Download the latest Hadoop Binary version from 3.4.0 its 921 mb in size.

Use link : <https://www.apache.org/dyn/closer.cgi/hadoop/common/hadoop-3.4.0/hadoop-3.4.0.tar.gz>

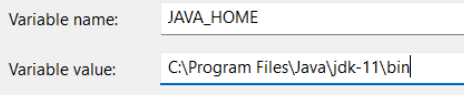
Download Java version 11.0.23 from oracle and log in to download.

Use link: <https://www.oracle.com/in/java/technologies/javase/jdk11-archive-downloads.html>

Download x64 installer.

Run the installer and follow the steps to install Java on your pc.

Make sure to set the Java Environment Variables by creating a new variable and also adding it to the path variable. (JAVA\_HOME)



Extract the previously downloaded Hadoop archive and copy the contents to your desired location.Extract the content in the C drive.

Download and Install Hadoop native IO binary from here.

To install, copy all the contents of the archive to the bin folder of your Hadoop.

Use link:<https://github.com/ruslanmv/How-to-install-Hadoop-on-Windows/tree/master/winutils>

Download the whole repository, then download the win utils file and the delete all file expect 3.4.0 x64

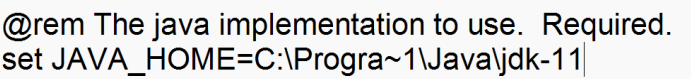
Copy the bin files of this folder and paste it to the previously downloaded hadoop folder in our C drive.

Set the Java Environment Variables by creating a new variable and also adding it to the path variable. (HADOOP\_HOME)



Add the path of both JAVA\_HOME AND HADOOP\_HOME in the Path variable of both user and system variables.

Set the JAVA\_HOME in the hadoop\_env.cmd file located inside the etc folder And save the changes.



Check the versions of both Java and Hadoop in cmd.

By typing , java -version and hadoop -version.

To verify all the above steps are completed successfully, open the bin(cmd) folder in terminal and run the command winutils.exe

Hadoop Configurations. All the files will be located in the etc folder of your Hadoop installation. Add the configuration for all the 4 files:

core-site.xml , hdfs-site.xml , mapred-site.xml , yarn-site.xml.

a) <configuration>

<property>

<name>fs.default.name</name>

<value>hdfs://0.0.0.0:19000</value>

</property>

</configuration>

b) Before configuring HDFS, make a folder called as data in your Hadoop installation. Make two subfolders named as namenode and datanode in that folder.

<configuration>

<property>

<name>dfs.replication</name>

<value>1</value>

</property>

<property>

<name>dfs.namenode.name.dir</name>

<value>/hadoop/hadoop-3.3.0/data/namenode</value>

</property>

<property>

<name>dfs.datanode.data.dir</name>

<value>/hadoop/hadoop-3.3.0/data/datanode</value>

</property>

</configuration>

c)<configuration>

<property>

<name>mapreduce.framework.name</name>

<value>yarn</value>

</property>

<property>

<name>mapreduce.application.classpath</name>

<value>%HADOOP\_HOME%/share/hadoop/mapreduce/\*,%HADOOP\_HOME%/share/hadoop/mapreduce/lib/\*,%HADOOP\_HOME%/share/hadoop/common/\*,%HADOOP\_HOME%/share/hadoop/common/lib/\*,%HADOOP\_HOME%/share/hadoop/yarn/\*,%HADOOP\_HOME%/share/hadoop/yarn/lib/\*,%HADOOP\_HOME%/share/hadoop/hdfs/\*,%HADOOP\_HOME%/share/hadoop/hdfs/lib/\*</value>

</property>

</configuration>

d) <configuration>

<property>

<name>yarn.nodemanager.aux-services</name>

<value>mapreduce\_shuffle</value>

</property>

<property>

<name>yarn.nodemanager.env-whitelist</name>

<value>JAVA\_HOME,HADOOP\_COMMON\_HOME,HADOOP\_HDFS\_HOME,HADOOP\_CONF\_DIR,CLASSPATH\_PREPEND\_DISTCACHE,HADOOP\_YARN\_HOME,HADOOP\_MAPRED\_HOME</value>

</property>

</configuration>

Initialise HDFS using the following command:

hdfs namenode -format

Start HDFS Daemons by going to the sbin(cmd) folder of Hadoop and running the command start-dfs.cmd

Verify HDFS web portal UI.

Use link:

<http://localhost:9870/dfshealth.html#tab-overview>

Start YARN Daemons by going to the sbin(cmd) folder of Hadoop and running the command start-yarn.cmd.

Use link and check :

<http://localhost:8088/cluster>