



P ① Perform setting up and Installing Hadoop in its two operating modes.

Ans.

- Pseudo distributed.
- Fully distributed

ii Use web based tools to monitor your Hadoop setup.

Solution: - Setting up and installing Hadoop in its two operating modes on a windows environment requires a series of steps. we will cover both modes, the installation of Hadoop & web-based tool to monitor the setup.

Prerequisites:

• Windows Subsystem for linux:-

As Hadoop is typically designed to run on linux-based system, installing Hadoop on windows directly can be cumbersome. Thus, we use WSL to create a linux-like environment on windows.

• Java:- Hadoop requires java to run. Make sure you have JDK installed.

• Hadoop binaries: Download the Hadoop binaries for your environment (e.g. from Apache Hadoop Release).



1. Pseudo-distributed Mode Installation :-

→ Step 1: Install WSL (if not already installed)

1. Enable WSL:

- Open PowerShell as Administrator and run:
- `dism.exe /online /enable-feature /featurename:Microsoft-Windows-Subsystem-Linux /all /norestart`

2. Install a linux distribution:

- from the Microsoft store, install Ubuntu or other preferred linux distribution.
- After installation, run it and setup the user credentials.

→ Step 2: Install Java in WSL

1. Open your WSL terminal and install Java using:
2. `sudo apt update`
3. `sudo apt install openjdk-8-jdk`
4. Verify the Java installation:
5. `java -version`

→ Step 3: Download & Extract Hadoop:-

1. Download Hadoop binaries from the Apache website
2. Extract the Hadoop tar file to a directory (ex: `/usr/local/hadoop`)
3. `tar -xzf hadoop-3.3.1.tar.gz -C /usr/local`
4. Set the environment variables in `.bashrc`:
5. `nano ~/.bashrc`

Add the following lines at the end of the file

```
export HADOOP_HOME=/usr/local/hadoop
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-
export PATH=$PATH:$HADOOP_HOME/bin:$HADOOP_HOME/sbin
```

6. Reload the .bashrc file:

7. Source ~/.bashrc.

→ Step 4: Configure Hadoop for Pseudo-distributed Mode

1. Edit the Hadoop Configuration Mode

core-site.xml: Located in \$HADOOP_HOME/etc/hadoop/core-site.xml

```
<configuration>
```

```
<property>
```

```
<name>fs-defaults</name>
```

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

```
</property>
```

```
</configuration>
```

hdfs-site.xml: Located in \$HADOOP_HOME/etc/hadoop/hdfs-site.xml

```
<configuration>
```

```
<property>
```

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

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

```
</property>
```

```
<property>
```

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




</property>

<property>

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

<value>file:///usr/local/hadoop/data/hdfs/datanode</value>

</property>

</configuration>

Mapred-site.xml: Located is \$HADOOP_HOME/etc/hadoop/
mapred-site.xml

<configuration>

<property>

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

<value>yarn</value>

</property>

</configuration>

<configuration>

<property>

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

<value>mapreduce-shuffle</value>

</property>

</configuration>

→ Step 5: Formate the HDFS File system

1. Initiculize HDFS running:

2. hdfs namenode-format

→ Step 6: Start Hadoop Daemons

1. start the Hadoop services.
2. start -dfs.sh
3. start-yarn.sh

→ Step 7: Verify the Installation

1. To check if the HDFS is running:
2. jps

you should see the following process:

- NameNode
- DataNode
- ResourceManager
- NodeManager.

→ Step 8: Monitor Hadoop using Web UI

1. HDFS Web UI: Access it at <http://localhost:50070/> for NameNode monitoring.
2. YARN Resource Manager UI: Access it at <http://localhost:8088/> to monitor YARN.

3. Web-based Tools to Monitor Hadoop setup:

1. HDFS Web UI:

- Navigate to <http://localhost:50070/> for NameNode web UI.
- You can view the health of the HDFS, file system status, & other HDFS statistics.



2. YARN Resource Manager UI:

- Navigate to <http://localhost:8088/> for YARN Resource Manager.
- This UI will show the status of running application, resource utilization, and job details.

3. Ganglia:

- you can set up Ganglia to monitor cluster performance (CPU, memory, disk, etc)
- follow Ganglia installation guides for your system.