

SET UP A SINGLE HADOOP CLUSTER AND SHOW THE PROCESS USING WEB UI

AIM: To set up a single hadoop cluster and show the process using web ui.

Installation steps:

Step 1: Download and install Java

Hadoop is built on Java, so you must have Java installed on your PC. You can get the most recent version of Java from the official website. After downloading, follow the installation wizard to install Java on your system.

JDK: <https://www.oracle.com/java/technologies/javase-downloads.html>

Step 2: Download Hadoop

Hadoop can be downloaded from the Apache Hadoop website. Make sure to have the latest stable release of Hadoop. Once downloaded, extract the contents to a convenient location.

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

Step 3: Set Environment Variables

You must configure environment variables after downloading and unpacking Hadoop. Launch the Start menu, type “Edit the system environment variables,” and select the result. This will launch the System Properties dialogue box. Click on “Environment Variables” button to open.

Click “New” under System Variables to add a new variable. Enter the variable

name “HADOOP_HOME” and the path to the Hadoop folder as the variable value. Then press “OK.”

Then, under System Variables, locate the “Path” variable and click “Edit.” Click “New” in the Edit Environment Variable window and enter “%HADOOP_HOME%bin” as the variable value. To close all the windows, use the “OK” button.

Step 4: Setup Hadoop

You must configure Hadoop in this phase by modifying several configuration files. Navigate to the “etc/hadoop” folder in the Hadoop folder. You must make changes to three files:

core-site.xml

hdfs-site.xml

mapred-site.xml

Open each file in a text editor and edit the following properties:

In core-site.xml

```
<configuration>
```

```
  <property>
```

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

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

```
  </property>
```

```
</configuration>
```

In hdfs-site.xml

```
<configuration>

  <property>

    <name>dfs.replication</name>

    <value>1</value>

  </property>

  <property>

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

    <value>C:/hadoop/hadoop/data/namenode</value>

  </property>

  <property>

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

    <value>C:/hadoop/hadoop/data/datanode</value>

  </property>

</configuration>
```

In mapred-site.xml

```
<configuration>

  <property>

    <name>mapred.job.tracker</name>

    <value>localhost:54311</value>

  </property>

</configuration>
```

Save the changes in each file.

Step 5: Format Hadoop NameNode

You must format the NameNode before you can start Hadoop. Navigate to the Hadoop bin folder using a command prompt. Execute this command:

```
hdfs namenode -format
```

Step 6: Start Hadoop

To start Hadoop, open a command prompt and navigate to the Hadoop bin folder. Run the following command:

```
start-dfs.cmd
```

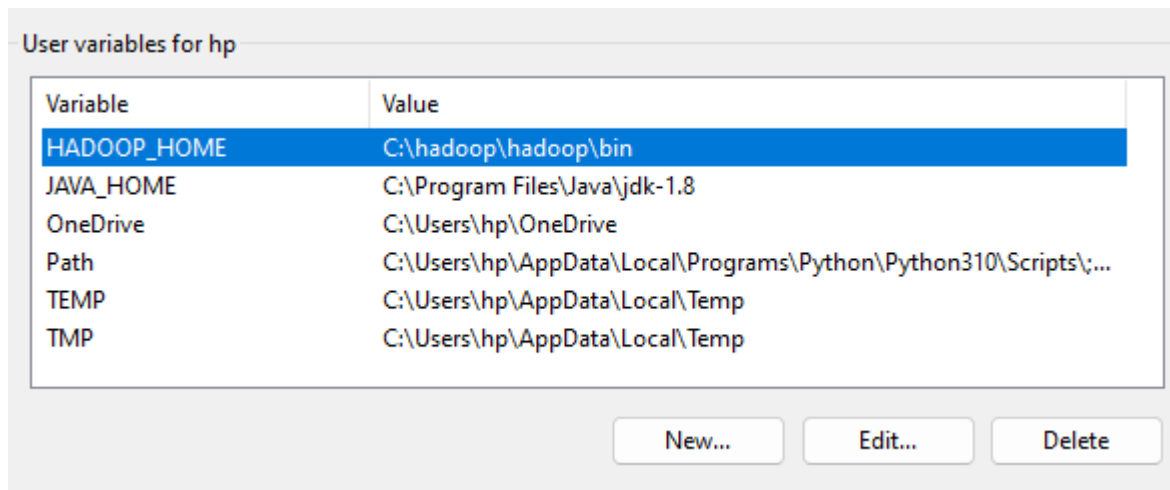
```
start-yarn.cmd
```

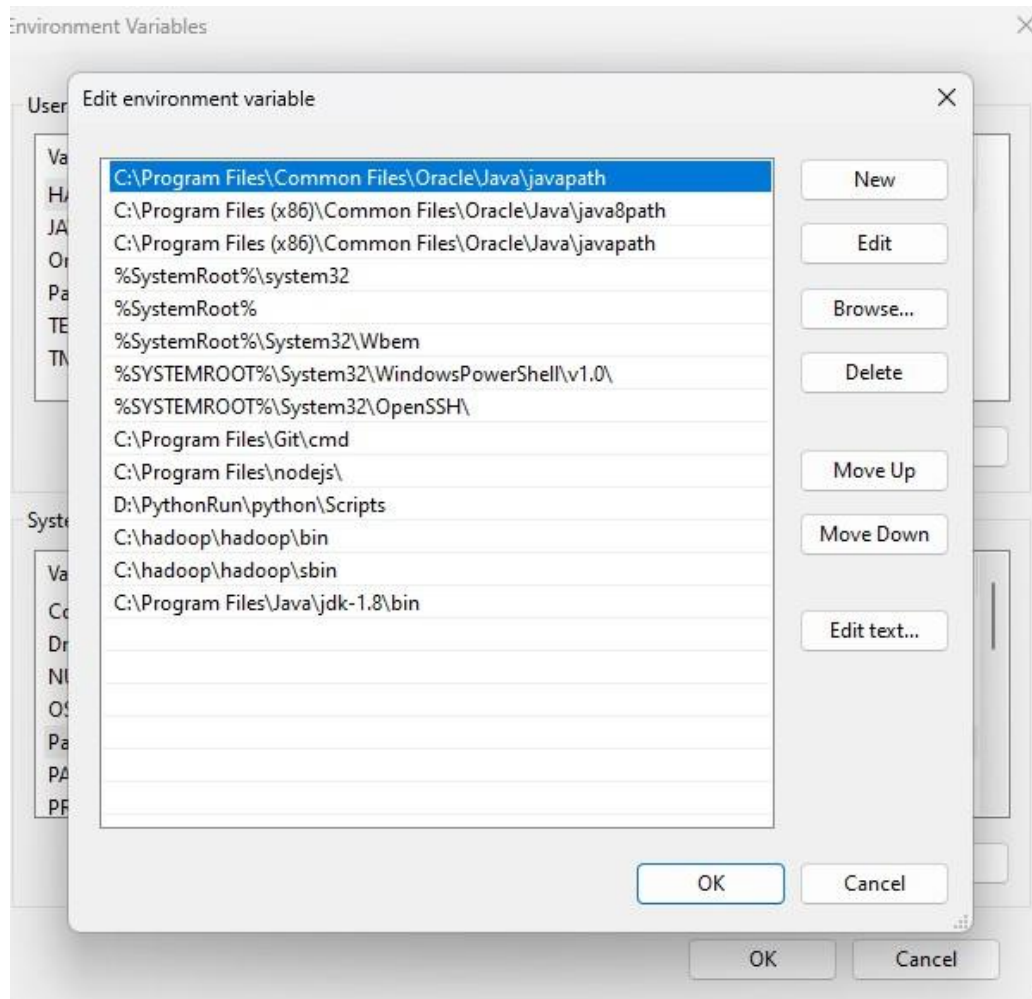
This command will start all the required Hadoop services, including the NameNode, DataNode, and JobTracker. Wait for a few minutes until all the services are started.

Installation

To ensure that Hadoop is properly installed, open a web browser and go to <http://localhost:9870>. This will launch the web interface for the Hadoop NameNode. You should see a page with Hadoop cluster information.

OUTPUT:





```
C:\Users\hp>hadoop version
Hadoop 3.3.6
Source code repository https://github.com/apache/hadoop.git -r 1be78238728da9266a4f88195058f08fd012bf9c
Compiled by ubuntu on 2023-06-18T08:22Z
Compiled on platform linux-x86_64
Compiled with protoc 3.7.1
From source with checksum 5652179ad55f76cb287d9c633bb53bbd
This command was run using /C:/hadoop/hadoop/share/hadoop/common/hadoop-common-3.3.6.jar
```

```
java version "1.8.0_421"
Java(TM) SE Runtime Environment (build 1.8.0_421-b09)
Java HotSpot(TM) 64-Bit Server VM (build 25.421-b09, mixed mode)
```

```
C:\Users\hp>hdfs namenode -format
2024-08-16 15:07:07,538 INFO namenode.NameNode: STARTUP_MSG:
/*****
STARTUP_MSG: Starting NameNode
STARTUP_MSG: host = DESKTOP-94HDGKA/192.168.0.102
STARTUP_MSG: args = [-format]
STARTUP_MSG: version = 3.3.6
STARTUP_MSG: classpath = C:\hadoop\hadoop\etc\hadoop;C:\hadoop\hadoop\share\hadoop\common;C:\hadoop\hadoop\share\hadoo
p\common\lib\animal-sniffer-annotations-1.17.jar;C:\hadoop\hadoop\share\hadoop\common\lib\audience-annotations-0.5.0.jar
;C:\hadoop\hadoop\share\hadoop\common\lib\avro-1.7.7.jar;C:\hadoop\hadoop\share\hadoop\common\lib\checker-qual-2.5.2.jar
;C:\hadoop\hadoop\share\hadoop\common\lib\commons-beanutils-1.9.4.jar;C:\hadoop\hadoop\share\hadoop\common\lib\commons-c
```

```
C:\Users\hp>start-dfs.cmd

C:\Users\hp>jps
10696 DataNode
17512 NameNode
16620 Jps
```

```
C:\Users\hp>start-yarn.cmd
starting yarn daemons

C:\Users\hp>jps
10696 DataNode
17512 NameNode
20456 Jps
264 NodeManager
9404 ResourceManager
```

Hadoop

Overview

Datanodes

Datanode Volume Failures

Snapshot

Startup Progress

Utilities

Overview 'localhost:9000' (✓active)

Started:	Fri Aug 16 15:08:58 +0530 2024
Version:	3.3.6, r1be78238728da9266a4f88195058f08fd012bf9c
Compiled:	Sun Jun 18 13:52:00 +0530 2023 by ubuntu from (HEAD detached at release-3.3.6-RC1)
Cluster ID:	CID-e4a3dea3-7721-4ab1-ab86-488205d901c7
Block Pool ID:	BP-160749255-192.168.0.102-1723778266263

Summary

Security is off.
Safemode is off.
7 files and directories, 3 blocks (3 replicated blocks, 0 erasure coded block groups) = 10 total filesystem object(s).
Heap Memory used 71.24 MB of 154 MB Heap Memory. Max Heap Memory is 889 MB.
Non Heap Memory used 52.72 MB of 54 MB Committed Non Heap Memory. Max Non Heap Memory is <unbounded>.

Configured Capacity:	231.71 GB
Configured Remote Capacity:	0 B
DFS Used:	513 B (0%)
Non DFS Used:	157.13 GB

RESULT:

Thus to set up a single hadoop cluster and show the process using web ui was completed successfully.