

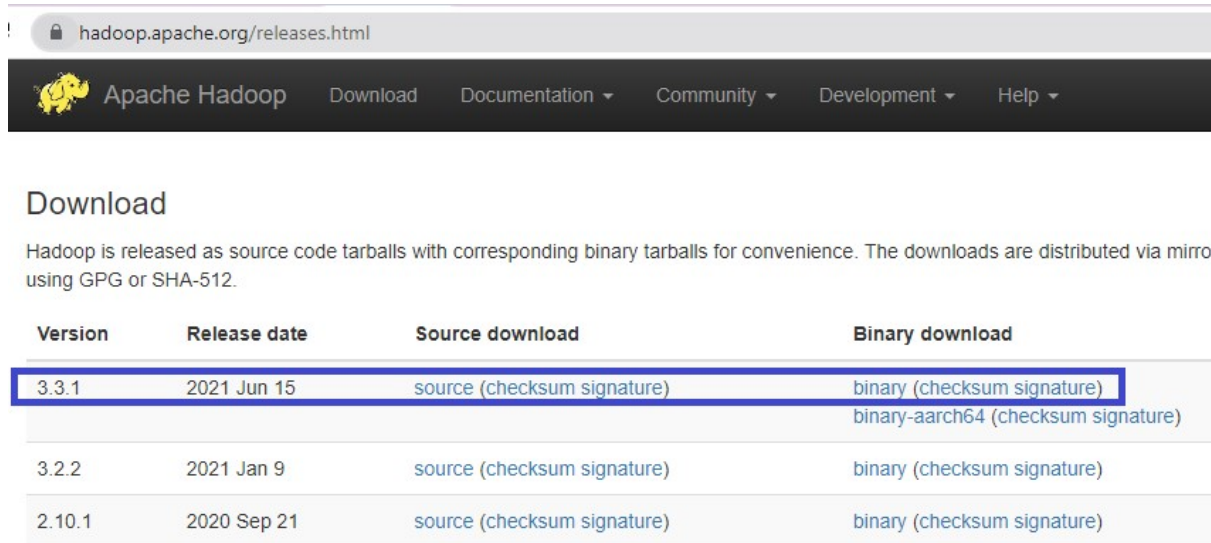
Hadoop Installation Step:

1. Hadoop Installation

Download and install the Hadoop latest version from the following official apache website.

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

Version: 3.3.1 Binary

A screenshot of the Apache Hadoop releases page. The browser address bar shows 'hadoop.apache.org/releases.html'. The page has a dark navigation bar with the Apache Hadoop logo and links for Download, Documentation, Community, Development, and Help. The main content area is titled 'Download' and contains a paragraph stating that Hadoop is released as source code tarballs with corresponding binary tarballs for convenience, distributed via mirrors using GPG or SHA-512. Below this is a table with four columns: Version, Release date, Source download, and Binary download. The first row, for version 3.3.1, is highlighted with a blue border. The 'Binary download' column for 3.3.1 lists 'binary (checksum signature)' and 'binary-aarch64 (checksum signature)'. The other rows show versions 3.2.2 and 2.10.1 with their respective release dates and download links.

Version	Release date	Source download	Binary download
3.3.1	2021 Jun 15	source (checksum signature)	binary (checksum signature) binary-aarch64 (checksum signature)
3.2.2	2021 Jan 9	source (checksum signature)	binary (checksum signature)
2.10.1	2020 Sep 21	source (checksum signature)	binary (checksum signature)

1.1 Click the “binary” link and save to local folder.

Example E:\Hadoop\hadoop-3.3.1.tar

1.2 Extract the tar file using WinRAR.

Note: Open the WinRAR with admin privilege and then extract

1.3 After extract, ensure your path like below

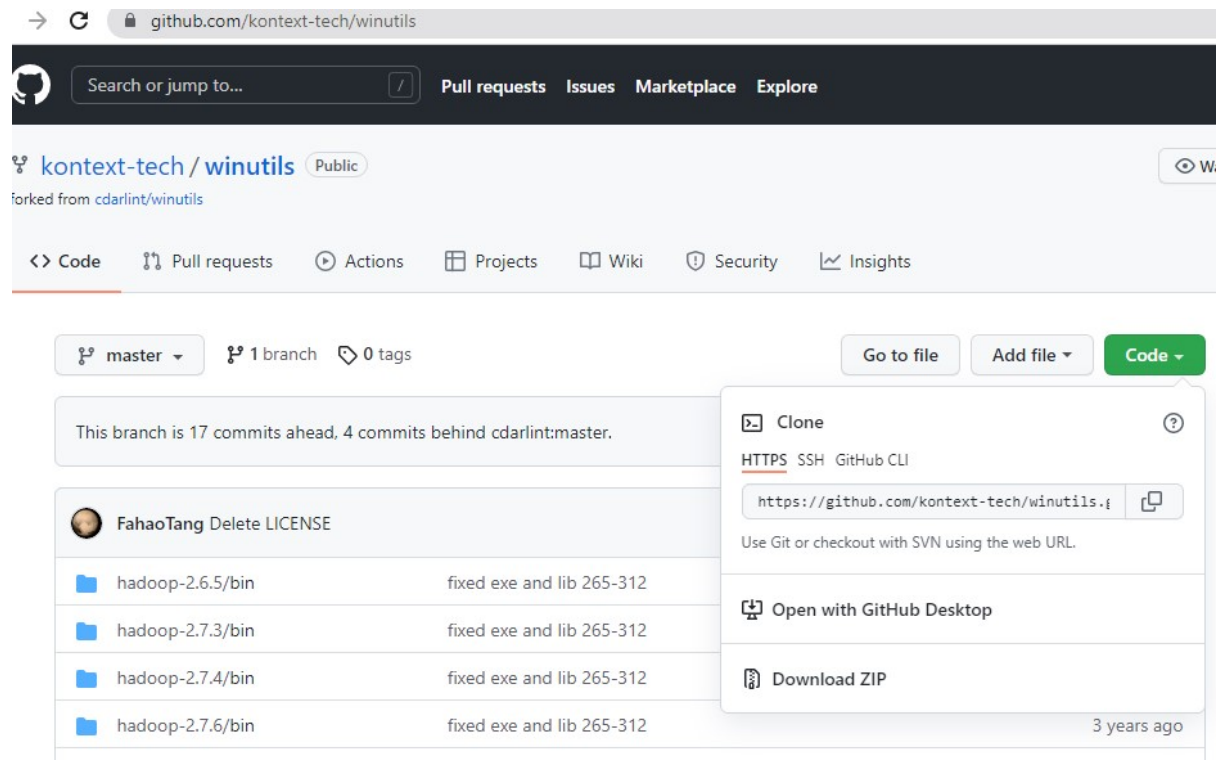
E:\Hadoop\hadoop-3.3.1\bin

2. Window Utilities Installation on Hadoop

Hadoop needs the Windows Utilities tools to run. It is not available in the default package downloaded. So download from the following URL and place it in the “E:\Hadoop\hadoop-3.3.1\bin” folder.

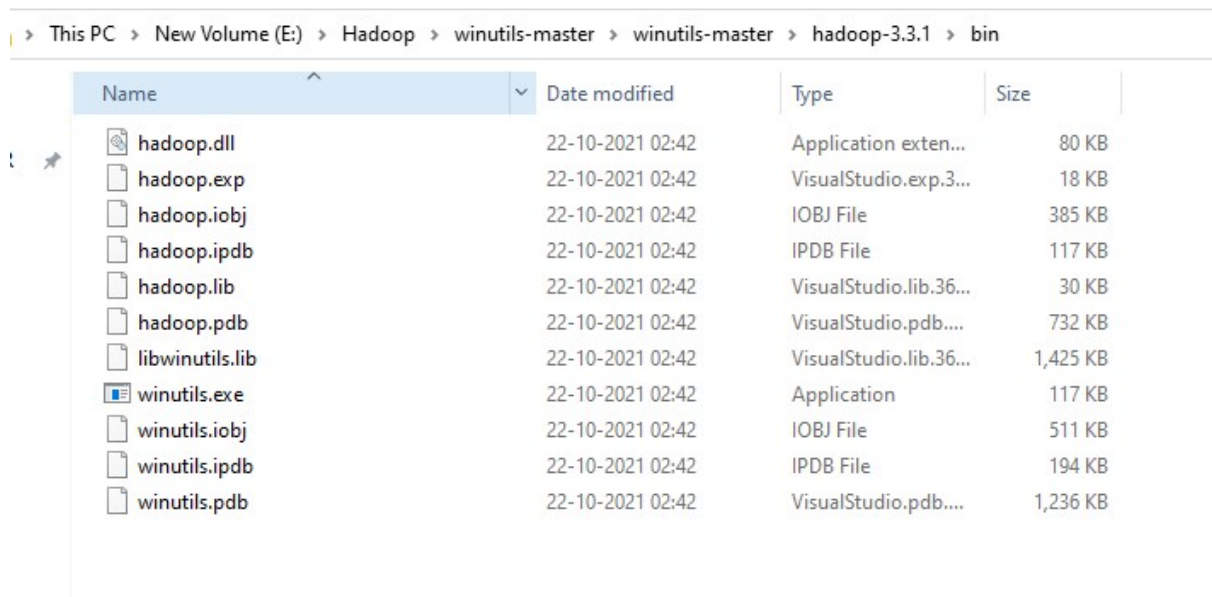
<https://github.com/kontext-tech/winutils>

Select the “Download Zip” option.



2.1 Extract the Zip files. And navigate the following path (Based on your downloaded location)

E:\Hadoop\winutils-master\winutils-master\hadoop-3.3.1\bin



Name	Date modified	Type	Size
hadoop.dll	22-10-2021 02:42	Application exten...	80 KB
hadoop.exp	22-10-2021 02:42	VisualStudio.exp.3...	18 KB
hadoop.iobj	22-10-2021 02:42	IOBJ File	385 KB
hadoop.ipdb	22-10-2021 02:42	IPDB File	117 KB
hadoop.lib	22-10-2021 02:42	VisualStudio.lib.36...	30 KB
hadoop.pdb	22-10-2021 02:42	VisualStudio.pdb....	732 KB
libwinutils.lib	22-10-2021 02:42	VisualStudio.lib.36...	1,425 KB
winutils.exe	22-10-2021 02:42	Application	117 KB
winutils.iobj	22-10-2021 02:42	IOBJ File	511 KB
winutils.ipdb	22-10-2021 02:42	IPDB File	194 KB
winutils.pdb	22-10-2021 02:42	VisualStudio.pdb....	1,236 KB

Select all the files, COPY and PASTE in the “E:\Hadoop\hadoop-3.3.1\bin” folder. After replace your Hadoop bin folder will be like below

» This PC » New Volume (E:) » Hadoop » hadoop-3.3.1 » bin

Name	Date modified	Type	Size
container-executor	15-06-2021 11:08	File	780 KB
hadoop	15-06-2021 10:44	File	9 KB
hadoop.cmd	15-06-2021 10:44	Windows Comma...	12 KB
hadoop.dll	22-10-2021 02:42	Application exten...	80 KB
hadoop.exp	22-10-2021 02:42	VisualStudio.exp.3...	18 KB
hadoop.iobj	22-10-2021 02:42	IOBJ File	385 KB
hadoop.ipdb	22-10-2021 02:42	IPDB File	117 KB
hadoop.lib	22-10-2021 02:42	VisualStudio.lib.36...	30 KB
hadoop.pdb	22-10-2021 02:42	VisualStudio.pdb....	732 KB
hdfs	15-06-2021 10:48	File	12 KB
hdfs.cmd	15-06-2021 10:48	Windows Comma...	8 KB
libwinutils.lib	22-10-2021 02:42	VisualStudio.lib.36...	1,425 KB
mapred	15-06-2021 11:09	File	7 KB
mapred.cmd	15-06-2021 11:09	Windows Comma...	7 KB
oom-listener	15-06-2021 11:08	File	33 KB
test-container-executor	15-06-2021 11:08	File	812 KB
winutils.exe	22-10-2021 02:42	Application	117 KB
winutils.iobj	22-10-2021 02:42	IOBJ File	511 KB
winutils.ipdb	22-10-2021 02:42	IPDB File	194 KB
winutils.pdb	22-10-2021 02:42	VisualStudio.pdb....	1,236 KB
yarn	15-06-2021 11:08	File	13 KB
yarn.cmd	15-06-2021 11:08	Windows Comma...	13 KB

3. Java Installation

Download the Java 8 from the following URL (May need to create the account and login in the Oracle website)

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

Select the [jdk-8u301-windows-x64.exe](#) and download.

→ ↻ [oracle.com/java/technologies/javase/javase8u211-later-archive-downloads.html](https://www.oracle.com/java/technologies/javase/javase8u211-later-archive-downloads.html)

ORACLE		
🔍 Products Industries Resources Support Events Developer Partners		
Solaris x64 Compressed Archive	92.66 MB	📄 jdk-8u301-solaris-x64.tar.gz
Windows x86 Installer	156.45 MB	📄 jdk-8u301-windows-i586.exe
Windows x64 Installer	169.46 MB	📄 jdk-8u301-windows-x64.exe

3.1 Install the Java

Note:

1. If your Java installation path contain empty space between then Hadoop will through error during start.

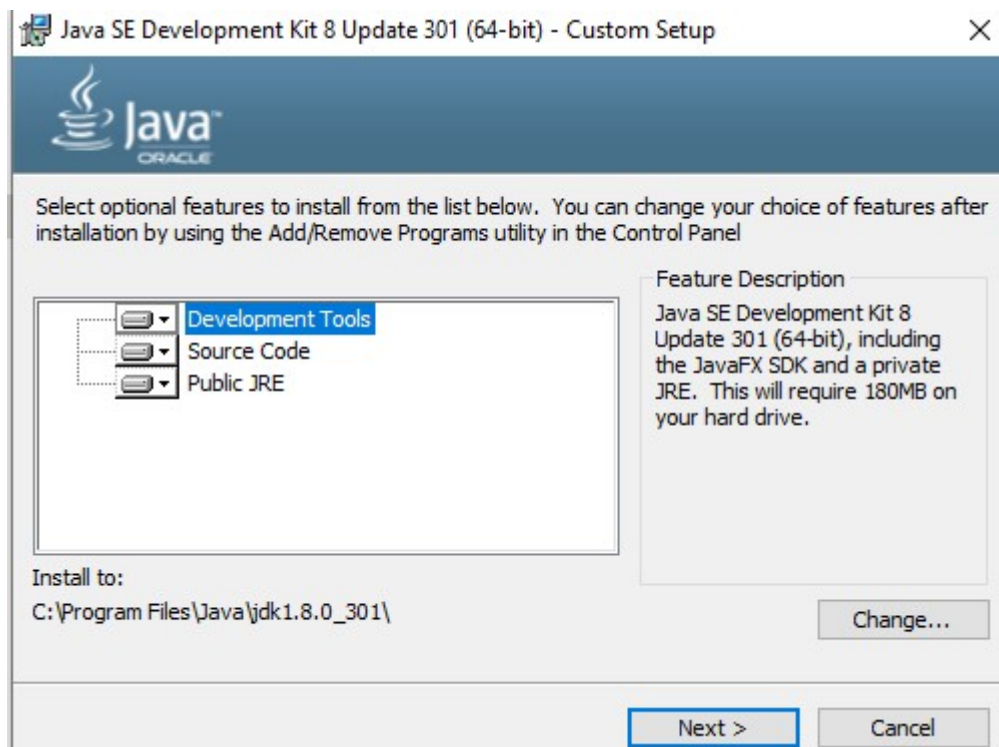
Example if you install Java in “C:\Program Files\Java” will through error. “C:\Java” is valid path

2. Same way if your Windows login User Name contains empty space then Hadoop through error while run.

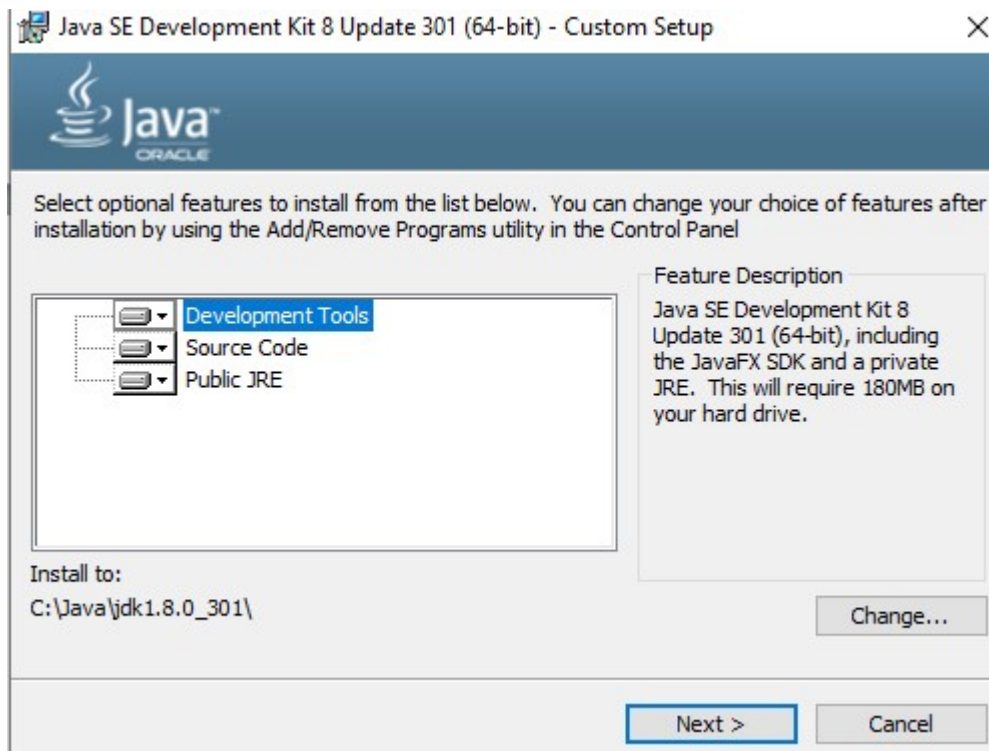
For example (SELVA KUMAR will throw the error. SELVAKUMAR is a valid user name)

Select the “C:\” Drive or any other drive to install. If Java already installed in C:\Program Files\Java”, then uninstall and then install again in C:\Java path.

So edit the path during Java installation and give the path without empty space.



Change to like below

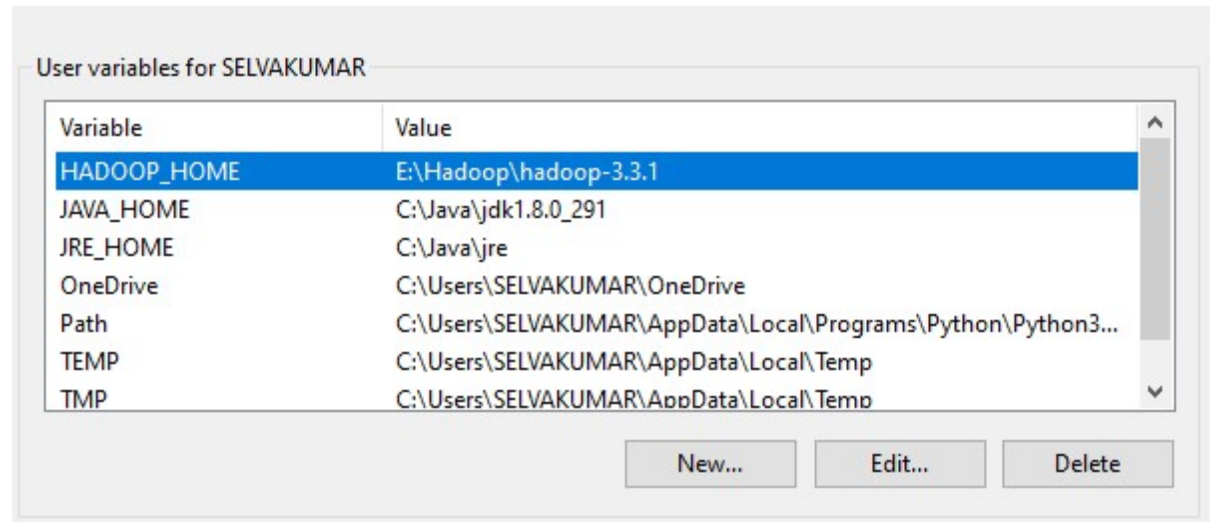


3.2 Similarly do the same for JRE.

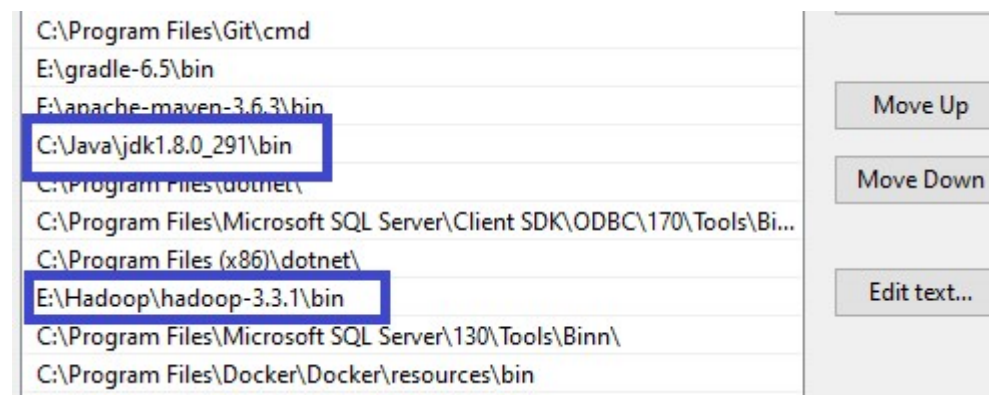
4. Set the Environment Variables

4.1 Set the HADOOP_HOME and JAVA_HOME based on your location installation.

Environment Variables



4.2 Add the bin location of Hadoop and Java to "Path" variable. This can be in User or System variables based on your preference.



5. Hadoop Configuration

5.1 core-site.xml

Go to the following file and add the following property values.

E:\Hadoop\hadoop-3.3.1\etc\hadoop\core-site.xml

Default:

```
<configuration>
```

```
</configuration>
```

After Modified:

```
<configuration>
```

```
  <property>
```

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

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

```
  </property>
```


```
</configuration>
```

5.2 hdfs-site.xml

Create the following folder

E:\Hadoop\node\data\dfs\data

E:\Hadoop\node\data\dfs\namespace_logs

This PC > New Volume (E:) > Hadoop > node > data > dfs			
Name		Date modified	Type
	data	17-12-2021 23:32	File folder
	namespace_logs	17-12-2021 23:32	File folder

Now, go to the following file and add the following property values.

E:\Hadoop\hadoop-3.3.1\etc\hadoop\hdfs-site.xml

Default:

```
<configuration>
```

```
</configuration>
```

After Modified:

```
<configuration>
```

```
  <property>
```

```
    <name>dfs.replication</name>
```

```
    <value>1</value>
```

```
  </property>
```

```
  <property>
```

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

```
    <value>file:///E:/Hadoop/node/data/dfs/namespace_logs</value>
```

```
  </property>
```

```
  <property>
```

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

```
    <value>file:///E:/Hadoop/node/data/dfs/data</value>
```

```
  </property>
```

```
</configuration>
```

5.3 mapred-site.xml

Go to the following file and add the following property values.

E:\Hadoop\hadoop-3.3.1\etc\hadoop\mapred-site.xml

Default:

```
<configuration>
```

```
</configuration>
```

After Modified:

```
<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/mapr  
educe/lib/*,%HADOOP_HOME%/share/hadoop/common/*,%HADOOP_HOME%/share/hadoop/com  
mon/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>
```

5.4 yarn-site.xml

Go to the following file and add the following property values.

E:\Hadoop\hadoop-3.3.1\etc\hadoop\yarn-site.xml

Default:

```
<configuration>
```

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

```
</configuration>
```

After Modified:

```
<configuration>
```

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

```
<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>
```

5. Name Node Format

Run the following command from command prompt

```
hdfs namenode -format
```

6. Start the Hadoop

Go to the following path in and execute the “hadoop-env.cmd” file

E:\Hadoop\hadoop-3.3.1\etc\hadoop>hadoop-env.cmd

```
E:\Hadoop\hadoop-3.3.1\etc\hadoop>hadoop-env.cmd
E:\Hadoop\hadoop-3.3.1\etc\hadoop>dir
Volume in drive E is New Volume
Volume Serial Number is 1C40-E556
```

Now, go to the following path in and execute the “start-all.cmd” file

E:\Hadoop\hadoop-3.3.1\sbin

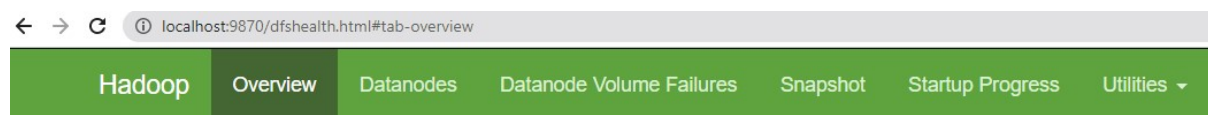
```
E:\Hadoop\hadoop-3.3.1\sbin>start-all.cmd
This script is Deprecated. Instead use start-dfs.cmd and start-yarn.cmd
starting yarn daemons
```

There are 4 windows will be started. The logs will be printed. Ensure there is no error or exception in the console. Wait till all log run print in the console.

7. Verify the Hadoop Health Status

7.1 Go the following URL in the browser

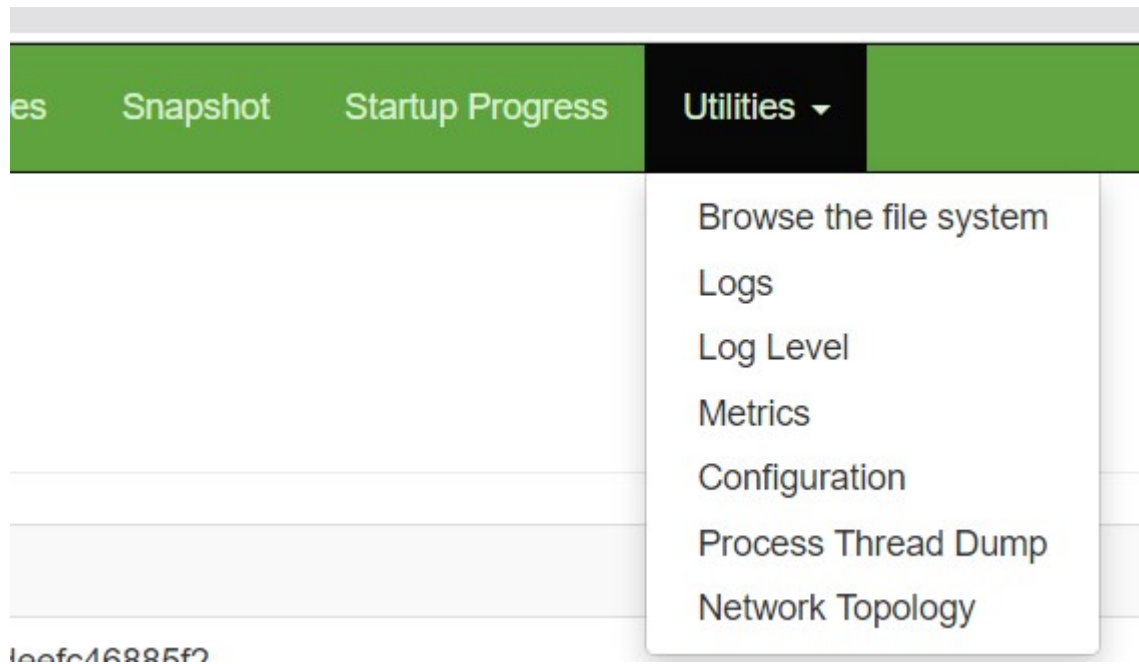
<http://localhost:9870/dfshealth.html#tab-overview>



Overview '0.0.0.0:19000' (✓active)

Started:	Sat Dec 18 00:59:06 +0530 2021
Version:	3.3.1, ra3b9c37a397ad4188041dd80621bdeefc46885f2
Compiled:	Tue Jun 15 10:43:00 +0530 2021 by ubuntu from (HEAD detached at release-3.3.1-RC3)
Cluster ID:	CID-d9103833-f75c-46ab-80ec-282d2627a59c
Block Pool ID:	BP-1743838207-192.168.56.1-1639763935962





7.2 You can browse the file system through the Utilities Menu.






localhost:9870/explorer.html#/

Hadoop Overview Datanodes Datanode Volume Failures Snapshot Startup Progress Utilities

Browse Directory

/ Go!    

Show 25 entries Search:

<input type="checkbox"/>	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name	
<input type="checkbox"/>	drwxr-xr-x	SELVAKUMAR	supergroup	0 B	Dec 17 23:34	0	0 B	input	
<input type="checkbox"/>	drwxr-xr-x	SELVAKUMAR	supergroup	0 B	Dec 17 23:35	0	0 B	output	
<input type="checkbox"/>	drwx-----	SELVAKUMAR	supergroup	0 B	Dec 17 23:35	0	0 B	tmp	

Showing 1 to 3 of 3 entries Previous 1 Next

8. Backup your Hadoop Configuration

Hadoop can crash some time. So it is backup your configuration and settings.

Also if want to distribute to other system or share to your team or classmates then you share the preconfigured backup Hadoop folder. This step helps to avoid the configuration every time and other system.

8.1 Create the zip file for the folder "E:\Hadoop\hadoop-3.3.1".

8.2 When need next time by you or friends, extract the backup zip file and modify the following value in **hdfs-site.xml** if required

```
<property>
```

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

```
    <value>file:///E:/Hadoop/node/data/dfs/namespace_logs</value>
```

```
</property>
```

```
<property>
```

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

```
    <value>file:///E:/Hadoop/node/data/dfs/data</value>
```

```
</property>
```

If you extracted to different drive in other system then change the path. For example if you placed your backup or extracted the zip files to "D" drive then change the drive value from "E" to "D" and straight away start the Hadoop.

3. Ensure HADOOP_HOME, JAVA_HOME environment variables and path value is configured.

4. Now you can directly start the Hadoop.

Reference:

<https://kontext.tech/column/hadoop/829/install-hadoop-331-on-windows-10-step-by-step-guide>

<https://kontext.tech/column/hadoop/377/latest-hadoop-321-installation-on-windows-10-step-by-step-guide>