**Hadoop Practical 1 & 2 hands on**

Step 1: download java 1.8

link : <https://www.oracle.com/java/technologies/downloads/#java8-windows>

keep the java file in c drive

Step 2: Download Hadoop 2.8.0 or 3rd version

link : <https://hadoop.apache.org/releases.html>

keept it in a separate directory/folder

Also go to search -> Edit envirement variable-> user variable(new) -> variable name-> HADOOP\_HOME , value -> C:\Users\sushi\Downloads\hadoop-2.8.0\hadoop-2.8.0\bin

add user variable -> variable name -> JAVA\_HOME and value-> c:\java

System variable-> click on path -> edit -> click new -> add Hadoop path and java path till bin

Step 3: go to hadoop / bin and copy the libraries

link: <https://onedrive.live.com/?authkey=%21AOXTzskGn%5FhvOSw&id=AB81AF6E7ADDA0B4%213422&cid=AB81AF6E7ADDA0B4>

Step 4: go to the hadoop/etc/hadoop/hadoop-env and change the java Home (give the path)

Step 5: go to the hadoop/etc/hadoop/core-site.xml and write below code in configuration tag

<configuration>

<property>

<name>fs.defaultFS</name>

<value>hdfs://localhost:9000</value>

</property>

</configuration>

Step 6: go to the hadoop/etc/hadoop/hdfs-site.xml and write below code in configuration tag

<configuration>

<property>

<name>dfs.replication</name>

<value>1</value>

</property>

<property>

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

<value>file:///C:/Users/sushi/Downloads/hadoop-2.8.0/hadoop-2.8.0/data/namenode</value>

</property>

<property>

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

<value>file:///C:/Users/sushi/Downloads/hadoop-2.8.0/hadoop-2.8.0/data/datanode</value>

</property>

</configuration>

Step 7: go to hadoop and create a folder called "data" and in that create 2 folders "namenode" and "datanode".

Step 8: go to the hadoop/etc/hadoop/mapred-site.xml and write below code in configuration tag

<configuration>

<property>

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

<value>yarn</value>

</property>

</configuration>

Step 9: go to the hadoop/etc/hadoop/yarn-site.xml and write below code in configuration tag

<configuration>

<property>

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

<value>mapreduce\_shuffle</value>

</property>

<property>

<name>yarn.nodemanager.auxservices.mapreduce.shuffle.class</name>

<value>org.apache.hadoop.mapred.ShuffleHandler</value>

</property>

<!-- Site specific YARN configuration properties -->

</configuration>

Step 10: Open cmd prompt as admin and type Hadoop to check , hdfs

and then follow the commands step by step

hdfs namenode -format

sbin\start-dfs

sbin\start-yarn

hdfs dfs -mkdir -p input

hdfs dfs -ls .

Step 11: create a text file with "Install Hadoop

Run Hadoop Wordcount Mapreduce Example" and save it in any local directory

i.e. c:\data\file1.txt

Step 12: follow the commands one by one

hdfs dfs -copyFromLocal C:\data\file1.txt input

hdfs dfs -ls input

hdfs dfs -cat input/file1.txt

Step 13: for the execution of mapreduce on wordcount , use share folder from hadoop to get the jar file and later run that jar

bin\yarn jar share\hadoop\mapreduce\hadoop-mapreduce-examples-2.8.0.jar wordcount input output

bin\hdfs dfs -cat output/\*

OutPut

Example 3

Hadoop 5

Install 2

Mapreduce 3

Run 3

Wordcount 2

dataset 1

weather 1