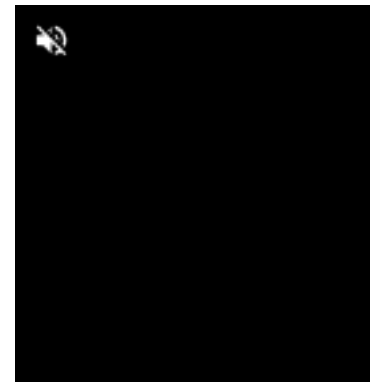


[HOME](#) [UBUNTU 18.04](#) [TUTORIALS](#) [FUNNY TOOLS](#) [FEEDBACK](#)[SUBMIT ARTICLE](#) [ABOUT US](#)

How to Setup Hadoop on Ubuntu 18.04 & 16.04 LTS

Written by [Rahul](#), Updated on April 9, 2019

💎 BIG-DATA | 💎 hadoop, hadoop setup



GOCLEAN steam ca

Professional manu
distributor wanted,
provided, various bi

Invest with Fast Direct Execution

Multi-awarded, multi-regulated broker

Our services involve a significant risk and can result in the loss of your invested capi

Apache Hadoop 3.1 have noticeable improvements any many bug fixes over the previous stable 3.0 releases. This version has many improvements in HDFS and MapReduce. This tutorial will help you to install and configure Hadoop 3.1.2 **Single-Node Cluster** on Ubuntu 18.04, 16.04 LTS and LinuxMint Systems. This article has been tested with Ubuntu 18.04 LTS.

POPULAR POSTS

[How To Install Java on Debian 10 \(Buster\)](#)

[How to Clone Specific Git Branch from Remote Repository](#)

[How to Check Laravel](#)

We use cookies to ensure that we give you the best experience on our website. If you continue to use this site we will assume that you are happy with it.

Debian 10/9/8

How to Install Slack on
Ubuntu 18.04 (Bionic)
Desktop

Step 1 – Prerequisites

Java is the primary requirement for running Hadoop on any system, So make sure you have Java installed on your system using the following command. If you don't have Java installed on your system, use one of the following links to install it first.

- [Install Oracle Java 11 on Ubuntu 18.04 LTS \(Bionic\)](#)
- [Install Oracle Java 11 on Ubuntu 16.04 LTS \(Xenial\)](#)

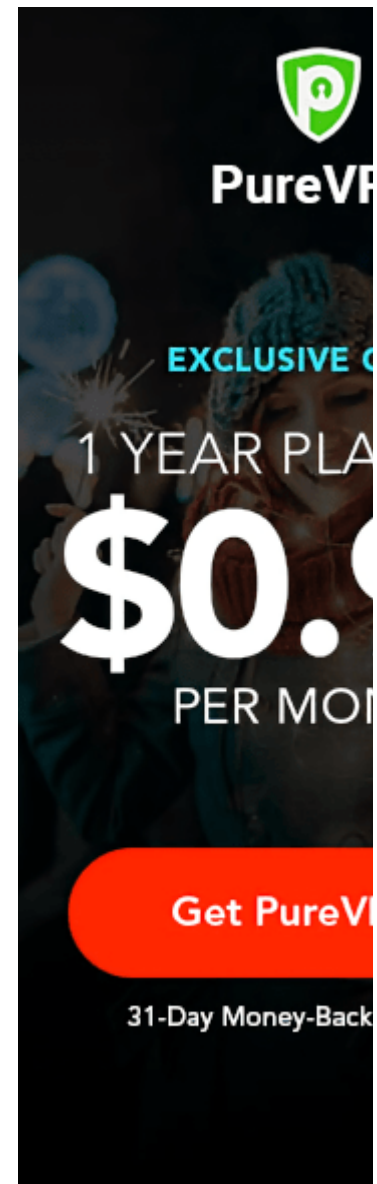
Step 2 – Create User for Hadoop

We recommend creating a normal (not root) account for Hadoop working. To create an account using the following command.

```
$ adduser hadoop
```

After creating the account, it also required to set up key-based ssh to its own account. To do this use execute following

We use cookies to ensure that we give you the best experience on our website. If you continue to use this site we will assume that you are happy with it.



```
$ ssh-keygen -t rsa -P '' -f ~/.ssh/id_rsa
$ cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys
$ chmod 0600 ~/.ssh/authorized_keys
```

Now, SSH to localhost with Hadoop user. This should not ask for the password but the first time it will prompt for adding RSA to the list of known hosts.

```
$ ssh localhost
$ exit
```

Step 3 – Download Hadoop Source Archive

In this step, download hadoop 3.1 source archive file using below command. You can also select alternate [download mirror](#) for increasing download speed.

```
$ cd ~
$ wget http://www-eu.apache.org/dist/hadoop/
$ tar xzf hadoop-3.1.2.tar.gz
$ mv hadoop-3.1.2 hadoop
```

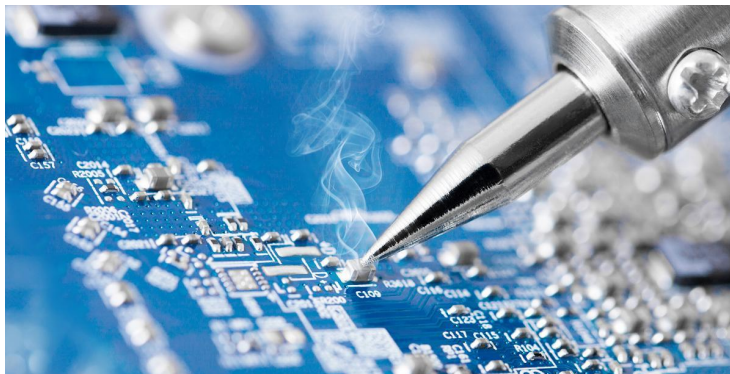
Step 4 – Setup Hadoop Pseudo-Distributed Mode

4.1. Setup Hadoop Environment Variables

We use cookies to ensure that we give you the best experience on our website. If you continue to use this site we will assume that you are happy with it.

following values at end of file.

```
export HADOOP_HOME=/home/hadoop/hadoop
export HADOOP_INSTALL=$HADOOP_HOME
export HADOOP_MAPRED_HOME=$HADOOP_HOME
export HADOOP_COMMON_HOME=$HADOOP_HOME
export HADOOP_HDFS_HOME=$HADOOP_HOME
export YARN_HOME=$HADOOP_HOME
export HADOOP_COMMON_LIB_NATIVE_DIR=$HADO
export PATH=$PATH:$HADOOP_HOME/sbin:$HADO
```



Best Performing FPGA Accelerating the Data

Ad Pair with our Intellectu
offerings for unmatched fu
dinigroup.com

[Learn more](#)

Then, apply the changes in the current running environment

```
$ source ~/.bashrc
```

Now edit **\$HADOOP_HOME/etc/hadoop/hadoop-env.sh** file and set **JAVA_HOME** environment variable. Change the JAVA path as per install on your system. This path may vary as per your operating system version and installation source. So make sure you are using the correct path.

```
$ vim $HADOOP_HOME/etc/hadoop/hadoop-env.sh
```

We use cookies to ensure that we give you the best experience on our website. If you continue to use this site we will assume that you are happy with it.

Ok

No

```
export JAVA_HOME=/usr/lib/jvm/java-11-ora
```

4.2. Setup Hadoop Configuration Files

Hadoop has many configuration files, which need to configure as per requirements of your Hadoop infrastructure. Let's start with the configuration with basic Hadoop single node cluster setup. first, navigate to below location

```
$ cd $HADOOP_HOME/etc/hadoop
```

Edit core-site.xml

```
<configuration>
<property>
  <name>fs.default.name</name>
  <value>hdfs://localhost:9000</value>
</property>
</configuration>
```

Edit hdfs-site.xml

```
<configuration>
<property>
  <name>dfs.replication</name>
  <value>1</value>
</property>

<property>
```

We use cookies to ensure that we give you the best experience on our website. If you continue to use this site we will assume that you are happy with it.

```
<property>
  <name>dfs.data.dir</name>
  <value>file:///home/hadoop/hadoopdata</value>
</property>
</configuration>
```

Edit mapred-site.xml

```
<configuration>
  <property>
    <name>mapreduce.framework.name</name>
    <value>yarn</value>
  </property>
</configuration>
```

Edit yarn-site.xml

```
<configuration>
  <property>
    <name>yarn.nodemanager.aux-services</name>
    <value>mapreduce_shuffle</value>
  </property>
</configuration>
```

4.3. Format Namenode

Now format the namenode using the following command, make sure that Storage directory is

```
$ hdfs namenode -format
```

Sample output:

We use cookies to ensure that we give you the best experience on our website. If you continue to use this site we will assume that you are happy with it.

```
STARTUP_MSG: Starting NameNode
STARTUP_MSG:   host = tecadmin/127.0.1.1
STARTUP_MSG:   args = [-format]
STARTUP_MSG:   version = 3.1.2
...
...
...
2018-05-02 17:52:13,717 INFO common.Stora
2018-05-02 17:52:13,806 INFO namenode.FSI
2018-05-02 17:52:14,161 INFO namenode.FSI
2018-05-02 17:52:14,224 INFO namenode.NNS
2018-05-02 17:52:14,282 INFO namenode.Nam
/*****
SHUTDOWN_MSG: Shutting down NameNode at t
*****
```

Step 5 – Start Hadoop Cluster

Let's start your Hadoop cluster using the scripts provided by Hadoop. Just navigate to your `$HADOOP_HOME/sbin` directory and execute scripts one by one.

```
$ cd $HADOOP_HOME/sbin/
```

Now execute **start-dfs.sh** script.

```
$ ./start-dfs.sh
```

Then execute **start-yarn.sh** script.

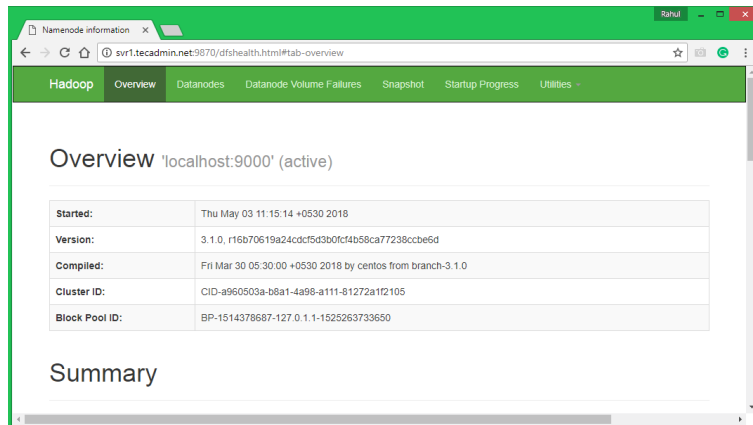
```
$ ./start-yarn.sh
```

We use cookies to ensure that we give you the best experience on our website. If you continue to use this site we will assume that you are happy with it.

Services in Browser

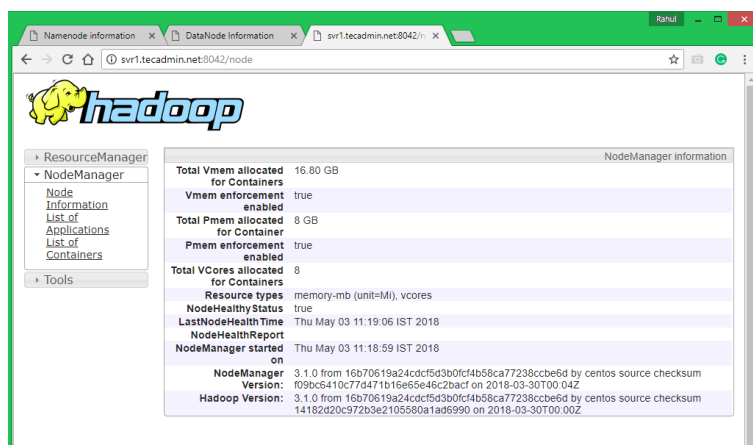
Hadoop NameNode started on default port 9870. Access your server on port 9870 in your favorite web browser.

```
http://svr1.tecadmin.net:9870/
```



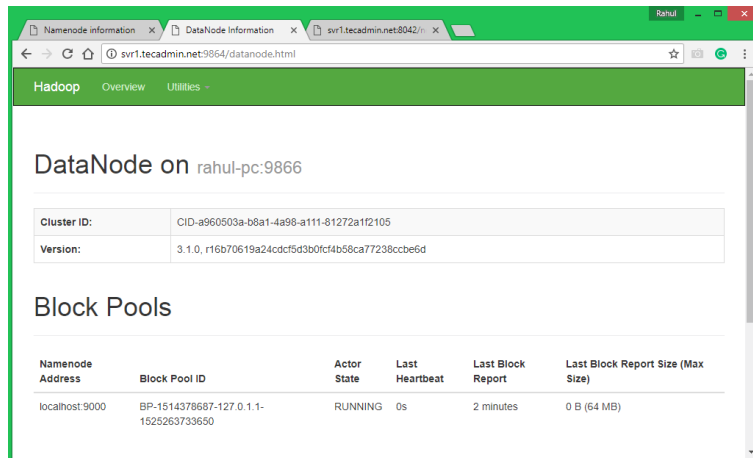
Now access port 8042 for getting the information about the cluster and all applications

```
http://svr1.tecadmin.net:8042/
```



We use cookies to ensure that we give you the best experience on our website. If you continue to use this site we will assume that you are happy with it.

<http://svr1.tecadmin.net:9864/>



Step 7 – Test Hadoop Single Node Setup

7.1. Make the HDFS directories required using following commands.

```
$ bin/hdfs dfs -mkdir /user
$ bin/hdfs dfs -mkdir /user/hadoop
```



Best Performing FPGA
Accelerating the Data

Ad Pair with our Intellectu
offerings for unmatched fu
dinigroup.com

[Learn more](#)

7.2. Copy all files from local file system
/var/log/httpd to hadoop distributed file

We use cookies to ensure that we give you the best experience on our website. If you continue to use this site we will assume that you are happy with it.

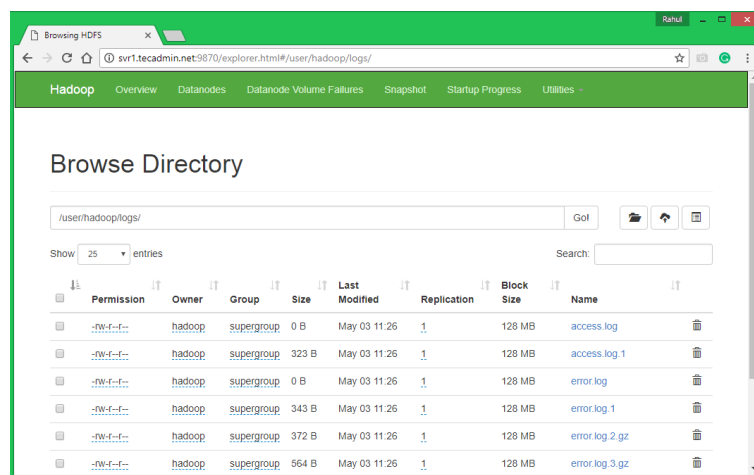
Ok

No

```
$ bin/hdfs dfs -put /var/log/apache2 logs
```

7.3. Browse Hadoop distributed file system by opening below URL in the browser. You will see an apache2 folder in the list. Click on the folder name to open and you will find all log files there.

```
http://svr1.tecadmin.net:9870/explorer.h
```



7.4 - Now copy logs directory for hadoop distributed file system to local file system.

```
$ bin/hdfs dfs -get logs /tmp/logs  
$ ls -l /tmp/logs/
```

You can also check [this tutorial](#) to run wordcount mapreduce job example using command line.

SPONSORED SEARCHES

We use cookies to ensure that we give you the best experience on our website. If you continue to use this site we will assume that you are happy with it.

OkNo

SPONSORED SEARCHES



SHARE IT!

**RAHUL**

I, Rahul Kumar am the founder and chief editor of TecAdmin.net. I am a Red Hat Certified Engineer (RHCE) and working as an IT professional since 2009..

RELATED POSTS

- **Hadoop Commands to Manage Files on HDFS**

August 26, 2016

- **HADOOP/HDFS ls: '': No such file or directory**

August 26, 2016

- **Hadoop – Namenode is in safe mode**

August 18, 2016

- **Hadoop – Running a Wordcount Mapreduce Example**

August 10, 2016

-

- **How to Setup Hadoop 3.1 on CentOS & Fedora**

November 12, 2015

We use cookies to ensure that we give you the best experience on our website. If you continue to use this site we will assume that you are happy with it.

COMMENTS ***NAME *****EMAIL *****WEBSITE**

I'm not a robot

reCAPTCHA
[Privacy](#) - [Terms](#)**SUBMIT**

Copyright © 2013-2019 [TecAdmin.net](#). All Rights Reserved. This site uses cookies. By using this website you agree with our [term and services](#)

We use cookies to ensure that we give you the best experience on our website. If you continue to use this site we will assume that you are happy with it.

Ok

No