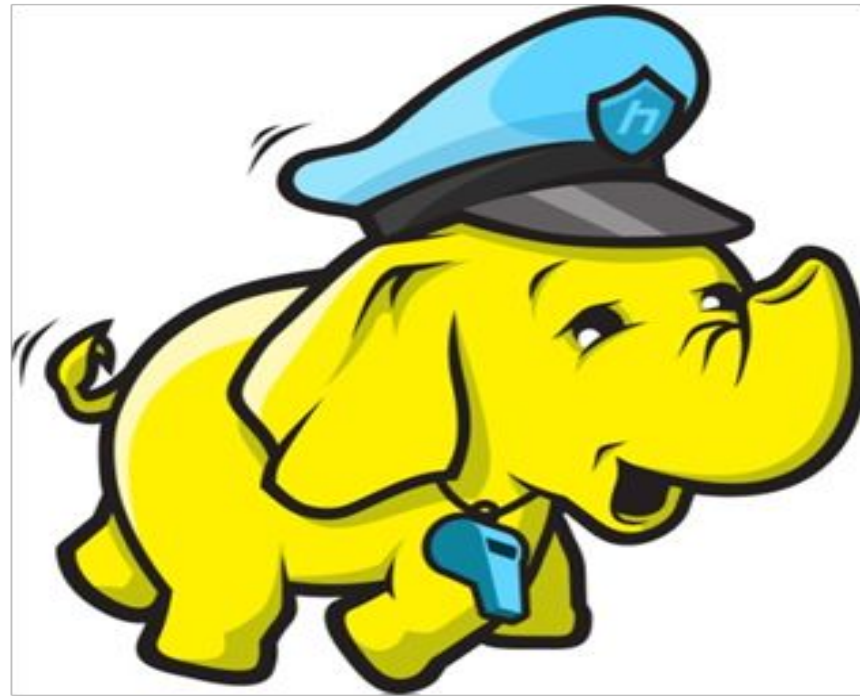


# Install Hadoop on Linux based OS

## Mac OS/Redhat/Ubuntu



For online Hadoop training, send mail to [neeraj.ymca.2k6@gmail.com](mailto:neeraj.ymca.2k6@gmail.com)

# Agenda

- Install Java & ssh on your Ubuntu
- Visit Apache Hadoop website
- Download Hadoop tar.gz file
- Untar the Hadoop tar.gz file
- Directory structure of Hadoop
- What is ssh (Secure Shell)
- How to configure password less ssh
- Configuration files in Hadoop
- Modify configuration files
- Format Namenode
- Start the services of Hadoop
- Verify Hadoop services
- Stop the services of Hadoop

# Install Java & ssh on your Ubuntu

**Run the below commands in terminal to install Java**

```
sudo add-apt-repository ppa:webupd8team/java  
sudo apt-get update  
sudo apt-get install oracle-java7-installer
```

**Run the below command in terminal to install ssh**


```
sudo apt-get install openssh-server openssh-client
```



**Make sure that your Ubuntu is connected to internet during execution of these commands**















# Download Hadoop tar.gz file

<http://www.motorlogy.com/apache/hadoop/common/hadoop-1.2.1/>

Index of /apache/hadoop/co... 


← [www.motorlogy.com/apache/hadoop/common/hadoop-1.2.1/](http://www.motorlogy.com/apache/hadoop/common/hadoop-1.2.1/)


## Index of /apache/hadoop/common/hadoop-1.2.1

Name	Last modified	Size	Description
 <a href="#">Parent Directory</a>	-	-	
 <a href="#">hadoop-1.2.1-1.i386.rpm</a>	22-Jul-2013 15:20	39M	
 <a href="#">hadoop-1.2.1-1.i386.rpm.mds</a>	22-Jul-2013 15:45	1.1K	
 <a href="#">hadoop-1.2.1-1.x86_64.rpm</a>	22-Jul-2013 15:30	35M	
 <a href="#">hadoop-1.2.1-1.x86_64.rpm.mds</a>	22-Jul-2013 15:45	1.1K	
 <a href="#">hadoop-1.2.1-bin.tar.gz</a>	22-Jul-2013 15:27	36M	
 <a href="#">hadoop-1.2.1-bin.tar.gz.mds</a>	22-Jul-2013 15:45	1.1K	
 <a href="#">hadoop-1.2.1.tar.gz</a>	22-Jul-2013 15:27	61M	
 <a href="#">hadoop-1.2.1.tar.gz.mds</a>	22-Jul-2013 15:45	958	
 <a href="#">hadoop_1.2.1-1_i386.deb</a>	22-Jul-2013 15:23	39M	
 <a href="#">hadoop_1.2.1-1_i386.deb.mds</a>	22-Jul-2013 15:45	1.1K	
 <a href="#">hadoop_1.2.1-1_x86_64.deb</a>	22-Jul-2013 15:33	35M	
 <a href="#">hadoop_1.2.1-1_x86_64.deb.mds</a>	22-Jul-2013 15:45	1.1K	
 <a href="#">releasenotes-1.2.1.html</a>	22-Jul-2013 15:15	519K	















Apache/2.2.22 Server at www.motorlogy.com Port 80

# Click on Save File

Index of /apache/hadoop/co... 

←  www.motorlogy.com/apache/hadoop/common/hadoop-1.2.1/


## Index of /apache/hadoop/common/hadoop-1.2.1

Name	Last modified	Size	Description
 <a href="#">Parent Directory</a>	-	-	-
 <a href="#">hadoop-1.2.1-1.i386.rpm</a>	22-Jul-2013 15:20	39M	
 <a href="#">hadoop-1.2.1-1.i386.rpm.mds</a>	22-Jul-2013 15:45	1.1K	
 <a href="#">hadoop-1.2.1-1.x86_64.rpm</a>	22-Jul-2013 15:45	39M	
 <a href="#">hadoop-1.2.1-1.x86_64.rpm.mds</a>	22-Jul-2013 15:45	1.1K	
 <a href="#">hadoop-1.2.1-bin.tar.gz</a>	22-Jul-2013 15:45	60.9 MB	
 <a href="#">hadoop-1.2.1-bin.tar.gz.mds</a>	22-Jul-2013 15:45	1.1K	
 <a href="#">hadoop-1.2.1.tar.gz</a>	22-Jul-2013 15:45	60.9 MB	
 <a href="#">hadoop-1.2.1.tar.gz.mds</a>	22-Jul-2013 15:45	1.1K	
 <a href="#">hadoop_1.2.1-1_i386.deb</a>	22-Jul-2013 15:45	39M	
 <a href="#">hadoop_1.2.1-1_i386.deb.mds</a>	22-Jul-2013 15:45	1.1K	
 <a href="#">hadoop_1.2.1-1_x86_64.deb</a>	22-Jul-2013 15:45	39M	
 <a href="#">hadoop_1.2.1-1_x86_64.deb.mds</a>	22-Jul-2013 15:45	1.1K	
 <a href="#">releasenotes-1.2.1.html</a>	22-Jul-2013 15:45	1.1K	

Apache/2.2.22 Server at www.motorlogy.com

**Opening hadoop-1.2.1.tar.gz**

You have chosen to open:

 **hadoop-1.2.1.tar.gz**  
which is a: Gzip archive (60.9 MB)  
from: http://www.motorlogy.com

**What should Firefox do with this file?**

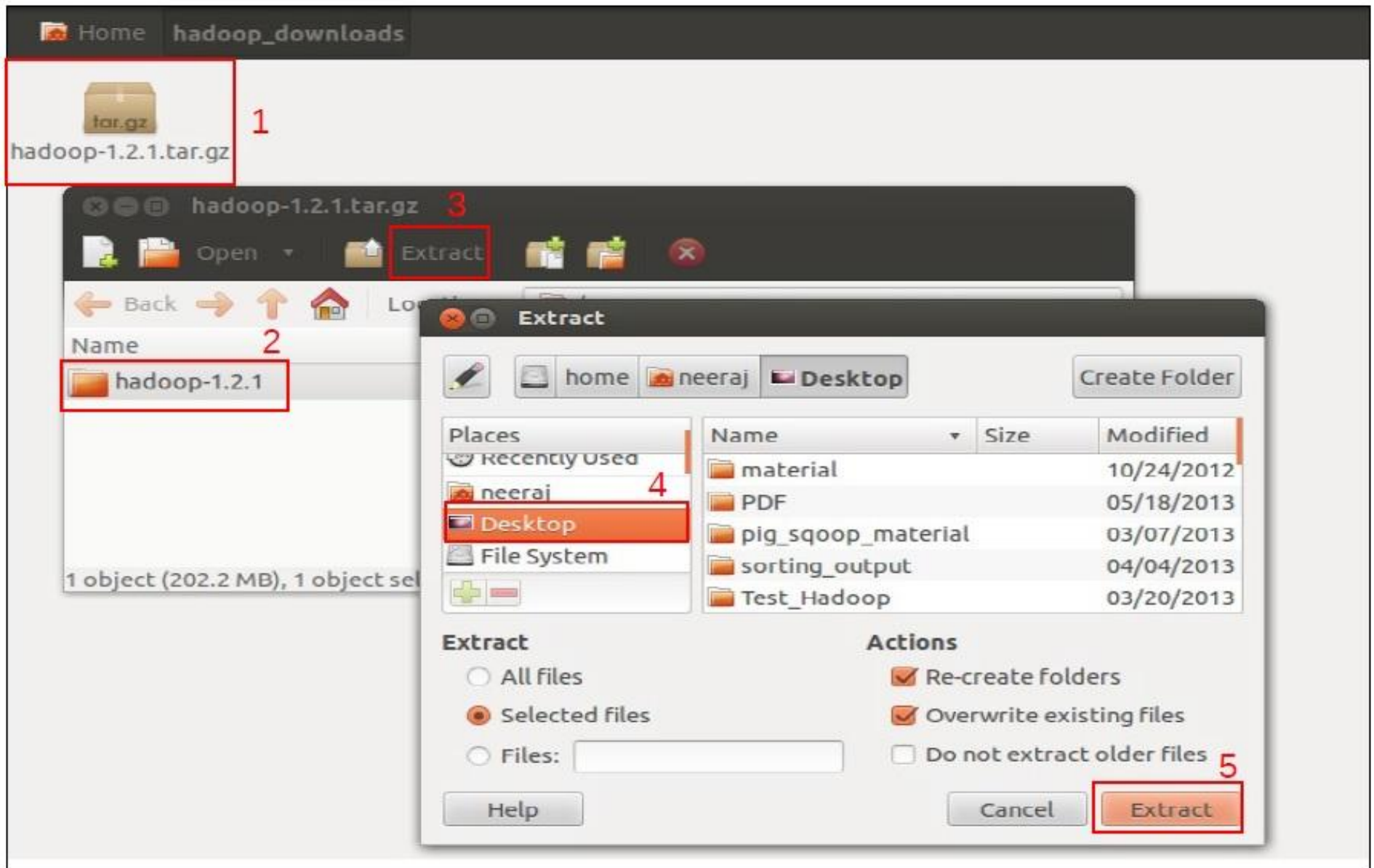
☐ Open with Archive Manager (default)

☒ **Save File**

☐ Do this automatically for files like this from now on.

Cancel **OK**

# Untar the Hadoop tar.gz file





# Directory structure of Hadoop

## **HADOOP\_HOME**

- bin [ contains scripts to start/stop Hadoop ]
- conf [ contains configuration files of Hadoop ]
- lib [ contains libraries to run Hadoop ]
- src [ contains the source code of Hadoop ]
- logs [ contains log files ]

# Configuration files in Hadoop

There are two types of important configuration files in Hadoop...

## **Default configuration files ( Read-only ) :**

src/core/core-default.xml  
src/hdfs/hdfs-default.xml  
src/mapred/mapred-default.xml

## **Site-specific configuration files ( can be modified ) :**

conf/core-site.xml  
conf/hdfs-site.xml  
conf/mapred-site.xml



# What is SSH ( Secure Shell )

Tool, used to remotely connect to other machine.

Namenode uses SSH for communicating with Datanodes.

Jobtracker uses SSH for communicating with Tasktrackers.

SSH must be installed, even if you are installing Hadoop on a single machine.

`ssh < IP of another machine >`

For example...

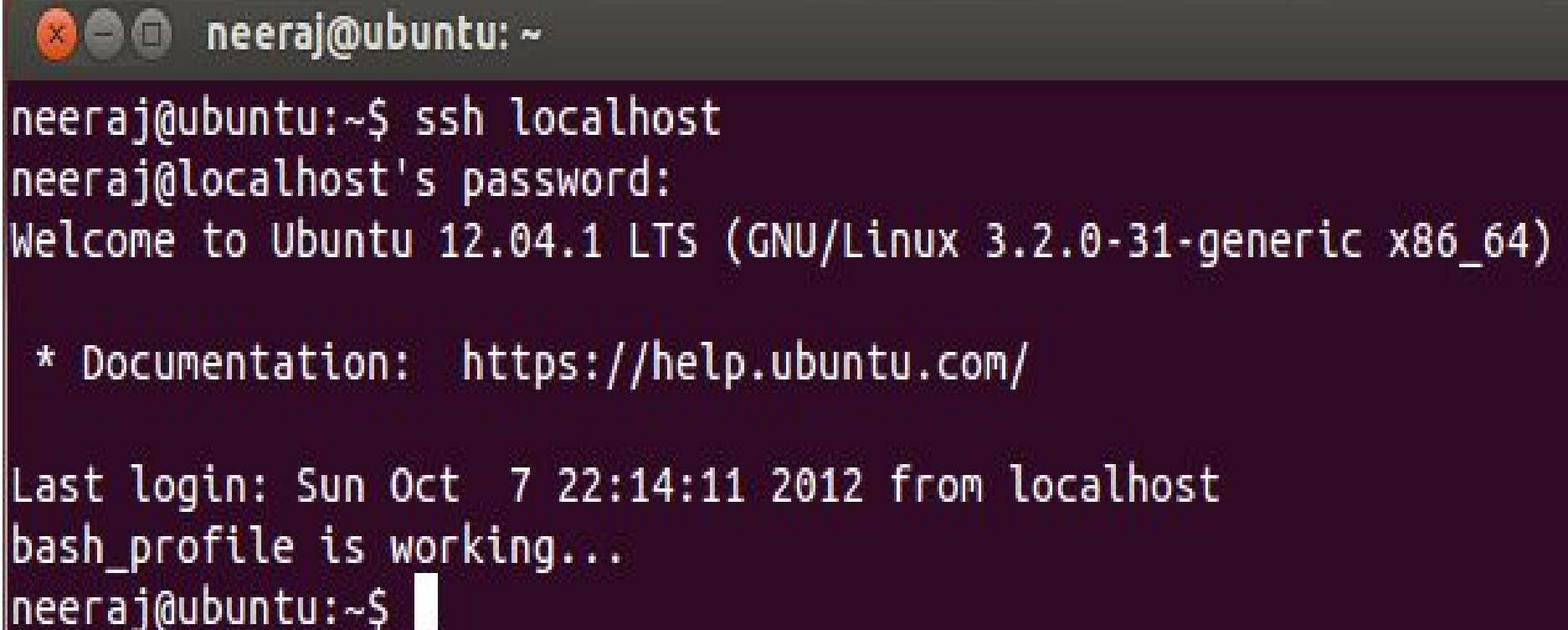
`ssh 192.168.342.187`

`ssh 127.0.0.1`

`ssh localhost`

Run the following command on your terminal

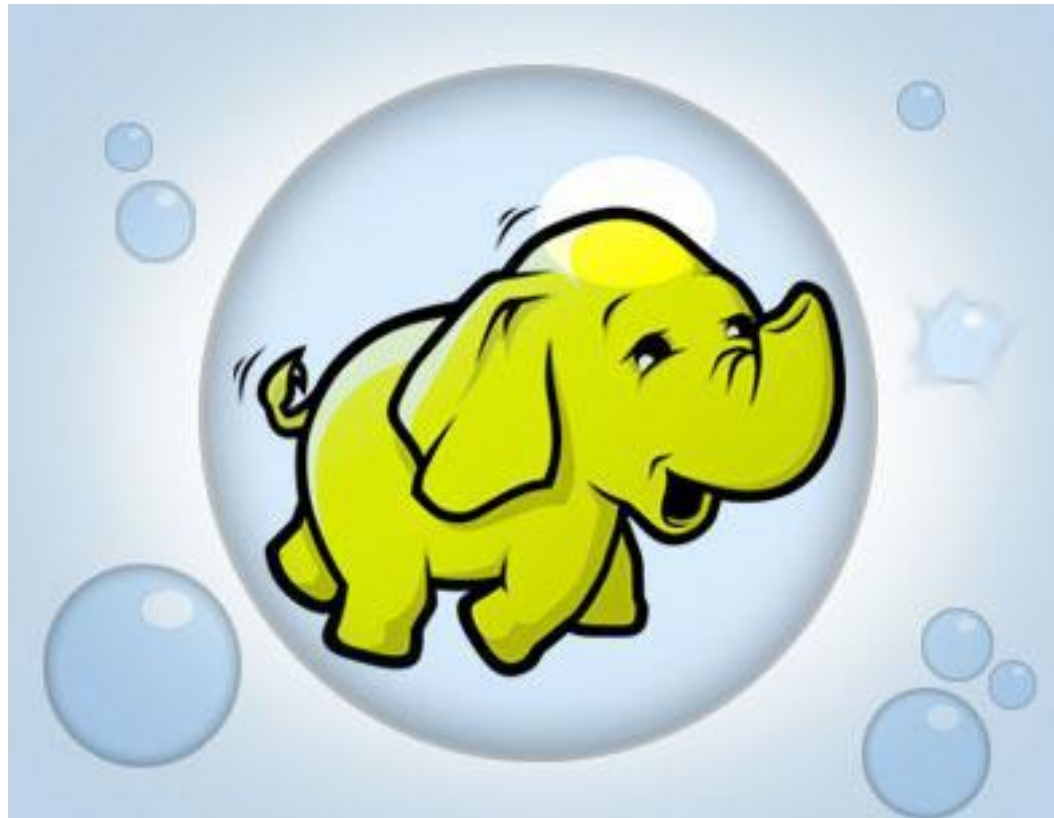
ssh localhost



A terminal window with a dark background and light text. The title bar at the top shows three window control icons (close, minimize, maximize) and the text 'neeraj@ubuntu: ~'. The terminal content shows the user 'neeraj' at 'ubuntu' running the command 'ssh localhost'. It prompts for a password, then displays a welcome message for Ubuntu 12.04.1 LTS, including the kernel version and architecture. It also provides a link to the documentation. Finally, it shows the last login time and that the bash profile is working, returning to the shell prompt.

```
neeraj@ubuntu: ~  
neeraj@ubuntu:~$ ssh localhost  
neeraj@localhost's password:  
Welcome to Ubuntu 12.04.1 LTS (GNU/Linux 3.2.0-31-generic x86_64)  
  
* Documentation:  https://help.ubuntu.com/  
  
Last login: Sun Oct  7 22:14:11 2012 from localhost  
bash_profile is working...  
neeraj@ubuntu:~$
```

...Thanks...



For online Hadoop training, send mail to [neeraj.ymca.2k6@gmail.com](mailto:neeraj.ymca.2k6@gmail.com)