

Aim - Demonstrate Hadoop Distributed File system

Theory -

Hadoop is a free, Java-based programming framework that supports the processing of large data sets in a distributed computing environment

Hadoop makes it possible to run application on system with thousands of nodes involving thousands of terabytes

Its distributed File system : Facilitate rapid data transfer rates among nodes & allows the system to continue operating uninterrupted in a case of a node failure

Hadoop was inspired by Google's Mapreduce, a software framework in which an application is broken down into numerous small parts. Any of these parts can be on any node in the cluster

Hadoop is used by

- Google
- Yahoo
- IBM

Preferred OS

- Windows
- Linux

Teacher's Sign.: _____

An imp part about Hadoop is software library. There is large library of programs that complement the base Hadoop framework & give companies the specific tools they need.

Hadoop is an open-source framework that allows to store and process big data in a distributed environment across clusters of computers using simple programming model.

Conclusion -

We understood the concepts of Hadoop & implemented all basic Hadoop commands successfully.

Teacher's Sign.: _____

1. List

```
[cloudera@quickstart ~]$ hadoop fs -ls
Found 6 items
drwxr-xr-x  - cloudera cloudera      0 2019-07-29 22:25 AMT
drwxr-xr-x  - cloudera cloudera      0 2019-08-08 03:13 Student
-rw-r--r--  1 cloudera cloudera     22 2019-07-29 22:41 file2.txt
-rw-r--r--  1 cloudera cloudera     14 2019-08-07 03:17 my.txt
drwxr-xr-x  - cloudera cloudera      0 2019-08-07 03:08 practice
drwxr-xr-x  - cloudera cloudera      0 2019-08-07 03:02 vesit
```

2. directory

```
[cloudera@quickstart ~]$ hadoop fs -mkdir richa
```

```
[cloudera@quickstart ~]$ hadoop fs -ls
Found 7 items
drwxr-xr-x  - cloudera cloudera      0 2019-07-29 22:25 AMT
drwxr-xr-x  - cloudera cloudera      0 2019-08-08 03:13 Student
-rw-r--r--  1 cloudera cloudera     22 2019-07-29 22:41 file2.txt
-rw-r--r--  1 cloudera cloudera     14 2019-08-07 03:17 my.txt
drwxr-xr-x  - cloudera cloudera      0 2019-08-07 03:08 practice
drwxr-xr-x  - cloudera cloudera      0 2019-08-12 21:44 richa
drwxr-xr-x  - cloudera cloudera      0 2019-08-07 03:02 vesit
```

3. Document

```
[cloudera@quickstart ~]$ cat -> richa_doc27.txt
RICHA
BHATIA
PIZZA
ICECREAM
```

4. Put

```
[cloudera@quickstart ~]$ hadoop fs -put richa doc27.txt richa
```

```
[cloudera@quickstart ~]$ hadoop fs -ls richa
Found 2 items
-rw-r--r--  1 cloudera cloudera     28 2019-08-12 21:44 richa/richa_doc.txt
-rw-r--r--  1 cloudera cloudera     28 2019-08-12 21:49 richa/richa_doc27.t
xt
```

```
[cloudera@quickstart ~]$ hadoop fs -cat richa/richa_doc27.txt
RICHA
BHATIA
PIZZA
ICECREAM
```

5. cp command

```
[cloudera@quickstart ~]$ hadoop fs -mkdir richacp
```

```
[cloudera@quickstart ~]$ hadoop fs -cp richa/richa_doc27.txt richacp/richa_doc27
.txt
```

```
[cloudera@quickstart ~]$ hadoop fs -ls richacp
Found 1 items
-rw-r--r--  1 cloudera cloudera          28 2019-08-12 21:52 richacp/richa_doc27
.txt
[cloudera@quickstart ~]$ hadoop fs -cat richacp/richa_doc27.txt
RICHIA
BHATIA
PIZZA
ICECREAM
```

6. get command

```
[cloudera@quickstart ~]$ hadoop fs -get richacp/richa_doc27.txt
```

```
[cloudera@quickstart ~]$ ls
abc.txt                kerberos              richa_doc.txt
a.txt                  lib                   Student
cloudera-manager      MatrixVectorMultiplication  Templates
cm_api.py              mine.txt              Videos
D17B65                mine.txt~            wc.jar
Desktop                Music                 WC_Mapper.class
doc.txt                my.txt~              WC_Mapper.java
Documents              OutputDir             WC_Reducer.class
Downloads              parcels               WC_Reducer.java
eclipse                Pictures              WC_Runner.java
enterprise-deployment.json  Public               WC_Runner.java~
express-deployment.json  richa                 wordcount
file1.txt              richa_doc27.txt       workspace
```

7. rm command

```
[cloudera@quickstart ~]$ hadoop fs -rm richa/richa_doc27.txt
Deleted richa/richa_doc27.txt
[cloudera@quickstart ~]$ hadoop fs -ls richa
Found 1 items
-rw-r--r--  1 cloudera cloudera          28 2019-08-12 21:44 richa/richa_doc.txt
[cloudera@quickstart ~]$
```