

1. Hadoop install & Config in Linux

Update Linux operationg system

```
(om@kali)-[~/Desktop/share]
$ sudo apt-get update
[sudo] password for om:
Sorry, try again.
[sudo] password for om:
0% [Working]
```

Installation of java

```
(om㉿kali)-[~/Desktop/share]
$ sudo apt-get install default-jdk
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
The following additional packages will be installed:
  default-jdk-headless libice-dev libpthread-stubs0-dev libsm-dev libx11-6
  libx11-data libx11-dev libxcb1 libxau-dev libxcb-damage0
  libxcb-dri2-0 libxcb-dri3-0 libxcb-glx0 libxcb-present0 libxcb-randr0
  libxcb-render0 libxcb-shape0 libxcb-shm0 libxcb-sync1 libxcb-xfixes0
  libxcb-xinerama0 libxcb-xinput0 libxcb-xkb1 libxcb1 libxcb1-dev
  libxdmcp-dev libxt-dev openjdk-11-jdk openjdk-11-jdk-headless
  x11proto-dev xorg-sgml-doctools xtrans-dev
Suggested packages:
  libice-doc libsm-doc libx11-doc libxcb-doc libxt-doc openjdk-11-demo
  openjdk-11-source visualvm
The following NEW packages will be installed:
  default-jdk default-jdk-headless libice-dev libpthread-stubs0-dev
  libsm-dev libx11-dev libxau-dev libxcb1-dev libxdmcp-dev libxt-dev
  openjdk-11-jdk openjdk-11-jdk-headless x11proto-dev xorg-sgml-doctools
  xtrans-dev
The following packages will be upgraded:
  libx11-6 libx11-data libx11-xcb1 libxcb-damage0 libxcb-dri2-0
  libxcb-dri3-0 libxcb-glx0 libxcb-present0 libxcb-randr0 libxcb-render0
  libxcb-shape0 libxcb-shm0 libxcb-sync1 libxcb-xfixes0 libxcb-xinerama0
  libxcb-xinput0 libxcb-xkb1 libxcb1
18 upgraded, 15 newly installed, 0 to remove and 1161 not upgraded.
Need to get 227 MB of archives.
After this operation, 239 MB of additional disk space will be used.
Do you want to continue? [Y/n] y
Get:1 http://http.kali.org/kali kali-rolling/main amd64 openjdk-11-jdk-headless amd64 11.0.16+8-1 [214 MB]
4% [1 openjdk-11-jdk-headless 11.8 MB/214 MB 6%]          991 kB/s 3min 36s
```

Create hadoop group and user

```
(om㉿kali)-[~]
$ sudo addgroup hadoop
addgroup: The group `hadoop' already exists.

(om㉿kali)-[~]
$ sudo adduser --ingroup hadoop hduser
Adding user `hduser' ...
Adding new user `hduser' (1001) with group `hadoop' ...
Creating home directory `/home/hduser' ...
Copying files from `/etc/skel' ...
New password:
Retype new password:
passwd: password updated successfully
Changing the user information for hduser
Enter the new value, or press ENTER for the default
  Full Name []:
  Room Number []:
  Work Phone []:
  Home Phone []:
  Other []:
Is the information correct? [Y/n] u
Is the information correct? [Y/n] y

(om㉿kali)-[~]
$ groups hduser
hduser : hadoop

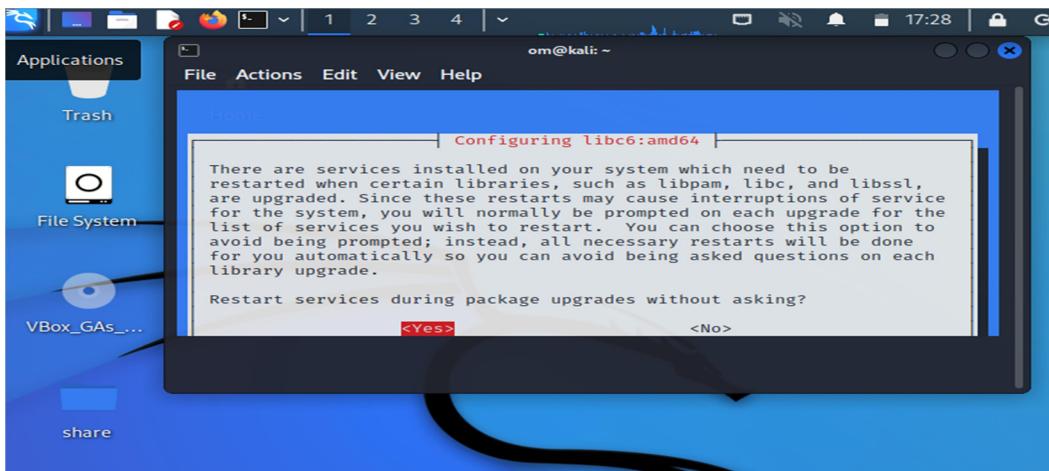
(om㉿kali)-[~]
```

Install ssh

Om Kant Sharma 2022

```
(om@kali)-[~]
$ sudo apt-get install ssh
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
The following additional packages will be installed:
  libc-bin libc-dev-bin libc-lion libc6 libc6-dev libc6-i386 libssl3
  locales openssh-client openssh-server openssh-sftp-server openssl
  runit-helper
Suggested packages:
  glibc-doc libnss-nis libnss-nisplus manpages-dev keychain libpam-ssh
  monkeyshphere ssh-askpass molly-guard ufw
Recommended packages:
  manpages-dev libc-devtools
The following NEW packages will be installed:
  ssh
The following packages will be upgraded:
  libc-bin libc-dev-bin libc-lion libc6 libc6-dev libc6-i386 libssl3
  locales openssh-client openssh-server openssh-sftp-server openssl
  runit-helper
13 upgraded, 1 newly installed, 0 to remove and 1148 not upgraded.
Need to get 17.5 MB of archives.
After this operation, 3,729 kB disk space will be freed.
Do you want to continue? [Y/n] y
8% [working]
```

```
(om@kali)-[~]
$ sudo apt-get install ssh
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
The following additional packages will be installed:
  libc-bin libc-dev-bin libc-lion libc6 libc6-dev libc6-i386 libssl3
  locales openssh-client openssh-server openssh-sftp-server openssl
  runit-helper
Suggested packages:
  glibc-doc libnss-nis libnss-nisplus manpages-dev keychain libpam-ssh
  monkeyshphere ssh-askpass molly-guard ufw
Recommended packages:
  manpages-dev libc-devtools
The following NEW packages will be installed:
  ssh
The following packages will be upgraded:
  libc-bin libc-dev-bin libc-lion libc6 libc6-dev libc6-i386 libssl3
  locales openssh-client openssh-server openssh-sftp-server openssl
  runit-helper
13 upgraded, 1 newly installed, 0 to remove and 1148 not upgraded.
Need to get 17.5 MB of archives.
After this operation, 3,729 kB disk space will be freed.
Do you want to continue? [Y/n] y
Get:1 http://kali.download/kali kali-rolling/main amd64 libc-lion all 2.35-4
[671 kB]
Get:2 http://kali.download/kali kali-rolling/main amd64 libc-dev-bin amd64 2.35-
4 [42.1 kB]
Get:3 http://kali.download/kali kali-rolling/main amd64 libc6-dev amd64 2.35-
4 [1,893 kB]
Get:4 http://kali.download/kali kali-rolling/main amd64 libc6-i386 amd64 2.35-
4 [2,446 kB]
Get:5 http://kali.download/kali kali-rolling/main amd64 locales all 2.35-4 [3
,895 kB]
Get:6 http://kali.download/kali kali-rolling/main amd64 libc6 amd64 2.35-4 [2
,738 kB]
Get:7 http://kali.download/kali kali-rolling/main amd64 libc-bin amd64 2.35-4
[617 kB]
Get:8 http://kali.download/kali kali-rolling/main amd64 libssl3 amd64 3.0.7-1
[2,008 kB]
72% [8 libssl3 1,216 kB/2,008 kB 61%] 351 kB/s 11s
```



Om Kant Sharma 2o22

```
└─(om㉿kali)-[~]
└─$ which ssh
/usr/bin/ssh

└─(om㉿kali)-[~]
└─$ which sshd
/usr/sbin/sshd

└─(om㉿kali)-[~]
└─$
```

```
└─(om㉿kali)-[~]
└─$ su hduser
Password:
└─(hduser㉿kali)-[/home/om]
└─$ █
```

```
└─(om㉿kali)-[~]
└─$ su hduser
Password:
└─(hduser㉿kali)-[/home/om]
└─$ cd
└─(hduser㉿kali)-[~]
└─$ pwd
/home/hduser

└─(hduser㉿kali)-[~]
└─$ ssh-keygen
Generating public/private rsa key pair.
Enter file in which to save the key (/home/hduser/.ssh/id_rsa):
Created directory '/home/hduser/.ssh'.
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
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:
SHA256:jpMLnGJi73Mzmd47ZRFGLULtpqcLV5e+K5Vu+OkAVI hduser@kali
The key's randomart image is:
+--- [RSA 3072] ---+
| |
| E + . |
| . .. o = |
| o S .. o . |
| + B .. O.B .o |
| + B ... o B.=oo |
| . +--+ .. o =+-+ |
| o=+ .. o .+++ |
+--- [SHA256] ---+
└─(hduser㉿kali)-[~]
└─$ █
```

```
| . +--+ .. o =+-+ |
| o=+ .. o .+++ |
+--- [SHA256] ---+

└─(hduser㉿kali)-[~]
└─$ cat $HOME/.ssh/id_rsa.pub >> $HOME/.ssh/authorized_keys

└─(hduser㉿kali)-[~]
└─$ ssh localhost
ssh: connect to host localhost port 22: Connection refused
```

```
[root@kali) [~]
# apt-get install ssh openssh-client openssh-server
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
openSSH-client is already the newest version (1:9.0p1-1+b2).
openSSH-server is already the newest version (1:9.0p1-1+b2).
ssh is already the newest version (1:9.0p1-1).
0 upgraded, 0 newly installed, 0 to remove and 1148 not upgraded.

[root@kali) [~]
# service ssh start
[root@kali) [~]
# ssh localhost
root@localhost's password:
Permission denied, please try again.
root@localhost's password:

File Actions Edit View Help
File (om@kali) [~]
$ passwd
Changing password for om.
Current password:
passwd: Authentication token manipulation error
passwd: password unchanged

(om@kali) [~]
$ passwd
Changing password for om.
Current password:
passwd: Authentication token manipulation error
passwd: password unchanged

(om@kali) [~]
$ passwd
Changing password for om.
Current password:
New password:
Retype new password:
The password has not been changed.
New password: ■
```

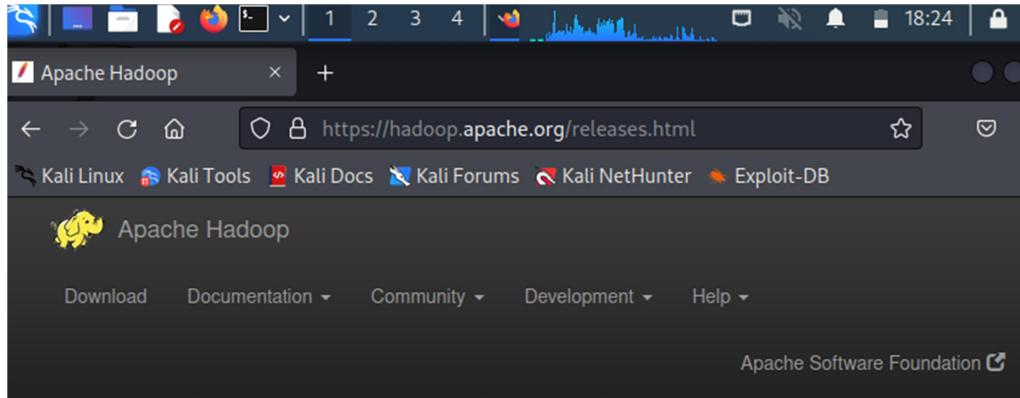
```
(om@kali) [~]
$ ssh localhost
The authenticity of host 'localhost (::1)' can't be established.
ED25519 key fingerprint is SHA256:mwcBk82aGaZxDd41im6iaXcz1pu0JvMRTbvRiDpxjes
.
This key is not known by any other names
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
Warning: Permanently added 'localhost' (ED25519) to the list of known hosts.
om@localhost's password:
Linux kali 5.18.0-kali5-amd64 #1 SMP PREEMPT_DYNAMIC Debian 5.18.5-1kali6 (20
22-07-07) x86_64

The programs included with the Kali GNU/Linux system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*copyright.

Kali GNU/Linux comes with ABSOLUTELY NO WARRANTY, to the extent
permitted by applicable law.

(om@kali) [~]
$ ■
```

Install Hadoop



To verify Hadoop releases using GPG:

<https://hadoop.apache.org/releases.html>

```
(om㉿kali)-[~]
$ ls
Desktop  Downloads  Music  Public  testpage
Documents  lsofD  Pictures  Templates  Videos

(om㉿kali)-[~]
$ cd Downloads
(om㉿kali)-[~/Downloads]
$ ls
hadoop-3.3.4.tar.gz

(om㉿kali)-[~/Downloads]
$ sudo cp hadoop-3.3.4.tar.gz /home/hduser
[sudo] password for om:

(om㉿kali)-[~/Downloads]
$
```

```
└─(om㉿kali)-[~]
$ ls
Desktop  Downloads  Music  Public  testpage
Documents  lsodF    Pictures  Templates  Videos

└─(om㉿kali)-[~]
$ cd Downloads

└─(om㉿kali)-[~/Downloads]
$ ls
hadoop-3.3.4.tar.gz

└─(om㉿kali)-[~/Downloads]
$ sudo cp hadoop-3.3.4.tar.gz /home/hduser

└─(om㉿kali)-[~/Downloads]
$ ls
hadoop-3.3.4.tar.gz

└─(om㉿kali)-[~/Downloads]
$ cd

└─(om㉿kali)-[~]
$ su hduser
Password:
└─(hduser㉿kali)-[/home/om]
$ cd

└─(hduser㉿kali)-[~]
$ ls
hadoop-3.3.4.tar.gz

└─(hduser㉿kali)-[~]
$ █
```

```
└─(hduser㉿kali)-[~]
$ ls
hadoop-3.3.4.tar.gz

└─(hduser㉿kali)-[~]
$ sudo hadoop-3.3.4.tar.gz
[sudo] password for hduser:
hduser is not in the sudoers file.
Add user 'hduser' to group `sudo' ...
Done.

└─(hduser㉿kali)-[~]
$ sudo -v
Sorry, user hduser may not run sudo on kali.

└─(hduser㉿kali)-[~]
$ su om
Password:
└─(om㉿kali)-[/home/hduser]
$ sudo adduser hduser sudo
Adding user 'hduser' to group `sudo' ...
Done.

└─(om㉿kali)-[/home/hduser]
$ su hduser
Password:
└─(hduser㉿kali)-[~]
$ ls
hadoop-3.3.4.tar.gz
```

NOW INSTALL HADOOP

```
└─(hduser㉿kali)-[~]
$ tar xvzf hadoop-3.3.4.tar.gz
hadoop-3.3.4/
hadoop-3.3.4/licenses-binary/
hadoop-3.3.4/licenses-binary/LICENSE-dust.txt
hadoop-3.3.4/licenses-binary/LICENSE-re2j.txt
hadoop-3.3.4/licenses-binary/LICENSE-slf4j.txt
hadoop-3.3.4/licenses-binary/LICENSE-jquery.txt
hadoop-3.3.4/licenses-binary/LICENSE-zstd-jni.txt
hadoop-3.3.4/licenses-binary/LICENSE-hsql.txt
hadoop-3.3.4/licenses-binary/LICENSE-datatables.txt
```

Om Kant Sharma 2022

```
(hduser㉿kali)-[~]
$ tar xvf hadoop-3.3.4.taz.gz
tar (child): hadoop-3.3.4.taz.gz: Cannot open: No such file or directory
tar (child): Error is not recoverable: exiting now
tar: Child returned status 2
tar: Error is not recoverable: exiting now

(hduser㉿kali)-[~]
$ ls
hadoop-3.3.4  hadoop-3.3.4.tar.gz
Desktop/hadoop-3.3.4/hadoop-3.3.4/hadoop-3.3.4.tar.gz

(hduser㉿kali)-[~]
$ cd hadoop-3.3.4

(hduser㉿kali)-[~/hadoop-3.3.4]
$ ls
bin      lib          licenses-binary  NOTICE.txt  share
etc      libexec     LICENSE.txt      README.txt
include  LICENSE-binary NOTICE-binary  sbin

(hduser㉿kali)-[~/hadoop-3.3.4]
$ sudo mv * /usr/local/hadoop
devices

(hduser㉿kali)-[~/hadoop-3.3.4]
$ ls
share

(hduser㉿kali)-[~/hadoop-3.3.4]
$
```

```
(hduser㉿kali)-[~/hadoop-3.3.4]
$ sudo mv * /usr/local/hadoop

(hduser㉿kali)-[~/hadoop-3.3.4] lib32 lib64 libexec
$ ls
bin etc lib libexec LICENSE-binary NOTICE-binary NOTICE.txt README.txt
LICENSE.txt

(hduser㉿kali)-[~/hadoop-3.3.4]
$ sudo chown -R hduser:hadoop /usr/local/hadoop

(hduser㉿kali)-[~/hadoop-3.3.4] sbin share
$ ls
sbin share

(hduser㉿kali)-[~/hadoop-3.3.4]
$ cd /usr/local/hadoop/
N
(hduser㉿kali)-[/usr/local/hadoop]
$ ls
bin include LICENSE-binary NOTICE-binary sbin
etc lib licenses-binary NOTICE.txt share
hadoop-3.3.4.tar.gz libexec LICENSE.txt README.txt

(hduser㉿kali)-[/usr/local/hadoop]
$
```

SETUP SOME CONFIGURATION FILES

```
(hduser㉿kali)-[~/usr/local/hadoop]
$ sudo nano ~/.bashrc
[sudo] password for hduser:

(hduser㉿kali)-[~/usr/local/hadoop]
$ sudo nano ~/.bashrc

(hduser㉿kali)-[~/usr/local/hadoop]
$
```

Add this configuration at the end

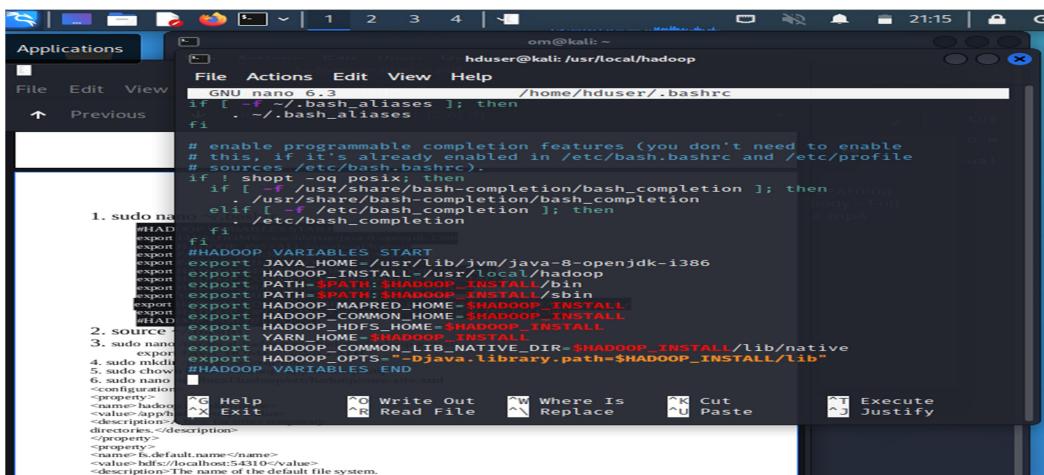
```
(hduser㉿kali)-[~/usr/local/hadoop]
$ sudo nano ~/.bashrc
[sudo] password for hduser:

(hduser㉿kali)-[~/usr/local/hadoop]
$ sudo nano ~/.bashrc

(hduser㉿kali)-[~/usr/local/hadoop]
$
```

#HADOOP VARIABLES START

```
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-i386
export HADOOP_INSTALL=/usr/local/hadoop
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
```



```

9. sudo chown -R hduser:hadoop /usr/local/hadoop
• Setup Configuration Files:- Open New Terminal Window:- update-alternatives
1. sudo nano ~/.bashrc
#HADOOP VARIABLES START
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-i386
export HADOOP_INSTALL=/usr/local/hadoop
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
2. source ~/.bashrc
3. sudo nano /usr/local/hadoop/etc/hadoop/hadoop-env.sh
    export JAVA_HOME=$JAVA PATH"
4. sudo mkdir -p /app/hadoop/tmp
5. sudo chown hduser:hadoop /hadoop/tmp

```

The terminal window shows the configuration of Hadoop variables in `~/.bashrc`. It includes environment variable assignments for `JAVA_HOME`, `HADOOP_INSTALL`, `PATH`, `HADOOP_MAPRED_HOME`, `HADOOP_COMMON_HOME`, `HADOOP_HDFS_HOME`, `YARN_HOME`, `HADOOP_COMMON_LIB_NATIVE_DIR`, and `HADOOP_OPTS`. The `HADOOP VARIABLES END` section is also present. Below this, the command `source ~/.bashrc` is run. A message from `update-alternatives` indicates there is one choice for the alternative java, providing `/usr/bin/java`. The terminal then lists the available java alternatives:

Selection	Path	Priority	Status
*	<code>/usr/lib/jvm/java-11-openjdk-amd64/bin/java</code>	1111	auto
ode	<code>/usr/lib/jvm/java-11-openjdk-amd64/bin/java</code>	1111	manual
1	<code>/usr/lib/jvm/java-11-openjdk-amd64/bin/java</code>	1111	manual
mode			

The terminal window shows the creation and modification of the `hadoop-env.sh` file. It starts with `sudo nano ~.bashrc`, followed by `-source ~.bashrc` which results in an error message: `-source: command not found`. Then, `sudo nano /usr/local/hadoop/etc/hadoop/hadoop-env.sh` is run, and finally `source ~.bashrc` is executed.

The terminal window shows the execution of the `hadoop-env.sh` file. It starts with `hduser@kali:[/usr/local/hadoop]`, followed by `$ sudo nano /usr/local/hadoop/etc/hadoop/hadoop-env.sh`, and ends with a prompt `$`.

Here enter java path

The screenshot shows a terminal window titled "hduser@kali: /usr/local/hadoop". The user has run the command "nano /etc/hadoop/hadoop-env.sh" to edit the Hadoop environment file. The file contains the following code:

```
GNU nano 6.3 /usr/local/hadoop/etc/hadoop/hadoop-env.sh *
#
# Unless required by applicable law or agreed to in writing, software
# distributed under the License is distributed on an "AS IS" BASIS,
# WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
# See the License for the specific language governing permissions and
# limitations under the license.

export JAVA_HOME=/usr/lib/jvm/java-11-openjdk-amd64/
# Set Hadoop-specific environment variables here.

1. sudo na
#HAD
^G Help ^C
^X Exit ^R Write Out
^W Where Is ^K Cut
Read File ^A Replace ^U Paste
^T Execute
^J Justify

2. source ~/.bashrc
```

Below the terminal, the command "\$ update-alternatives --config java" is shown, listing two Java alternatives:

Selection	Path	Priority
*	/usr/lib/jvm/java-11-openjdk-amd64/bin/java	111
1	/usr/lib/jvm/java-11-openjdk-amd64/bin/java	111

A message at the bottom of the screen says "Press <enter> to keep the current choice[*], or type selection number".

```
[hduser@kali]-[~/Desktop]
$ nano /usr/local/hadoop/etc/hadoop/hadoop-env.sh

[hduser@kali]-[~/Desktop]
$ sudo mkdir -p /app/hadoop/tmp

[hduser@kali]-[~/Desktop]
$ sudo chown hduser:hadoop /app/hadoop/tmp

[hduser@kali]-[~/Desktop]
$ sudo nano /usr/local/hadoop/etc/hadoop/core-site.xml
```

Om Kant Sharma 2022

```
export HADOOP_OPTS="-Djava.library.path=$HADOOP_HOME/lib"
#HADOOP VARIABLES END
2. source ~/bashrc
3. sudo nano /usr/local/hadoop/etc/hadoop/core-site.xml
   You may obtain a copy of the License at
   http://www.apache.org/licenses/LICENSE-2.0
Unless required by applicable law or agreed to in writing,
distributed under the License is distributed on an "AS IS"
WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express
or implied. See the License for the specific language governing
permissions and limitations under the License. See accompanying LICENSE file.
→

<configuration>
<property>
<name>hadoop.tmp.dir</name>
<value>/app/hadoop/tmp</value>
<description>A base for other temporary
directories.</description>
</property>
<property>
<name>fs.default.name</name>
<value>hdfs://localhost:54310</value>
<description>The name of the default file system.
A URI whose
scheme and authority determine the FileSystem implementation.
The
uri's scheme determines the config property (fs.SCHEME.impl)
naming
the FileSystem implementation class.
used to
determine the host, port, etc. for a filesystem.</description>
</property>
</configuration>
7. sudo nano /usr/local/hadoop/etc/hadoop/mapred-site.xml
<configuration>
<property>
<name>mapred.job.tracker</name>
<value>localhost:54311</value>
<description>The host and port that the MapReduce job
runs
at.
If "local", then jobs are run in-process as a single map
</description>
</property>
</configuration>
```

Ctrl+o then enter then ctrl+x

```
(hduser㉿kali)-[~/usr/local/hadoop]
$ sudo nano /usr/local/hadoop/etc/hadoop/mapred-site.xml
[sudo] password for hduser:
(hduser㉿kali)-[~/usr/local/hadoop]
$ sudo nano /usr/local/hadoop/etc/hadoop/mapred-site.xml
```

The screenshot shows a Linux desktop interface. In the top bar, there are icons for file operations like minimize, maximize, and close, followed by window numbers 1, 2, 3, and 4. Below the bar, a window titled "Minimize all open windows and show the desktop" is visible. To its right is a terminal window titled "hduser@kali: /usr/local/hadoop". The terminal shows the following content:

```

GNU nano 6.3      /usr/local/hadoop/etc/hadoop/mapred-site.xml
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<!--
Licensed under the Apache License, Version 2.0 (the "License");
you may not use this file except in compliance with the License.
See the License for the specific language governing permissions and
limitations under the License. See accompanying LICENSE file.
-->

<!-- Put site-specific property overrides in this file. -->

<configuration>
<property>
<name>mapred.job.tracker</name>
<value>localhost:54311</value>
<description>The host and port that the MapReduce job tracker runs
at.
If "local", then jobs are run in-process as a single map
and reduce task.
</description>
</property>
</configuration>

```

At the bottom of the terminal window, there are several keyboard shortcuts:

- ^C Help
- ^O Write Out
- ^R Read File
- ^W Where Is
- ^K Cut
- ^X Exit
- ^A Replace
- ^U Paste
- ^T Execute
- [*], or type selection number: []
- Read 30 lines
- ^J Justify

Ctrl+o then enter then ctrl+x for came out from the file

Then next

The screenshot shows a terminal window with the following command history:

```

[sudo] password for hduser:
[hduser@kali ~] $ sudo nano /usr/local/hadoop/etc/hadoop/mapred-site.xml
[hduser@kali ~] $ sudo mkdir -p /usr/local/hadoop_store/hdfs/namenode

```

Below the terminal, a table shows system processes:

	Priority	Status
jdk-amd64/bin/java	1111	auto m
jdk-amd64/bin/java	1111	manual

Om Kant Sharma 2022

```
8. sudo mkdir -p /usr/local/hadoop_store/hdfs/namenode  
9. sudo mkdir -p /usr/local/hadoop_store/hdfs/datanode  
10. sudo chown -R hdfs:hadoop /usr/local/hadoop_store  
11. sudo nano /usr/local/hadoop/etc/hadoop/hdfs-site.xml  
<configuration>  
<property>
```

```
8. sudo mkdir -p /usr/local/hadoop_store/hdfs/namenode
9. sudo mkdir -p /usr/local/hadoop_store/hdfs/datanode
10. sudo chown -R hdfs:hadoop /usr/local/hadoop_store
11. sudo nano /usr/local/hadoop/etc/hadoop/hdfs-site.xml
<configuration>
<property>
<name>dfs.replication</name>
<value>1</value>
<description>Default block replication.
The actual number of replications can be specified when the
file is created.
The default is used if replication is not specified in create
time.
```

```
8. sudo mkdir -p /usr/local/hadoop_store/hdfs/namenode
9. sudo mkdir -n /usr/local/hadoop_store/hdfs/datanode
10. sudo chown -R hdfs:hadoop /usr/local/hadoop_store
11. sudo nano /usr/local/hadoop/etc/hadoop/hdfs-site.xml
<configuration>
  <property>
    <name>dfs.replication</name>
    <value>1</value>
    <description>Default block replication.
The actual number of replications can be specified when the
file is created.
The default is used if replication is not specified in create
time.
    </description>
  </property>
</configuration>
```

```

8. sudo mkdir -p /usr/local/hadoop_p/store/hdfs/namenode
9. sudo mkdir -p /usr/local/hadoop/store/hdfs/datanode
10. sudo chown -R hdfs:hadoop /usr/local/hadoop_store
11. sudo nano /usr/local/hadoop/etc/hadoop/p/hdfs-site.xml
<configuration>
<property>
<name>dfs.replication</name>
<value>1</value>
<description>Default block replication.
The actual number of replications can be specified when the file is created.
The default is used if replication is not specified in create time.
</description>
</property>
<property>
<name>dfs.block.size</name>
<value>1048576</value>
<description>Default block size.
The actual block size can be specified when the file is created.
The default is used if block size is not specified in create time.
</description>
</property>

```

```
</configuration>
```

```
8. sudo mkdir -p /usr/local/hadoop_store/hdfs/namenode  
9. sudo mkdir -p /usr/local/hadoop_store/hdfs/datanode  
10. sudo chown -R hdfs:hadoop /usr/local/hadoop_store  
11. sudo nano /usr/local/hadoop/etc/hadoop/hdfs-site.xml  
<configuration>  
  <property>  
    <name>dfs.replication</name>  
    <value>1</value>  
  </property>  
  <description>Default block replication.  
The actual number of replications can be specified when the  
file is created.  
The default is used if replication is not specified in create  
time.  
  </description>  
  <property>  
    <name>dfs.block.size</name>  
    <value>1048576</value>  
  </property>
```

```
→ sudo nano /usr/local/hadoop/etc/hadoop/mapred-site.xml  
└── (hduser㉿kali)-[~/usr/local/hadoop]  
$ sudo mkdir -p /usr/local/hadoop_store/hdfs/namenode  
└── (hduser㉿kali)-[~/usr/local/hadoop]  
$ sudo mkdir -p /usr/local/hadoop_store/hdfs/datanode
```

```
[hadoop@kali ~]$ sudo mkdir -p /usr/local/hadoop_store/hdfs/namenode  
[hadoop@kali ~]$ sudo mkdir -p /usr/local/hadoop_store/hdfs/datanode  
[hadoop@kali ~]$ sudo chown -R hadoop:hadoop /usr/local/hadoop_store
```

```
[sudo] password for hduser: [ ]
```

```
[hduser@kali]~$ sudo mkdir -p /usr/local/hadoop_store/hdfs/datanode  
[hduser@kali]~$ sudo chown -R hduser:hadoop /usr/local/hadoop_store  
[sudo] password for hduser:  
[hduser@kali]~$ sudo nano /usr/local/hadoop/etc/hadoop/hdfs-site.xml
```

```
See the License for the specific language governing permissions and  
limitations under the License. See accompanying LICENSE file.  
→  
←— Put site-specific property overrides in this file. →  
  
<configuration>  
  
</configuration> va (providing /usr/bin/java).  
  
priority Status  
-----  
jdk-amd64/bin/java 1111 auto_m  
  
jdk-amd64/bin/java 1111 manual  
  
[*], or type selection number: 0
```

```

8. sudo mkdir -p /usr/local/hadoop_store/hdfs/namenode
9. sudo mkdir -p /usr/local/hadoop_store/hdfs/datanode
10. sudo chown -R hduser:hadoop /usr/local/hadoop_store
11. sudo nano /usr/local/hadoop/etc/hadoop/hdfs-site.xml
<configuration>
<property>
<name>dfs.replication</name>
<value>1</value>
<description>Default block replication.
The actual number of replications can be specified when the
file is created.
The default is used if replication is not specified in create
time.
</description>
</property>
<property>
<name>dfs.block.size</name>
<value>1048576</value>
</property>
<property>
<name>dfs.namenode.name.dir</name>
<value>file:/usr/local/hadoop_store/hdfs/namenode</value>
</property>
<property>
<name>dfs.datanode.data.dir</name>
<value>file:/usr/local/hadoop_store/hdfs/datanode</value>
</property>
</configuration>

```

```

<value>file:/usr/local/hadoop_store/nodes/datanode</value>
</property>
</configuration>

12. hadoop namenode -format
13. su mainuser
14. cd /usr/local/hadoop/sbin
15. ls
16. sudo su hduser
17. start-dfs.sh
18. start-yarn.sh
19. jps
20 stop-yarn.sh
21. stop-dfs.sh

```

```

<!-- Put site-specific property overrides in this file. -->

<configuration>
<property>
<name>dfs.replication</name>
<value>1</value>
<description>Default block replication.
The actual number of replications can be specified when the
file is created.
The default is used if replication is not specified in create
time.
</description>
</property>
<property>
<name>dfs.block.size</name>
<value>1048576</value>
</property>
<property>
<name>dfs.namenode.name.dir</name>
<value>file:/usr/local/hadoop_store/hdfs/namenode</value>
</property>
<property>
<name>dfs.datanode.data.dir</name>
<value>file:/usr/local/hadoop_store/hdfs/datanode</value>
</property>
</configuration>

[ Read 44 lines ]
^G Help      ^O Write Out   ^W Where Is    ^K Cut        ^T Execute
^X Exit      ^R Read File   ^\ Replace     ^U Paste      ^J Justify

```

```

File Actions Edit View Help

[(hduser㉿ kali)-[/usr/local/hadoop]] $ hadoop namenode -format

```

```

/***** SHUTDOWN_MSG: Shutting down NameNode at kali/127.0.1.1 *****
***** or type selection number: [ ] *****
[(hduser㉿ kali)-[/usr/local/hadoop]] $ su om

```

Om Kant Sharma 2022

The terminal window shows the following command history:

```
hduser@kali: /usr/local/hadoop/sbin
[om@kali:~/Desktop] $ cd /usr/local/hadoop/sbin
[om@kali:~/Desktop] $ ls
distribute-exclude.sh    start-all.sh      stop-balancer.sh
FederationStateStore    start-balancer.sh  stop-dfs.cmd
hadoop-daemon.sh        start-dfs.cmd    stop-dfs.sh
hadoop-daemons.sh       start-dfs.sh     stop-secure-dns.sh
httpfs.sh                start-secure-dns.sh stop-yarn.cmd
kms.sh                   start-yarn.cmd   stop-yarn.sh
mr-jobhistory-daemon.sh start-yarn.sh    workers.sh
refresh-namenodes.sh    stop-all.cmd    yarn-daemon.sh
start-all.cmd            stop-all.sh     yarn-daemons.sh

[om@kali:~/Desktop] $ sudo su hduser
[sudo] password for om:
[hduser@kali:~/Desktop] $ start-dfs.sh
Starting namenodes on [localhost]
localhost: Warning: Permanently added 'localhost' (ED25519) to the list of known hosts.
Starting datanodes
Starting secondary namenodes [kali]
kali: Warning: Permanently added 'kali' (ED25519) to the list of known hosts.
2022-11-06 22:46:36,958 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform ... using builtin-java classes where applicable

[hduser@kali:~/Desktop] $
```

The PDF viewer on the left shows the following content:

```
<name> dfs.datanode.data.dir</name>
<value> file:/usr/local/hadoop_store/hdfs/datanode</value>
</configuration>

12. hadoop namenode -format
13. su mainuser
14. cd /usr/local/hadoop/sbin
15. ls
16. sudo su hduser
17. start-dfs.sh
18. start-yarn.sh
19. jps
20 stop-yarn.sh
21. stop-dfs.sh

• Setting up Trash-
1. touch example.desktop
hadoop fs -put example.desktop /
2. Go to web UI
◦ http://localhost:9870/
3. cd /usr/local/hadoop/etc/hadoop/
4. sudo nano core-site.xml
<property>
<name>fs.trash.interval</name>
<value>3</value>
</property>
<property>
<name>fs.trash.checkpoint.interval</name>
<value>1</value>
</property>

22. hadoop fs –chmod -R 755 /user
```

The terminal window shows the following command history:

```
hduser@kali: /usr/local/hadoop/sbin
File Actions Edit View Help
start-all.cmd          stop-all.sh      yarn-daemons.sh

[om@kali:~/Desktop] $ sudo su hduser
[sudo] password for om:
[hduser@kali:~/Desktop] $ start-dfs.sh
Starting namenodes on [localhost]
localhost: Warning: Permanently added 'localhost' (ED25519) to the list of known hosts.
Starting datanodes
Starting secondary namenodes [kali]
kali: Warning: Permanently added 'kali' (ED25519) to the list of known hosts.
2022-11-06 22:46:36,958 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform ... using builtin-java classes where applicable

[hduser@kali:~/Desktop] $ start-yarn.sh
Starting resourcemanager
Starting nodemanagers

[hduser@kali:~/Desktop] $ jps
77905 SecondaryNameNode
78483 ResourceManager
78596 NodeManager
78939 Jps
77564 NameNode
77694 DataNode

[hduser@kali:~/Desktop] $
```

The PDF viewer on the left shows the following content:

```
<name> dfs.datanode.data.dir</name>
<value> file:/usr/local/hadoop_store/hdfs/datanode</value>
</configuration>

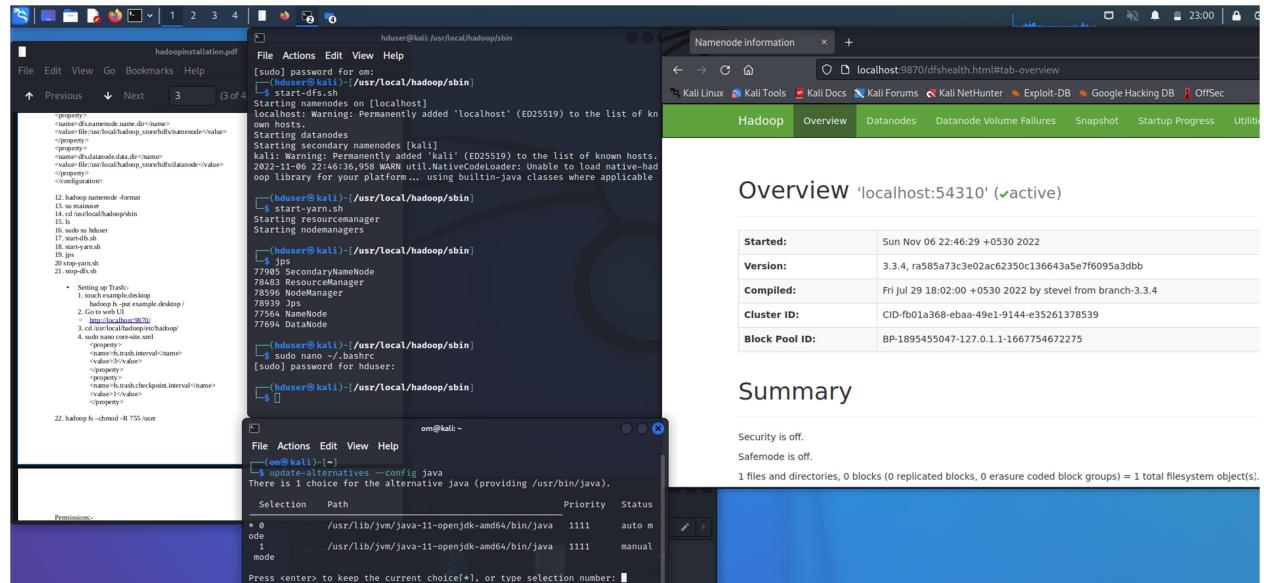
12. hadoop namenode -format
13. su mainuser
14. cd /usr/local/hadoop/sbin
15. ls
16. sudo su hduser
17. start-dfs.sh
18. start-yarn.sh
19. jps
20 stop-yarn.sh
21. stop-dfs.sh

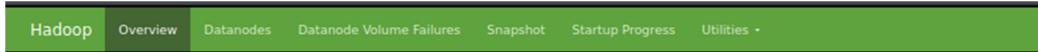
• Setting up Trash-
1. touch example.desktop
hadoop fs -put example.desktop /
2. Go to web UI
◦ http://localhost:9870/
3. cd /usr/local/hadoop/etc/hadoop/
4. sudo nano core-site.xml
<property>
<name>fs.trash.interval</name>
<value>3</value>
</property>
<property>
<name>fs.trash.checkpoint.interval</name>
<value>1</value>
</property>

22. hadoop fs –chmod -R 755 /user
```

After this – <http://localhost:9870/>

Open this link in Browser, you can see hadoop dashboard.





Overview 'localhost:54310' (✓active)

Started:	Sun Nov 06 22:46:29 +0530 2022
Version:	3.3.4, ra585a73c3e02ac62350c136643a5e7f6095a3dbb
Compiled:	Fri Jul 29 18:02:00 +0530 2022 by stevel from branch-3.3.4
Cluster ID:	CID-fb01a368-ebaa-49e1-9144-e35261378539
Block Pool ID:	BP-1895455047-127.0.1.1-1667754672275

Summary

Security is off.
Safemode is off.
1 files and directories, 0 blocks (0 replicated blocks, 0 erasure coded block groups) = 1 total filesystem object(s).
Heap Memory used 77.32 MB of 122 MB Heap Memory. Max Heap Memory is 984 MB.
Non Heap Memory used 67.24 MB of 69.56 MB Committed Non Heap Memory. Max Non Heap Memory is <unbounded>.

Configured Capacity:	23.5 GB
Configured Remote Capacity:	0 B
DFS Used:	28 KB (0%)
Non DFS Used:	14.58 GB
DFS Remaining:	7.7 GB (32.77%)
Block Pool Used:	28 KB (0%)
DataNodes usages% (Min/Median/Max/stdDev):	0.00% / 0.00% / 0.00% / 0.00%
Live Nodes	1 (Decommissioned: 0, In Maintenance: 0)
Dead Nodes	0 (Decommissioned: 0, In Maintenance: 0)
Decommissioning Nodes	0
Entering Maintenance Nodes	0
Total Datanode Volume Failures	0 (0 B)

/

Need some configuration to show trash folder like

Setting up trash folder

<https://stackoverflow.com/questions/53229221/terminal-error-zsh-permission-denied-startup-sh>

```
(om㉿kali)-[~/usr/local/hadoop]
$ cd /usr/local/hadoop/etc/hadoop

(om㉿kali)-[~/usr/local/hadoop/etc/hadoop]
$ sudo nano core-site.xml

(om㉿kali)-[~/usr/local/hadoop/etc/hadoop]
$
```

```
nadoop is -put example.desktop /
2. Go to web UI
  o http://localhost:9870/
3. cd /usr/local/hadoop/etc/hadoop
4. sudo nano core-site.xml

<property>
<name>fs.trash.interval</name>
<value>3</value>
</property>
<property>
<name>fs.trash.checkpoint.interval</name>
<value>1</value>
</property>

22. hadoop fs -chmod -R 755 /user

<configuration>
<property>
<name>fs.trash.interval</name>
<value>3</value>
</property>
<property>
<name>fs.trash.checkpoint.interval</name>
<value>1</value>
</property>
</configuration>
```

```
(om㉿kali)-[~/usr/local/hadoop/etc/hadoop]
$ sudo nano hdfs-site.xml

(om㉿kali)-[~/usr/local/hadoop/etc/hadoop]
$
```

```
Permissions:-
```

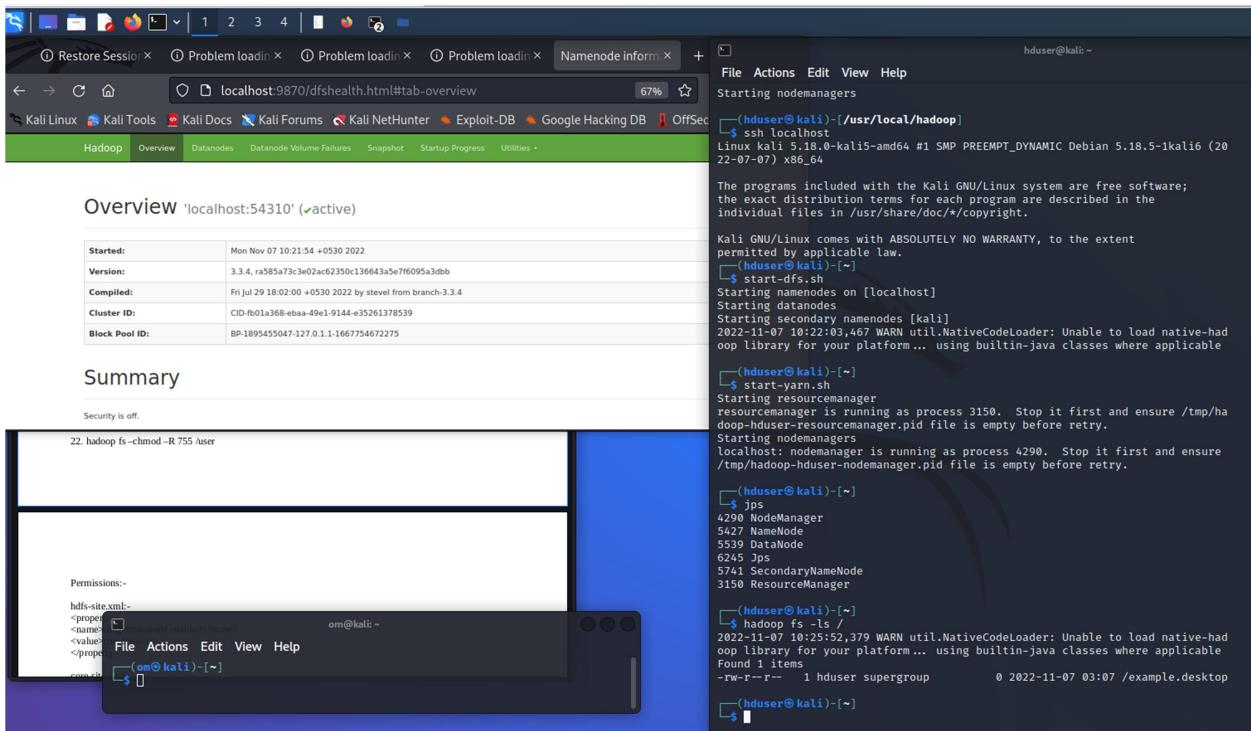
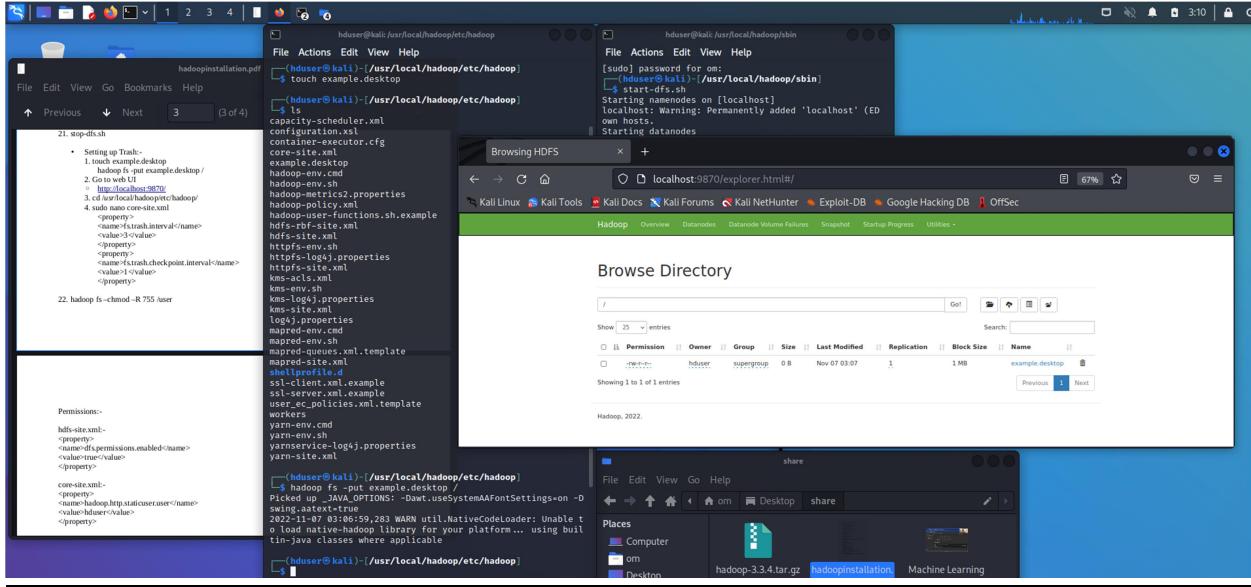
```
<description>
</property>
<property>
<name>dfs.permissions.enabled</name>
<value>true</value>
</property>

core-site.xml:-
```

```
<property>
<name>hadoop.http.staticuser.user</name>
<value>hduser</value>
</property>
```

```
<property>
<name>dfs.block.size</name>
<value>1048576</value>
</property>
<property>
<name>dfs.namenode.name.dir</name>
<value>file:/usr/local/hadoop_store/hdfs/namenode</value>
</property>
<property>
<name>dfs.datanode.data.dir</name>
<value>file:/usr/local/hadoop_store/hdfs/datanode</value>
</property>
<property>
<name>dfs.permissions.enabled</name>
<value>true</value>
</property>
</configuration>
```

Om Kant Sharma 2022



Hadoop installation and configuration's Excusion

```
└──(kali㉿ kali)-[~]
└─$ cd Desktop

└──(kali㉿ kali)-[~/Desktop]
└─$ mkdir share

└──(kali㉿ kali)-[~/Desktop]
└─$ sudo mount -t vboxsf kali share
[sudo] password for kali:

└──(kali㉿ kali)-[~/Desktop]
└─$ cd /share
cd: no such file or directory: /share

└──(kali㉿ kali)-[~/Desktop]
└─$ ls
share

└──(kali㉿ kali)-[~/Desktop]
└─$ cd share

└──(kali㉿ kali)-[~/Desktop/share]
└─$ ls
'ALL HAOOP INST.pdf'      'Machine Learning for
Everybody - Full Course.mp4'
hadoop-3.3.4.tar.gz        pwdofDs
hadoopinstallation.pdf     splunk-7.1.0-2e75b3406c5b-
linux-2.6-amd64.deb

└──(kali㉿ kali)-[~/Desktop/share]
└─$ sudo apt-get update
[sudo] password for kali:
Hit:1 http://kali.download/kali kali-rolling InRelease
Reading package lists... Done
```

└─(kali㉿kali)-[~/Desktop/share]

└─\$ **sudo apt-get dist-upgrade**

└─(kali㉿kali)-[~/Desktop/share]

└─\$ **sudo apt-get install default-jdk**

```
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
The following package was automatically installed and is no longer
required:
  openjdk-11-jre
Use 'sudo apt autoremove' to remove it.
The following additional packages will be installed:
  default-jdk-headless default-jre default-jre-headless java-common
    libc-bin libc-dev-bin libc-110n libc6 libc6-dev libc6-i386 libice-
dev
```

```
openjdk-amd64/bin/jhsdb to provide /usr/bin/jhsdb (jhsdb) in auto mode
Setting up libxt-dev:amd64 (1:1.2.1-1) ...
Setting up default-jdk (2:1.17-73) ...
```

└─(kali㉿kali)-[~/Desktop/share]

└─\$ **which java**

```
/usr/bin/java
```

└─(kali㉿kali)-[~/Desktop/share]

└─\$ **sudo addgroup hadoop**

```
Adding group `hadoop' (GID 1001) ...
```

```
Done.
```

└─(kali㉿kali)-[~/Desktop/share]

└─\$ **sudo adduser --ingroup hadoop hduser**

```
Adding user `hduser' ...
Adding new user `hduser' (1001) with group `hadoop' ...
Creating home directory `/home/hduser' ...
Copying files from `/etc/skel' ...
New password:
Retype new password:
passwd: password updated successfully
Changing the user information for hduser
Enter the new value, or press ENTER for the default
  Full Name []:
  Room Number []:
```

```
Work Phone []:
Home Phone []:
Other []:
Is the information correct? [Y/n] y
```

└─(kali㉿kali)-[~/Desktop/share]

└─\$ **groups hduser**

```
hduser : hadoop
```

└─(kali㉿kali)-[~/Desktop/share]

└─\$ **sudo apt-get install ssh**

```
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
The following package was automatically installed and is no longer required:
  openjdk-11-jre
Use 'sudo apt autoremove' to remove it.
The following additional packages will be installed:
  libssl3 openssh-client openssh-server openssh-sftp-
Get:1 http://kali.download/kali kali-rolling/main amd64 libssl3 amd64 3.0.7-1 [2,008
kB]
-----
-----Unpacking openssl (3.0.7-1)
over (3.0.4-2) ...
.
ssh.socket is a disabled or a static unit not running, not starting it.
Setting up ssh (1:9.0p1-1) ...
Processing triggers for man-db (2.10.2-1) ...
Processing triggers for kali-menu (2022.3.1) ...
Processing triggers for libc-bin (2.35-4) ...
```

└─(kali㉿kali)-[~/Desktop/share]

└─\$ **ssh-keygen**

```
Generating public/private rsa key pair.
Enter file in which to save the key (/home/kali/.ssh/id_rsa):
Created directory '/home/kali/.ssh'.
```

```
Enter passphrase (empty for no passphrase):
```

```
Enter same passphrase again:
```

```
Your identification has been saved in /home/kali/.ssh/id_rsa
```

```
Your public key has been saved in /home/kali/.ssh/id_rsa.pub
```

```
The key fingerprint is:
```

```
SHA256:GC2N7Znq2+P3g1CsO0BWR+RtUwaTUHEwRvo2FSqqamw kali@kali
```

```
The key's randomart image is:
```

```
+---[RSA 3072]---+
```

```
 |      o++X==  |
 |      =...+.B . |
 |      +.+o+ = . |
 |      o= +o+ o  |
 |      o. So +   |
 |      .oo . .   |
 |      . o. o .  |
 |      Eo .+ o . |
```

```
|   o. ooo+ ... |
+--- [SHA256] ---+
```

```
└─(kali㉿ kali)-[~/Desktop/share]
└─$ cat $HOME/.ssh/id_rsa.pub >> $HOME/.ssh/authorized_keys
```

```
└─(kali㉿ kali)-[~/Desktop/share]
└─$ ssh localhost
ssh: connect to host localhost port 22: Connection refused
```

```
└─(kali㉿ kali)-[~/Desktop/share]
└─$ service ssh start
Failed to start ssh.service: Access denied
See system logs and 'systemctl status ssh.service' for details.
```

```
└─(kali㉿ kali)-[~/Desktop/share]
└─$ apt-get install ssh openssh-client openssh-server
E: Could not open lock file /var/lib/dpkg/lock-frontend - open (13: Permission denied)
E: Unable to acquire the dpkg frontend lock (/var/lib/dpkg/lock-frontend), are you root?
```

```
└─(kali㉿ kali)-[~/Desktop/share]
└─$ sudo cp hadoop-3.3.4.tar.gz /home/hduser
```

```
└─(kali㉿ kali)-[~/Desktop/share]
└─$ su hduser
Password:
└─(hduser㉿ kali)-[/home/kali/Desktop/share]
└─$ cd
└─(hduser㉿ kali)-[~]
└─$ sudo cp hadoop-3.3.4.tar.gz /home/hduser
[sudo] password for hduser:
hduser is not in the sudoers file.
└─(hduser㉿ kali)-[~]
└─$ su hduser
Password:
└─(hduser㉿ kali)-[~]
└─$ sudo cp hadoop-3.3.4.tar.gz /home/hduser
[sudo] password for hduser:
kaSorry, try again.
[sudo] password for hduser:
hduser is not in the sudoers file.
└─(hduser㉿ kali)-[~]
└─$ su kali
Password:
└─(kali㉿ kali)-[/home/hduser]
└─$ cd
```

```
└─(kali㉿ kali)-[~]
└─$ sudo adduser hduser sudo
Adding user `hduser' to group `sudo' ...
Done.
```

```
└──(kali㉿ kali)-[~]
```

```
└─$ su hduser
```

Password:

```
└──(hduser㉿ kali)-[/home/kali]
```

```
└─$ ls
```

```
Desktop Downloads Music Public Videos  
Documents hadoop-3.3.4.tar.gz Pictures Templates
```

```
└──(hduser㉿ kali)-[/home/kali]
```

```
└─$ sudo cp hadoop-3.3.4.tar.gz /home/hduser
```

[sudo] password for hduser:

```
└──(hduser㉿ kali)-[/home/kali]
```

```
└─$ cd
```

```
└──(hduser㉿ kali)-[~]
```

```
└─$ ls
```

hadoop-3.3.4.tar.gz

```
└──(hduser㉿ kali)-[~]
```

```
└─$ sudo tar xvzf hadoop-3.3.4.tar.gz
```

```
hadoop-3.3.4/  
hadoop-3.3.4/licenses-binary/  
hadoop-3.3.4/licenses-binary/LICENSE-dust.txt  
hadoop-3.3.4/licenses-binary/LICENSE-re2j.txt  
hadoop-3.3.4/licenses-binary/LICENSE-slf4j.txt  
hadoop-3.3.4/licenses-binary/LICENSE-jquery.txt  
nfs/images/newwindow.png  
hadoop-3.3.4/share/doc/hadoop/hdfs-nfs/images/h3.jpg
```

```
└──(hduser㉿ kali)-[~]
```

```
└─$ ls
```

hadoop-3.3.4 hadoop-3.3.4.tar.gz

```
└──(hduser㉿ kali)-[~]
```

```
└─$ cd hadoop-3.3.4
```

```
└─(hduser㉿ kali)-[~/hadoop-3.3.4]
└─$ ls
bin      lib          licenses-binary  NOTICE.txt
share
etc      libexec      LICENSE.txt       README.txt
include  LICENSE-binary NOTICE-binary    sbin
```

```
└─(hduser㉿ kali)-[~/hadoop-3.3.4]
└─$ sudo mkdir -p /usr/local/hadoop
```

```
└─(hduser㉿ kali)-[~/hadoop-3.3.4]
└─$ sudo mv * /usr/local/hadoop
```

```
└─(hduser㉿ kali)-[~/hadoop-3.3.4]
└─$ cd /usr/local/hadoop
```

```
└─(hduser㉿ kali)-[/usr/local/hadoop]
└─$ ls
bin      lib          licenses-binary  NOTICE.txt  share
etc      libexec      LICENSE.txt       README.txt
include  LICENSE-binary NOTICE-binary    sbin
```

```
└─(hduser㉿ kali)-[/usr/local/hadoop]
└─$ sudo chown -R hduser:hadoop /usr/local/hadoop
```

NEXT CMD OPEN IN ANOTHER TERMINAL

```
└─(hduser㉿ kali)-[~]
└─$ update-alternatives --config java
There are 2 choices for the alternative java (providing /usr/bin/java).
      Selection    Path                                  Priority      Status
      * 0          /usr/lib/jvm/java-17-openjdk-amd64/bin/java  1711      auto mode
                    /usr/lib/jvm/java-11-openjdk-amd64/bin/java  1111      manual
mode
      2          /usr/lib/jvm/java-17-openjdk-amd64/bin/java  1711      manual
mode

Press <enter> to keep the current choice[*], or type selection number: ^C
```

└─(hduser㉿ kali)-[~]

└─\$ sudo nano /usr/local/hadoop/etc/hadoop/hadoop-env.sh

After licence drop this line-

export JAVA_HOME=/usr/lib/jvm/java-17-openjdk-amd64

Or if in root then at the end-

└─(root㉿ kali)-[~]

└─\$ sudo nano /usr/local/hadoop/etc/hadoop/hadoop-env.sh

export JAVA_HOME=/usr/lib/jvm/java-17-openjdk-amd64

if in root then at the end-

```
export HDFS_NAMENODE_USER=root
export HDFS_DATANODE_USER=root
export HDFS_SECONDARYNAMENODE_USER=root
export YARN_RESOURCEMANAGER_USER=root
export YARN_NODEMANAGER_USER=root
```

└─(hduser㉿ kali)-[~]

└─\$ sudo nano ~/.bashrc

At end in opened file paste blow code -

```
#HADOOP VARIABLES START
```

```
export JAVA_HOME=/usr/lib/jvm/java-17-openjdk-amd64
export HADOOP_INSTALL=/usr/local/hadoop
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
```

└─(hduser㉿ kali)-[~]

└─\$ source ~/.bashrc

```
└─(hduser㉿ kali)-[/usr/local/hadoop]
└─$ sudo mkdir -p /app/hadoop/tmp

└─(hduser㉿ kali)-[/usr/local/hadoop]
└─$ sudo chown hduser:hadoop /app/hadoop/tmp

└─(hduser㉿ kali)-[/usr/local/hadoop]
└─$ cd

└─(hduser㉿ kali)-[~]
└─$ sudo nano /usr/local/hadoop/etc/hadoop/core-site.xml

<configuration>
<property>
<name>hadoop.tmp.dir</name>
<value>/app/hadoop/tmp</value>
<description>A base for other temporary
directories.</description>
</property>
<property>
<name>fs.default.name</name>
<value>hdfs://localhost:54310</value>
<description>The name of the default file system.
A URI whose
scheme and authority determine the FileSystem implementation.
The
uri's scheme determines the config property (fs.SCHEME.impl)
naming
the FileSystem implementation class.
used to
The uri's authority is
determine the host, port, etc. for a filesystem.</description>
</property>
</configuration>
```

```
└── (hduser㉿ kali) - [ ~ ]  
└─$ sudo nano /usr/local/hadoop/etc/hadoop/mapred-site.xml
```

<configuration>

<property>

<name>mapred.job.tracker</name>

<value>localhost:54311</value>

<description>The host and port that the MapReduce job tracker runs

at.

If "local", then jobs are run in-process as a single map and reduce task.

</description>

</property>

</configuration>

```
└── (hduser㉿ kali)-[~]  
└─$ sudo mkdir -p /usr/local/hadoop_store/hdfs/datanode
```

```
└── (hduser㉿ kali)-[~]
```

```
└─$ sudo mkdir -p /usr/local/hadoop_store/hdfs/namenode
```

```
└── (hduser㉿ kali)-[~]
```

```
└─$ sudo chown -R hduser:hadoop /usr/local/hadoop_store
```

```
└── (hduser㉿ kali) - [ ~ ]  
└─$ sudo nano /usr/local/hadoop/etc/hadoop/hdfs-site.xml
```

```
<configuration>  
<property>  
<name>dfs.replication</name>  
<value>1</value>  
<description>Default block replication.  
The actual number of replications can be specified when the  
file is created.  
The default is used if replication is not specified in create  
time.  
</description>  
</property>  
<property>  
<name>dfs.block.size</name>  
<value>1048576</value>  
</property>  
<property>  
<name>dfs.namenode.name.dir</name>  
<value>file:/usr/local/hadoop_store/hdfs/namenode</value>  
</property>  
<property>  
<name>dfs.datanode.data.dir</name>  
<value>file:/usr/local/hadoop_store/hdfs/datanode</value>  
</property>  
</configuration>
```

```
└──(hduser㉿ kali)-[~]
└─$ start-dfs.sh
Starting namenodes on [localhost]
localhost: ssh: connect to host localhost port 22: Connection refused
Starting datanodes
localhost: ssh: connect to host localhost port 22: Connection refused
Starting secondary namenodes [kali]
kali: ssh: connect to host kali port 22: Connection refused
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
2022-11-13 11:20:43,303 WARN util.NativeCodeLoader: Unable to load native-hadoop
library for your platform... using builtin-java classes where applicable

└──(hduser㉿ kali)-[~]
└─$ sudo su

└──(root㉿ kali)-[/home/hduser]
└─# cd

└──(root㉿ kali)-[~]
└─# service ssh start

└──(root㉿ kali)-[~]
└─# su hduser

└──(hduser㉿ kali)-[/root]
└─$ cd

└──(hduser㉿ kali)-[~]
└─$ ssh localhost
The authenticity of host 'localhost (::1)' can't be established.
ED25519 key fingerprint is SHA256:HVWS/WDZyTNqWzrWTK88Ze5dijcqJ6qIUhgK71DFuzY.
This key is not known by any other names
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
Warning: Permanently added 'localhost' (ED25519) to the list of known hosts.
hduser@localhost's password:
Linux kali 5.18.0-kali5-amd64 #1 SMP PREEMPT_DYNAMIC Debian 5.18.5-1kali6 (2022-07-07)
x86_64

The programs included with the Kali GNU/Linux system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*/*copyright.

Kali GNU/Linux comes with ABSOLUTELY NO WARRANTY, to the extent
permitted by applicable law.

└──(hduser㉿ kali)-[~]
└─$ start-dfs.sh
Starting namenodes on [localhost]
localhost: hduser@localhost: Permission denied (publickey,password).
Starting datanodes
localhost: hduser@localhost: Permission denied (publickey,password).
Starting secondary namenodes [kali]
kali: Warning: Permanently added 'kali' (ED25519) to the list of known hosts.
kali: hduser@kali: Permission denied (publickey,password).
2022-11-13 11:22:13,007 WARN util.NativeCodeLoader: Unable to load native-hadoop
library for your platform... using builtin-java classes where applicable
```

└──(hduser㉿ kali)-[~]

└─\$ ssh-keygen

```
Generating public/private rsa key pair.
Enter file in which to save the key (/home/hduser/.ssh/id_rsa):
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
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:
SHA256:/WM8/ZBUGde+bWhhfqbDXE+QyW6FZr4G2wWyM1P5mv4 hduser@kali
The key's randomart image is:
+---[RSA 3072]---+
|          ... |
|         . .o |
|        ..*. |
|       . . #.o |
|      S . @.Bo |
|     o*oBoX |
|    *%+. |
|   ..o@o. |
|    o.oE|
+---[SHA256]---+
```

└──(hduser㉿ kali)-[~]

└─\$ start-dfs.sh

```
Starting namenodes on [localhost]
localhost: hduser@localhost: Permission denied (publickey,password).
Starting datanodes
localhost: hduser@localhost: Permission denied (publickey,password).
Starting secondary namenodes [kali]
kali: hduser@kali: Permission denied (publickey,password).
2022-11-13 11:24:26,830 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
```

└──(hduser㉿ kali)-[~]

└─\$ sudo su

[sudo] password for hduser:

└──(root㉿ kali)-[/home/hduser]

└─# service ssh start

└──(root㉿ kali)-[/home/hduser]

└─# su hduser

└──(hduser㉿ kali)-[~]

└─\$ cat \$HOME/.ssh/id_rsa.pub >> \$HOME/.ssh/authorized_keys

└──(hduser㉿ kali)-[~]

└─\$ ssh localhost

```
Linux kali 5.18.0-kali5-amd64 #1 SMP PREEMPT_DYNAMIC Debian 5.18.5-1kali6
(2022-07-07) x86_64
The programs included with the Kali GNU/Linux system are free software;
```

```
the exact distribution terms for each program are described in the  
individual files in /usr/share/doc/*/*copyright.
```

```
Kali GNU/Linux comes with ABSOLUTELY NO WARRANTY, to the extent  
permitted by applicable law.
```

```
Last login: Sun Nov 13 11:21:47 2022 from ::1
```

```
└─(hduser㉿ kali)-[~]
```

```
└─$ start-dfs.sh
```

```
Starting namenodes on [localhost]
```

```
localhost: WARNING: /usr/local/hadoop/logs does not exist. Creating.
```

```
Starting datanodes
```

```
Starting secondary namenodes [kali]
```

```
2022-11-13 11:28:39,083 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your  
platform... using builtin-java classes where applicable
```

```
└─(hduser㉿ kali)-[~]
```

```
└─$ start-yarn.sh
```

```
Starting resourcemanager
```

```
Starting nodemanagers
```

```
└─(hduser㉿ kali)-[~]
```

```
└─$ jps
```

```
47686 Jps
```

```
47548 NodeManager
```

```
47021 DataNode
```

```
47439 ResourceManager
```

```
47183 SecondaryNameNode
```

```
└─(hduser㉿ kali)-[~]
```

```
└─$ hadoop namenode -format
```

```
WARNING: Use of this script to execute namenode is deprecated.
```

```
WARNING: Attempting to execute replacement "hdfs namenode" instead.
```

```
2022-11-13 11:31:27,673 INFO namenode.NameNode: STARTUP_MSG:
```

```
*****
```

```
STARTUP_MSG: Starting NameNode
```

```
STARTUP_MSG: host = kali/127.0.1.1
```

```
STARTUP_MSG: args = [-format]
```

```
STARTUP_MSG: version = 3.3.4
```

```
STARTUP_MSG: classpath = 1.jar:/usr/local/hadoop/share/hadoop/common/lib/protobuf-java-
```

```
2.8.9.jar:/usr/local/hadoop/share/hadoop/common/lib/accessors-smart-
```

```
=====
```

```
=====
```

```
2022-11-13 11:31:28,339 INFO util.GSet: VM type = 64-bit
```

```
2022-11-13 11:31:28,340 INFO util.GSet: 1.0% max memory 984 MB = 9.8 MB
```

```
2022-11-13 11:31:28,340 INFO util.GSet: capacity = 2^20 = 1048576 entries
```

```
2022-11-13 11:31:28,341 INFO namenode.FSDirectory: ACLs
```

```
file /usr/local/hadoop_store/hdfs/namenode/current/fsimage.ckpt_00000000000000000000 of size 401  
bytes saved in 0 seconds .
```

```
2022-11-13 11:31:28,594 INFO namenode.NNStorageRetentionManager: Going to retain 1 images with  
txid >= 0
```

```
2022-11-13 11:31:28,615 INFO namenode.FSNamesystem: Stopping services started for active state  
2022-11-13 11:31:28,615 INFO namenode.FSNamesystem: Stopping services started for standby state
```

```
2022-11-13 11:31:28,624 INFO namenode.FSImage: FSImageSaver clean checkpoint: txid=0 when meet
shutdown.
2022-11-13 11:31:28,624 INFO namenode.NameNode: SHUTDOWN_MSG:
/*****SHUTDOWN_MSG: Shutting down NameNode at kali/127.0.1.1
*****/
```

└──(hduser㉿kali)-[~]

└─\$ **stop-all.sh**

WARNING: Stopping all Apache Hadoop daemons as hduser
in 10 seconds.

WARNING: Use CTRL-C to abort.

Stopping namenodes on [localhost]

Stopping datanodes

Stopping secondary namenodes [kali]

```
2022-11-13 11:32:03,922 WARN util.NativeCodeLoader:
Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
```

Stopping nodemanagers

Stopping resourcemanager

└──(hduser㉿kali)-[~]

└─\$ **start-all.sh**

WARNING: Attempting to start all Apache Hadoop daemons as hduser in 10
seconds.

WARNING: This is not a recommended production deployment configuration.

WARNING: Use CTRL-C to abort.

Starting namenodes on [localhost]

Starting datanodes

Starting secondary namenodes [kali]

```
2022-11-13 11:32:58,298 WARN util.NativeCodeLoader: Unable to load native-
hadoop library for your platform... using builtin-java classes where
applicable
```

Starting resourcemanager

Starting nodemanagers

└──(hduser㉿kali)-[~]

└─\$ **jps**

50306 SecondaryNameNode

52261 Jps

49991 NameNode

50632 NodeManager

50110 DataNode

50510 ResourceManager

<http://localhost:9870/>

HADOOP TRASH

```
└──(hduser㉿ kali)-[~]
    └─$ touch example.desktop
```

```
└──(hduser㉿ kali)-[~]
    └─$ ls
example.desktop hadoop-3.3.4 hadoop-3.3.4.tar.gz
```

```
└──(hduser㉿ kali)-[~]
    └─$ hadoop fs -put example.desktop /
```

```
2022-11-13 16:23:00,539 WARN util.NativeCodeLoader: Unable to load native-
hadoop library for your platform... using builtin-java classes where
applicable
```

```
└──(hduser㉿ kali)-[~]
    └─$ jps
50306 SecondaryNameNode
49991 NameNode
50632 NodeManager
123519 Jps
50110 DataNode
50510 ResourceManager
```

```
└──(hduser㉿ kali)-[~]
    └─$ ls
example.desktop hadoop-3.3.4 hadoop-3.3.4.tar.gz
```

```
└──(hduser㉿kali)-[~]
    └─$ cd /usr/local/hadoop/etc/hadoop
```

```
└──(hduser㉿kali)-[/usr/local/hadoop/etc/hadoop]
    └─$ ls
capacity-scheduler.xml          kms-log4j.properties
configuration.xsl                kms-site.xml
container-executor.cfg           log4j.properties
core-site.xml                   mapred-env.cmd
hadoop-env.cmd                  mapred-env.sh
hadoop-env.sh                   mapred-queues.xml.template
hadoop-metrics2.properties      mapred-site.xml
hadoop-policy.xml               shellprofile.d
hadoop-user-functions.sh.example ssl-client.xml.example
hdfs-rbf-site.xml               ssl-server.xml.example
hdfs-site.xml                   user_ec_policies.xml.template
```

```
httpfs-env.sh          workers
httpfs-log4j.properties   yarn-env.cmd
httpfs-site.xml        yarn-env.sh
kms-acls.xml          yarnservice-log4j.properties
kms-env.sh            yarn-site.xml
```

└──(hduser㉿kali)-[~/usr/local/hadoop/etc/hadoop]

└─\$ sudo nano core-site.xml

[sudo] password for hduser: At the end

```
<property>
<name>fs.trash.interval</name>
<value>3</value>
</property>
<property>
<name>fs.trash.checkpoint.interval</name>
<value>1</value>
</property>
```

└──(hduser㉿kali)-[~/usr/local/hadoop/etc/hadoop]

└─\$ sudo nano hdfs-site.xml

```
<property>
<name>dfs.permissions.enabled</name>
<value>true</value>
</property>
```

└──(hduser㉿kali)-[~/usr/local/hadoop/etc/hadoop]

└─\$ sudo nano core-site.xml

```
<property>
<name>hadoop.http.staticuser.user</name>
<value>hduser</value>
</property>
```

└──(hduser㉿kali)-[~/usr/local/hadoop/etc/hadoop]

└─\$ stop-all.sh

```
WARNING: Stopping all Apache Hadoop daemons as hduser  
in 10 seconds.  
WARNING: Use CTRL-C to abort.  
Stopping namenodes on [localhost]  
Stopping datanodes  
Stopping secondary namenodes [kali]  
2022-11-13 16:32:26,229 WARN util.NativeCodeLoader:  
Unable to load native-hadoop library for your  
platform... using builtin-java classes where applicable  
Stopping nodemanagers  
Stopping resourcemanager
```

└─(hduser㉿kali)-[/usr/local/hadoop/etc/hadoop]

└─\$ start-all.sh

```
WARNING: Attempting to start all Apache Hadoop daemons  
as hduser in 10 seconds.
```

```
WARNING: This is not a recommended production  
deployment configuration.
```

```
WARNING: Use CTRL-C to abort.
```

```
Starting namenodes on [localhost]
```

```
Starting datanodes
```

```
Starting secondary namenodes [kali]
```

```
2022-11-13 16:33:22,936 WARN util.NativeCodeLoader:
```

```
Unable to load native-hadoop library for your  
platform... using builtin-java classes where applicable
```

```
Starting resourcemanager
```

```
Starting nodemanagers
```

└─(hduser㉿kali)-[/usr/local/hadoop/etc/hadoop]

└─\$ hadoop fs -rm /example.desktop

```
2022-11-13 16:36:52,986 WARN util.NativeCodeLoader:
```

```
Unable to load native-hadoop library for your  
platform... using builtin-java classes where applicable
```

```
2022-11-13 16:36:53,605 INFO fs.TrashPolicyDefault:
```

```
Moved: 'hdfs://localhost:54310/example.desktop' to  
trash at:
```

```
hdfs://localhost:54310/user/hduser/.Trash/Current/examp  
le.desktop
```

```
└─(hduser㉿ kali)-[~/usr/local/hadoop/etc/hadoop]
└─$ hadoop fs -ls
2022-11-13 16:40:30,579 WARN util.NativeCodeLoader:
Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
Found 1 items
drwx-----  - hduser supergroup          0 2022-11-13
16:40 .Trash
```

```
└─(hduser㉿ kali)-[~/usr/local/hadoop/etc/hadoop]
└─$ touch example.omk
```

```
└─(hduser㉿ kali)-[~/usr/local/hadoop/etc/hadoop]
└─$ hadoop fs -put example.omk /
2022-11-13 16:47:01,430 WARN util.NativeCodeLoader:
Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
```

```
└─(hduser㉿ kali)-[~/usr/local/hadoop/etc/hadoop]
└─$ hadoop fs -rm /example.omk
2022-11-13 16:48:17,532 WARN util.NativeCodeLoader:
Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
2022-11-13 16:48:18,183 INFO fs.TrashPolicyDefault:
Moved: 'hdfs://localhost:54310/example.omk' to trash
at:
hdfs://localhost:54310/user/hduser/.Trash/Current/examp
le.omk
```

```
└─(hduser㉿ kali)-[~/usr/local/hadoop/etc/hadoop]
└─$ touch example.omkant
```

```
└─(hduser㉿ kali)-[~/usr/local/hadoop/etc/hadoop]
└─$ hadoop fs -put example.omkant /
2022-11-13 16:50:13,669 WARN util.NativeCodeLoader:
Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
```

```
└─(hduser㉿ kali) - [/usr/local/hadoop/etc/hadoop]
```

```
└─$ jps
```

```
126691 DataNode
132295 Jps
126599 NameNode
127161 ResourceManager
127272 NodeManager
126878 SecondaryNameNode
```

```
└─(hduser㉿ kali) - [/usr/local/hadoop/etc/hadoop]
```

```
└─$ stop-dfs.sh
```

```
Stopping namenodes on [localhost]
Stopping datanodes
Stopping secondary namenodes [kali]
2022-11-13 16:51:03,323 WARN util.NativeCodeLoader:
Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
```

```
└─(hduser㉿ kali) - [/usr/local/hadoop/etc/hadoop]
```

```
└─$ jps
```

```
135749 jps
```

Installing Hadoop on Ubuntu 20.04

- *Install java*

1. *sudo apt-get update*
2. *sudo apt-get install default-jdk*

Check java is installed or not

- *Create Hadoop Group and user*

1. *sudo addgroup hadoop*
2. *sudo adduser --ingroup hadoop hduser*

3. groups hduser

- *Install SSH*

1. *sudo apt-get install ssh*

check ssh and sshd

- *Create and setuid SSH*

1. *su hduser*

2. ssh-keygen

3. *cat \$HOME/.ssh/id_rsa.pub >> \$HOME/.ssh/authorized_keys*

```
4. ssh localhost
• Install Hadoop
Download Hadoop
1. tar xvzf hadoop-3.3.4.tar.gz
2. sudo -v
3. su mainuser
4. sudo adduser hduser sudo
5. su hduser
6. sudo mkdir -p /usr/local/hadoop
7. cd hadoop-3.3.4
8. sudo mv * /usr/local/hadoop
9. sudo chown -R hduser:hadoop /usr/local/hadoop
• Setup Configuration Files:-
Open New Terminal Window:- update-alternatives --config java
1. sudo nano ~/.bashrc
#HADOOP VARIABLES START
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-i386
export HADOOP_INSTALL=/usr/local/hadoop
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
2. source ~/.bashrc
3. sudo nano /usr/local/hadoop/etc/hadoop/hadoop-env.sh
export JAVA_HOME="JAVA PATH"
4. sudo mkdir -p /app/hadoop/tmp
5. sudo chown hduser:hadoop /app/hadoop/tmp
6. sudo nano /usr/local/hadoop/etc/hadoop/core-site.xml
<configuration>
<property>
<name>hadoop.tmp.dir</name>
<value>/app/hadoop/tmp</value>
<description>A base for other temporary
directories.</description>
<property>
<property>
<name>fs.default.name</name>
<value>hdfs://localhost:54310</value>
<description>The name of the default file system.
A URI whose
scheme and authority determine the FileSystem implementation.
The
uri's scheme determines the config property (fs.SCHEME.impl)
naming
the FileSystem implementation class.
used to
The uri's authority is
determine the host, port, etc. for a filesystem.</description>
<property>
</configuration>
7. sudo nano /usr/local/hadoop/etc/hadoop/mapred-site.xml
<configuration>
<property>
<name>mapred.job.tracker</name>
<value>localhost:54311</value>
<description>The host and port that the MapReduce job tracker
```

*runs
at.
If "local", then jobs are run in-process as a single map
and reduce task.*

</description>

<property>

<configuration>

8. sudo mkdir -p /usr/local/hadoop_store/hdfs/namenode

9. sudo mkdir -p /usr/local/hadoop_store/hdfs/datanode

10. sudo chown -R hduser:hadoop /usr/local/hadoop_store

11. sudo nano /usr/local/hadoop/etc/hadoop/hdfs-site.xml

<configuration>

<property>

<name>dfs.replication</name>

<value>1</value>

<description>Default block replication.
The actual number of replications can be specified when the
file is created.
The default is used if replication is not specified in create
time.

</description>

<property>

<property>

<name>dfs.block.size</name>

<value>1048576</value>

<property>

<property>

<name>dfs.namenode.name.dir</name>

<value>file:/usr/local/hadoop_store/hdfs/namenode</value>

<property>

<property>

<name>dfs.datanode.data.dir</name>

<value>file:/usr/local/hadoop_store/hdfs/datanode</value>

<property>

<configuration>

12. hadoop namenode -format

13. su mainuser(om)

14. cd /usr/local/hadoop/sbin

15. ls

16. sudo su hduser

17. start-dfs.sh

18. start-yarn.sh

19. jps

20 stop-yarn.sh

21. stop-dfs.sh

• Setting up Trash:-

1. touch example.desktop

hadoop fs -put example.desktop /

2. Go to web UI

◦ <http://localhost:9870/>

3. cd /usr/local/hadoop/etc/hadoop/

4. sudo nano core-site.xml

<property>

<name>fs.trash.interval</name>

<value>3</value>

<property>

<property>

<name>fs.trash.checkpoint.interval</name>

<value>1</value>

<property>

22. hadoop fs -chmod -R 755 /user

Permissions:-

```
hdfs-site.xml:-  
<property>  
<name>dfs.permissions.enabled</name>  
<value>true</value>  
</property>  
core-site.xml:-  
<property>  
<name>hadoop.http.staticuser.user</name>  
<value>hduser</value>  
</property>
```



Hadoop installation done

Hadoop Clustering:-

Restart Hadoop 1, 2

Om Kant Sharma 2022

kali2 [Running] - Oracle VM VirtualBox

File Machine View Input Devices Help

File System Home share ALL HADOOP... hadoopinst...

```
(hduser@kali2) ~$ su hduser
Password:
(hduser@kali2) ~$ cd
(hduser@kali2) ~$ ./start-all.sh
WARNING: Attempting to start all Apache Hadoop daemons as hduser in 10 seconds.
WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [localhost]
localhost: ssh: connect to host localhost port 22: Connection refused
Starting datanodes
localhost: ssh: connect to host localhost port 22: Connection refused
Starting secondary namenodes [kali2]
kali2: ssh: connect to host kali2 port 22: Connection refused
Picked up JAVA OPTIONS: -Dawt.useSystemAAFontSettings=true
2022-11-11 08:51:21,893 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Starting resourcemanager
Starting nodemanagers
localhost: ssh: connect to host localhost port 22: Connection refused
(hduser@kali2) ~$ service ssh start
(hduser@kali2) ~$ ./start-all.sh
WARNING: Attempting to start all Apache Hadoop daemons as hduser in 10 seconds.
WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [localhost]
localhost: ssh: connect to host localhost port 22: Connection refused
Starting datanodes
localhost: ssh: connect to host localhost port 22: Connection refused
Starting secondary namenodes [kali2]
kali2: ssh: connect to host kali2 port 22: Connection refused
Picked up JAVA OPTIONS: -Dawt.useSystemAAFontSettings=true
2022-11-11 08:53:56,529 WARN util.NativeCodeLoader:
```

Browsing HDFS x +

localhost:9870/explorer.html#/

Kali Linux Kali Tools Kali Docs Kali Forums Kali NetHunter Exploit-DB Google Hacking DB OffSec

Hadoop

Overview Datanodes Datanode Volume Failures Snapshot Startup Progress Utilities

Browse Directory

/

Show 25 entries Search:

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r--r--	hduser	supergroup	0 B	Nov 10 15:28	1	1 MB	example

Showing 1 to 1 of 1 entries

Previous Next

Demo Kali [Running] - Oracle VM VirtualBox

File Machine View Input Devices Help

File System Home share testfile hadoopinst...

```
(hduser@kali) ~$ ./start-all.sh
Command 'start' not found, did you mean:
  command 'start' from deb coreutils
  command 'restart' from deb-x11-utils
  command 'start' from deb-minitools
  command 'kstart' from deb-kde-clients
Try: sudo apt install <deb name>

(hduser@kali) ~$ ./start-all.sh
WARNING: Attempting to start all Apache Hadoop daemons as hduser in 10 seconds.
WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [localhost]
localhost: ssh: connect to host localhost port 22: Connection refused
Starting datanodes
localhost: ssh: connect to host localhost port 22: Connection refused
Starting secondary namenodes [kali]
kali: ssh: connect to host kali port 22: Connection refused
Picked up JAVA OPTIONS: -Dawt.useSystemAAFontSettings=true
2022-11-11 09:12:19,546 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Starting resourcemanager
Starting nodemanagers
localhost: ssh: connect to host localhost port 22: Connection refused
(hduser@kali) ~$ service ssh start
(hduser@kali) ~$ ./start-all.sh
WARNING: Attempting to start all Apache Hadoop daemons as hduser in 10 seconds.
WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [localhost]
localhost: ssh: connect to host localhost port 22: Connection refused
Starting datanodes
localhost: ssh: connect to host localhost port 22: Connection refused
Starting secondary namenodes [kali]
kali: ssh: connect to host kali port 22: Connection refused
Picked up JAVA OPTIONS: -Dawt.useSystemAAFontSettings=true
2022-11-11 09:13:10,818 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Starting resourcemanager: Resource manager is running as process 2002. Stop it first and ensure /tmp/hadoop-hduser-resourcemanager.pid file is empty before retry.
Starting nodemanagers
```

Browsing HDFS x +

localhost:9870/explorer.html#/

Kali Linux Kali Tools Kali Docs Kali Forums Kali NetHunter Exploit-DB Google Hacking DB OffSec

Hadoop Overview Datanodes Datanode Volume Failures Snapshot Startup Progress Utilities

Browse Directory

/

Show 25 entries Search:

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r--r--	hduser	supergroup	0 B	Nov 07 03:07	1	1 MB	example.desktop

Showing 1 to 1 of 1 entries

Previous Next

Om Kant Sharma 2022

