

## **EXP. 22: PERFORM THE BASIC CONFIGURATION SETUP FOR INSTALLING HADOOP 2.X LIKE CREATING THE HDUSER AND SSH LOCALHOST**

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

### **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 xvzf 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:

```
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:~$
```

The screenshot displays the Hadoop web interface in a Mozilla Firefox browser window. The address bar shows 'localhost:8088/cluster/cluster'. The page title is 'About the Cluster'. On the left, there is a sidebar with a 'Cluster' menu containing options like 'About', 'Nodes', 'Node Labels', 'Applications', and a 'Tools' button. The main content area is titled 'About the Cluster' and contains several sections:

**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 Configuration Details:**

- 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

**RESULT:** To Perform The Basic Configuration Setup For Installing Hadoop 2.X Like Creating The Hduser And Ssh Localhost Have Successfully Implemented