Overview of Hadoop HDFS Commands

Hadoop HDFS Commands: Explanation and Examples

HDFS commands are used to interact with the Hadoop distributed file system from the command line. Commands usually start with:

```
hadoop fs -<command>
```

or

```
hdfs dfs -<command>
```

1. Checking Hadoop Version

- Purpose: Verifies Hadoop installation and version.
- Command:

```
hadoop version
```

- Example:

```
Hadoop 3.3.1
Subversion https://...
Compiled by user on 2021-01-15
```

2. Create Directory (`mkdir`)

- Purpose: Create a new directory in HDFS.
- Syntax:

```
hadoop fs -mkdir /directory_path
```

- Example:

```
hadoop fs -mkdir /user/mydata
```

- Illustration: Creates `/user/mydata` directory in HDFS.

3. List Files/Directories (`ls`)

- Purpose: Lists files and directories inside a path.
- Syntax:

```
hadoop fs -ls /path
```

- Example:

```
hadoop fs -ls /user
```

- Output:

```
Found 2 items
drwxr-xr-x   - user supergroup          0 2025-08-01
12:30 /user/mydata
```

```
-rw-r--r--   3 user supergroup       1024 2025-08-01
10:00 /user/file1.txt
```

4. Copy File from Local to HDFS (`put`)

- Purpose: Upload files from local filesystem to HDFS.
- Syntax:

```
hadoop fs -put /local/path/file /hdfs/path/
```

- Example:

```
hadoop fs -put sample.txt /user/mydata/
```

5. Copy File from HDFS to Local (`get`)

- Purpose: Download files from HDFS to local file system.
- Syntax:

```
hadoop fs -get /hdfs/file /local/destination
```

- Example:

```
hadoop fs -get /user/mydata/sample.txt ./
```

6. Display File Contents (`cat`)

- Purpose: Prints content of a file on the command line.
- Syntax:

```
hadoop fs -cat /hdfs/path/file
```

- Example:

```
hadoop fs -cat /user/mydata/sample.txt
```

- Output:

```
Content of sample.txt appears here
```

7. Move or Rename (`mv`)

- Purpose: Move or rename files/directories in HDFS.
- Syntax:

```
hadoop fs -mv /source/path /destination/path
```

-
- Example:

```
hadoop fs -mv /user/mydata/sample.txt
/user/mydata/sample_old.txt
```

8. Copy Files Inside HDFS (`cp`)

- Purpose: Copies files/directories inside HDFS.
- Syntax:

```
hadoop fs -cp /source/path /destination/path
```

- Example:

```
hadoop fs -cp /user/mydata/sample_old.txt
/user/mydata/backup/sample.txt
```

9. Remove File or Directory (`rm`)

- Purpose: Deletes file or directory from HDFS. Use `-r` for recursive delete.
- Syntax:

```
hadoop fs -rm /hdfs/file
hadoop fs -rm -r /hdfs/directory
```

- Example:

```
hadoop fs -rm /user/mydata/sample_old.txt

hadoop fs -rm -r /user/mydata/backup
```

10. Check Disk Usage (`du`)

- Purpose: Displays size of a file or directory in HDFS.
- Syntax:

```
hadoop fs -du /path
```

- Example:

```
hadoop fs -du /user/mydata
```

- Output:

```
1024  /user/mydata/sample.txt
```

Quick Reference Table

Command	Purpose	Sample Command
`version`	Check Hadoop version	`hadoop version`
`mkdir`	Create directory	`hadoop fs -mkdir /user/mydata`
`ls`	List files/directories	`hadoop fs -ls /user`
`put`	Upload file (local -> HDFS)	`hadoop fs -put local.txt /user/mydata/`
`get`	Download file (HDFS -> local)	`hadoop fs -get /user/mydata/local.txt ./`
`cat`	Display file contents	`hadoop fs -cat /user/mydata/file.txt`
`mv`	Move or rename files/directories	`hadoop fs -mv /old/path /new/path`

`cp`	Copy files/directories	`hadoop fs -cp /source/path /dest/path`
`rm`	Remove files/directories	`hadoop fs -rm file` or `hadoop fs -rm -r dir`
`du`	Show disk space usage	`hadoop fs -du /user/mydata`

Illustration: Example Use Case

1. Create directory:

   ```
   hadoop fs -mkdir /user/anna
   ```

2. Upload file:

   ```
   hadoop fs -put /home/anna/data.csv /user/anna/
   ```

3. List files:

   ```
   hadoop fs -ls /user/anna
   ```

4. View file content:

   ```
   hadoop fs -cat /user/anna/data.csv
   ```

5. Copy file:

```
hadoop fs -cp /user/anna/data.csv
/user/anna/data_backup.csv
```

6. Delete a file:

```
hadoop fs -rm /user/anna/data_backup.csv
```

7.

These commands provide the foundational HDFS operations for daily Hadoop cluster work.