

ROLLNO:19UITE015

REGNO:920419205008

NAME:K.DEEPIKA

SUBJECT:IT8711-FOSS AND CLOUD COMPUTING LAB

EXERCISE NO:9

PROCEDURAL STEPS

Step 1: Download the following Packages

1. hadoop-3.2.2.tar.gz

Link: <https://www.apache.org/dyn/closer.cgi/hadoop/common/hadoop-3.2.2/hadoop-3.2.2.tar.gz>

2. 7zip to unzip the tar.gz file

Link: <https://www.7-zip.org/download.html>

3. Java 8 (JDK-8U333 & JRE-8U333)

Link: <https://www.oracle.com/java/technologies/javase/javase8u211-later-archive-downloads.html>

4. Hadoop dll Files

Link: <https://github.com/cdarlint/winutils/archive/refs/heads/master.zip>

5. hadoop-hdfs-3.2.2.jar

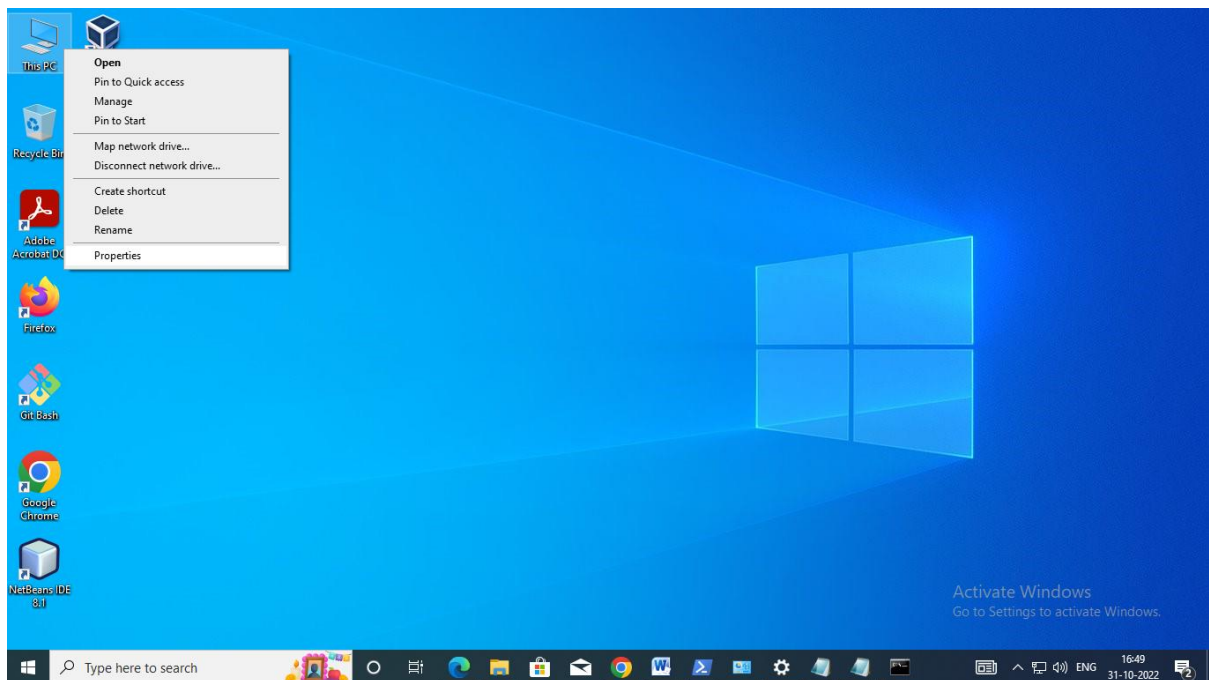
Link: <https://jar-download.com/artifacts/org.apache.hadoop/hadoop-hdfs/3.2.2/source-code>

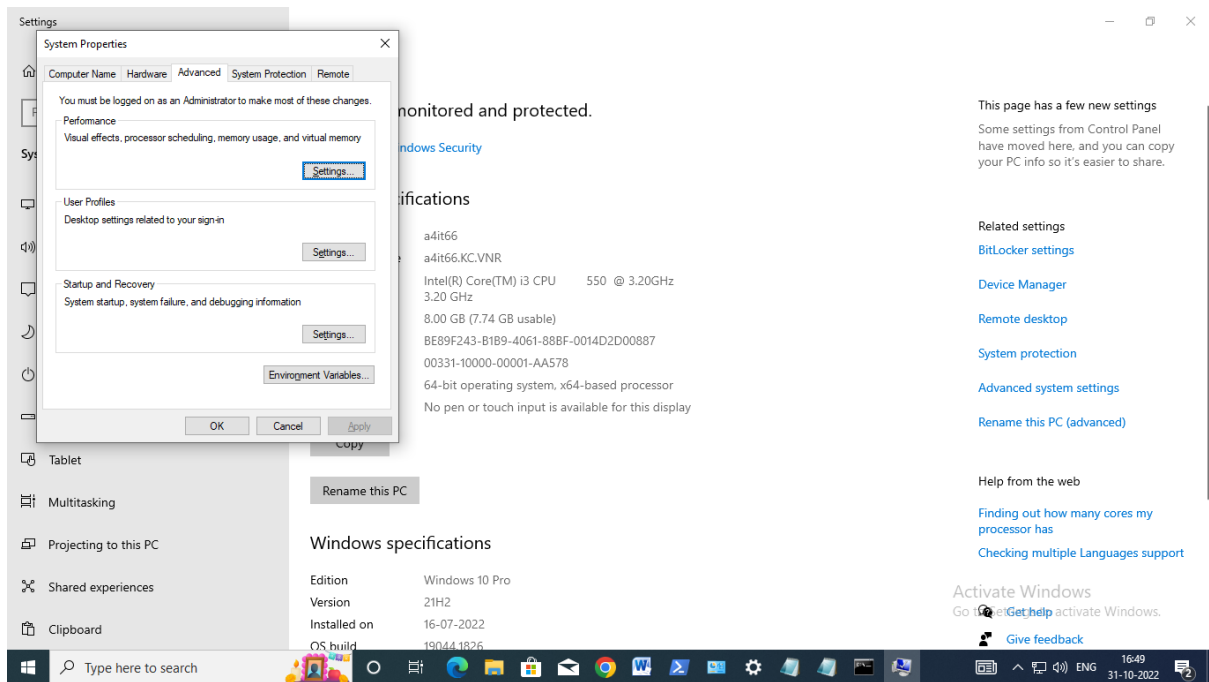
Step 2: Install / Extract the following Packages:

1. Install Java 8 in the location **c:\JAVA**. Also create two folders namely **"jdk1.8.0_333"** and **"jre1.8.0_333"**.
2. Create a folder **"hadoop-env"** in **D:**
3. Paste the downloaded **"hadoop-3.2.2.tar.gz"** into **"d:\hadoop-env"**
4. Unzip the **"hadoop-3.2.2.tar.gz"** using 7zip (do the unzip function two times)
5. Copy the downloaded Hadoop dll files to the location **"D:\hadoop-env\hadoop-3.2.2\bin"**
6. Copy the downloaded **hadoop-hdfs-3.2.2.jar** to **"D:\hadoop-env\hadoop-3.2.2\share\hadoop\hdfs"**
7. Edit the **JAVA_HOME = C:\JAVA\jdk1.8.0_333** in the path **"D:\hadoop-env\hadoop-3.2.2\etc\hadoop\hadoop-env.cmd"**

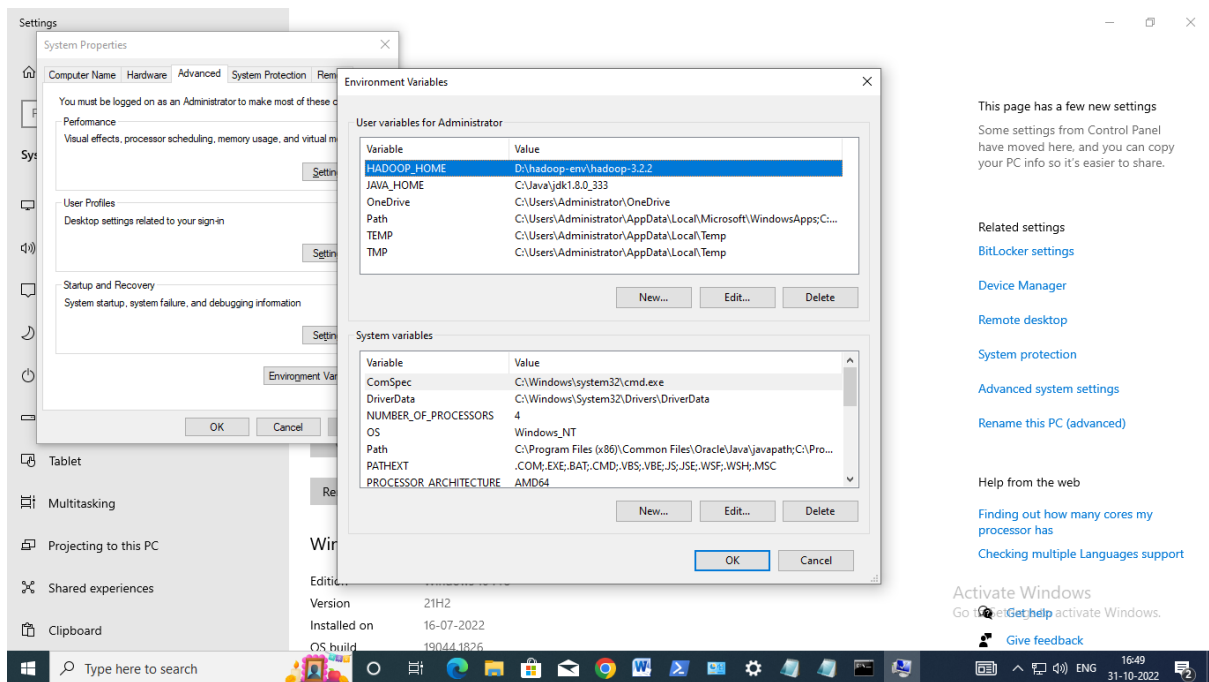
Step 3: Set the path for Java and Hadoop in system environment variables

To edit environment variables, go to Control Panel → System and Security → System (or) right-click “This PC” → Properties (My Computer icon) and click on the “Advanced system settings” link.



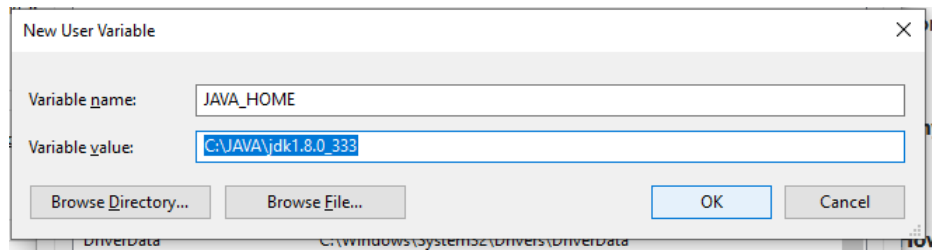


In the “Environment Variables” dialog, press the “New” button to add a new variable.

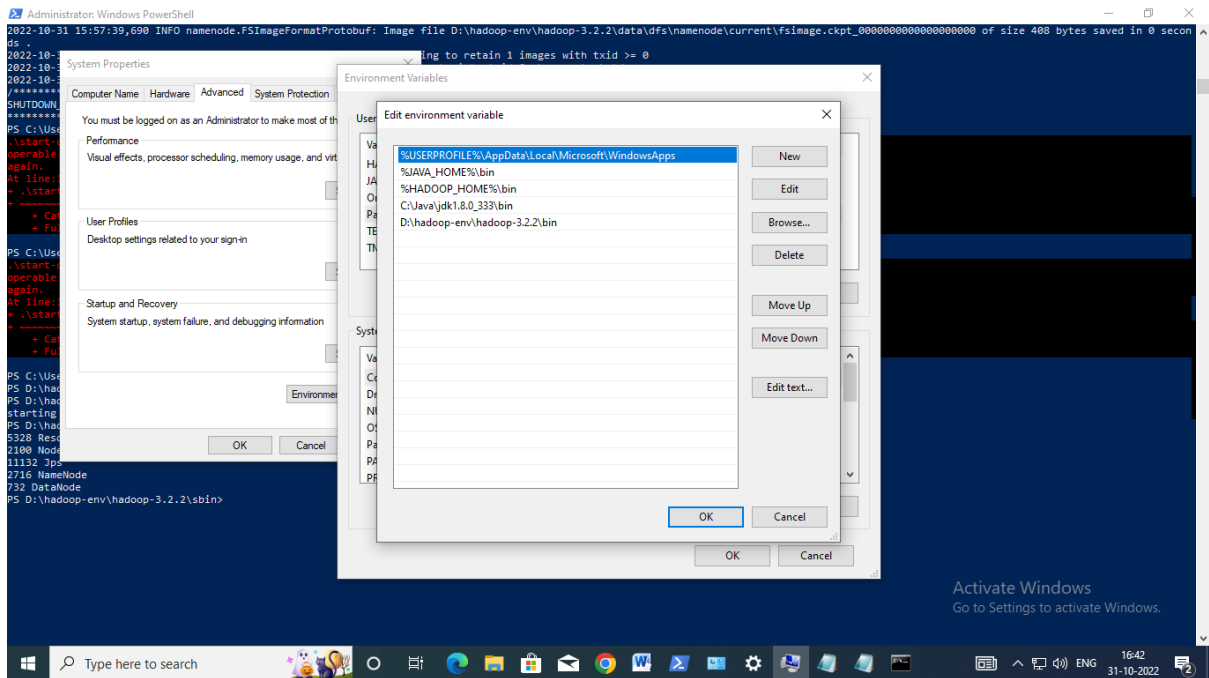
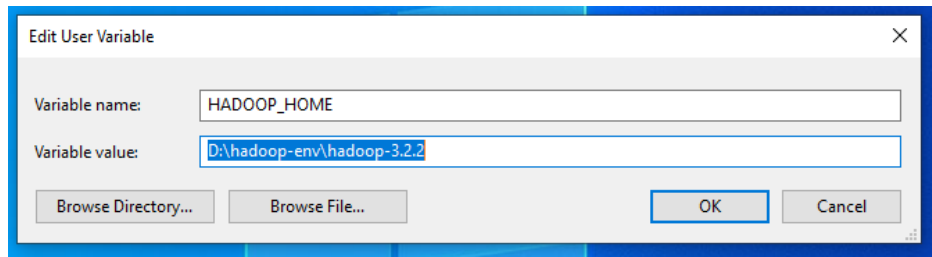


Now, Edit the PATH variable to add the Java and Hadoop binaries paths as shown in the following screenshots.

JAVA_HOME: JDK installation folder path



HADOOP_HOME: Hadoop installation folder path



Step 4: Open Command Prompt as Administrator and run the following command:

hadoop -version

File Name: “hdfs-site.xml”

Add the following properties within the <configuration></configuration> element:

```
<property>
<name>dfs.replication</name>
<value>1</value>
</property>
<property>
<name>dfs.namenode.name.dir</name>
<value>file:///D:/hadoop-env/hadoop-3.2.2/data/dfs/namenode</value>
</property>
<property>
<name>dfs.datanode.data.dir</name>
<value>file:///D:/hadoop-env/hadoop-3.2.2/data/dfs/datanode</value>
</property>
```

File Name: “core-site.xml”

Add the following properties within the <configuration></configuration> element:

```
<property>
<name>fs.default.name</name>
<value>hdfs://localhost:9820</value>
</property>
```

File Name: “mapred-site.xml”

Add the following properties within the <configuration></configuration> element:

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

File Name: “yarn-site.xml”

Add the following properties within the <configuration></configuration> element:

```
<property>
```

$$\langle \text{property} \rangle$$

```
.\start-dfs.cmd
```



```
Apache Hadoop Distribution - hadoop datanode
DEPRECATED: Use of this script to execute hdfs command is deprecated.
Instead use the hdfs command for it.
2022-10-31 16:00:11,185 INFO datanode.DataNode: STARTUP_MSG:
/*****
STARTUP_MSG: Starting DataNode
STARTUP_MSG: host = a41t66/192.168.56.1
STARTUP_MSG: args = []
STARTUP_MSG: version = 3.2.2
STARTUP_MSG: classpath = D:\hadoop-env\hadoop-3.2.2\etc\hadoop:D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\access
sors-smart-1.2.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\animal-sniffer-annotations-1.17.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\asm-5.0.4.j
ar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\audience-annotations-0.5.0.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\avro-1.7.7.jar;D:\hadoop-env\had
oop-3.2.2\share\hadoop\common\lib\checker-qual-2.5.2.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\commons-beanutils-1.9.4.jar;D:\hadoop-env\hadoop-3.2.2\share
\hadoop\common\lib\commons-cli-1.2.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\commons-codec-1.11.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\comm
ons-collections-3.2.2.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\commons-compress-1.19.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\commons-config
uration2-2.11.1.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\commons-io-2.5.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\commons-lang3-3.7.jar;D:\had
oop-env\hadoop-3.2.2\share\hadoop\common\lib\commons-logging-1.1.3.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\commons-math3-3.1.1.jar;D:\hadoop-env\hadoop-3
.2.2\share\hadoop\common\lib\commons-net-3.6.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\commons-text-1.4.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\
lib\curator-client-2.13.0.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\curator-framework-2.13.0.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\curator
-recipes-2.13.0.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\dnsjava-2.1.7.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\error-prone-annotations-2.2
.0.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\failureaccess-1.0.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\gson-2.2.4.jar;D:\hadoop-env\hadoop-3.
2.2\share\hadoop\common\lib\guava-27.0-jre.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\hadoop-annotations-3.2.2.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\c
ommon\lib\hadoop-auth-3.2.2.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\htrace-core4-4.1.0-incubating.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\
httpclient-4.5.13.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\httpcore-4.4.13.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\j2objc-annotations-1.1.j
ar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\jackson-annotations-2.0.10.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\jackson-core-2.9.10.jar;D:\hadoo
p-env\hadoop-3.2.2\share\hadoop\common\lib\jackson-core-asl-1.9.13.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\jackson-databind-2.9.10.4.jar;D:\hadoop-env\had
oop-3.2.2\share\hadoop\common\lib\jackson-jaxrs-1.9.13.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\jackson-mapper-asl-1.9.13.jar;D:\hadoop-env\hadoop-3.2.2\
share\hadoop\common\lib\jackson-xc-1.9.13.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\javax.activation-api-1.2.0.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\
common\lib\javax.servlet-api-3.1.0.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\jaxb-api-2.2.11.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\jaxb-im
pl-2.2.3-1.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\jcip-annotations-1.0-1.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\jersey-core-1.19.jar;D:\
hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\jersey-json-1.19.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\jersey-server-1.19.jar;D:\hadoop-env\hadoop-3.2.
2\share\hadoop\common\lib\jersey-servlet-1.19.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\jetty-son-1.1.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib
\jetty-http-9.4.20.v20190813.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\jetty-io-9.4.20.v20190813.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\jet
ty-security-9.4.20.v20190813.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\jetty-server-9.4.20.v20190813.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib
\jetty-servlet-9.4.20.v20190813.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\jetty-util-9.4.20.v20190813.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib
\jetty-webapp-9.4.20.v20190813.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\json-smart-2.3.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\jsp-api-2.1.jar;D:\hadoop-env\
hadoop-3.2.2\share\hadoop\common\lib\jsr305-3.0.2.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\jsr311-api-1.1.1.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\co
mon\lib\jul-to-slf4j-1.7.25.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\kerb-admin-1.0.1.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\kerb-client-
1.0.1.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\kerb-common-1.0.1.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\kerb-core-1.0.1.jar;D:\hadoop-env\h
adoop-3.2.2\share\hadoop\common\lib\kerb-crypto-1.0.1.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\kerb-identity-1.0.1.jar;D:\hadoop-env\hadoop-3.2.2\share\h
adoop\common\lib\kerb-server-1.0.1.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\kerby-asn1-1.0.1.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\kerby-identity-1.0.1.jar;D:\h
adoop-env\hadoop-3.2.2\share\hadoop\common\lib\kerby-pkix-1.0.1.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\kerby-util-1.0.1.jar;D:\hadoop-env\hadoop-3.2.2\sh
are\hadoop\common\lib\kerby-xdr-1.0.1.jar;D:\hadoop-env\hadoop-3.2.2\share\hadoop\common\lib\listenablefuture-9999.0-empty-to-avoid-conflict-with-guava.jar;D:\hadoop-e
```

Next, Start the Hadoop Yarn service using the following command:

./start-yarn.cmd

```
Administrator: Windows PowerShell
PS C:\Users\Administrator> cd D:\hadoop-env\hadoop-3.2.2\sbin
PS D:\hadoop-env\hadoop-3.2.2\sbin> .\start-dfs.cmd
PS D:\hadoop-env\hadoop-3.2.2\sbin> .\start-yarn.cmd
starting yarn daemons
PS D:\hadoop-env\hadoop-3.2.2\sbin> jps
5328 ResourceManager
2180 NodeManager
11132 Jps
2716 NameNode
732 DataNode
PS D:\hadoop-env\hadoop-3.2.2\sbin>
```

Two command prompt windows will open (one for the **resource manager** and one for the **node manager**) as follows:

14636 Jps

```
Administrator: Windows PowerShell
PS C:\Users\Administrator> cd D:\hadoop-env\hadoop-3.2.2\sbin
PS D:\hadoop-env\hadoop-3.2.2\sbin> .\start-dfs.cmd
PS D:\hadoop-env\hadoop-3.2.2\sbin> .\start-yarn.cmd
starting yarn daemons
PS D:\hadoop-env\hadoop-3.2.2\sbin> jps
5328 ResourceManager
2100 NodeManager
11132 Jps
2716 NameNode
732 DataNode
PS D:\hadoop-env\hadoop-3.2.2\sbin>
```

Output

Hadoop Web UI

<http://localhost:9870/dfshealth.html>

Hadoop Overview

Overview 'localhost:9820' (active)

Started:	Mon Oct 31 16:00:19 +0530 2022
Version:	3.2.2, r7a3bc90b05f257c8ace2f76d74264906f0f7a932
Compiled:	Sun Jan 03 14:56:00 +0530 2021 by hexiaoqiao from branch-3.2.2
Cluster ID:	CID-9d314296-acea-4485-bbec-89b4e4407716
Block Pool ID:	BP-1107299340-192.168.56.1-1667212058215

Summary

Security is off.
Safemode is off.
1 files and directories, 0 blocks (0 replicated blocks, 0 erasure coded block groups) = 1 total filesystem object(s).
Heap Memory used 80.65 MB of 189.5 MB Heap Memory. Max Heap Memory is 889 MB.
Non Heap Memory used 49.06 MB of 50.21 MB Committed Non Heap Memory. Max Non Heap Memory is <unbounded>.

<http://localhost:9864/datanode.html>

New tabNamenode informationWelcomeWelcome to Microsoft EdgeDataNode Informationlocalhost:9864/datanode.html

HadoopOverviewUtilities

DataNode on a4it66.KC.VNR:9866

Cluster ID:	CID-9d314296-acea-4485-bbec-89b4e4407716
Version:	3.2.2, r7a3bc90b05f257c8ace2f76d74264906f07a932

Block Pools

Namenode Address	Block Pool ID	Actor State	Last Heartbeat	Last Block Report	Last Block Report Size (Max Size)
localhost:9820	BP-1107299340-192.168.56.1-1667212058215	RUNNING	1s	an hour	0 B (64 MB)

Volume Information


Directory	Storage Type	Capacity Used	Capacity Left	Capacity Reserved	Reserved Space for Replicas	Blocks
D:\hadoop-env\hadoop-3.2.2\data\dfs\datanode	DISK	321 B	224.9 GB	0 B	0 B	0

Activate Windows
Go to Settings to activate Windows.

Type here to search

<http://localhost:8088/cluster>

New tabNamenode informationWelcomeWelcome to Microsoft EdgeDataNode InformationAll Applicationslocalhost:8088/cluster



All Applications

Cluster

AboutNodesNode LabelsApplicationsNEWNEW_SAVINGSUBMITTEDACCEPTEDRUNNINGFINISHEDFAILEDKILLED

Scheduler

Tools

Cluster Metrics

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Used Resources	Total Resources	Reserved Resources
0	0	0	0	0	<memory:0, vCores:0>	<memory:8192, vCores:8>	<memory:0, vCores:0>

Cluster Nodes Metrics

Active Nodes	Decommissioning Nodes	Decommissioned Nodes	Lost Nodes	Unhealthy Nodes
1	0	0	0	0

Scheduler Metrics

Scheduler Type	Scheduling Resource Type	Minimum Allocation	Maximum Allocation
Capacity Scheduler	[memory-mb (unit=Mi), vcores]	<memory:1024, vCores:1>	<memory:8192, vCores:4>

Show 20 entries

ID	User	Name	Application Type	Queue	Application Priority	StartTime	LaunchTime	FinishTime	State	FinalStatus	Running Containers	Allocated CPU VCores	Allocated Memory MB	Allocated GPUs	Reserved CPU VCores
No data available in table															

Showing 0 to 0 of 0 entries

Activate Windows
Go to Settings to activate Windows.

Type here to search