**7.Aim:Study and configure hadoop for big data**

## Step 1 — Installing Java

To get started, we'll update our package list:

* sudo apt-get update

Next, we'll install OpenJDK, the default Java Development Kit on Ubuntu 16.04.

* sudo apt-get install default-jdk

Once the installation is complete, let's check the version.

* java -version

## Step 2 — Installing Hadoop

On the server, we'll use wget to fetch it:

* wget http://apache.mirrors.tds.net/hadoop/common/hadoop-2.7.3/hadoop-2.7.3.tar.gz

Again, we'll right-click to copy the file location, then use wget to transfer the file:

* wget https://dist.apache.org/repos/dist/release/hadoop/common/hadoop-2.7.3/hadoop-2.7.3.tar.gz.mds

Then run the verification:

* shasum -a 256 hadoop-2.7.3.tar.gz

Compare this value with the SHA-256 value in the .mds file:

* cat hadoop-2.7.3.tar.gz.mds
* tar -xzvf hadoop-2.7.3.tar.gz

Finally, we'll move the extracted files into /usr/local, the appropriate place for locally installed software. Change the version number, if needed, to match the version you downloaded.

* sudo mv hadoop-2.7.3 /usr/local/hadoop

With the software in place, we're ready to configure its environment.

## Step 3 — Configuring Hadoop's Java Home

To find the default Java path

* readlink -f /usr/bin/java | sed "s:bin/java::"

To begin, open hadoop-env.sh:

* sudonano /usr/local/hadoop/etc/hadoop/hadoop-env.sh

Then, choose one of the following options:

### Option 1: Set a Static Value

#export JAVA\_HOME=${JAVA\_HOME}

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

. . .

### Option 2: Use Readlink to Set the Value Dynamically

#export JAVA\_HOME=${JAVA\_HOME}

export JAVA\_HOME=$(readlink -f /usr/bin/java | sed "s:bin/java::")

## Step 4 — Running Hadoop

Now we should be able to run Hadoop:

* $ /usr/local/hadoop/bin/hadoop
* $ mkdir ~/input
* $ cp /usr/local/hadoop/etc/hadoop/\*.xml ~/input