

```
Administrator Command Prompt
Microsoft Windows [Version 10.0.14393]
(c) 2016 Microsoft Corporation. All rights reserved.

C:\WINDOWS\system32>cd ..
C:\Windows>cd ..
C:\>cd Hadoop-2.8.0\sbin
C:\Hadoop-2.8.0\sbin>start-all.cmd
This script is deprecated. Instead use start-dfs.cmd and start-yarn.cmd
starting yarn daemons
C:\Hadoop-2.8.0\sbin>

Apache Hadoop Distribution - hadoop: namenode
17/07/17/07
17/07/17/07 Apache Hadoop Distribution - hadoop: datanode
8:01:17/07
17/07/17/07 Apache Hadoop Distribution - yarn: resourcemanager
17/07/17/07
For 17/07/17/07
17/07/17/07 Apache Hadoop Distribution - yarn: nodemanager
17/07/17/07
17/07/17/07 Jul 20 17/07/20 17:54:28 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
-ea4117/07/Jul 20 17/07/20 17:54:31 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
17/07/hadoopINFO: 17/07/20 17:54:34 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
The n17/07/Jul 20 17/07/20 17:54:34 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
he m185417/07/Jul 20 17/07/20 17:54:37 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
17/07/17/07/Jul 20 17/07/20 17:54:40 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
17/07/20 17/07/20 17:54:43 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
17/07/20 17/07/20 17:54:43 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
7-754117/07/with 17/07/20 17:54:46 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
075, 168.6817/07/20 17:54:49 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
908417/07/INFO: 17/07/20 17:54:52 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
17/07/16663217/07/20 17:54:55 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
17/07/17/07/Jul 20 17/07/20 17:54:58 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
17/07/17/07/Jul 20 17/07/20 17:55:01 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
17/07/17/07/INFO: 17/07/20 17:55:05 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
17/07/54d07/07/20 17:55:08 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
17/07/17/07/20 17:55:11 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
17/07/with 17/07/20 17:55:14 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
17/07/17/07/20 17:55:17 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
17/07/67-adey: 1617/07/20 17:55:17 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
The n17/07/17/07/20 17:55:20 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
he m1838ae17/07/20 17:55:23 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
17/07/17/07/17/07/20 17:55:26 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
17/07/67/CAC17/07/20 17:55:29 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
17/07/67/20 17:55:32 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
17/07/17/07/20 17:55:35 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
17/07/17/07/20 17:55:38 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
or RPTCT17/07/20 17:55:41 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
17/07/17/07/20 17:55:44 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
17/07/17/07/20 17:55:47 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
cores: 17/07/20 17:55:50 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
17/07/20 17:55:54 WARN util.SysInfoFskIndows: Expected split length of sysinfo to be 11. Got 7
```

1. Create an input directory in HDFS.

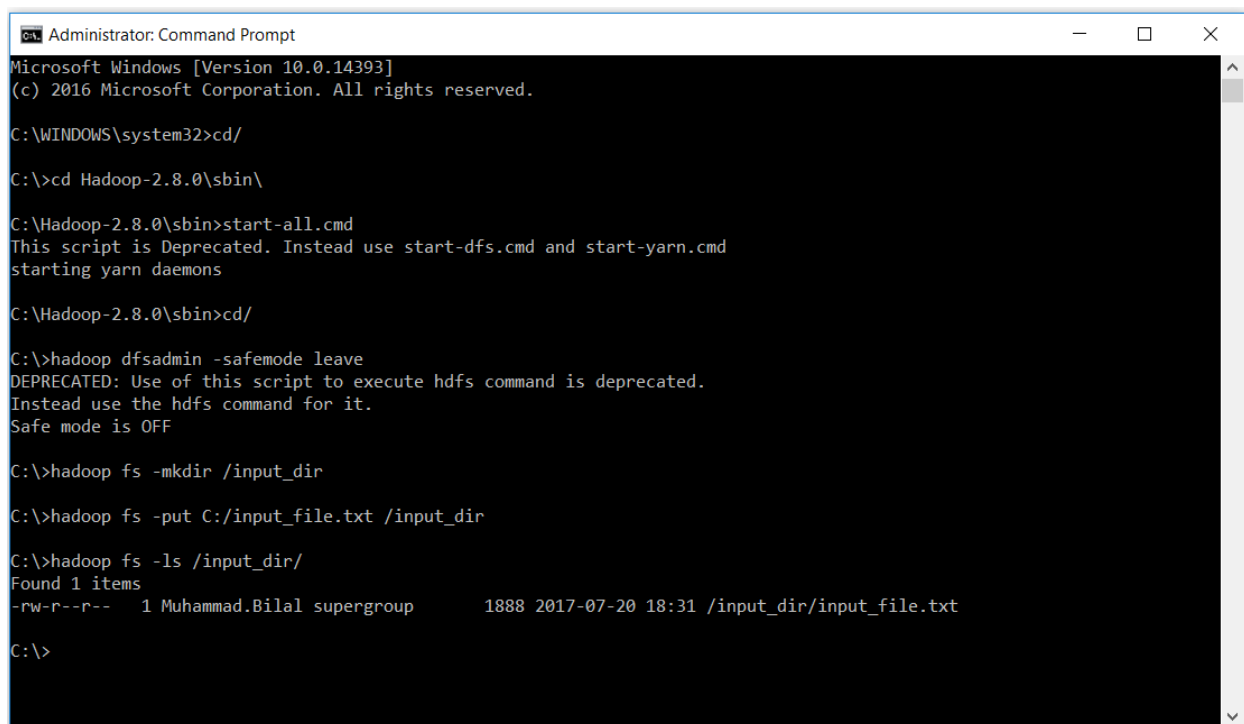
```
hadoop fs -mkdir /input_dir
```

1. Copy the input text file named input\_file.txt in the input directory (input\_dir) of HDFS.

```
hadoop fs -put C:/input_file.txt /input_dir
```

1. Verify input\_file.txt available in HDFS input directory (input\_dir).

```
hadoop fs -ls /input_dir/
```



```
Administrator: Command Prompt
Microsoft Windows [Version 10.0.14393]
(c) 2016 Microsoft Corporation. All rights reserved.

C:\WINDOWS\system32>cd/

C:\>cd Hadoop-2.8.0\sbin\

C:\Hadoop-2.8.0\sbin>start-all.cmd
This script is Deprecated. Instead use start-dfs.cmd and start-yarn.cmd
starting yarn daemons

C:\Hadoop-2.8.0\sbin>cd/

C:\>hadoop dfsadmin -safemode leave
DEPRECATED: Use of this script to execute hdfs command is deprecated.
Instead use the hdfs command for it.
Safe mode is OFF

C:\>hadoop fs -mkdir /input_dir

C:\>hadoop fs -put C:/input_file.txt /input_dir

C:\>hadoop fs -ls /input_dir/
Found 1 items
-rw-r--r--  1 Muhammad.Bilal supergroup    1888 2017-07-20 18:31 /input_dir/input_file.txt

C:\>
```

1. Verify content of the copied file.

```
hadoop dfs -cat /input_dir/input_file.txt
```

```
C:\>hadoop fs -ls /input_dir/
Found 1 items
-rw-r--r-- 1 Muhammad.Bilal supergroup 1888 2017-07-20 18:31 /input_dir/input_file.txt

C:\>hadoop dfs -cat /input_dir/input_file.txt
DEPRECATED: Use of this script to execute hdfs command is deprecated.
Instead use the hdfs command for it.
23 23 27 43 24 25 26 26 26 26 25 26 25
26 27 28 28 28 30 31 31 31 30 30 30 29
31 32 32 32 33 34 35 36 36 34 34 34 34
39 38 39 39 39 41 42 43 40 39 38 38 40
38 39 39 39 39 41 41 41 28 40 39 39 45
23 23 27 43 24 25 26 26 26 26 25 26 25
26 27 28 28 28 30 31 31 31 30 30 30 29
31 32 32 32 33 34 35 36 36 34 34 34 34
39 38 39 39 39 41 42 43 40 39 38 38 40
38 39 39 39 39 41 41 41 28 40 39 39 45
23 23 27 43 24 25 26 26 26 26 25 26 25
26 27 28 28 28 30 31 31 31 30 30 30 29
31 32 32 32 33 34 35 36 36 34 34 34 34
39 38 39 39 39 41 42 43 40 39 38 38 40
38 39 39 39 39 41 41 41 28 40 39 39 45
23 23 27 43 24 25 26 26 26 26 25 26 25
26 27 28 28 28 30 31 31 31 30 30 30 29
31 32 32 32 33 34 35 36 36 34 34 34 34
39 38 39 39 39 41 42 43 40 39 38 38 40
38 39 39 39 39 41 41 41 28 40 39 39 45
```

1. Run MapReduceClient.jar and also provide input and out directories.

```
hadoop jar C:/MapReduceClient.jar wordcount /input_dir /output_dir
```

```
Administrator: Command Prompt
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=1999
HDFS: Number of bytes written=120
HDFS: Number of read operations=6
HDFS: Number of large read operations=0
HDFS: Number of write operations=2

Job Counters
  Launched map tasks=1
  Launched reduce tasks=1
  Data-local map tasks=1
  Total time spent by all maps in occupied slots (ms)=2180
  Total time spent by all reduces in occupied slots (ms)=2442
  Total time spent by all map tasks (ms)=2180
  Total time spent by all reduce tasks (ms)=2442
  Total vcore-milliseconds taken by all map tasks=2180
  Total vcore-milliseconds taken by all reduce tasks=2442
  Total megabyte-milliseconds taken by all map tasks=2232320
  Total megabyte-milliseconds taken by all reduce tasks=2506080

Map-Reduce Framework
  Map input records=30
  Map output records=390
  Map output bytes=2730
  Map output materialized bytes=195
  Input split bytes=111
  Combine input records=390
  Combine output records=21
  Reduce input groups=21
  Reduce shuffle bytes=195
  Reduce input records=21
  Reduce output records=21
  Spilled Records=42
  Shuffled Maps =1
  Failed Shuffles=0
  Merged Map outputs=1
  GC time elapsed (ms)=70
  CPU time spent (ms)=764
  Physical memory (bytes) snapshot=471478272
  Virtual memory (bytes) snapshot=619420888
  Total committed heap usage (bytes)=353894400

Shuffle Errors
  BAD_ID=0
  CONNECTION=0
  IO_ERROR=0
  WRONG_LENGTH=0
  WRONG_MAP=0
  WRONG_REDUCE=0

File Input Format Counters
  Bytes Read=1888
File Output Format Counters
  Bytes Written=120

C:\>
```

## 1. Verify content for generated output file.

```
hadoop dfs -cat /output_dir/*
```

```
C:\>hadoop dfs -cat /output_dir/*
DEPRECATED: Use of this script to execute hdfs command is deprecated.
Instead use the hdfs command for it.
23      12
24      6
25      18
26      36
27      12
28      24
29      6
30      24
31      24
32      18
33      6
34      30
35      6
36      12
38      24
39      66
40      18
41      24
42      6
43      12
45      6

C:\>
```

## Some Other usefull commands

To leave Safe mode

```
hadoop dfsadmin -safemode leave
```

To Delete file from HDFS directory

```
hadoop fs -rm -r /iutput_dir/input_file.txt
```

To Delete directory from HDFS directory

```
hadoop fs -rm -r /iutput_dir
```

```
C:\>hadoop dfsadmin -safemode leave
DEPRECATED: Use of this script to execute hdfs command is deprecated.
Instead use the hdfs command for it.
Safe mode is OFF

C:\>hadoop fs -rm -r /input_dir/input_file.txt
Deleted /input_dir/input_file.txt

C:\>hadoop fs -rm -r /input_dir
Deleted /input_dir

C:\>
```

## Practice

### Part I

#### 1) Make directory

```
hadoop fs -mkdir /input_dir
```

#### 2) Assign our input file to hadoop for processing

```
hadoop fs -put c:\WordCount\input_file.txt /input_dir
```

#### 3) Execute the process in Hadoop

```
hadoop jar c:\WordCount\MapReduceClient.jar wordcount /input_dir /output_dir
```

#### 4) View the Results

```
hadoop dfs -cat /output_dir/*
```

### Part II

#### a) Remove a folder

```
hadoop fs -rmdir /input_dir
```

#### b) Remove a file

```
hadoop fs -rm /input_dir/*
```

c) Listing the contents of folder

```
hadoop fs -ls /input_dir
```

d) View the contents of file in Hadoop folder

```
hadoop dfs -cat /output_dir/*
```