

```
Administrator: Windows PowerShell
2022-11-01 14:29:25,093 INFO util.GSet: Computing capacity for map InodeMap
2022-11-01 14:29:25,093 INFO util.GSet: VM type = 64-bit
2022-11-01 14:29:25,096 INFO util.GSet: 1.0% max memory 889 MB = 8.9 MB
2022-11-01 14:29:25,098 INFO util.GSet: capacity = 2^20 = 1048576 entries
2022-11-01 14:29:25,088 INFO namenode.FSDirectory: ACLs enabled? false
2022-11-01 14:29:25,088 INFO namenode.FSDirectory: POSIX ACL inheritance enabled? true
2022-11-01 14:29:25,091 INFO namenode.FSDirectory: Xattrs enabled? true
2022-11-01 14:29:25,093 INFO namenode.NameNode: Caching file names occurring more than 10 times
2022-11-01 14:29:25,104 INFO snapshot.SnapshotManager: Loaded config captureOpenFiles: false, skipCaptureAccessTimeOnlyChange: false, snapshotDiffAllowSnapRootDescendant: true, maxSnapshotLimit: 65536
2022-11-01 14:29:25,109 INFO snapshot.SnapshotManager: SkipList is disabled
2022-11-01 14:29:25,116 INFO util.GSet: Computing capacity for map cachedBlocks
2022-11-01 14:29:25,116 INFO util.GSet: VM type = 64-bit
2022-11-01 14:29:25,118 INFO util.GSet: 0.25% max memory 889 MB = 2.2 MB
2022-11-01 14:29:25,119 INFO util.GSet: capacity = 2^18 = 262144 entries
2022-11-01 14:29:25,138 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.window.num.buckets = 10
2022-11-01 14:29:25,138 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.num.users = 10
2022-11-01 14:29:25,139 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.windows.minutes = 1,5,25
2022-11-01 14:29:25,148 INFO namenode.FSNamesystem: Retry cache on namenode is enabled
2022-11-01 14:29:25,149 INFO namenode.FSNamesystem: Retry cache will use 0.03 of total heap and retry cache entry expiry time is 600000 millis
2022-11-01 14:29:25,155 INFO util.GSet: Computing capacity for map NameNodeRetryCache
2022-11-01 14:29:25,155 INFO util.GSet: VM type = 64-bit
2022-11-01 14:29:25,157 INFO util.GSet: 0.02999999999929447746% max memory 889 MB = 273.1 KB
2022-11-01 14:29:25,157 INFO util.GSet: capacity = 2^15 = 32768 entries
Re-format filesystem in Storage Directory root= D:\hadoop-env\hadoop-3.2.2\data\dfs\namenode; location= null ? (Y or N) Y
2022-11-01 14:29:33,702 INFO namenode.FSImage: Allocated new BlockPoolId: BP-1378187979-192.168.56.1-1667293173677
2022-11-01 14:29:33,703 INFO common.Storage: Will remove files: [D:\hadoop-env\hadoop-3.2.2\data\dfs\namenode\current\fsimage_00000000000000000000.mds, D:\hadoop-env\hadoop-3.2.2\data\dfs\namenode\current\fsimage_00000000000000000000.mds, D:\hadoop-env\hadoop-3.2.2\data\dfs\namenode\current\VERSION]
2022-11-01 14:29:33,805 INFO common.Storage: Storage directory D:\hadoop-env\hadoop-3.2.2\data\dfs\namenode has been successfully formatted.
2022-11-01 14:29:33,870 INFO namenode.FSImageFormatProtobuf: Saving image file D:\hadoop-env\hadoop-3.2.2\data\dfs\namenode\current\fsimage.ckpt_00000000000000000000 using no compression
2022-11-01 14:29:34,029 INFO namenode.FSImageFormatProtobuf: Image file D:\hadoop-env\hadoop-3.2.2\data\dfs\namenode\current\fsimage.ckpt_00000000000000000000 of size 404 bytes saved in 0 second
2022-11-01 14:29:34,067 INFO namenode.NNStorageRetentionManager: Going to retain 1 images with txid >= 0
2022-11-01 14:29:34,077 INFO namenode.FSImage: FSImageSaver clean checkpoint: txid=0 when meet shutdown.
2022-11-01 14:29:34,077 INFO namenode.NameNode: SHUTDOWN_MSG:
/*****
SHUTDOWN_MSG: Shutting down NameNode at A41T44/192.168.56.1
*****/
PS D:\hadoop-env\hadoop-3.2.2\sbin> .\start-dfs.cmd
PS D:\hadoop-env\hadoop-3.2.2\sbin> .\start-yarn.cmd
starting yarn daemons
PS D:\hadoop-env\hadoop-3.2.2\sbin> jps
3252 NodeManager
5000 NameNode
5008 ResourceManager
8264 DataNode
6364 Jps
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -mkdir /user
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -put D:/data.txt /user
PS D:\hadoop-env\hadoop-3.2.2\sbin>
Activate Windows
Go to Settings to activate Windows.
```

```
Administrator: Windows PowerShell
2022-11-01 14:29:25,093 INFO namenode.NameNode: Caching file names occurring more than 10 times
2022-11-01 14:29:25,104 INFO snapshot.SnapshotManager: Loaded config captureOpenFiles: false, skipCaptureAccessTimeOnlyChange: false, snapshotDiffAllowSnapRootDescendant: true, maxSnapshotLimit: 65536
2022-11-01 14:29:25,109 INFO snapshot.SnapshotManager: SkipList is disabled
2022-11-01 14:29:25,116 INFO util.GSet: Computing capacity for map cachedBlocks
2022-11-01 14:29:25,116 INFO util.GSet: VM type = 64-bit
2022-11-01 14:29:25,118 INFO util.GSet: 0.25% max memory 889 MB = 2.2 MB
2022-11-01 14:29:25,119 INFO util.GSet: capacity = 2^18 = 262144 entries
2022-11-01 14:29:25,138 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.window.num.buckets = 10
2022-11-01 14:29:25,138 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.num.users = 10
2022-11-01 14:29:25,139 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.windows.minutes = 1,5,25
2022-11-01 14:29:25,148 INFO namenode.FSNamesystem: Retry cache on namenode is enabled
2022-11-01 14:29:25,149 INFO namenode.FSNamesystem: Retry cache will use 0.03 of total heap and retry cache entry expiry time is 600000 millis
2022-11-01 14:29:25,155 INFO util.GSet: Computing capacity for map NameNodeRetryCache
2022-11-01 14:29:25,155 INFO util.GSet: VM type = 64-bit
2022-11-01 14:29:25,157 INFO util.GSet: 0.02999999999929447746% max memory 889 MB = 273.1 KB
2022-11-01 14:29:25,157 INFO util.GSet: capacity = 2^15 = 32768 entries
Re-format filesystem in Storage Directory root= D:\hadoop-env\hadoop-3.2.2\data\dfs\namenode; location= null ? (Y or N) Y
2022-11-01 14:29:33,702 INFO namenode.FSImage: Allocated new BlockPoolId: BP-1378187979-192.168.56.1-1667293173677
2022-11-01 14:29:33,703 INFO common.Storage: Will remove files: [D:\hadoop-env\hadoop-3.2.2\data\dfs\namenode\current\fsimage_00000000000000000000.mds, D:\hadoop-env\hadoop-3.2.2\data\dfs\namenode\current\fsimage_00000000000000000000.mds, D:\hadoop-env\hadoop-3.2.2\data\dfs\namenode\current\VERSION]
2022-11-01 14:29:33,805 INFO common.Storage: Storage directory D:\hadoop-env\hadoop-3.2.2\data\dfs\namenode has been successfully formatted.
2022-11-01 14:29:33,870 INFO namenode.FSImageFormatProtobuf: Saving image file D:\hadoop-env\hadoop-3.2.2\data\dfs\namenode\current\fsimage.ckpt_00000000000000000000 using no compression
2022-11-01 14:29:34,029 INFO namenode.FSImageFormatProtobuf: Image file D:\hadoop-env\hadoop-3.2.2\data\dfs\namenode\current\fsimage.ckpt_00000000000000000000 of size 404 bytes saved in 0 second
2022-11-01 14:29:34,067 INFO namenode.NNStorageRetentionManager: Going to retain 1 images with txid >= 0
2022-11-01 14:29:34,077 INFO namenode.FSImage: FSImageSaver clean checkpoint: txid=0 when meet shutdown.
2022-11-01 14:29:34,077 INFO namenode.NameNode: SHUTDOWN_MSG:
/*****
SHUTDOWN_MSG: Shutting down NameNode at A41T44/192.168.56.1
*****/
PS D:\hadoop-env\hadoop-3.2.2\sbin> .\start-dfs.cmd
PS D:\hadoop-env\hadoop-3.2.2\sbin> .\start-yarn.cmd
starting yarn daemons
PS D:\hadoop-env\hadoop-3.2.2\sbin> jps
3252 NodeManager
5000 NameNode
5008 ResourceManager
8264 DataNode
6364 Jps
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -mkdir /user
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -put D:/data.txt /user
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -ls /user/
Found 1 items
-rw-r--r-- 1 19uite01 supergroup 43 2022-11-01 15:06 /user/data.txt
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -cat /user/data.txt
This is Venkatesh Prabhu from Final Year II
PS D:\hadoop-env\hadoop-3.2.2\sbin>
Activate Windows
Go to Settings to activate Windows.
```

```
Administrator: Windows PowerShell
PS D:\hadoop-env\hadoop-3.2.2\sbin> jps
5252 NodeManager
5600 NameNode
5608 ResourceManager
8264 DataNode
6364 Jps
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -mkdir /user
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -put D:\data.txt /user
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -ls /user/
Found 1 items
-rw-r--r- 1 19uite001 supergroup          43 2022-11-01 15:06 /user/data.txt
PS D:\hadoop-env\hadoop-3.2.2\sbin> hdfs dfs -cat /user/data.txt
This is Venkatesh Prabhu from Final Year IT
PS D:\hadoop-env\hadoop-3.2.2\sbin> hadoop jar D:\hadoop-env\hadoop-3.2.2\share\hadoop\mapreduce\hadoop-mapreduce-examples-3.2.2.jar wordcount /user /out
2022-11-01 15:09:42,761 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032
2022-11-01 15:09:44,099 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/19uite001/.staging/job_1667293411803_0001
2022-11-01 15:09:44,618 INFO input.FileInputFormat: Total input files to process : 1
2022-11-01 15:09:45,533 INFO mapreduce.JobSubmitter: number of splits:1
2022-11-01 15:09:45,875 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1667293411803_0001
2022-11-01 15:09:45,879 INFO mapreduce.JobSubmitter: Executing with tokens: []
2022-11-01 15:09:46,382 INFO conf.Configuration: resource-types.xml not found
2022-11-01 15:09:46,383 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.
2022-11-01 15:09:47,169 INFO impl.YarnClientImpl: Submitted application application_1667293411803_0001
2022-11-01 15:09:47,293 INFO mapreduce.Job: The url to track the job: http://fd41144:8080/proxy/application_1667293411803_0001/
2022-11-01 15:09:47,295 INFO mapreduce.Job: Running job: job_1667293411803_0001
2022-11-01 15:10:05,757 INFO mapreduce.Job: Job job_1667293411803_0001 running in uber mode : false
2022-11-01 15:10:05,761 INFO mapreduce.Job: map 0% reduce 0%
2022-11-01 15:10:13,967 INFO mapreduce.Job: map 100% reduce 0%
2022-11-01 15:10:22,093 INFO mapreduce.Job: map 100% reduce 100%
2022-11-01 15:10:27,166 INFO mapreduce.Job: Job job_1667293411803_0001 completed successfully
2022-11-01 15:10:27,375 INFO mapreduce.Job: Counters: 54
  File System Counters
    FILE: Number of bytes read=98
    FILE: Number of bytes written=471545
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=143
    HDFS: Number of bytes written=60
    HDFS: Number of read operations=8
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=2
    HDFS: Number of bytes read erasure-coded=0
  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)=5113
    Total time spent by all reduces in occupied slots (ms)=5425
    Total time spent by all map tasks (ms)=5113
    Total time spent by all reduce tasks (ms)=5425
  Map-Reduce Framework
    Map input records=1
    Map output records=8
    Map output bytes=76
    Map output materialized bytes=98
    Input split bytes=100
    Combine input records=8
    Combine output records=8
    Reduce input groups=8
    Reduce shuffle bytes=98
    Reduce input records=8
    Reduce output records=8
    Spilled Records=16
    Shuffled Maps =1
    Failed Shuffles=0
    Merged Map outputs=1
    GC time elapsed (ms)=162
    CPU time spent (ms)=2045
    Physical memory (bytes) snapshot=506421248
    Virtual memory (bytes) snapshot=736817152
    Total committed heap usage (bytes)=360710144
    Peak Map Physical memory (bytes)=305922046
    Peak Map Virtual memory (bytes)=421195776
    Peak Reduce Physical memory (bytes)=202018816
    Peak Reduce Virtual memory (bytes)=317112320
  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=43
  File Output Format Counters
    Bytes Written=60
PS D:\hadoop-env\hadoop-3.2.2\sbin>
```

```
Administrator: Windows PowerShell
HDFS: Number of bytes read erasure-coded=0
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)=5113
  Total time spent by all reduces in occupied slots (ms)=5425
  Total time spent by all map tasks (ms)=5113
  Total time spent by all reduce tasks (ms)=5425
  Total vcore-milliseconds taken by all map tasks=5113
  Total vcore-milliseconds taken by all reduce tasks=5425
  Total megabyte-milliseconds taken by all map tasks=5235712
  Total megabyte-milliseconds taken by all reduce tasks=5555200
Map-Reduce Framework
  Map input records=1
  Map output records=8
  Map output bytes=76
  Map output materialized bytes=98
  Input split bytes=100
  Combine input records=8
  Combine output records=8
  Reduce input groups=8
  Reduce shuffle bytes=98
  Reduce input records=8
  Reduce output records=8
  Spilled Records=16
  Shuffled Maps =1
  Failed Shuffles=0
  Merged Map outputs=1
  GC time elapsed (ms)=162
  CPU time spent (ms)=2045
  Physical memory (bytes) snapshot=506421248
  Virtual memory (bytes) snapshot=736817152
  Total committed heap usage (bytes)=360710144
  Peak Map Physical memory (bytes)=305922046
  Peak Map Virtual memory (bytes)=421195776
  Peak Reduce Physical memory (bytes)=202018816
  Peak Reduce Virtual memory (bytes)=317112320
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=43
File Output Format Counters
  Bytes Written=60
PS D:\hadoop-env\hadoop-3.2.2\sbin>
```

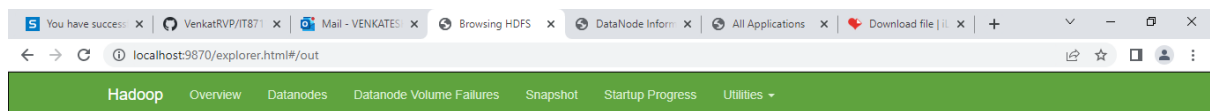
Name : ANANTHAN C P

Roll No: 19UITE017

Ex. No : 10

Word Count program using MapReduce

```
Administrator: Windows PowerShell
Total committed heap usage (bytes)=360710144
Peak Map Physical memory (bytes)=305922048
Peak Map Virtual memory (bytes)=421195776
Peak Reduce Physical memory (bytes)=202018816
Peak Reduce Virtual memory (bytes)=317112320
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=43
File Output Format Counters
Bytes Written=60
PS D:\hadoop-env\hadoop-3.2.2\sbin> hadoop fs -cat /out/*
final 1
IT 1
Prabhu 1
This 1
Venkatesh 1
Year 1
from 1
is 1
PS D:\hadoop-env\hadoop-3.2.2\sbin>
```



Browse Directory

/out

Go!

Show 25 entries

Search:

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r--r--	19uite001	supergroup	0 B	Nov 01 15:10	1	128 MB	_SUCCESS
-rw-r--r--	19uite001	supergroup	60 B	Nov 01 15:10	1	128 MB	part-r-00000

Showing 1 to 2 of 2 entries

Previous 1 Next

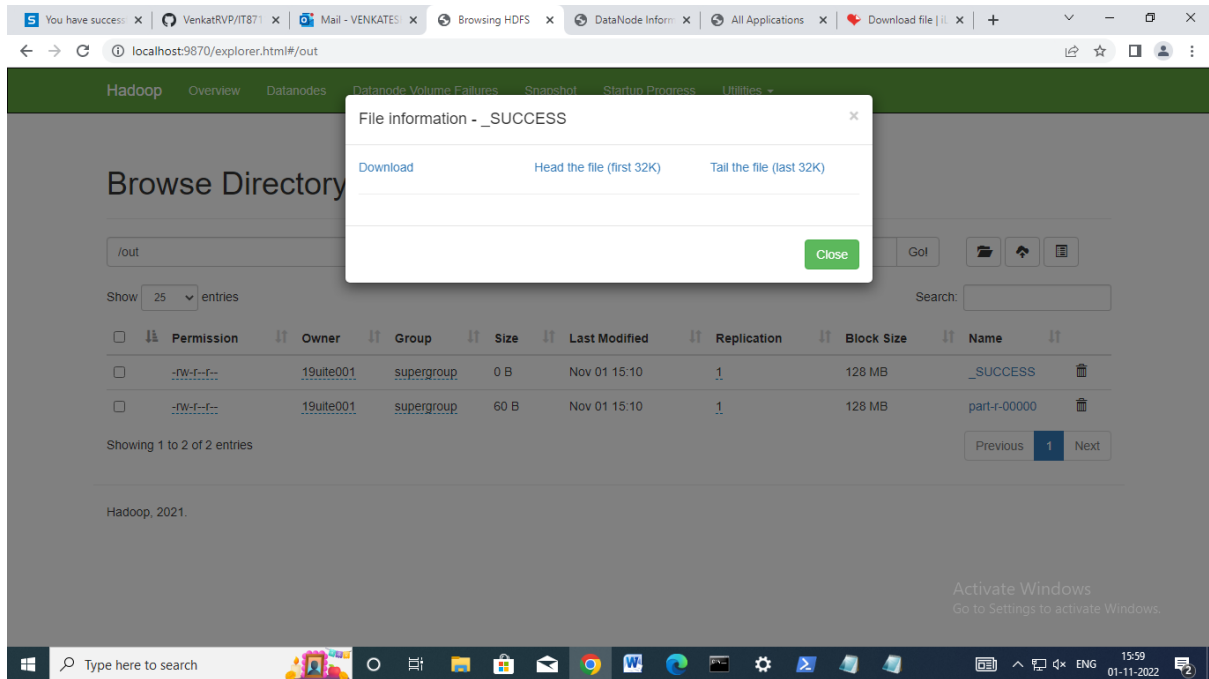
Hadoop, 2021.



Name : ANANTHAN C P
Ex. No : 10

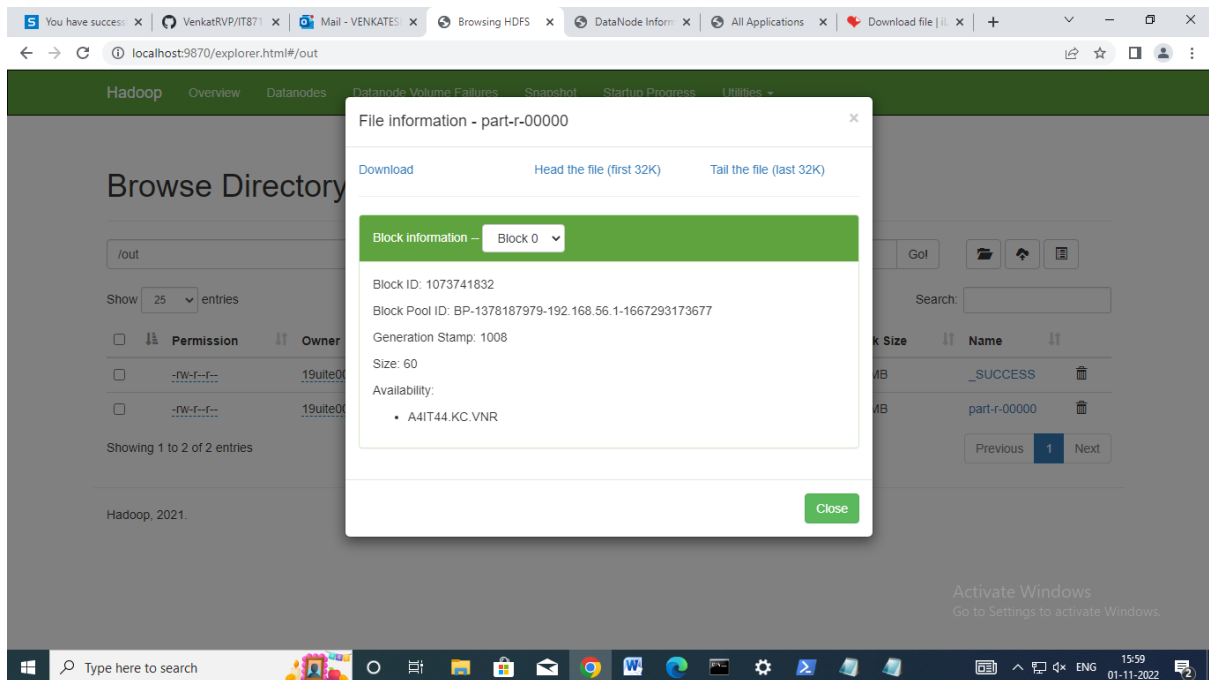
Word Count program using MapReduce

Roll No: 19UITE017



The screenshot shows the Hadoop File Browser interface at localhost:9870/explorer.html#/out. A modal window titled "File information - _SUCCESS" is open, displaying options to "Download", "Head the file (first 32K)", and "Tail the file (last 32K)". The background shows a table of files in the /out directory.

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r--r--	19uite001	supergroup	0 B	Nov 01 15:10	1	128 MB	_SUCCESS
-rw-r--r--	19uite001	supergroup	60 B	Nov 01 15:10	1	128 MB	part-r-00000



The screenshot shows the Hadoop File Browser interface at localhost:9870/explorer.html#/out. A modal window titled "File information - part-r-00000" is open, displaying options to "Download", "Head the file (first 32K)", and "Tail the file (last 32K)". The modal also shows "Block information" for Block 0.

Block information -- Block 0

Block ID: 1073741832
Block Pool ID: BP-1378187979-192.168.56.1-1667293173677
Generation Stamp: 1008
Size: 60
Availability:
• A4IT44.KC.VNR

Name : ANANTHAN C P
Ex. No : 10

Roll No: 19UITE017

Word Count program using MapReduce

The screenshot displays the Hadoop Distributed File System (HDFS) Explorer interface. The main window shows the directory structure, with the file 'out' selected. A modal window titled 'Block information' is open, displaying details for Block 0. The block information includes the Block ID (1073741832), Block Pool ID (BP-1378187979-192.168.56.1-1667293173677), Generation Stamp (1008), Size (60), and Availability (A4IT44.KC.VNR). Below the block information, the 'File contents' section shows the output of the Word Count program, listing words and their frequencies: Final (1), IT (1), Prabhu (1), This (1), Venkatesh (1), Year (1), from (1), and is (1). The background interface shows the 'Browse Directory' view with a search bar and a table of files. The Windows taskbar at the bottom indicates the system time as 16:00 on 01-11-2022.

Block information -- Block 0

Block ID: 1073741832
Block Pool ID: BP-1378187979-192.168.56.1-1667293173677
Generation Stamp: 1008
Size: 60
Availability:
• A4IT44.KC.VNR

File contents

Final 1
IT 1
Prabhu 1
This 1
Venkatesh 1
Year 1
from 1
is 1

Close