

## HDFS @ Phonepe

### Task 1:

- Setup a single node hadoop cluster.

1. Hadoop Framework uses Java so First install Java using the below command:

```
hadoop@hadoop-VirtualBox:~$ sudo apt install default-jdk
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following additional packages will be installed:
  ca-certificates-java default-jdk-headless default-jre default-jre-headless
  fonts-dejavu-extra java-common libatk-wrapper-java libatk-wrapper-java-jni
  libice-dev libpthread-stubs0-dev libsm-dev libx11-dev libxau-dev
  libxcb1-dev libxdmcp-dev libxt-dev openjdk-11-jdk openjdk-11-jdk-headless
  openjdk-11-jre openjdk-11-jre-headless x11proto-core-dev 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 fonts-ipafont-gothic fonts-ipafont-mincho
  fonts-wqy-microhei | fonts-wqy-zenhei
```

2. Create a User on the machine with sudo privileges

```
Sudo adduser --ingroup hadoop hduser
sudo usermod -aG sudo hduser
```

```
Setting up libX11-dev:amd64 (1:1.6.3-1) ...
hadoop@hadoop-VirtualBox:~$ 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
hadoop@hadoop-VirtualBox:~$ sudo usermod -aG sudo hduser
```

3. Install SSH ad create a private key on the user created.

```
hadoop@hadoop-VirtualBox:~$ sudo apt install ssh
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following additional packages will be installed:
  ncurses-term openssh-server openssh-sftp-server ssh-import-id
Suggested packages:
  molly-guard monkeysphere ssh-askpass
The following NEW packages will be installed:
  ncurses-term openssh-server openssh-sftp-server ssh-import-id
hadoop@hadoop-VirtualBox:~$ su hduser
Password:
To run a command as administrator (user "root"), use "sudo <command>".
See "man sudo_root" for details.
```

```
hduser@hadoop-VirtualBox:/home/hadoop$ cd
hduser@hadoop-VirtualBox:~$ 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:M9F+V4LTyjALuUvoetMVwrrwT9s/8oTRcMDDf4BylE2s hduser@hadoop-VirtualBox
The key's randomart image is:
+---[RSA 3072]---+
|   o +.. |
|   o * X  |
| o+ + E + . |
| .*=.* * o  |
| ..S=o.+ o  |
| . .o+oo o  |
| ...o ++ .  |
| .o . .+o . |
| ... . . . . |
+---[SHA256]---+
hduser@hadoop-VirtualBox:~$
```

```
hduser@hadoop-VirtualBox:~$ cat $HOME/.ssh/id_rsa.pub >> $HOME/.ssh/authorized_keys
hduser@hadoop-VirtualBox:~$ ssh localhost
The authenticity of host 'localhost (127.0.0.1)' can't be established.
ECDSA key fingerprint is SHA256:g5B21gw5SXVfnKf6zl7l4VHRyd/fjU1zmMFF/N63Y28.
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
Warning: Permanently added 'localhost' (ECDSA) to the list of known hosts.
Enter passphrase for key '/home/hduser/.ssh/id_rsa':
Welcome to Ubuntu 20.04.4 LTS (GNU/Linux 5.13.0-35-generic x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:    https://landscape.canonical.com
 * Support:       https://ubuntu.com/advantage

0 updates can be applied immediately.

Your Hardware Enablement Stack (HWE) is supported until April 2025.
*** System restart required ***

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

Ubuntu comes with ABSOLUTELY NO WARRANTY, to the extent permitted by
applicable law.
```

Installing Hadoop:

1. Download hadoop and copy the folder on the hduser.

2. Extract the hadoop tar.gz file.

```
tar xzvf <hadoop-file>
```

3. Create a directory and move the hadoop folder onto this directory

```
hduser@hadoop-VirtualBox:~$ sudo mkdir -p /usr/local/hadoop
hduser@hadoop-VirtualBox:~$ cd hadoop-2.10.1
hduser@hadoop-VirtualBox:~/hadoop-2.10.1$ ls
bin  include  libexec  NOTICE.txt  sbin
etc  lib      LICENSE.txt  README.txt  share
hduser@hadoop-VirtualBox:~/hadoop-2.10.1$ sudo mv * /usr/local/hadoop
hduser@hadoop-VirtualBox:~/hadoop-2.10.1$
```

4. Provide the ownership privileges to the below directory to the hduser from the hadoop user.

```
hduser@hadoop-VirtualBox:~/hadoop-2.10.1$ sudo chown -R hduser:hadoop /usr/loc
1/hadoop
```

5. Set the Configuration files to configure the single node hadoop cluster.

```
hduser@hadoop-VirtualBox:~$ sudo nano .bashrc
hduser@hadoop-VirtualBox:~$ source .bashrc
hduser@hadoop-VirtualBox:~$ sudo nano /usr/local/hadoop/etc/hadoop/hadoop-env.s
h
```

```
hduser@hadoop-VirtualBox:~$ sudo nano /usr/local/hadoop/etc/hadoop/core-site.xml
hduser@hadoop-VirtualBox:~$ sudo mkdir -p /app/hadoop/tmp
hduser@hadoop-VirtualBox:~$ sudo chown hduser:hadoop /app/hadoop/tmp
hduser@hadoop-VirtualBox:~$ sudo nano /usr/local/hadoop/etc/hadoop/core-site.xml
hduser@hadoop-VirtualBox:~$ cp /usr/local/hadoop/etc/hadoop/mapred-site.xml.tem
plate /usr/local/hadoop/etc/hadoop/mapred-site.xml
hduser@hadoop-VirtualBox:~$ sudo nano /usr/local/hadoop/etc/hadoop/mapred-site.
xml
```

.bashrc file:

The screenshot shows a terminal window titled "hduser@hadoop-VirtualBox: ~". It has two tabs: "GNU nano 4.8" and ".bashrc". The ".bashrc" tab is active, displaying the contents of the .bashrc file. The file contains shell script code for setting up the environment, including variable exports for Hadoop components like Java, Hadoop Install, PATH, and specific Hadoop modules.

```
GNU nano 4.8 .bashrc

if [ -f ~/.bash_aliases ]; then
    . ~/.bash_aliases
fi

# enable programmable completion features (you don't need to enable
# this, if it's already enabled in /etc/bash.bashrc and /etc/profile
# sources /etc/bash.bashrc).
if ! shopt -oq posix; then
    if [ -f /usr/share/bash-completion/bash_completion ]; then
        . /usr/share/bash-completion/bash_completion
    elif [ -f /etc/bash_completion ]; then
        . /etc/bash_completion
    fi
fi

#HADOOP VARIABLES START
export JAVA_HOME=/usr/lib/jvm/java-11-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
```

/usr/local/hadoop/etc/hadoop/mapred-site.xml

The screenshot shows a terminal window titled "hduser@hadoop-VirtualBox: ~". It has two tabs: "GNU nano 4.8" and "/usr/local/hadoop/etc/hadoop/mapred-site.xml". The "/usr/local/hadoop/etc/hadoop/mapred-site.xml" tab is active, displaying the XML configuration file. The file includes a license notice and a configuration section for the MapReduce job tracker, specifying its host and port.

```
GNU nano 4.8 /usr/local/hadoop/etc/hadoop/mapred-site.xml
<?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.
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, 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. 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>
```

/usr/local/hadoop/etc/hadoop/core-site.xml

The screenshot shows a terminal window titled "hduser@hadoop-VirtualBox: ~". It displays the contents of the core-site.xml file. The file starts with a license notice and then defines properties for HDFS and the default file system.

```
GNU nano 4.8          /usr/local/hadoop/etc/hadoop/core-site.xml
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. See accompanying LICENSE file.
-->

<!-- Put site-specific property overrides in this 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. The uri's authority is
used to
determine the host, port, etc. for a filesystem.</description>
</property>
</configuration>
```

/usr/local/hadoop/etc/hadoop/hadoop-env.sh

The screenshot shows a terminal window titled "hduser@hadoop-VirtualBox: ~". It displays the contents of the hadoop-env.sh file. The file includes a license notice and then sets environment variables for Hadoop, specifically defining the JAVA\_HOME variable.

```
GNU nano 4.8          /usr/local/hadoop/etc/hadoop/hadoop-env.sh
# Licensed to the Apache Software Foundation (ASF) under one
# or more contributor license agreements. See the NOTICE file
# distributed with this work for additional information
# regarding copyright ownership. The ASF licenses this file
# to you under the Apache License, Version 2.0 (the
# "License"); you may not use this file except in compliance
# with the License. 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, 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.

# Set Hadoop-specific environment variables here.

# The only required environment variable is JAVA_HOME. All others are
# optional. When running a distributed configuration it is best to
# set JAVA_HOME in this file, so that it is correctly defined on
# remote nodes.

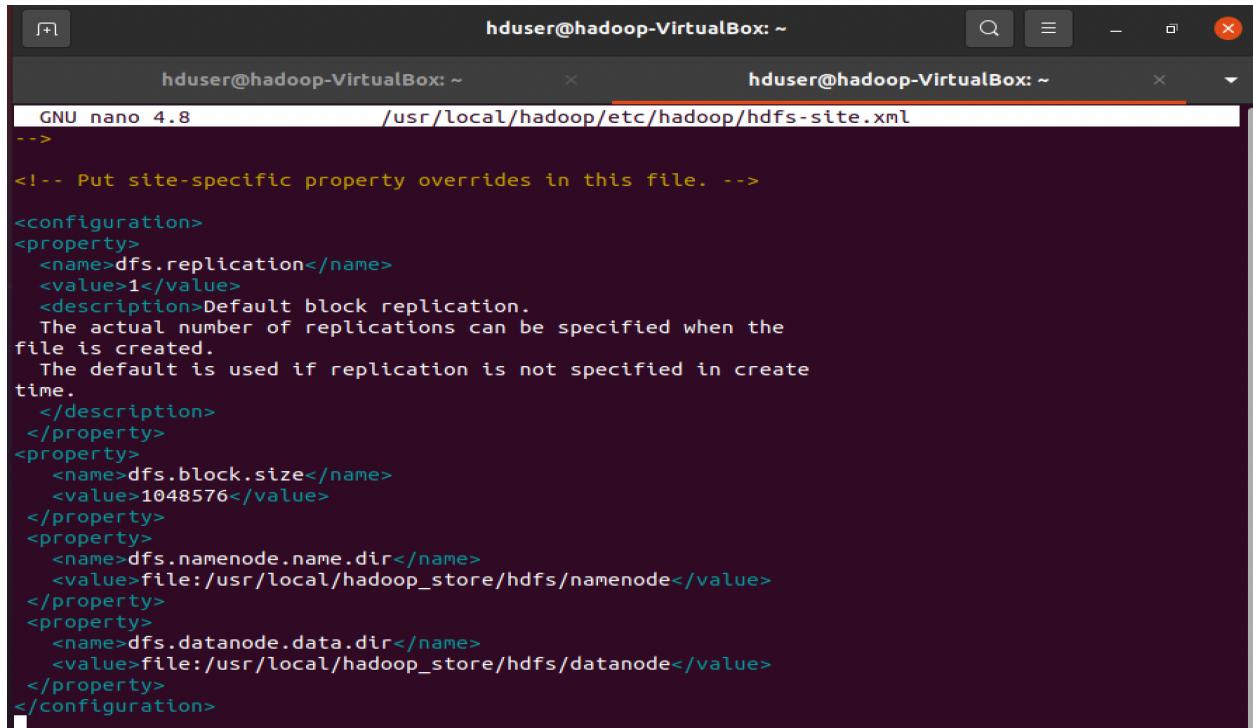
# The java implementation to use.
export JAVA_HOME=/usr/lib/jvm/java-11-openjdk-amd64

# The jsvc implementation to use. Jsvc is required to run secure datanodes
# that bind to privileged ports to provide authentication of data transfer
# protocol. Jsvc is not required if SASL is configured for authentication of
[ Read 117 lines ]
```

6. Create Directories for datanode and namenode and give ownership to hduser.

```
hduser@hadoop-VirtualBox:~$ sudo mkdir -p /usr/local/hadoop_store/hdfs/namenode
hduser@hadoop-VirtualBox:~$ sudo mkdir -p /usr/local/hadoop_store/hdfs/datanode
hduser@hadoop-VirtualBox:~$ sudo chown -R hduser:hadoop
chown: missing operand after 'hduser:hadoop'
Try 'chown --help' for more information.
hduser@hadoop-VirtualBox:~$ sudo chown -R hduser:hadoop /usr/local/hadoop_store
hduser@hadoop-VirtualBox:~$ sudo nano /usr/local/hadoop/etc/hadoop/hdfs-site.xml
l
hduser@hadoop-VirtualBox:~$
```

Do the following configuration on the hdfs-site.xml  
/usr/local/hadoop/etc/hadoop/hdfs-site.xml



```
GNU nano 4.8          /usr/local/hadoop/etc/hadoop/hdfs-site.xml
-->

<!-- 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>
```

7. Format the Hadoop Filesystem

The hadoop file system needs to be formatted so that we can start to use it.

Hadoop namenode -format

This should be used only for the first time we will be start using the hadoop. If executed again , it will destroy all the data on hdfs.

```
hduser@hadoop-VirtualBox:~$ hadoop namenode -format
DEPRECATED: Use of this script to execute hdfs command is deprecated.
Instead use the hdfs command for it.

22/03/10 19:35:25 INFO namenode.NameNode: STARTUP_MSG:
/*****STARTUP_MSG: Starting NameNode*****
STARTUP_MSG: host = hadoop-VirtualBox/127.0.1.1
STARTUP_MSG: args = [-format]
STARTUP_MSG: version = 2.10.1
STARTUP_MSG: classpath = /usr/local/hadoop/etc/hadoop:/usr/local/hadoop/share/hadoop/common/lib/commons-lang-2.6.jar:/usr/local/hadoop/share/hadoop/common/lib/commons-net-3.1.jar:/usr/local/hadoop/share/hadoop/common/lib/curator-recipes-2.13.0.jar:/usr/local/hadoop/share/hadoop/common/lib/commons-lang3-3.4.jar:/usr/local/hadoop/share/hadoop/common/lib/guava-11.0.2.jar:/usr/local/hadoop/share/hadoop/common/lib/servlet-api-2.5.jar:/usr/local/hadoop/share/hadoop/common/lib/mockito-all-1.8.5.jar:/usr/local/hadoop/share/hadoop/common/lib/hamcrest-core-1.3.jar:/usr/local/hadoop/share/hadoop/common/lib/netty-3.10.6.Final.jar:/usr/local/hadoop/share/hadoop/common/lib/snappy-java-1.0.5.jar:/usr/local/hadoop/share/hadoop/common/lib/xmlenc-0.52.jar:/usr/local/hadoop/share/hadoop/common/lib/jersey-server-1.9.jar:/usr/local/hadoop/share/hadoop/common/lib/commons-math3-3.1.1.jar:/usr/local/hadoop/share/hadoop/common/lib/audience-annotations-0.5.0.jar:/usr/local/hadoop/share/hadoop/common/lib/jaxb-api-2.2.2.jar:/usr/local/hadoop/share/hadoop/common/lib/junit-4.11.jar:/usr/local/hadoop/share/hadoop/common/lib/jsch-0.1.55.jar:/usr/local/hadoop/share/hadoop/common/lib/jsp-api-2.1.jar:/usr/local/hadoop/share/hadoop/common/lib/commons-compress-1.19.jar:/usr/local/hadoop/share/hadoop/common/lib/commons-beanutils-1.9.4.jar:/usr/local/hadoop/share/hadoop/common/lib/paramater-2.3.jar:/usr/local/hadoop/share/hadoop/common
```

## 8. Starting Hadoop

Use start-dfs.sh to start the hadoop distributed file system.

Use jps command to check if all the hadoop daemons like datanode , namenode and secondary namenode are running or not.

```
hduser@hadoop-VirtualBox:~$ cd /usr/local/hadoop/sbin
hduser@hadoop-VirtualBox:/usr/local/hadoop/sbin$ start-dfs.sh
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication
.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-
2.10.1.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop
.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflec-
tive access operations
WARNING: All illegal access operations will be denied in a future release
22/03/10 19:36:44 WARN util.NativeCodeLoader: Unable to load native-hadoop libr-
ary for your platform... using builtin-java classes where applicable
Starting namenodes on [localhost]
Enter passphrase for key '/home/hduser/.ssh/id_rsa':
localhost: starting namenode, logging to /usr/local/hadoop/logs/hadoop-hduser-n
amenode-hadoop-VirtualBox.out
localhost: WARNING: An illegal reflective access operation has occurred
localhost: WARNING: Illegal reflective access by org.apache.hadoop.security.aut-
hentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/h
adoop-auth-2.10.1.jar) to method sun.security.krb5.Config.getInstance()
localhost: WARNING: Please consider reporting this to the maintainers of org.ap
ache.hadoop.security.authentication.util.KerberosUtil
localhost: WARNING: Use --illegal-access=warn to enable warnings of further ill-
egal reflective access operations
localhost: WARNING: All illegal access operations will be denied in a future re
lease
Enter passphrase for key '/home/hduser/.ssh/id_rsa':
```

```
hduser@hadoop-VirtualBox:/usr/local/hadoop/sbin$ jps
2753 SecondaryNameNode
2373 NameNode
2549 DataNode
2873 Jps
```

GUI Interface of hadoop cluster look like this. To view this go to localhost:50070 on your browser.

**Overview** 'localhost:54310' (active)

<b>Started:</b>	Thu Mar 10 19:37:05 +0530 2022
<b>Version:</b>	2.10.1, r1827467c9a56f133025f28557bfc2c562d78e816
<b>Compiled:</b>	Mon Sep 14 18:47:00 +0530 2020 by centos from branch-2.10.1
<b>Cluster ID:</b>	CID-847afe5a-b64e-416c-b4d9-bcacd934d84b
<b>Block Pool ID:</b>	BP-1513769911-127.0.1.1-1646921126082

## Summary

### Task 2:

Perform the following operation post cluster setup and share the result in a google doc.

- List Files.

For listing the files in hdfs:

- hdfs dfs -ls / : This lists all the files /directories for the destination path.
- hdfs dfs -ls /file:/// : This shows all the hdfs system files.
- hdfs dfs -ls -R /hadoop : This shows the files in the recursively in the given directories and subdirectories.

Create File:

```
hduser@hadoop-VirtualBox:~$ hdfs dfs -mkdir /test
```

List File:

```
hduser@hadoop-VirtualBox:~$ hadoop fs -ls /
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.10.1.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
22/03/10 19:42:54 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 1 items
drwxr-xr-x - hduser supergroup          0 2022-03-10 19:40 /test
```

```
hduser@hadoop-VirtualBox:~$ hadoop fs -ls / file:/// 
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.10.1.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
22/03/10 19:41:22 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 1 items
drwxr-xr-x - hduser supergroup          0 2022-03-10 19:40 /test
Found 26 items
drwxr-xr-x - root   root      4096 2022-03-10 19:27 file:///app
drwxr-xr-x - root   root    36864 2022-03-10 19:18 file:///bin
drwxr-xr-x - root   root      4096 2022-03-10 18:46 file:///boot
drwxrwxr-x - root   root      4096 2022-03-10 18:35 file:///cdrom
drwxr-xr-x - root   root     4080 2022-03-10 19:20 file:///dev
drwxr-xr-x - root   root    12288 2022-03-10 19:18 file:///etc
drwxr-xr-x - root   root      4096 2022-03-10 18:48 file:///home
drwxr-xr-x - root   root      4096 2022-03-10 18:49 file:///lib
drwxr-xr-x - root   root      4096 2022-02-23 14:17 file:///lib32
drwxr-xr-x - root   root      4096 2022-03-10 18:45 file:///lib64
drwxr-xr-x - root   root      4096 2022-02-23 14:17 file:///libx32
drwx----- - root   root   16384 2022-03-10 18:33 file:///lost+found
drwxr-xr-x - root   root      4096 2022-03-10 19:17 file:///media
drwxr-xr-x - root   root      4096 2022-02-23 14:17 file:///mnt
```

- Read/Write Files.

Create a file /hello

```
hduser@hadoop-VirtualBox:~$ cat >hello
Hello hi
^C
hduser@hadoop-VirtualBox:~$ ls
hadoop-2.10.1  hadoop-2.10.1.tar.gz  hello
hduser@hadoop-VirtualBox:~$ hadoop fs -put hello /
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.10.1.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
22/03/10 21:06:49 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
```

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r--r--	hduser	supergroup	9 B	Mar 10 21:06	1	1 MB	hello
drwxr-xr-x	hduser	supergroup	0 B	Mar 10 19:40	0	0 B	test

The GUI currently shows that we have two files on the hdfs system.

Create one file on the local machine and use it to write into the hdfs file..

```
hduser@hadoop-VirtualBox:~$ cat > hel2
HEYYYYYY!
^C
```

Read file:

- hdfs dfs -cat /test : This will display the content of HDFS file test.

```
hduser@hadoop-VirtualBox:~$ hdfs dfs -cat /hello
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.10.1.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
22/03/10 23:44:23 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Hello hi
```

Write file:

- hdfs dfs -appendToFile /localsystemfile /hdfsfile - appends the content of the local system file to a hdfs file.

```
hduser@hadoop-VirtualBox:~$ pwd
/home/hduser
hduser@hadoop-VirtualBox:~$ hdfs dfs -appendToFile /home/hduser/hel2 /hello
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.10.1.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
22/03/10 23:55:03 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
hduser@hadoop-VirtualBox:~$ hdfs dfs -cat /hello
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.10.1.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
22/03/10 23:55:10 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Hello hi
HEYYYYYY!
```

- **Upload/Download Files.**

To Upload a file , create a file and upload it on the hadoop file system using : hadoop fs -put hello /

```
hduser@hadoop-VirtualBox:~$ cat >hello
Hello hi
^C
hduser@hadoop-VirtualBox:~$ ls
hadoop-2.10.1  hadoop-2.10.1.tar.gz  hello
hduser@hadoop-VirtualBox:~$ hadoop fs -put hello /
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.10.1.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
22/03/10 21:06:49 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
```

Now you'll be able to see it on hdfs GUI board.

To download a file from the hadoop file system use : hdfs dfs -get /filenametodownload /directorywheretoputthedownloadfile

```

hduser@hadoop-VirtualBox:~$ hdfs dfs -get /test ~
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.10.1.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
22/03/10 21:08:06 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-j
ava classes where applicable
hduser@hadoop-VirtualBox:~$ ls
hadoop-2.10.1  hadoop-2.10.1.tar.gz  hello  test
hduser@hadoop-VirtualBox:~$ 

```

## ● File Management.

File Management Commands:

### 1. copy command

```
hdfs dfs -cp /source /destination
```

The file Hey is created same as hello

```

hduser@hadoop-VirtualBox:~$ hdfs dfs -cp /hello /hey
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/
mmon/lib/hadoop-auth-2.10.1.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
22/03/11 01:43:41 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes when
applicable

```

<input type="checkbox"/>	-rw-r--r--	hduser	supergroup	19 B	Mar 11 01:43	1	1 MB	hey	
--------------------------	------------	--------	------------	------	--------------	---	------	-----	--

### 2. Move command

```
hdfs dfs -mv /source /destination
```

```

hduser@hadoop-VirtualBox:~$ hdfs dfs -mv /hello /hdfs
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/
mmon/lib/hadoop-auth-2.10.1.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
22/03/11 01:44:08 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes when
applicable

```

<input type="checkbox"/>	-rw-r--r--	hduser	hadoop	19 B	Mar 10 23:55	1	1 MB	hdfs	
--------------------------	------------	--------	--------	------	--------------	---	------	------	--

Hello file is moved to hdfs now.

### 3. Remove command

```
hdfs dfs -rm /source
```

```

hduser@hadoop-VirtualBox:~$ hdfs dfs -rm /hey
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/
mmon/lib/hadoop-auth-2.10.1.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
22/03/11 01:45:41 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes when
applicable
Deleted /hey

```

This Removes the file named /hey .

#### 4. Make Directory Command

hdfs dfs -mkdir /filename

```
hduser@hadoop-VirtualBox:~$ hdfs dfs -mkdir /filename
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/mmon/lib/hadoop-auth-2.10.1.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
22/03/11 01:46:23 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
```

This makes the directory of the name filename.

□	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name	⋮
□	drwxr-xr-x	hduser	supergroup	0 B	Mar 11 01:46	0	0 B	filename	trash

#### 5. Remove Directory Command

hdfs dfs -rmdir /filename

```
hduser@hadoop-VirtualBox:~$ hdfs dfs -rmdir /filename
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/mmon/lib/hadoop-auth-2.10.1.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
22/03/11 01:47:00 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
```

This removes the directory.

#### 6. Create a file

hdfs fs -touchz /filename

```
hduser@hadoop-VirtualBox:~$ hdfs dfs -touchz /file
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/mmon/lib/hadoop-auth-2.10.1.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
22/03/11 01:47:32 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
```

This creates a file with the required name.

□	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name	⋮
□	-rw-r--r--	hduser	supergroup	0 B	Mar 11 01:47	1	1 MB	file	trash
□	-rw-r--r--	hdbs	hadoop	19 B	Mar 10 23:55	1	1 MB	hdfs	trash
□	-rw-r--r--	hduser	supergroup	0 B	Mar 10 23:49	1	1 MB	hello1	trash

- **Ownership and Validation.**

Ownership Privileges can be provided by setting up these commands.

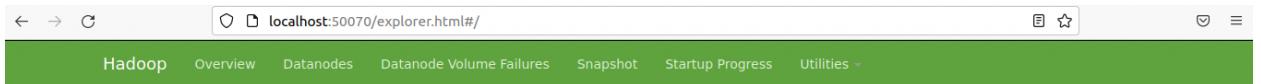
hdfs fs -chown hdbs:hdfs /filename : Changes owner of the file.

```
hduser@hadoop-VirtualBox:~$ hdfs dfs -chown -R hdfs:hdfs /test
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.10.1.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
22/03/10 23:26:25 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
```

**hdfs fs -chgrp -R <group-name> <filename>** : Changes the group of the file.

**Hdfs dfs -chmod 755 filename** : Changes permission of the file

```
hduser@hadoop-VirtualBox:~$ hdfs dfs -chgrp -R hadoop hello
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.10.1.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
22/03/11 01:30:35 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
hduser@hadoop-VirtualBox:~$ hdfs dfs -chmod 755 hello
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.10.1.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
22/03/11 01:31:03 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
```



## Browse Directory

Browse Directory									
<input style="width: 200px; margin-right: 10px;" type="text" value="/"/> <input type="button" value="Go!"/> <input type="button" value="New"/> <input type="button" value="Upload"/> <input type="button" value="Delete"/>									
Show	25	entries	<input style="width: 150px; margin-right: 10px;" type="text" value="Search:"/> <input type="button" value="Search"/>						
	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name	
	-rw-r--r--	hdfs	hadoop	9 B	Mar 10 21:06	1	1 MB	hello	
	drwxr-xr-x	hdfs	hadoop	0 B	Mar 10 19:40	0	0 B	test	

Showing 1 to 2 of 2 entries

[Previous](#) [1](#) [Next](#)

Hadoop, 2020.

There is this checksum command which can be used to dump the checksum information for files that match the file pattern .

```
hduser@hadoop-VirtualBox:~$ hdfs dfs -checksum /hello
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.10.1.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
22/03/10 23:56:14 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
/hello MD5-of-0MD5-of-512CRC32C 00000200000000000000000018950a8143602d44ed206b888eab45f
hduser@hadoop-VirtualBox:~$
```