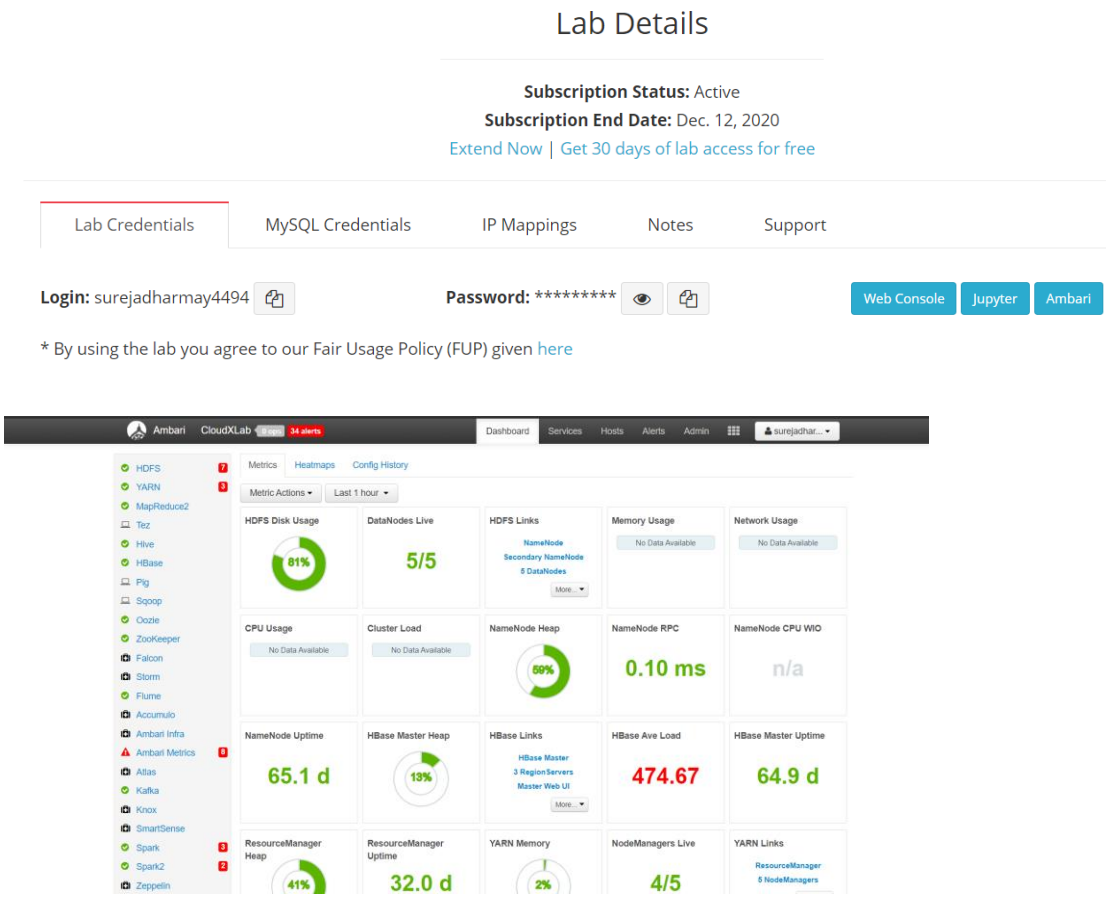
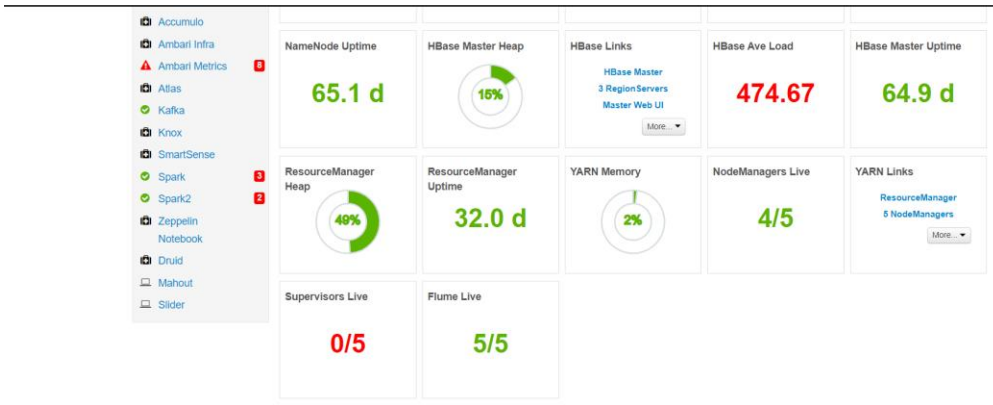
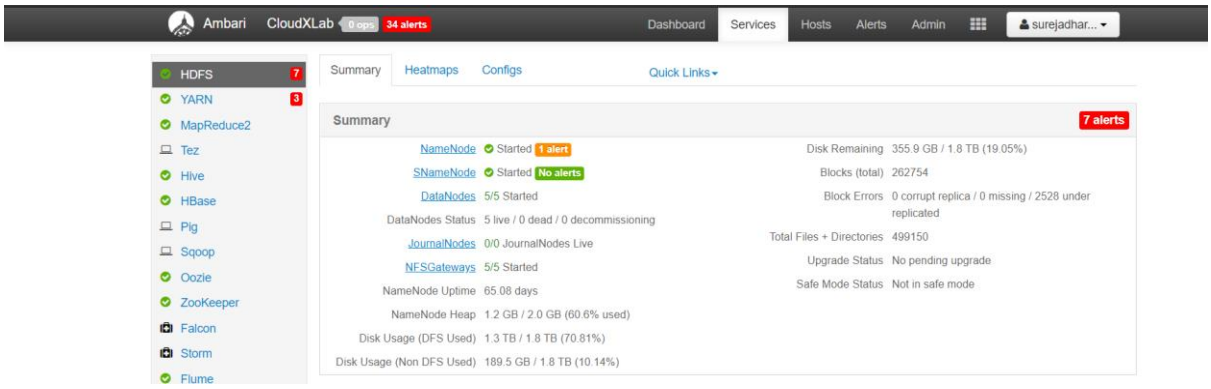
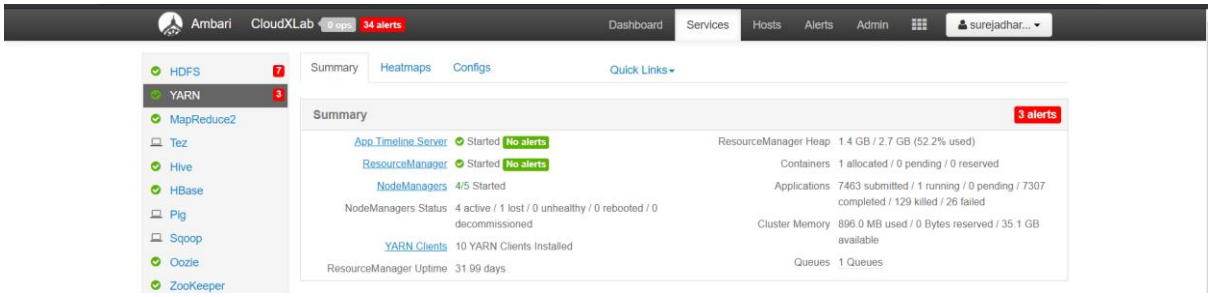
Practical-7

TASK : **Login into CloudxLab and Open Ambari and check your server status and other configurations available.**

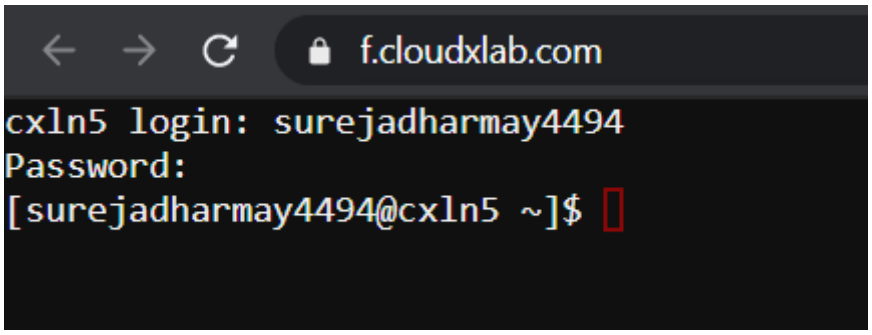




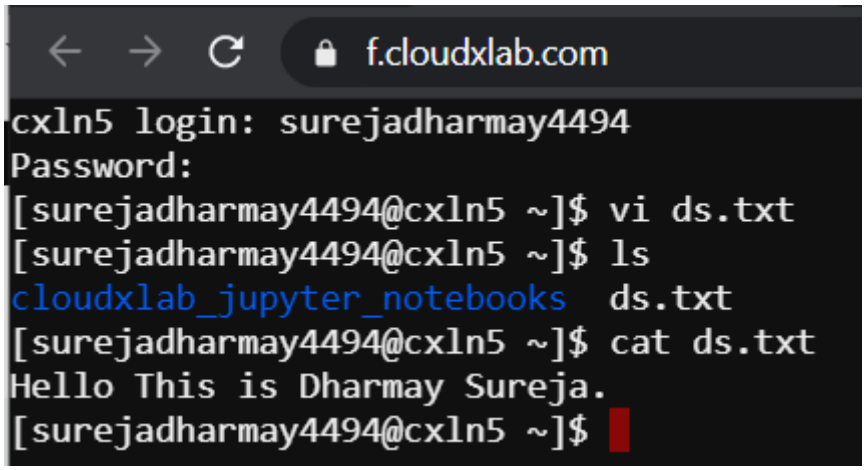

Ambari
CloudXLab
0 ops
34 alerts
Dashboard
Services
Hosts
Alerts
Admin
surejadhar...
HDFS 7
YARN 3
MapReduce2
Tez
Hive
HBase
Pig
Sqoop
Oozie
ZooKeeper
Falcon
Storm
Flume
Summary
Heatmaps
Configs
Quick Links
Summary
7 alerts
NameNode Started 1 alert
SNameNode Started No alerts
DataNodes 5/5 Started
DataNodes Status 5 live / 0 dead / 0 decommissioning
JournalNodes 0/0 JournalNodes Live
NFSGateways 5/5 Started
NameNode Uptime 65.08 days
NameNode Heap 1.2 GB / 2.0 GB (60.6% used)
Disk Usage (DFS Used) 1.3 TB / 1.8 TB (70.81%)
Disk Usage (Non DFS Used) 189.5 GB / 1.8 TB (10.14%)
Disk Remaining 355.9 GB / 1.8 TB (19.05%)
Blocks (total) 262754
Block Errors 0 corrupt replica / 0 missing / 2528 under replicated
Total Files + Directories 499150
Upgrade Status No pending upgrade
Safe Mode Status Not in safe mode


Ambari
CloudXLab
0 ops
34 alerts
Dashboard
Services
Hosts
Alerts
Admin
surejadhar...
HDFS 7
YARN 3
MapReduce2
Tez
Hive
HBase
Pig
Sqoop
Oozie
ZooKeeper
Summary
Heatmaps
Configs
Quick Links
Summary
3 alerts
App Timeline Server Started No alerts
ResourceManager Started No alerts
NodeManagers 4/5 Started
NodeManagers Status 4 active / 1 lost / 0 unhealthy / 0 rebooted / 0 decommissioned
YARN Clients 10 YARN Clients Installed
ResourceManager Uptime 31.99 days
ResourceManager Heap 1.4 GB / 2.7 GB (52.2% used)
Containers 1 allocated / 0 pending / 0 reserved
Applications 7463 submitted / 1 running / 0 pending / 7307 completed / 129 killed / 26 failed
Cluster Memory 896.0 MB used / 0 Bytes reserved / 35.1 GB available
Queues 1 Queues

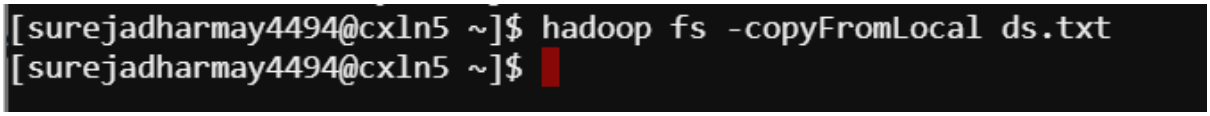
TASK :- Using CloudxLab Linux Console .



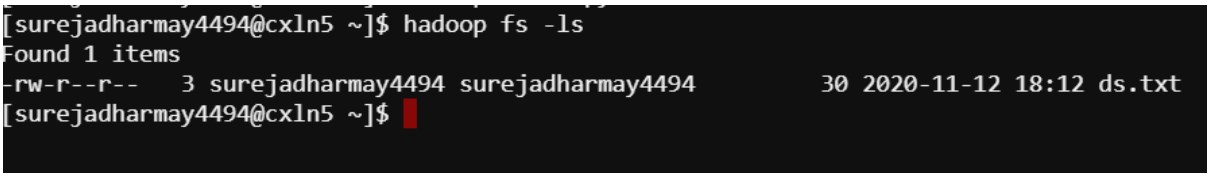
1. Create File and save some content in it.



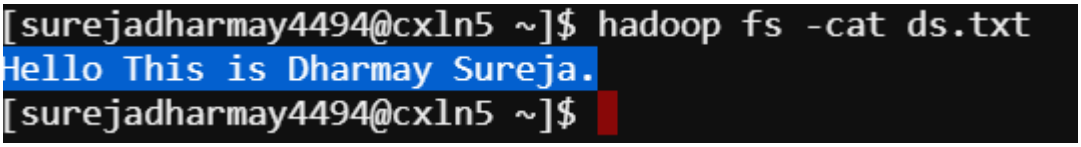
2. Transfer it into the HDFS.



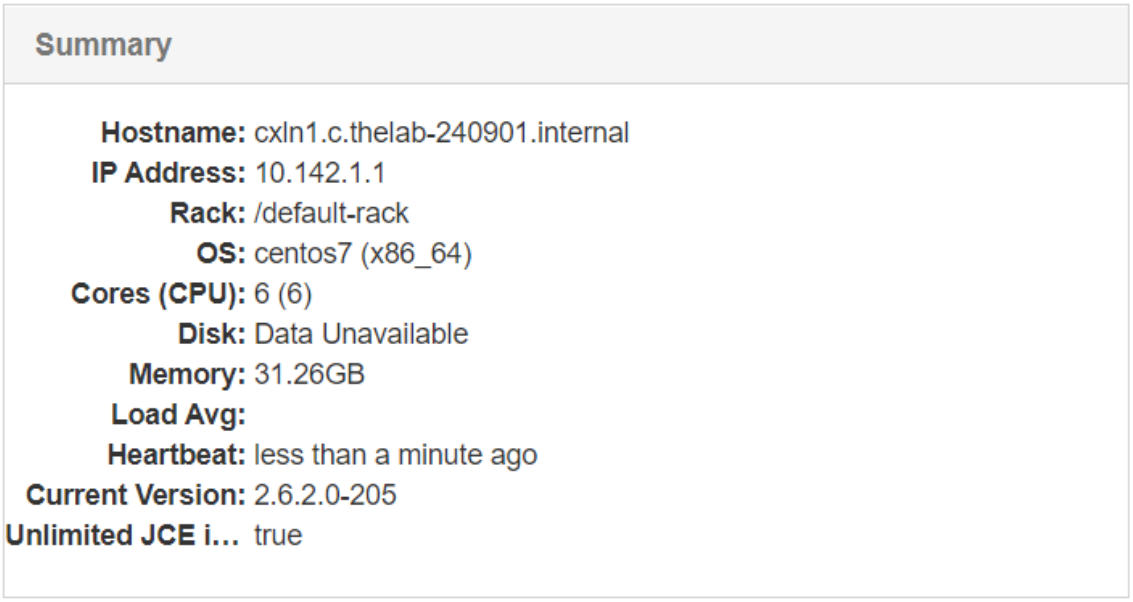
3. Verify if it is available in HDFS.



4. Check the Content of the file (Display it on screen)



TASK :- Access the files which you have stored in the HDFS and list out the IP addresses of the nodes of the blocks in which the file is stored, and check its replication factor. Change the Replication factor other than by default you observed and Give a Reason what happens when you Increase and decrease the replication factor of the file stored.



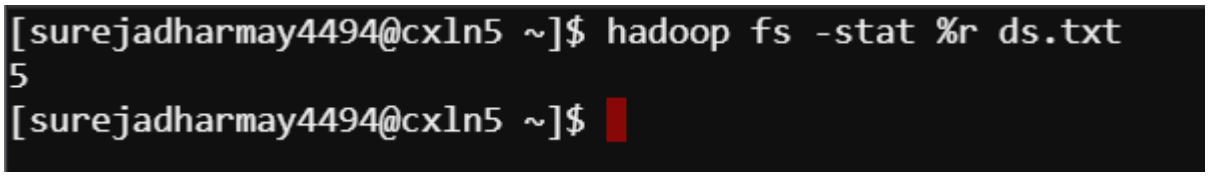
[Default Summary]

```
[surejadharmay4494@cxln5 ~]$ hadoop fs -stat %r ds.txt
3
[surejadharmay4494@cxln5 ~]$
```

Changing the replication factor to 5

```
[surejadharmay4494@cxln5 ~]$ hadoop fs -setrep -w 5 ds.txt
Replication 5 set: ds.txt
Waiting for ds.txt ..... done
[surejadharmay4494@cxln5 ~]$
```

Checking the RF.



Reason : When the replication factor is changed depending upon the increase or decrease of the RF, the copies of the file is made.

TASK : Perform the following tasks.

1. Create directory.

```
[surejadharmay4494@cxln5 ~]$ mkdir ds-test
[surejadharmay4494@cxln5 ~]$ ls
cloudxlab_jupyter_notebooks  ds-test  ds.txt
[surejadharmay4494@cxln5 ~]$
```

2. Create File

```
[surejadharmay4494@cxln5 ~]$ touch filev.txt
[surejadharmay4494@cxln5 ~]$ ls
cloudxlab_jupyter_notebooks  ds-test  ds.txt  filev.txt
```

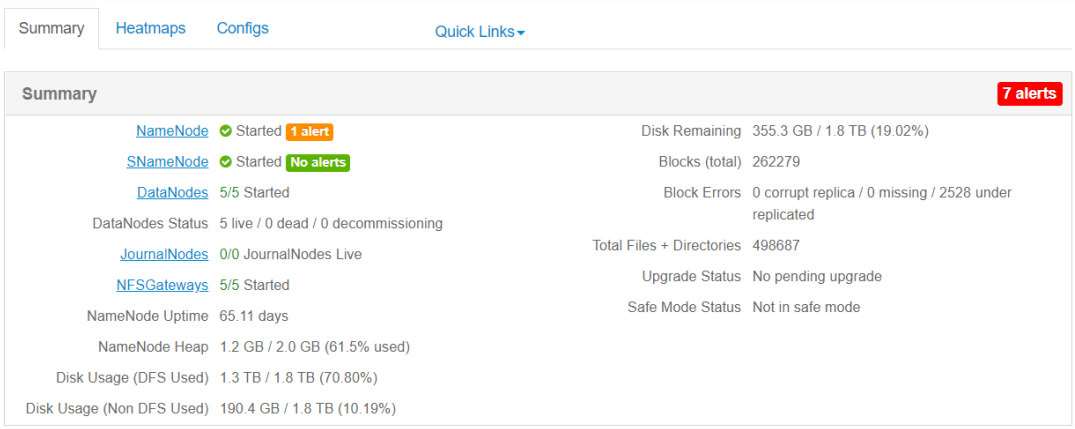
3. Transfer new Created file into created directory .

```
[surejadharmay4494@cxln5 ~]$ mv filev.txt ds-test/
[surejadharmay4494@cxln5 ~]$ ls
cloudxlab_jupyter_notebooks  ds-test  ds.txt
[surejadharmay4494@cxln5 ~]$ ls ds-test/
filev.txt
[surejadharmay4494@cxln5 ~]$
```

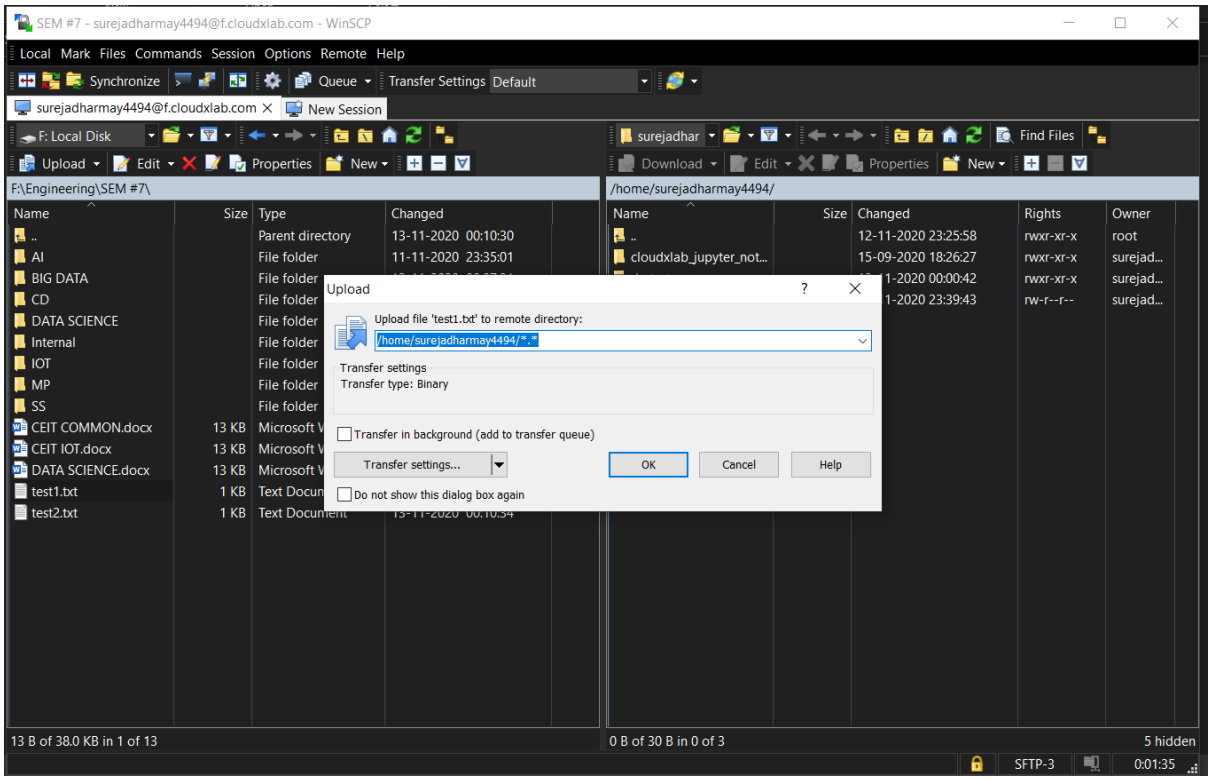
4. Change the replication factor of the file and the whole directory .

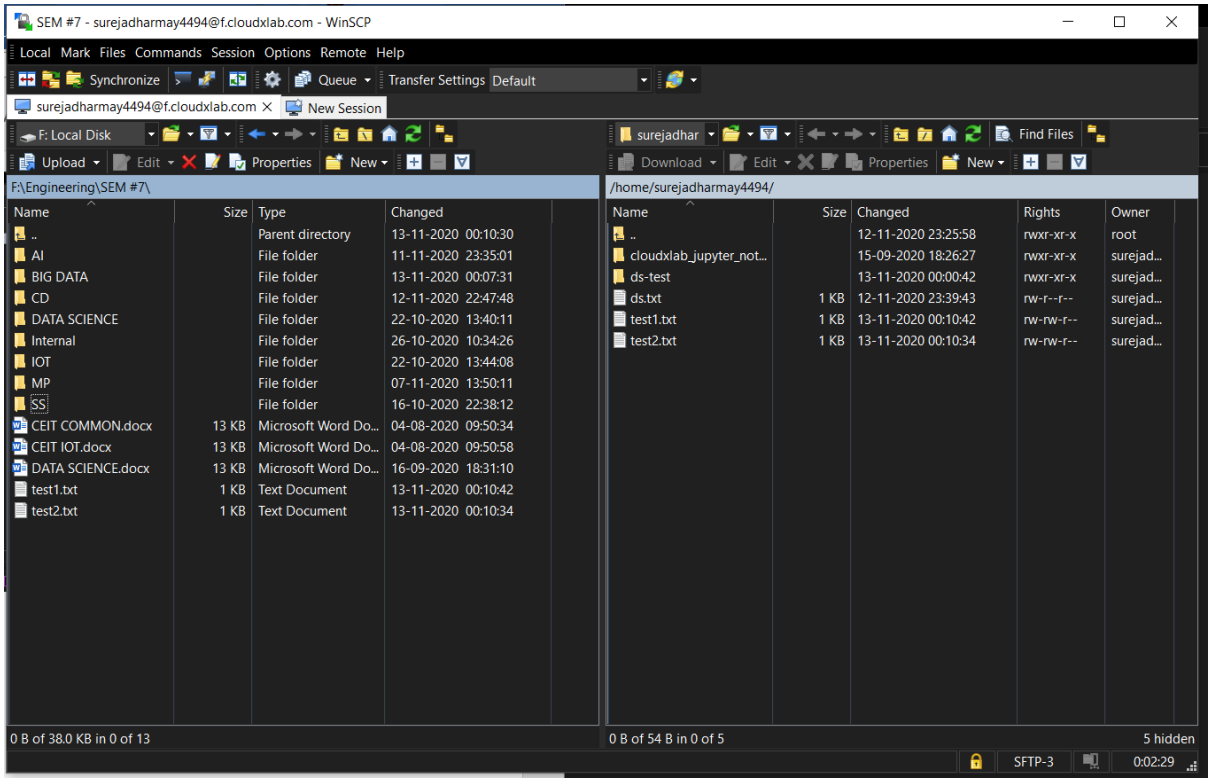
```
[surejadharmay4494@cxln5 ~]$ hadoop fs -put ds-test/
[surejadharmay4494@cxln5 ~]$ hadoop fs -stat %r ds-test/
0
[surejadharmay4494@cxln5 ~]$ hadoop fs -stat %r ds-test/filev.txt
3
[surejadharmay4494@cxln5 ~]$ hadoop fs -setrep -w 2 ds-test/
Replication 2 set: ds-test/filev.txt
Waiting for ds-test/filev.txt ... done
[surejadharmay4494@cxln5 ~]$ hadoop fs -setrep -w 8 ds-test/filev.txt
Replication 8 set: ds-test/filev.txt
Waiting for ds-test/filev.txt ... done
[surejadharmay4494@cxln5 ~]$
```

5. Check the list of files and directories in HDFS (Not in Linux Console).



<u>TASK</u> : Upload the files (at least 2) from your local machines to Hadoop cluster console and then Copy it to the HDFS .





TASK :- Perform the following commands and write about what output it gives you.

1. ls – list out the contents

```
[surejadharmay4494@cxln5 ~]$ ls
cloudxlab_jupyter_notebooks  ds-test  ds.txt  test1.txt  test2.txt
[surejadharmay4494@cxln5 ~]$
```

2. pwd – shows the current directory

```
[surejadharmay4494@cxln5 ~]$ pwd
/home/surejadharmay4494
[surejadharmay4494@cxln5 ~]$
```

3. help – list out help commands

```
[surejadharmay4494@cxln5 ~]$ help
GNU bash, version 4.2.46(2)-release (x86_64-redhat-linux-gnu)
These shell commands are defined internally.  Type `help' to see this list.
Type `help name' to find out more about the function `name'.
Use `info bash' to find out more about the shell in general.
Use `man -k' or `info' to find out more about commands not in this list.

A star (*) next to a name means that the command is disabled.

 job_spec [&]                            history [-c] [-d offset] [n] or histo>
 (( expression ))                        if COMMANDS; then COMMANDS; [ elif CO>
 . filename [arguments]                  jobs [-lnprs] [jobspec ...] or jobs ->
 :                                       kill [-s sigspec | -n signum | -sigsp>
 [ arg... ]                              let arg [arg ...]
 [[ expression ]]                        local [option] name[=value] ...
 alias [-p] [name[=value] ... ]          logout [n]
 bg [job_spec ...]                       mapfile [-n count] [-O origin] [-s co>
 bind [-lpvsPVS] [-m keymap] [-f filena> popd [-n] [+N | -N]
 break [n]                               printf [-v var] format [arguments]
 builtin [shell-builtin [arg ...]]       pushd [-n] [+N | -N | dir]
 caller [expr]                           pwd [-LP]
 case WORD in [PATTERN [| PATTERN]...) > read [-ers] [-a array] [-d delim] [-i>
 cd [-L|[-P [-e]]] [dir]                 readarray [-n count] [-O origin] [-s >
 command [-pVv] command [arg ...]        readonly [-aAf] [name[=value] ...] or>
 compgen [-abcdefgjksuv] [-o option]  [> return [n]
 complete [-abcdefgjksuv] [-pr] [-DE] [> select NAME [in WORDS ... ;] do COMMA>
 compopt [-o|+o option] [-DE] [name ...> set [-abefhkmnptuvxBCHP] [-o option-n>
```

4. ls --la : - list out the contents with extra information

```
[surejadharmay4494@cxln5 ~]$ ls -la
total 1028
drwx------     5 surejadharmay4494 surejadharmay4494    4096 Nov 12 18:42 .
drwxr-xr-x 26590 root              root              1007616 Nov 12 17:55 ..
-rw-r--r--     1 surejadharmay4494 surejadharmay4494      18 Oct 30  2018 .bash_log
out
-rw-r--r--     1 surejadharmay4494 surejadharmay4494     193 Oct 30  2018 .bash_pro
file
-rw-r--r--     1 surejadharmay4494 surejadharmay4494     231 Oct 30  2018 .bashrc
drwxr-xr-x     2 surejadharmay4494 surejadharmay4494    4096 Sep 15 12:56 cloudxlab
_jupyter_notebooks
drwxr-xr-x     2 surejadharmay4494 surejadharmay4494    4096 Nov 12 18:30 ds-test
-rw-r--r--     1 surejadharmay4494 surejadharmay4494      30 Nov 12 18:09 ds.txt
drwxr-xr-x     2 surejadharmay4494 surejadharmay4494    4096 Nov 12 18:12 .oracle_j
re_usage
-rw-rw-r--     1 surejadharmay4494 surejadharmay4494      13 Nov 12 18:40 test1.txt
-rw-rw-r--     1 surejadharmay4494 surejadharmay4494      11 Nov 12 18:40 test2.txt
-rw-------     1 surejadharmay4494 surejadharmay4494     596 Nov 12 18:09 .viminfo
[surejadharmay4494@cxln5 ~]$
```

<u>TASK</u> :- Get the list of all directories and files in the HDFS and Remove the file you created and directories also .

```
[surejadharmay4494@cxln5 ~]$ hdfs dfs -ls
Found 2 items
drwxr-xr-x   - surejadharmay4494 surejadharmay4494          0 2020-11-12 18:32 ds-t
est
-rw-r--r--   5 surejadharmay4494 surejadharmay4494         30 2020-11-12 18:12 ds.t
xt
[surejadharmay4494@cxln5 ~]$
```

```
[surejadharmay4494@cxln5 ~]$ hadoop fs -rm ds-test/*
20/11/12 18:48:04 INFO fs.TrashPolicyDefault: Moved: 'hdfs://cxln1.c.thelab-240901.interna
l:8020/user/surejadharmay4494/ds-test/filev.txt' to trash at: hdfs://cxln1.c.thelab-240901
.internal:8020/user/surejadharmay4494/.Trash/Current/user/surejadharmay4494/ds-test/filev.
txt
[surejadharmay4494@cxln5 ~]$ hadoop fs -ls ds-test/
[surejadharmay4494@cxln5 ~]$
```

```
[surejadharmay4494@cxln5 ~]$ hadoop fs -ls ds-test/
[surejadharmay4494@cxln5 ~]$ hadoop fs -rmdir ds-test/
h[surejadharmay4494@cxln5 ~]$ hadoop fs -rm testv.txt
rm: `testv.txt': No such file or directory
[surejadharmay4494@cxln5 ~]$ hadoop fs -rm filev.txt
rm: `filev.txt': No such file or directory
[surejadharmay4494@cxln5 ~]$ hadoop fs -rm test1.txt
rm: `test1.txt': No such file or directory
[surejadharmay4494@cxln5 ~]$ hdfs dfs -ls
Found 2 items
drwx------   - surejadharmay4494 surejadharmay4494          0 2020-11-12 18:48 .Trash
-rw-r--r--   5 surejadharmay4494 surejadharmay4494         30 2020-11-12 18:12 ds.txt
[surejadharmay4494@cxln5 ~]$ hadoop fs -rm ds.txt
20/11/12 18:50:27 INFO fs.TrashPolicyDefault: Moved: 'hdfs://cxln1.c.thelab-240901.interna
l:8020/user/surejadharmay4494/ds.txt' to trash at: hdfs://cxln1.c.thelab-240901.internal:8
020/user/surejadharmay4494/.Trash/Current/user/surejadharmay4494/ds.txt
[surejadharmay4494@cxln5 ~]$ hfs dfs -ls
-bash: hfs: command not found
[surejadharmay4494@cxln5 ~]$ hdfs dfs -ls
Found 1 items
drwx------   - surejadharmay4494 surejadharmay4494          0 2020-11-12 18:48 .Trash
[surejadharmay4494@cxln5 ~]$
```