**Exp.1 Downloading and installing Hadoop, Understanding different Hadoop modes, Startup scripts, Configuration files.**

**Aim:**

To Download and install Hadoop, Understanding different Hadoop modes, Startup scripts, Configuration files.

**Procedure:**

**1. Install WSL and a Linux Distribution**

If you haven’t already, install WSL and a Linux distribution (such as Ubuntu):

* **Install WSL:** Open PowerShell as Administrator and run:

powershell

Copy code

wsl --install

* **Install a Linux Distribution:** You can install Ubuntu or another distribution from the Microsoft Store.

**2. Update and Install Required Packages**

Once you have WSL set up, open your WSL terminal and update your package lists:

bash

Copy code

sudo apt update

sudo apt upgrade

Install Java Development Kit (JDK) and other required packages:

bash

Copy code

sudo apt install openjdk-8-jdk wget ssh pdsh

**3. Download and Install Hadoop**

1. **Download Hadoop:**

Visit the [Apache Hadoop releases page](https://hadoop.apache.org/releases.html) to get the latest version URL, or use the command below for Hadoop 3.2.1:

bash

Copy code

wget https://downloads.apache.org/hadoop/common/hadoop-3.2.1/hadoop-3.2.1.tar.gz

1. **Extract Hadoop:**

bash

Copy code

tar -xzf hadoop-3.2.1.tar.gz

1. **Move Hadoop to /opt:**

bash

Copy code

sudo mv hadoop-3.2.1 /opt/hadoop

1. **Set Environment Variables:**

Edit your profile file (~/.bashrc or ~/.zshrc) to set the Hadoop environment variables:

bash

Copy code

nano ~/.bashrc

Add the following lines:

bash

Copy code

export HADOOP\_HOME=/opt/hadoop

export JAVA\_HOME=/usr/lib/jvm/java-8-openjdk-amd64

export PATH=$HADOOP\_HOME/bin:$HADOOP\_HOME/sbin:$PATH

Apply the changes:

bash

Copy code

source ~/.bashrc

**4. Configure Hadoop**

1. **Edit core-site.xml:**

Open the configuration file:

bash

Copy code

nano $HADOOP\_HOME/etc/hadoop/core-site.xml

Add the following configuration:

xml

Copy code

<configuration>

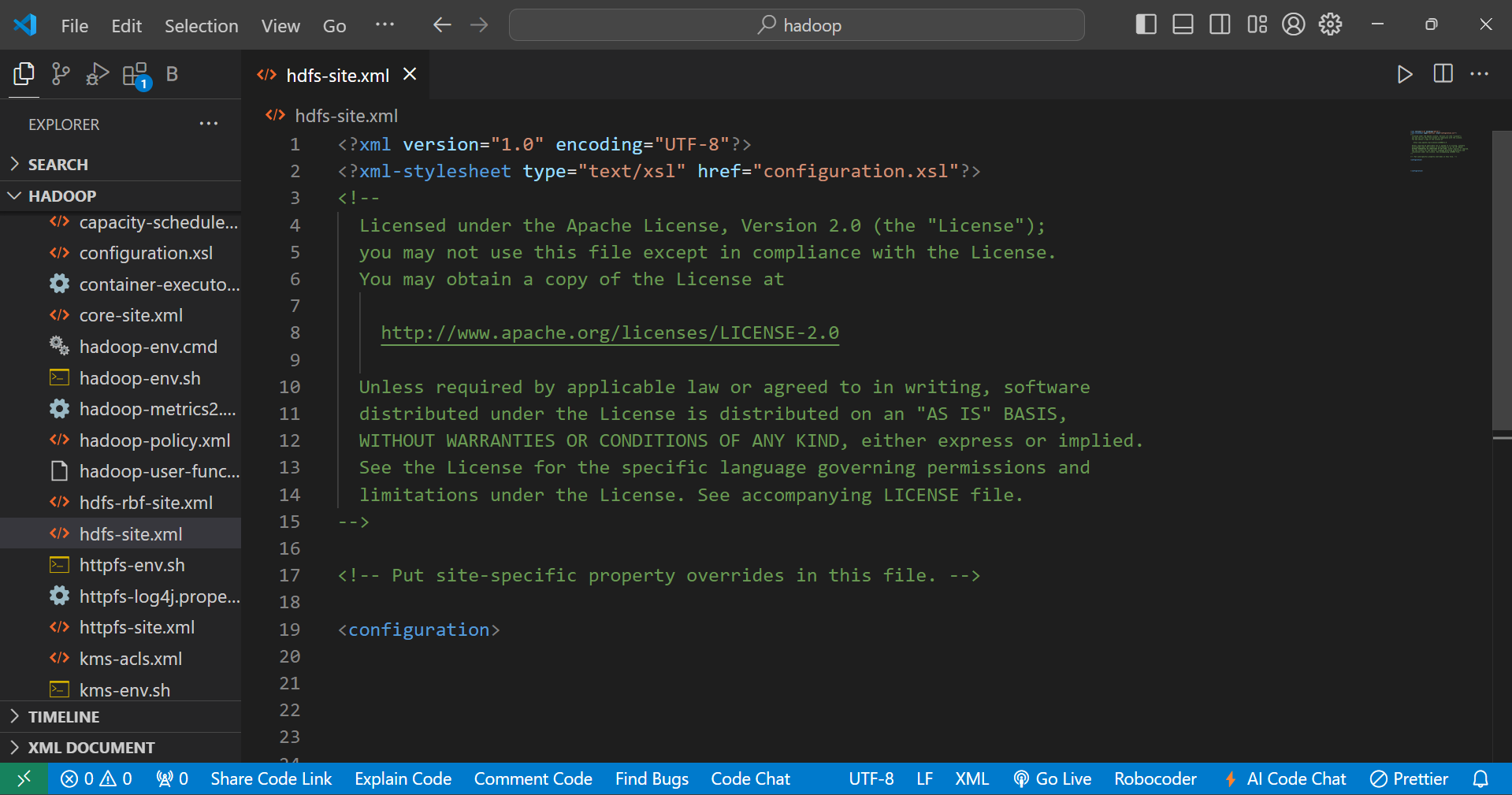
<property>

<name>fs.defaultFS</name>

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

</property>

</configuration>



1. **Edit hdfs-site.xml:**

Open the configuration file:

bash

Copy code

nano $HADOOP\_HOME/etc/hadoop/hdfs-site.xml

Add the following configuration:

xml

Copy code

<configuration>

<property>

<name>dfs.replication</name>

<value>1</value>

</property>

<property>

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

<value>/home/$(whoami)/hadoop/dfs/name</value>

</property>

<property>

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

<value>/home/$(whoami)/hadoop/dfs/data</value>

</property>

</configuration>

1. **Edit mapred-site.xml:**

Create a new file if it doesn’t exist:

bash

Copy code

cp $HADOOP\_HOME/etc/hadoop/mapred-site.xml.template $HADOOP\_HOME/etc/hadoop/mapred-site.xml

Edit the file:

bash

Copy code

nano $HADOOP\_HOME/etc/hadoop/mapred-site.xml

Add the following configuration:

xml

Copy code

<configuration>

<property>

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

<value>yarn</value>

</property>

</configuration>

1. **Edit yarn-site.xml:**

Open the configuration file:

bash

Copy code

nano $HADOOP\_HOME/etc/hadoop/yarn-site.xml

Add the following configuration:

xml

Copy code

<configuration>

<property>

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

<value>mapreduce\_shuffle</value>

</property>

<property>

<name>yarn.nodemanager.aux-services.mapreduce\_shuffle.class</name>

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

</property>

</configuration>

**5. Format the Namenode**

Format the HDFS filesystem (only needed for a new setup):

bash

Copy code

hdfs namenode -format

**6. Start Hadoop**

Start Hadoop services:

bash

Copy code

start-dfs.sh

start-yarn.sh

**7. Verify Installation**

Check the status of Hadoop services:

bash

Copy code

jps

You should see the following processes running:

* NameNode
* DataNode
* ResourceManager
* NodeManager

**8. Access Hadoop Web UI**

You can access the Hadoop web interfaces in your browser:

* **NameNode Web UI:** http://localhost:9870
* **ResourceManager Web UI:** <http://localhost:8088>

