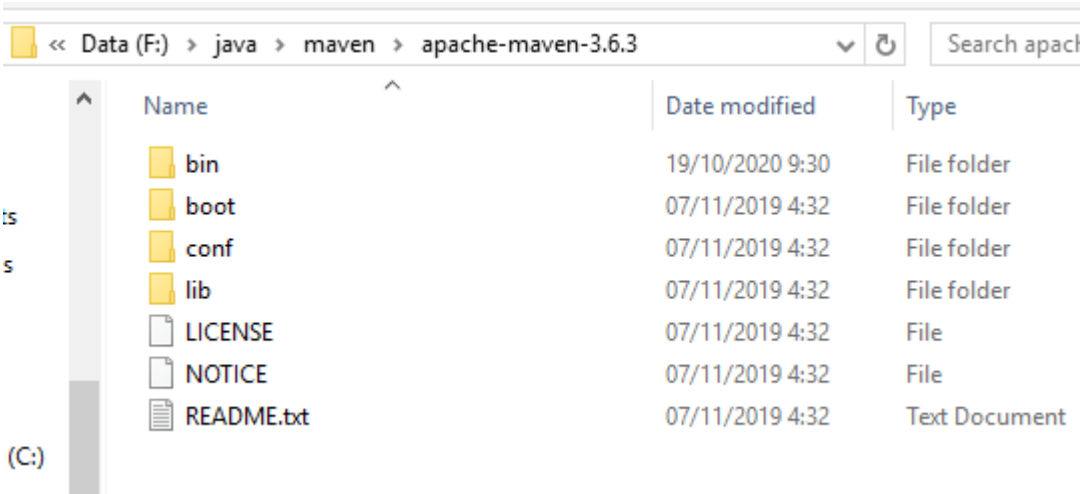
Konfigurasikan JAVA_HOME pada System Path



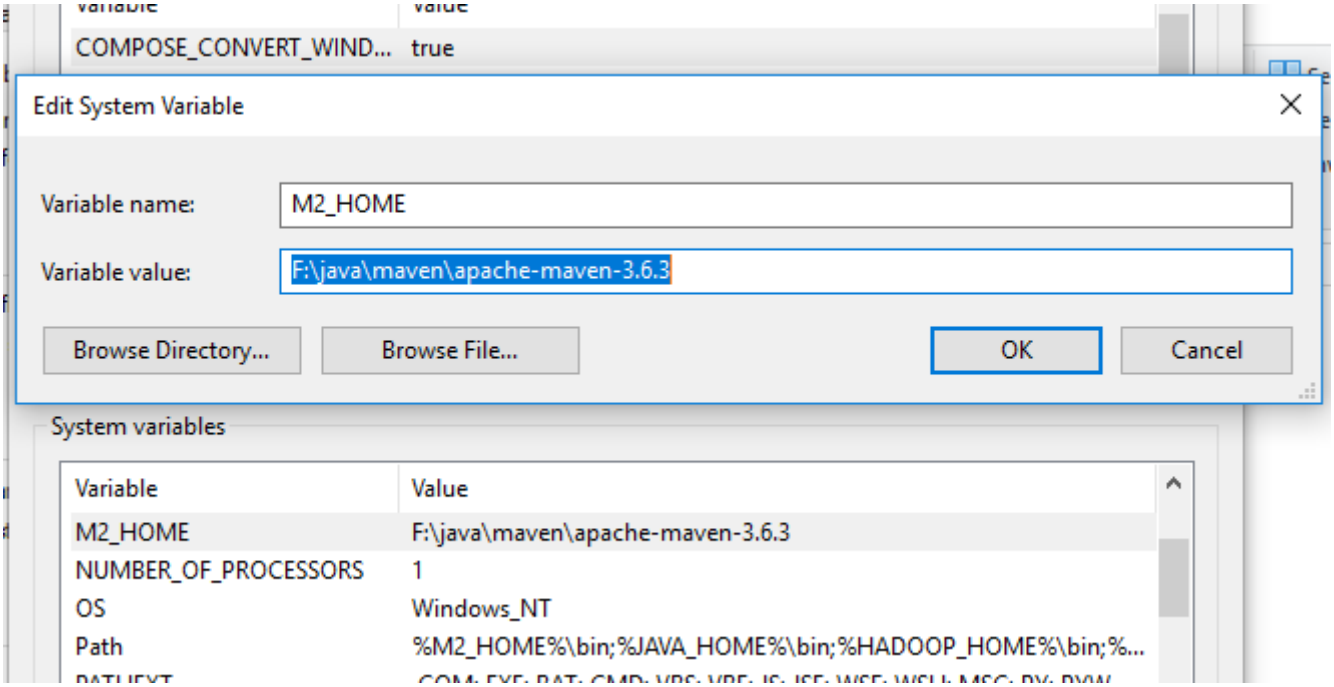
Instalasi dan Konfigurasi Apache Maven

Download Apache Maven dari Apache Website dan instalasikan ke folder Anda misalkan ke F:\java\maven\apache-maven-3.6.3

```
https://downloads.apache.org/maven/maven-3/3.6.3/binaries/apache-maven-3.6.3-bin.zip
```



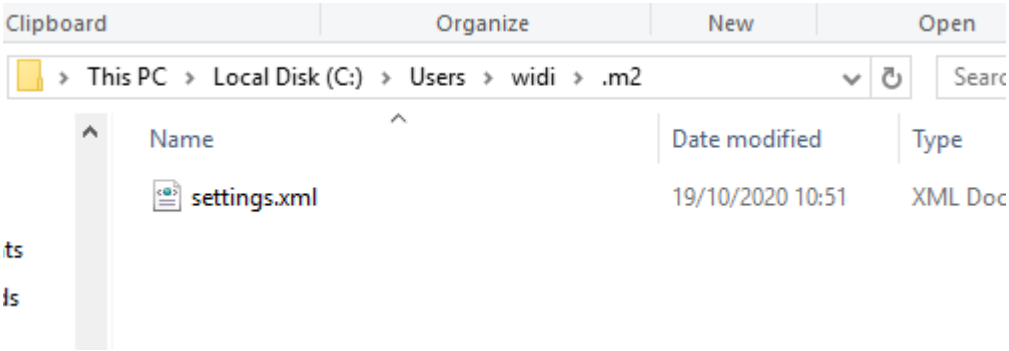
Konfigurasikan Maven Home (M2_HOME)



Edit file settings.xml pada baris 55 isikan kode berikut

```
<localRepository>F:/java/maven/repo</localRepository>
```

Buat folder C:/User/{namaAnda}/.m2 , copy kan file settings.xml ke folder tersebut

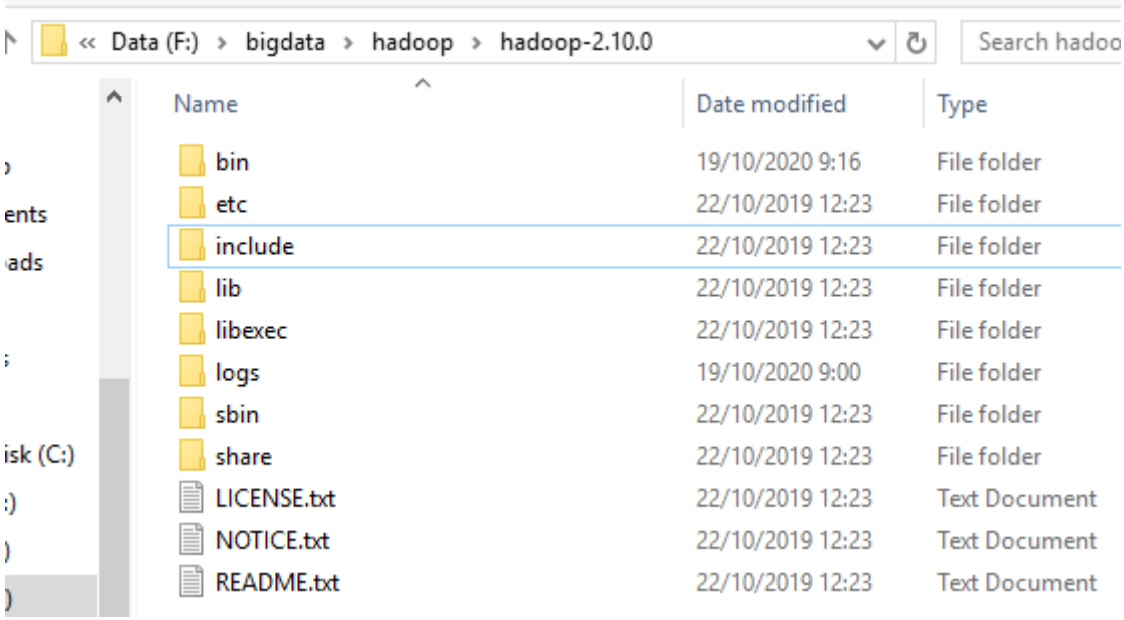


Instalasi dan Konfigurasi Hadoop

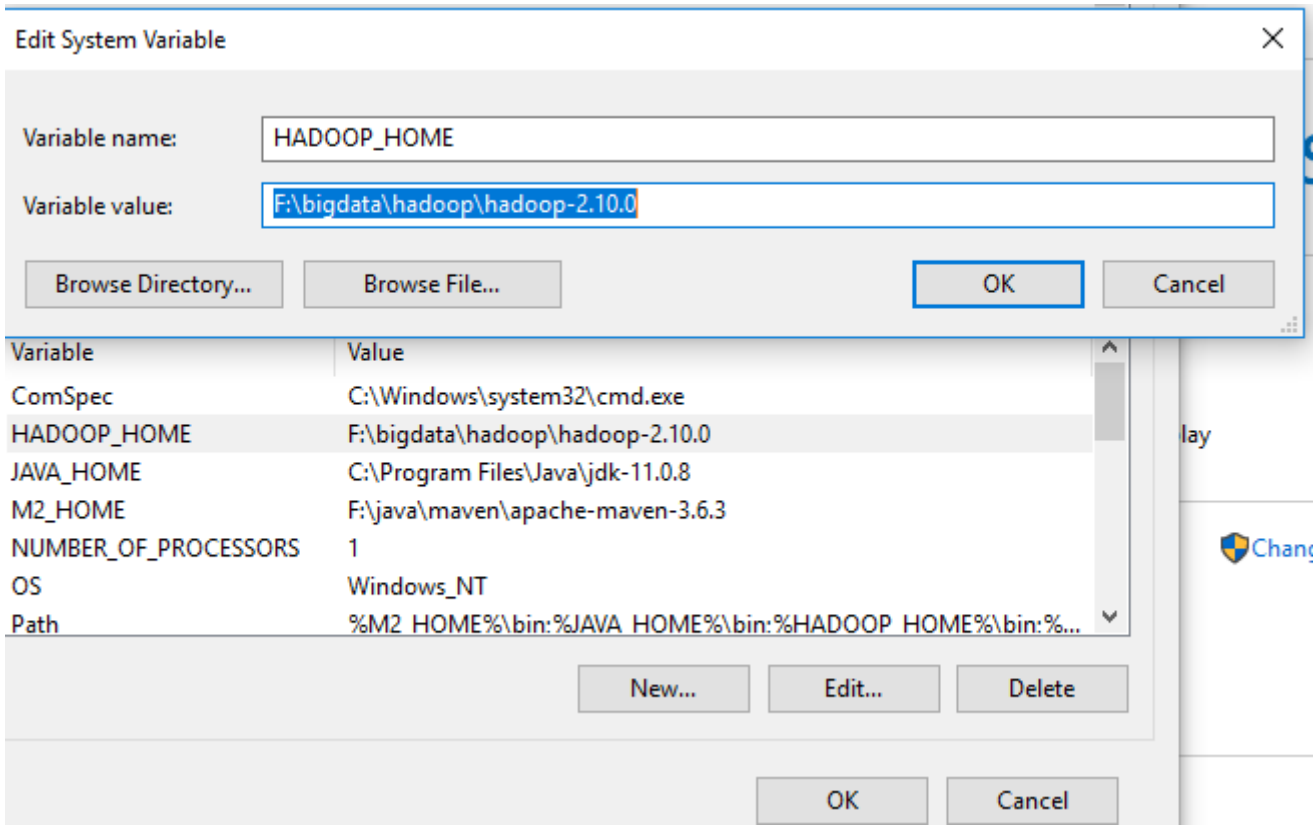
Download HADOOP dari Apache Website

```
https://downloads.apache.org/hadoop/core/hadoop-2.10.1/hadoop-2.10.1.tar.gz
```

Ekstrak ke folder Anda yang akan menjadi HADOOP_HOME misalkan ke F:\bigdata\hadoop\hadoop-2.10.0 sehingga terdapat struktur direktori seperti berikut



Konfigurasikan HADOOP_HOME pada System Environment



Konfigurasikan etc/hadoop/hdfs-site.xml

```
<configuration>
<property>
    <name>dfs.namenode.name.dir</name>
    <value>F:\bigdata\data\hadoop-content\hdfs\namenode</value>
  </property>

  <property>
    <name>dfs.datanode.data.dir</name>
    <value>F:\bigdata\data\hadoop-content\hdfs\datanode</value>
```

```
  </property>
</configuration>
```

Edit etc/hadoop/core-site.xml:

```
<configuration>
 <property>
        <name>fs.defaultFS</name>
        <value>hdfs://localhost:9000</value>
    </property>
    <property>
        <name>hadoop.tmp.dir</name>
        <value>F:\bigdata\data\hadoop-content\hdfs\tmp</value>
    </property>
</configuration>
```

edit hadoop-env.cmd

```
set JAVA_HOME="C:\\PROGRA~1\\Java\\jdk-11.0.8"
```

Edit mapred-site.xml

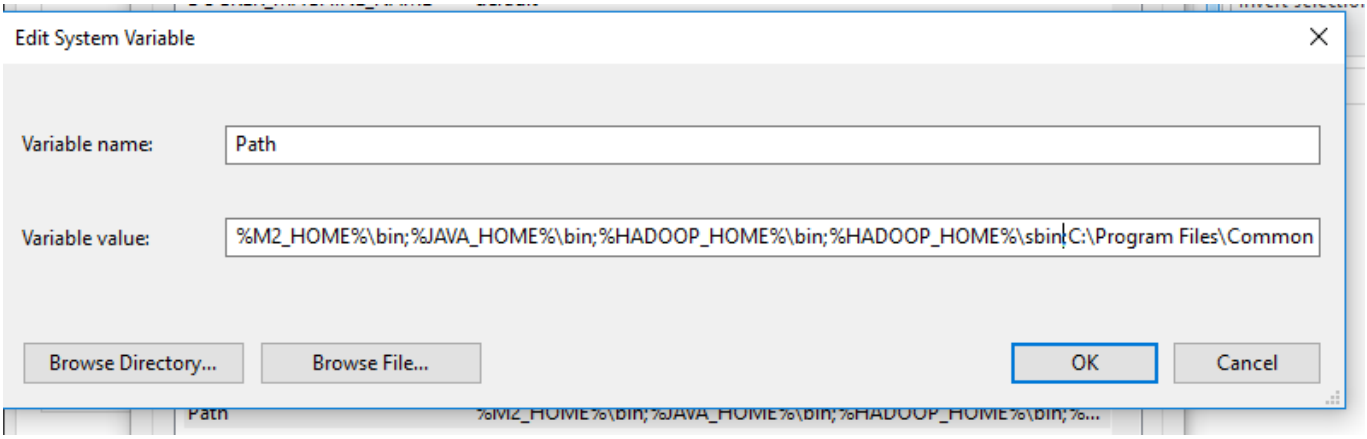
```
<configuration>
    <property>
        <name>mapreduce.framework.name</name>
        <value>yarn</value>
        <description>MapReduce framework name</description>
    </property>
</configuration>
```

Edit yarn-site.xml

```
<configuration>
    <property>
        <name>yarn.nodemanager.aux-services</name>
            <value>mapreduce_shuffle</value>
        <description>Yarn Node Manager Aux Service</description>
    </property>
</configuration>
```

Konfigurasi Path

Konfigurasikan M2_HOME/bin, JAVA_HOME/bin, HADOOP_HOME/bin, dan HADOOP_HOME/sbin agar didaftarkan pada system Path



Restart Sistem

Format File System

Format HDFS pertama kali

```
hdfs namenode -format
hdfs datanode -format
```

Start Hadoop

Start Hadoop PseudoCluster = Namenode + Datanode

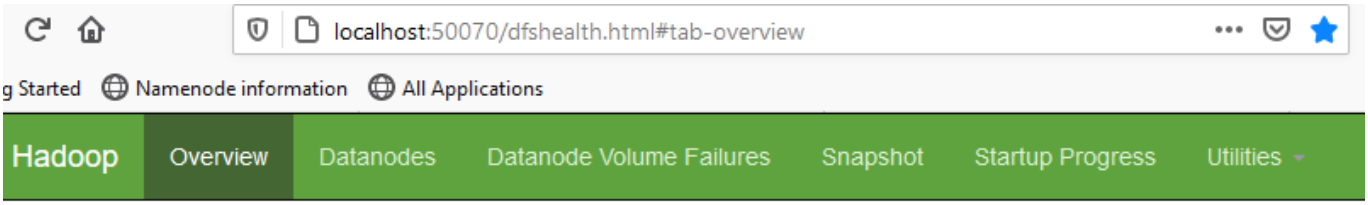
```
start-dfs.cmd
```

Cek log Hadoop sehingga start HDFS selesai

Cek HDFS melalui browser.

```
http://localhost:50070/
```

Sehingga tampil seperti berikut.



Overview 'localhost:9000' (active)

Started:	Mon Oct 19 19:45:46 -0700 2020
Version:	2.10.0, re2f1f118e465e787d8567dfa6e2f3b72a0eb9194
Compiled:	Tue Oct 22 12:10:00 -0700 2019 by jhung from branch-2.10.0
Cluster ID:	CID-cf5bde14-4f31-49a4-b83c-745e57fab0d3
Block Pool ID:	BP-1986357193-192.168.99.1-1603124439208

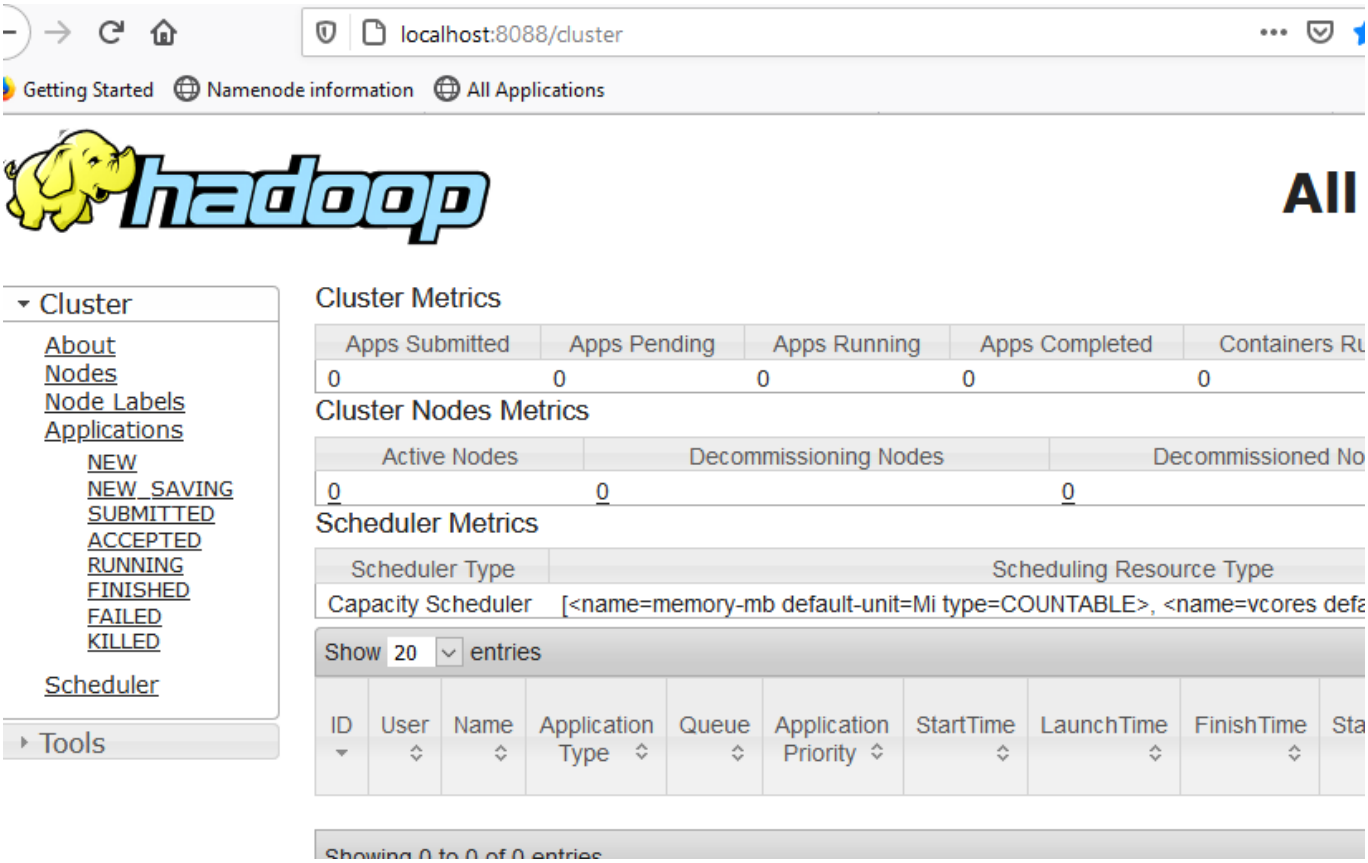
Summary

Security is off.

Cek Cluster melalui browser

```
http://localhost:8088/cluster
```

Sehingga tampil seperti berikut.



Stop Hadoop PseudoCluster

```
stop-dfs.cmd
```

Stop yarn

```
stop-yarn.cmd
```

Percobaan menggunakan operasi Hadoop

Coba buat file melalui command prompt

```
hdfs dfs -mkdir /user
hdfs dfs -mkdir /user/widi
```

```
hdfs dfs -mkdir input
hdfs dfs -put etc/hadoop/*.xml input
hadoop jar share/hadoop/mapreduce/hadoop-mapreduce-examples-3.3.0.jar grep input
output 'dfs[a-z.]+'
hdfs dfs -cat output/*
```