

EXP. 9: PERFORM THE BASIC CONFIGURATION SETUP FOR INSTALLING HADOOP 2.X LIKE CREATING THE HDUSER AND SSH LOCALHOST

AIM:

PROCEDURE:

Step 1 – System Update

```
$ sudo apt-get update
```

Step 2 – Install Java and Set JAVA_HOME

```
//This first thing to do is to setup the webupd8 ppa on your system. Run the following command and proceed.
```

```
$ sudo apt-add-repository ppa:webupd8team/java
```

```
$ sudo apt-get update
```

```
//After setting up the ppa repository, update the package cache as well.
```

```
//Install the Java 8 installer
```

```
$ sudo apt-get install oracle-java8-installer
```

```
// After the installation is finished, Oracle Java is setup. Run the java command again to check the version and vendor.
```

[or]

```
$ sudo apt-get install default-jdk
```

```
$ java -version
```

Step 3 – Add a dedicated Hadoop user

```
$ sudo addgroup hadoop
```

```
$ sudo adduser --ingroup hadoop hduser
```

```
// Add hduser to sudo user group
```

```
$ sudo adduser hduser sudo
```

Step 4 – Install SSH and Create Certificates

```
$ sudo apt-get install ssh
```

```
$ su hduser
```

```
$ ssh-keygen -t rsa -P ""
```

```
// Set Environmental variables
```

```
$ cat $HOME/.ssh/id_rsa.pub >> $HOME/.ssh/authorized_keys
```

Step 5 – Check if SSH works

```
$ ssh localhost
```

Step 6 – Install Hadoop

```
// Extract Hadoop-2.7.2
```

```
$ sudo tar xvfz hadoop-2.7.2.tar.gz
```

```
// Create a folder 'hadoop' in /usr/local
```

```
$ sudo mkdir -p /usr/local/hadoop
```

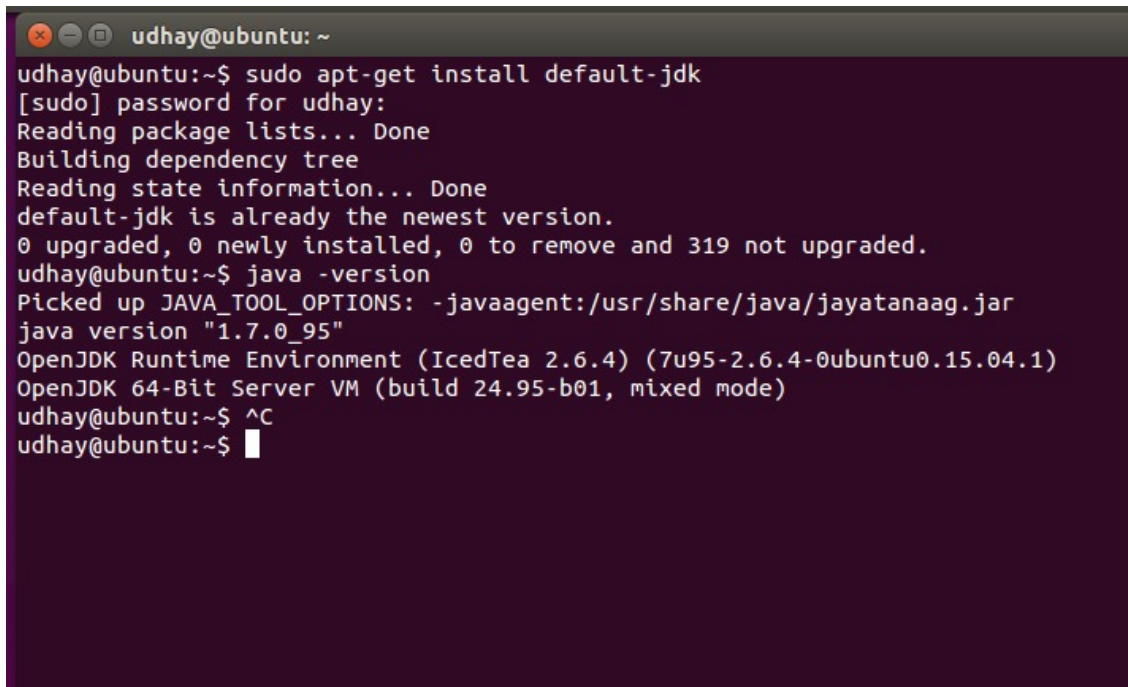
```
// Move the Hadoop folder to /usr/local/hadoop
```

```
$ sudo mv hadoop-2.7.2 /usr/local/hadoop
```

```
// Assigning read and write access to Hadoop folder
```

```
$ sudo chown -R hduser:hadoop /usr/local/hadoop
```

Implementation:

A terminal window titled 'udhay@ubuntu: ~' with a dark purple background. The terminal shows the following commands and output:

```
udhay@ubuntu:~$ sudo apt-get install default-jdk
[sudo] password for udhay:
Reading package lists... Done
Building dependency tree
Reading state information... Done
default-jdk is already the newest version.
0 upgraded, 0 newly installed, 0 to remove and 319 not upgraded.
udhay@ubuntu:~$ java -version
Picked up JAVA_TOOL_OPTIONS: -javaagent:/usr/share/java/jayatanaag.jar
java version "1.7.0_95"
OpenJDK Runtime Environment (IcedTea 2.6.4) (7u95-2.6.4-0ubuntu0.15.04.1)
OpenJDK 64-Bit Server VM (build 24.95-b01, mixed mode)
udhay@ubuntu:~$ ^C
udhay@ubuntu:~$
```

```
udhay@ubuntu:~$ sudo apt-get install ssh
Reading package lists... Done
Building dependency tree
Reading state information... Done
ssh is already the newest version.
0 upgraded, 0 newly installed, 0 to remove and 319 not upgraded.
udhay@ubuntu:~$ su hduser
Password:
hduser@ubuntu:/home/udhay$
```

```
udhay@ubuntu:~$ su hduser
Password:
hduser@ubuntu:/home/udhay$ ssh-keygen -t rsa -P ""
Generating public/private rsa key pair.
Enter file in which to save the key (/home/hduser/.ssh/id_rsa):
/home/hduser/.ssh/id_rsa already exists.
Overwrite (y/n)? y
Your identification has been saved in /home/hduser/.ssh/id_rsa.
Your public key has been saved in /home/hduser/.ssh/id_rsa.pub.
The key fingerprint is:
09:0f:15:f2:b2:b7:5e:11:1a:6c:d3:2f:c3:09:02:15 hduser@ubuntu
The key's randomart image is:
+---[RSA 2048]---+
|      ..E.o.      |
|      . = .       |
|      = B o       |
|      O B +       |
|      . S * .     |
|      . . +       |
|      . .         |
|      . .         |
|      .           |
+-----+
hduser@ubuntu:/home/udhay$
```

```
hduser@ubuntu:/home/udhay$ cat $HOME/.ssh/id_rsa.pub >> $HOME/.ssh/authorized_keys
hduser@ubuntu:/home/udhay$ ssh localhost
Welcome to Ubuntu 15.04 (GNU/Linux 3.19.0-84-generic x86_64)

* Documentation:  https://help.ubuntu.com/

Last login: Thu Jul 15 22:00:14 2021 from localhost
hduser@ubuntu:~$
```

Home

Clone of Ubuntu 64-bit

About the Cluster - Mozilla Firefox

Restore Session

About the Cluster

Namenode information

localhost:8088/cluster/cluster

Search

About the Cluster

Cluster

About

Nodes

Node Labels

Applications

NEW

NEW SAVING

SUBMITTED

ACCEPTED

RUNNING

FINISHED

FAILED

KILLED

Scheduler

Tools

Cluster Metrics

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total	Memory Reserved	VCores Used	VCores Total	VCores Reserved	Active Nodes
0	0	0	0	0	0 B	8 GB	0 B	0	8	0	1

Scheduler Metrics

Scheduler Type	Scheduling Resource Type	Minimum Allocation
Capacity Scheduler	[MEMORY]	<memory:1024, vCores:1>

Cluster ID:

1626414170591

ResourceManager state:

STARTED

ResourceManager HA state:

active

ResourceManager HA zookeeper connection state:

ResourceManager HA is not enabled.

ResourceManager RMStateStore:

org.apache.hadoop.yarn.server.resourcemanager.recovery.NullRMStateStore

ResourceManager started on:

Thu Jul 15 22:42:50 -0700 2021

ResourceManager version:

2.7.2 from b165c4fe8a74265c792ce23f546c64604acf0e41 by jenkins source checksum 2016-01-26T00:16Z

Hadoop version:

2.7.2 from b165c4fe8a74265c792ce23f546c64604acf0e41 by jenkins source checksum 2016-01-26T00:08Z

Activate Windows

RESULT:

```
1 @echo off
2 %rem Licensed to the Apache Software Foundation (ASF) under one or more
3 %rem contributor license agreements. See the NOTICE file distributed with
4 %rem this work for additional information regarding copyright ownership.
5 %rem The ASF licenses this file to You under the Apache License, Version 2.0
6 %rem (the "License"); you may not use this file except in compliance with
7 %rem the License. You may obtain a copy of the License at
8 %rem
9 %rem http://www.apache.org/licenses/LICENSE-2.0
10 %rem
11 %rem Unless required by applicable law or agreed to in writing, software
12 %rem distributed under the License is distributed on an "AS IS" BASIS,
13 %rem WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
14 %rem See the License for the specific language governing permissions and
15 %rem limitations under the License.
16
17 %rem Set Hadoop-specific environment variables here.
18
19 %rem The only required environment variable is JAVA_HOME. All others are
20 %rem optional. When running a distributed configuration it is best to
21 %rem set JAVA_HOME in this file, so that it is correctly defined on
22 %rem remote nodes.
23
24 %rem The java implementation to use. Required.
25 set JAVA_HOME=C:\Java\jdk1.8.0_241
26
27 %rem The javc implementation to use. Jsvc is required to run secure datanodes.
28 %rem set JSVC_HOME=%JSVC_HOME%
29
30 %rem set HADOOP_CONF_DIR=
31
32 %rem Extra Java CLASSPATH elements. Automatically insert capacity-scheduler.
33 if exist %HADOOP_HOME%\contrib\capacity-scheduler (
34   if not defined HADOOP_CLASSPATH (
35     set HADOOP_CLASSPATH=%HADOOP_HOME%\contrib\capacity-scheduler\*.jar
36   ) else (
37     set HADOOP_CLASSPATH=%HADOOP_CLASSPATH%;%HADOOP_HOME%\contrib\capacity-scheduler\*.jar
38   )
39 )
```

