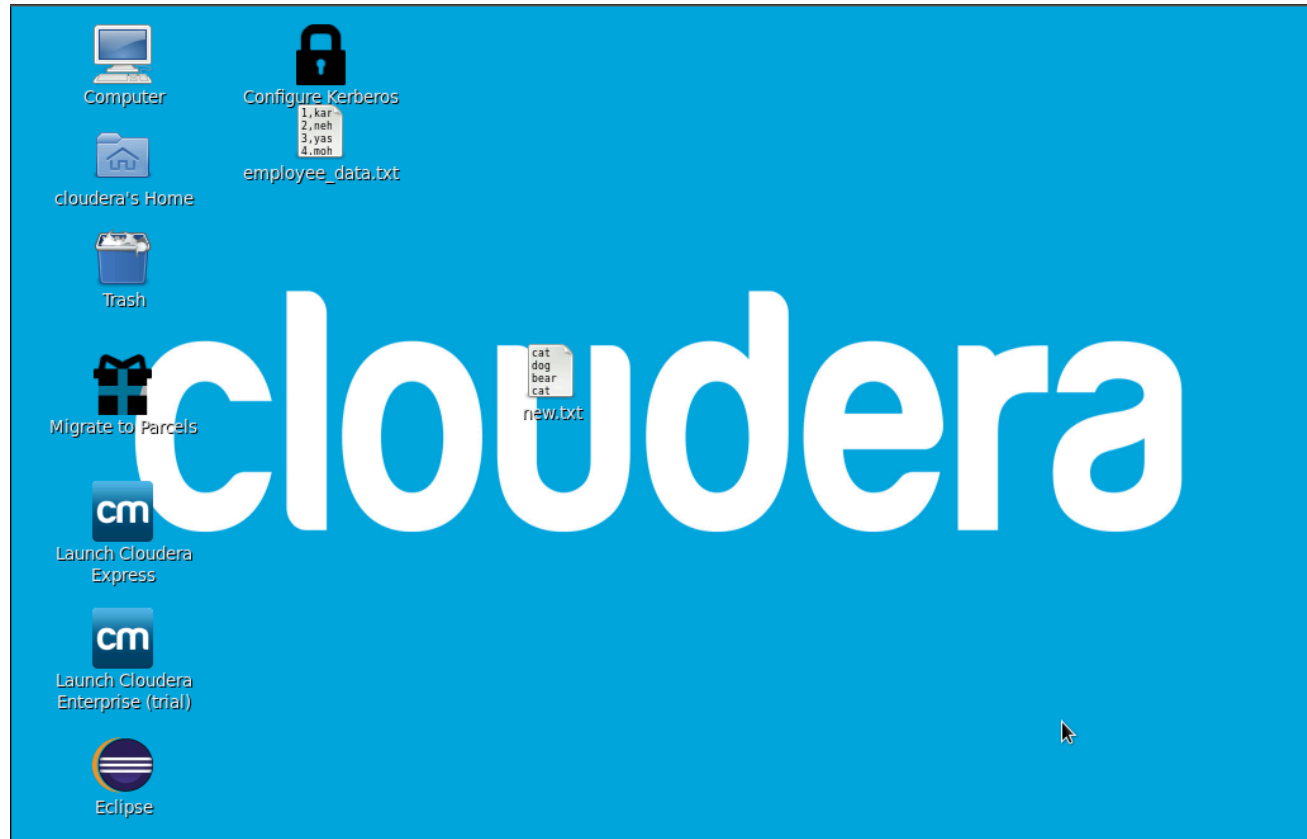
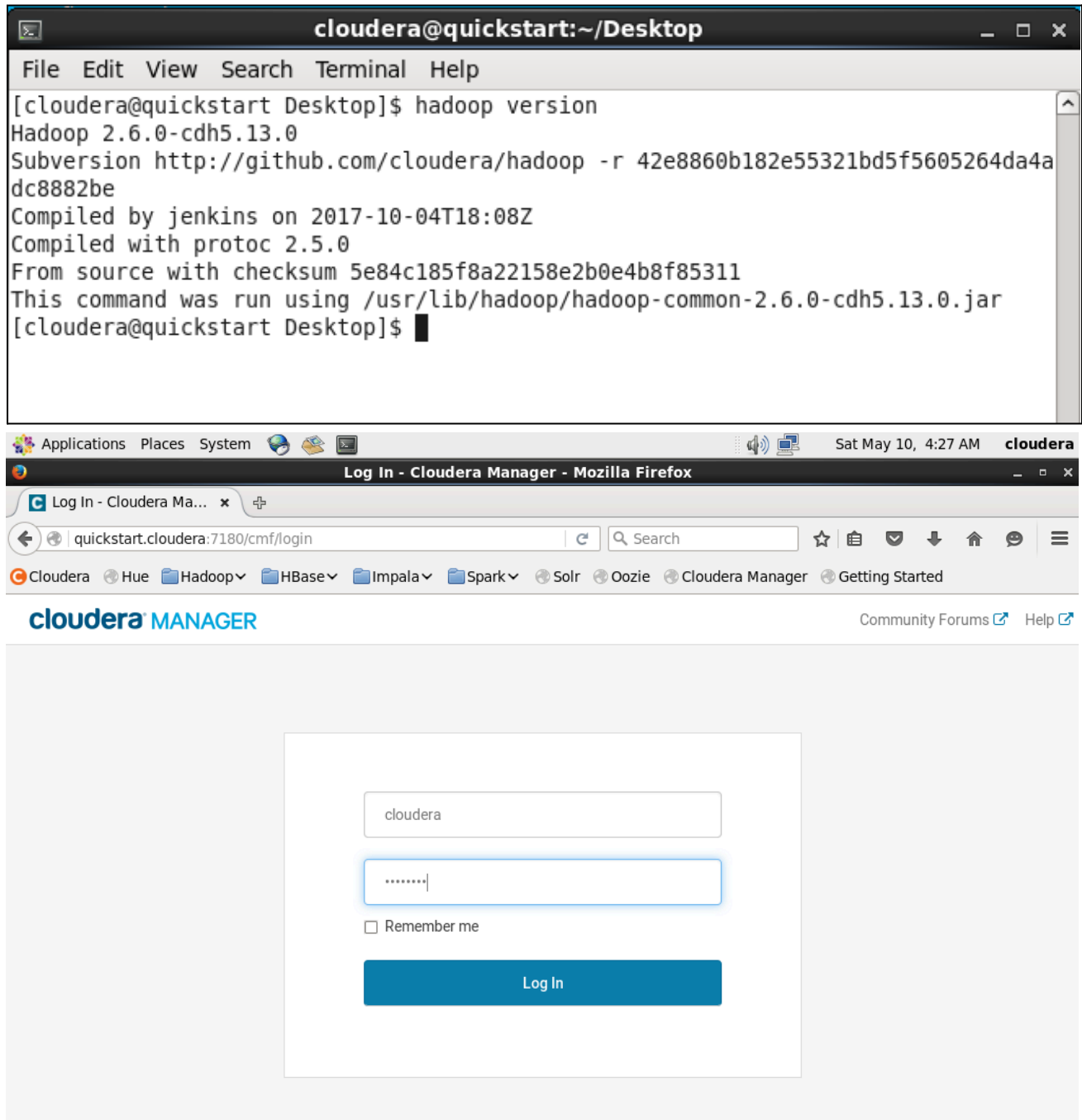


Program No:	4
Roll No :	1545
Title of Program :	Introduction to HDFS
Objective :	Understand HDFS architecture and illustrate HDFS

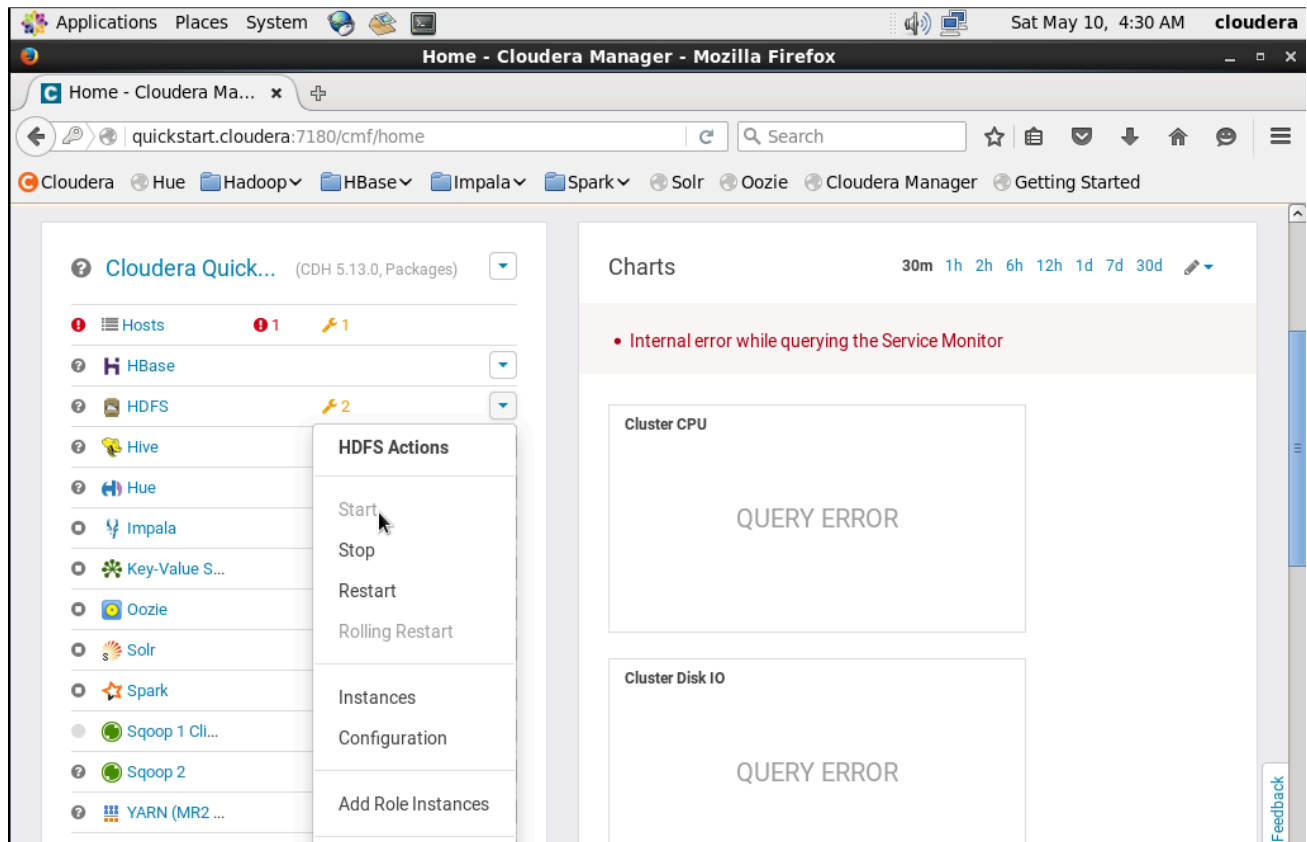




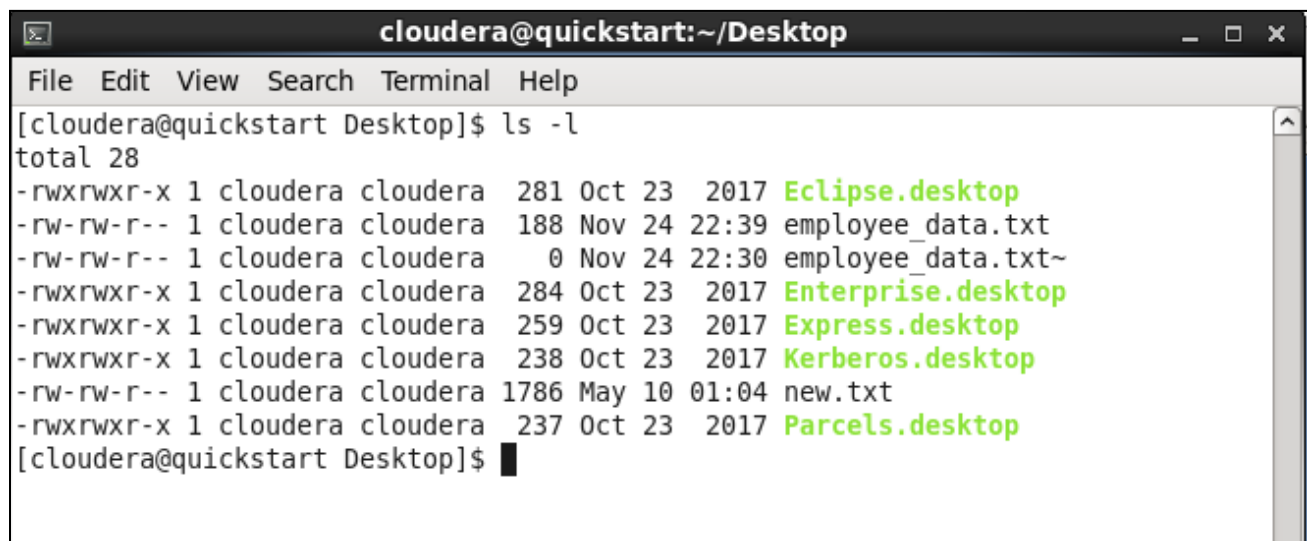
The image shows a desktop environment with a terminal window and a web browser. The terminal window, titled "cloudera@quickstart:~/Desktop", displays the output of the "hadoop version" command:

```
[cloudera@quickstart Desktop]$ hadoop version
Hadoop 2.6.0-cdh5.13.0
Subversion http://github.com/cloudera/hadoop -r 42e8860b182e55321bd5f5605264da4a
dc8882be
Compiled by jenkins on 2017-10-04T18:08Z
Compiled with protoc 2.5.0
From source with checksum 5e84c185f8a22158e2b0e4b8f85311
This command was run using /usr/lib/hadoop/hadoop-common-2.6.0-cdh5.13.0.jar
[cloudera@quickstart Desktop]$
```

The web browser, Mozilla Firefox, shows the "Log In - Cloudera Manager" page. The address bar displays "quickstart.cloudera:7180/cm/login". The page features a login form with a username field containing "cloudera", a password field with masked characters, a "Remember me" checkbox, and a "Log In" button. The Cloudera Manager logo and navigation links are visible at the top of the page.



The screenshot shows the Cloudera Manager web interface in a Mozilla Firefox browser. The address bar shows the URL `quickstart.cloudera:7180/cmf/home`. The interface includes a top navigation bar with links to Cloudera, Hue, Hadoop, HBase, Impala, Spark, Solr, Oozie, Cloudera Manager, and Getting Started. On the left, there is a sidebar with a search bar and a list of services: Hosts, HBase, HDFS, Hive, Hue, Impala, Key-Value S..., Oozie, Solr, Spark, Sqoop 1 Cli..., Sqoop 2, and YARN (MR2 ...). A context menu for HDFS is open, showing actions: Start, Stop, Restart, Rolling Restart, Instances, Configuration, and Add Role Instances. The main content area displays charts for Cluster CPU and Cluster Disk IO, both showing a "QUERY ERROR". A red alert message at the top of the charts section reads: "Internal error while querying the Service Monitor".



The screenshot shows a terminal window titled `cloudera@quickstart:~/Desktop`. The terminal output shows the result of the `ls -l` command, listing files in the current directory:

```
[cloudera@quickstart Desktop]$ ls -l
total 28
-rwxrwxr-x 1 cloudera cloudera 281 Oct 23 2017 Eclipse.desktop
-rw-rw-r-- 1 cloudera cloudera 188 Nov 24 22:39 employee_data.txt
-rw-rw-r-- 1 cloudera cloudera 0 Nov 24 22:30 employee_data.txt~
-rwxrwxr-x 1 cloudera cloudera 284 Oct 23 2017 Enterprise.desktop
-rwxrwxr-x 1 cloudera cloudera 259 Oct 23 2017 Express.desktop
-rwxrwxr-x 1 cloudera cloudera 238 Oct 23 2017 Kerberos.desktop
-rw-rw-r-- 1 cloudera cloudera 1786 May 10 01:04 new.txt
-rwxrwxr-x 1 cloudera cloudera 237 Oct 23 2017 Parcels.desktop
[cloudera@quickstart Desktop]$
```

```
cloudera@quickstart:~/mydata
File Edit View Search Terminal Help
[cloudera@quickstart ~]$ mkdir mydata
[cloudera@quickstart ~]$ cd mydata
[cloudera@quickstart mydata]$ vi input.txt
[cloudera@quickstart mydata]$ ls -l
total 4
-rw-rw-r-- 1 cloudera cloudera 90 May 10 04:46 input.txt
[cloudera@quickstart mydata]$ cat input.txt
deer bear bear car
bear car river
river river car
bear deer car river
deer car river bear
[cloudera@quickstart mydata]$
```

```
cloudera@quickstart:~/mydata
File Edit View Search Terminal Help
[cloudera@quickstart mydata]$ hdfs dfs -ls
Found 5 items
drwx----- - cloudera cloudera      0 2025-05-10 01:52 .staging
-rw-r--r-- 1 cloudera cloudera      188 2024-11-24 23:04 employee_data.txt
-rw-r--r-- 1 cloudera cloudera    1557 2025-05-10 01:35 input.txt
drwxr-xr-x - cloudera cloudera      0 2025-05-10 00:51 mydir
drwxr-xr-x - cloudera cloudera      0 2025-05-10 01:36 wordcount
[cloudera@quickstart mydata]$
```

make directory

```
[cloudera@quickstart mydata]$ hdfs dfs -mkdir shivam
[cloudera@quickstart mydata]$ hdfs dfs -ls
Found 6 items
drwx----- - cloudera cloudera      0 2025-05-10 01:52 .staging
-rw-r--r-- 1 cloudera cloudera      188 2024-11-24 23:04 employee_data.txt
-rw-r--r-- 1 cloudera cloudera    1557 2025-05-10 01:35 input.txt
drwxr-xr-x - cloudera cloudera      0 2025-05-10 00:51 mydir
drwxr-xr-x - cloudera cloudera      0 2025-05-10 04:53 shivam
drwxr-xr-x - cloudera cloudera      0 2025-05-10 01:36 wordcount
[cloudera@quickstart mydata]$
```

list files and directory in hdfs

```
[cloudera@quickstart mydata]$ hdfs dfs -ls shivam/
Found 1 items
-rw-r--r--  1 cloudera cloudera           0 2025-05-10 05:07 shivam/f1.txt
[cloudera@quickstart mydata]$ █
```

copy file from local to hdfs

```
[cloudera@quickstart mydata]$ hdfs dfs -copyFromLocal input.txt /user/cloudera/shivam
[cloudera@quickstart mydata]$ hdfs dfs -ls
Found 6 items
drwx-----  - cloudera cloudera           0 2025-05-10 01:52 .staging
-rw-r--r--  1 cloudera cloudera          188 2024-11-24 23:04 employee_data.txt
-rw-r--r--  1 cloudera cloudera       1557 2025-05-10 01:35 input.txt
drwxr-xr-x  - cloudera cloudera           0 2025-05-10 00:51 mydir
drwxr-xr-x  - cloudera cloudera           0 2025-05-10 05:15 shivam
drwxr-xr-x  - cloudera cloudera           0 2025-05-10 01:36 wordcount
[cloudera@quickstart mydata]$ █
```

```
[cloudera@quickstart mydata]$ cat > a.txt
hello world
^C
[cloudera@quickstart mydata]$ cat a.txt
hello world
[cloudera@quickstart mydata]$ █
```

copy file from local to hdfs

```
[cloudera@quickstart mydata]$ hdfs dfs -put a.txt /user/cloudera/shivam
[cloudera@quickstart mydata]$ hdfs dfs -ls shivam
Found 3 items
-rw-r--r--  1 cloudera cloudera          12 2025-05-10 05:19 shivam/a.txt
-rw-r--r--  1 cloudera cloudera           0 2025-05-10 05:07 shivam/f1.txt
-rw-r--r--  1 cloudera cloudera          90 2025-05-10 05:15 shivam/input.txt
[cloudera@quickstart mydata]$ █
```

move file from one directory to another in hdfs

```
[cloudera@quickstart ~]$ hdfs dfs -ls shivam
Found 3 items
-rw-r--r--  1 cloudera cloudera          12 2025-05-10 05:19 shivam/a.txt
-rw-r--r--  1 cloudera cloudera           0 2025-05-10 05:07 shivam/f1.txt
-rw-r--r--  1 cloudera cloudera          90 2025-05-10 05:15 shivam/input.txt
[cloudera@quickstart ~]$ hdfs dfs -mv /user/cloudera/shivam/a.txt /user/cloudera/shivam2
[cloudera@quickstart ~]$ hdfs dfs -ls shivam2
Found 1 items
-rw-r--r--  1 cloudera cloudera          12 2025-05-10 05:19 shivam2/a.txt
[cloudera@quickstart ~]$ █
```

copy all files from one directory to another in hdfs

```
[cloudera@quickstart ~]$ hdfs dfs -ls shivam
Found 2 items
-rw-r--r--  1 cloudera cloudera          0 2025-05-10 05:07 shivam/f1.txt
-rw-r--r--  1 cloudera cloudera        90 2025-05-10 05:15 shivam/input.txt
[cloudera@quickstart ~]$ hdfs dfs -cp /user/cloudera/shivam /user/cloudera/shivam2
[cloudera@quickstart ~]$ hdfs dfs -ls shivam2
Found 2 items
-rw-r--r--  1 cloudera cloudera        12 2025-05-10 05:19 shivam2/a.txt
drwxr-xr-x  - cloudera cloudera          0 2025-05-10 05:35 shivam2/shivam
[cloudera@quickstart ~]$ hdfs dfs -ls shivam2/shivam
Found 2 items
-rw-r--r--  1 cloudera cloudera          0 2025-05-10 05:35 shivam2/shivam/f1.txt
-rw-r--r--  1 cloudera cloudera        90 2025-05-10 05:35 shivam2/shivam/input.txt
[cloudera@quickstart ~]$ █
```

du - give size of each file

dus - give total size of directory

```
[cloudera@quickstart ~]$ hdfs dfs -du /user/cloudera/shivam
0  0  /user/cloudera/shivam/f1.txt
90 90 /user/cloudera/shivam/input.txt
[cloudera@quickstart ~]$ hdfs dfs -dus /user/cloudera/shivam
dus: DEPRECATED: Please use 'du -s' instead.
90 90 /user/cloudera/shivam
```

remove file

```
[cloudera@quickstart ~]$ hdfs dfs -rmr /user/cloudera/shivam/f1.txt
rmr: DEPRECATED: Please use 'rm -r' instead.
25/05/10 05:41:23 INFO fs.TrashPolicyDefault: Moved: 'hdfs://quickstart.cloudera:8020/user/cloudera/shivam/f1.txt' to trash at: hdfs://quickstart.cloudera:8020/user/cloudera/.Trash/Current/user/cloudera/shivam/f1.txt
```