Introduction to Big Data

ECAP456

Dr. Rajni Bhalla

Associate Professor

Learning Outcomes

After this lecture, you will be able to

- learn basic commands in HDFS

Commands in hdfs

- jps
- This command is a part of Java since v1.5.0.

  jps stands for Java Virtual Machine Process Status Tool.
- In our practice, when we start the single node cluster following processes are must be up and running:
  - Name Node

    Data Node

    Resource Manager

    Node Manager

Commands in hdfs

- jps is a tool to check, whether expected Hadoop processes are up and in running state or not.

Commands in hdfs

- The syntax is

```
jps
```

Commands in hdfs

- ls
- Using the ls command, we can check for the directories in HDFS.
- List the HDFS contents

Hadoop Prefix

Every command in Hadoop have prefix:

hadoop fs

Or

hdfs dfs

Basic Commands

List the contents that are in Hadoop

```
hadoop fs -ls
```

Commands in hdfs

- Display the contents of directory
- Syntax

```
$hadoop fs –ls directoryname
```

Commands in hdfs

- Create a directory in hdfs
- Syntax

```
hadoop fs –mkdir abc
```

Commands in hdfs

- Verify a directory in hdfs
- Syntax

```
hadoop fs –ls
```

Commands in hdfs

- Every command start with (-)hyphen in hdfs.
- Syntax to create a file

$hadoop fs –touchz file1.txt

This command will create an empty file in home location.

Create a file of zero length.

Commands in hdfs

- Copy from local file system into hdfs directory
- Firstly create a simple file.

gedit f1.txt

$hadoop fs –copyFromLocal f1.txt file1

Or

put(Copy single src, or multiple srcs from local file system to the destination file system).

Verify
$hadoop fs -ls

Commands in hdfs

- Create a file
- To check either f1.txt is available or not. To check the contents of the file.

$hadoop fs –cat abc/f1.txt

Commands in hdfs

- Sending a file from hdfs to local file system
- Syntax

Hadoop fs –copyToLocal abc/f1.txt ~/desktop/

Verify either it is available in desktop location or not:

cd Desktop/

ls

Commands in hdfs

Verify either it's the same file

cat f1.txt

Commands in hdfs

- Copy Command
- Sending file from hdfs to hdfs directory
- Syntax is

hadoop fs –mkdir abc1

Hadoop fs –cp abc/f1.txt abc1/

–cp abc/f1.txt - Source

abc1/ - Destination

Commands in hdfs

Verify

$hadoop –fs –cat abc1/f1.txt

Commands in hdfs

- Move the directory from one hdfs to other hdfs
- Firstly create another directory with abc2

hadoop fs –mkdir abc2

Hadoop fs –mv abc/f1.txt abc2/

Commands in hdfs

Check the content with cat command

hadoop fs –cat abc2/f1.txt

Commands in hdfs

Which directory is taking more space?

hadoop fs –du abc1

-du	-dus
Display size of each and every file	Display total size


That's all for now...