**INSTALLING HADOOP 2.7.3**

#Download the latest stable Hadoop using wget from one of the Apache mirrors.

root@ip-172-31-23-142:/home/ubuntu# wget https://archive.apache.org/dist/hadoop/core/hadoop-2.7.3/hadoop-2.7.3.tar.gz

--2020-04-05 13:02:47-- https://archive.apache.org/dist/hadoop/core/hadoop- 2.7.3/hadoop-2.7.3.tar.gz

Resolving archive.apache.org (archive.apache.org)... 163.172.17.199

Connecting to archive.apache.org (archive.apache.org)|163.172.17.199|:443... connected. HTTP request sent, awaiting response... 200 OK

Length: 214092195 (204M) [application/x-gzip] Saving to: 'hadoop-2.7.3.tar.gz'

hadoop-2.7.3.tar.gz 100%[===========================================================

=================>] 204.17M 15.4MB/s in 19s

2020-04-05 13:03:07 (10.9 MB/s) - 'hadoop-2.7.3.tar.gz' saved [214092195/214092195]

root@ip-172-31-23-142:/home/ubuntu# tar -xvf hadoop-2.7.3.tar.gz

Create a directory where the hadoop will store its data. We will set this directory path in hdfs-site.

root@ip-172-31-23-142:/home/ubuntu# Java -version

#Add the Hadoop related environment variables in your bash file. root@ip-172-31-23-142:/home/ubuntu#vi ~/.bashrc

#Copy and paste these environment variables. #java variables

export JAVA\_HOME=/usr/lib/jvm/java-8-openjdk-amd64/ export JRE\_HOME=/usr/lib/jvm/java-8-openjdk-amd64/jre export PATH:$JAVA\_HOME:$JRE\_HOME/bin

#Hadoop variables

export HADOOP\_HOME=/home/ubuntu/hadoop-2.7.3

export HADOOP\_COMMON\_LIB\_NATIVE\_DIR=$HADOOP\_HOME/lib/native export HADOOP\_MAPRED\_HOME=$HADOOP\_HOME

export HADOOP\_COMMON\_HOME=$HADOOP\_HOME export HADOOP\_HDFS\_HOME=$HADOOP\_HOME

export YARN\_HOME=$HADOOP\_HOME

export HADOOP\_OPTS="-Djava.library.path=$HADOOP\_HOME/lib" export PATH=$PATH:$HADOOP\_HOME/bin:$HADOOP\_HOME/sbin

#Save and exit and use this command to refresh the bash settings. root@ip-172-31-23-142:/home/ubuntu# source ~/.bashrc

# Setting hadoop environment for password less ssh access. Password less SSH Configuration is a mandatory installation requirement. However it is more useful in a distributed environment.

root@ip-172-31-23-142:/home/ubuntu# ssh-keygen -t rsa -P ''

root@ip-172-31-23-142:/home/ubuntu# cat $HOME/.ssh/id\_rsa.pub >> root@ip-172-31- 23-142:/home/ubuntu# $HOME/.ssh/authorized\_keys

## check password less ssh access to localhost root@ip-172-31-23-142:/home/ubuntu# ssh localhost #exit from inner localhost shell

root@ip-172-31-23-142:/home/ubuntu# exit

# Set the hadoop config files. We need to set the below files in order for hadoop to function properly.

* core-site.xml
* hadoop-env.sh
* yarn-site.xml
* hdfs-site.xml
* mapred-site.xml

# go to directory where all the config files are present (cd /home/ubuntu/hadoop- 2.6.0/etc/Hadoop)

* Copy and paste the below configurations in core-site.xml ##Add the following text between the configuration tabs.

<property>

<name>hadoop.tmp.dir</name>

Thank you for using [www.freepdfconvert.com](http://www.freepdfconvert.com/) service!

Only two pages are converted. Please Sign Up to convert all pages. <https://www.freepdfconvert.com/membership>