

How to install Apache Hadoop on Ubuntu ?

Apache Hadoop Tutorial – We shall learn to install Apache Hadoop on Ubuntu. Java is a prerequisite to run Hadoop.

Install Apache Hadoop on Ubuntu

Following is a step by step guide to Install Apache Hadoop on Ubuntu

Install Java

Hadoop is an open-source framework written in Java. So, for Hadoop to run on your computer, you should install Java in prior.

Open a terminal and run the following command :

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

To verify the installation of Java, run the following command in the terminal :

```
$ java -version
```

The output for the command would be as shown below.

```
hadoopuser@tutorialkart:~# java -version
openjdk version "1.8.0_131"
OpenJDK Runtime Environment (build 1.8.0_131-8u131-b11-0ubuntu1.16.04.2-b11)
OpenJDK 64-Bit Server VM (build 25.131-b11, mixed mode)
```

Install Hadoop

Download latest Hadoop binary package from [<http://hadoop.apache.org/releases.html>].

Look for latest stable release (not in alpha channel) and click on binary link provided for the release.

Apache Hadoop Releases

hadoop.apache.org/releases.html

Apache > Hadoop >



Search with Apache Solr

Last Published: 08/06/2017 13:55:10

Top Wiki

About

- Welcome
- Releases
- Download
 - Release Notes
- Release Versioning
- Mailing Lists
- Issue Tracking
- Who We Are?
- Who Uses Hadoop?
- Buy Stuff
- Sponsorship
- Thanks
- Privacy Policy
- Bylaws
- Committer criteria
- License
- Documentation
- Related Projects

Apache Hadoop Releases

Download

Hadoop is released as source code tarballs with corresponding binary tarballs for convenience. The downloads are distributed via mirror sites and should be checked for tampering using GPG or SHA-256.

| Version | Release Date | Tarball | GPG | SHA-256 |
|------------------------------|------------------|------------------------|---------------------------|-------------------------------------|
| 3.0.0-alpha4 | 07 July, 2017 | source | signature | checksum file |
| 2.8.1 | 08 June, 2017 | binary | signature | checksum file |
| 2.7.4 | 04 August, 2017 | source | signature | 0748C0E2 519382F2.. |
| 2.6.5 | 08 October, 2016 | source | signature | B5BE5275 78EF2C85.. |
| | | binary | signature | D52B8CE8 446F4C10.. |
| | | binary | signature | 8F791BFC FA5BB7C7.. |
| | | source | signature | 3A843F18 73D9951A.. |
| | | binary | signature | 001AD18D 4B6D0FE5.. |

To verify Hadoop releases using GPG:

Click on the first mirror link

Apache Download Mir

www.apache.org/dyn/closer.cgi/hadoop



Google Custom

The Apache Way

Contribute

ASF Sponsors

We suggest the following mirror site for your download:

<http://www-us.apache.org/dist/hadoop/common/hadoop-2.8.1/hadoop-2.8.1.tar.gz>

Other mirror sites are suggested below. Please use the backup mirrors only to download PGP and MD5 signatures to [verify your downloads](#) or if no other mirrors are working.

Copy the downloaded tar file to /usr/lib/ and untar.

```
$ sudo cp hadoop-2.8.1.tar.gz /usr/lib/
$ sudo tar zxf hadoop-2.8.1.tar.gz
$ sudo rm hadoop-2.8.1.tar.gz
```

Provide the password if asked.

Set Java and Hadoop Path

Make sure you have the PATHs set up for Java and Hadoop in bashrc file. Open a Terminal and run the following command to edit bashrc file.

```
$ sudo nano ~/.bashrc
```

Paste the following entries at the end of .bashrc file.

```
#HADOOP VARIABLES START
export JAVA_HOME=/usr/lib/jvm/default-java/jre
export HADOOP_INSTALL=/usr/lib/hadoop-2.8.1
export PATH=$PATH:$HADOOP_INSTALL/bin
export PATH=$PATH:$HADOOP_INSTALL/sbin
export HADOOP_MAPRED_HOME=$HADOOP_INSTALL
export HADOOP_COMMON_HOME=$HADOOP_INSTALL
export HADOOP_HDFS_HOME=$HADOOP_INSTALL
export YARN_HOME=$HADOOP_INSTALL
export HADOOP_COMMON_LIB_NATIVE_DIR=$HADOOP_INSTALL/lib/native
export HADOOP_OPTS="-Djava.library.path=$HADOOP_INSTALL/lib"
#HADOOP VARIABLES END
```

Run Hadoop

After setting up the path for Hadoop and Java, you may run the hadoop command, from anywhere, using the terminal.

```
$ hadoop
```

The output would be as shown below :

```
Usage: hadoop [--config confdir] [COMMAND | CLASSNAME]
  CLASSNAME          run the class named CLASSNAME
or
  where COMMAND is one of:
  fs                 run a generic filesystem user client
```

```
version          print the version
jar              run a jar file
                 note: please use "yarn jar" to launch
                 YARN applications, not this command.
checknative [-a|-h] check native hadoop and compression libraries availability
distcp          copy file or directories recursively
archive -archiveName NAME -p * create a hadoop archive
classpath       prints the class path needed to get the
                 Hadoop jar and the required libraries
credential      interact with credential providers
daemonlog       get/set the log level for each daemon
trace           view and modify Hadoop tracing settings
```

Most commands print help when invoked w/o parameters.

Conclusion

In this [Apache Hadoop Tutorial](#), we have successfully installed Hadoop on Ubuntu. In subsequent tutorials, we shall look into HDFS and MapReduce and start with Word Count Example in Hadoop.

Learn Apache Hadoop

◆ [Hadoop Tutorial](#)

⇒ [Install Hadoop on Ubuntu](#)

◆ [Hadoop MapReduce 1.0](#)