

```
C:\Windows\System32\cmd.exe
Microsoft Windows [Version 10.0.22631.4460]
(c) Microsoft Corporation. All rights reserved.

C:\LK\6th SEM\Parallel and distribution systems\labsheet>code .

C:\LK\6th SEM\Parallel and distribution systems\labsheet>docker-compose up -d
time="2024-11-25T18:42:35+05:30" level=warning msg="C:\LK\6th SEM\Parallel and distribution systems\labsheet\docker-compose.yml: the attribute 'version'
is obsolete, it will be ignored, please remove it to avoid potential confusion"
[*] Running 7/7
✓Network labsheet_default Created 0.0s
✓Volume "labsheet_historyserver-data" Created 0.0s
✓Volume "labsheet_namenode-data" Created 0.0s
✓Volume "labsheet_datanode-data" Created 0.0s
✓Container namenode Start... 0.6s
✓Container datanode Start... 0.8s
✓Container historyserver Started 1.0s

C:\LK\6th SEM\Parallel and distribution systems\labsheet>docker cp "C:\LK\6th SEM\Parallel and distribution systems\labsheet\sample.txt" namenode:/sample.tx
t
Successfully copied 2.56kB to namenode:/sample.txt

C:\LK\6th SEM\Parallel and distribution systems\labsheet>docker exec -it namenode ls /
KEYS dev hadoop lib mnt root sample.txt sys var
bin entrypoint.sh hadoop-data lib64 opt run sbin tmp
boot etc home media proc run.sh srv usr

What's next:
Try Docker Debug for seamless, persistent debugging tools in any container or image → docker debug namenode
Learn more at https://docs.docker.com/go/debug-cli/

C:\LK\6th SEM\Parallel and distribution systems\labsheet>docker exec -it namenode hdfs dfs -mkdir -p /input

What's next:
Try Docker Debug for seamless, persistent debugging tools in any container or image → docker debug namenode
Learn more at https://docs.docker.com/go/debug-cli/

C:\LK\6th SEM\Parallel and distribution systems\labsheet>docker exec -it namenode hdfs dfs -put /sample.txt /input
2024-11-25 13:16:03,540 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false

What's next:
Try Docker Debug for seamless, persistent debugging tools in any container or image → docker debug namenode
Learn more at https://docs.docker.com/go/debug-cli/

C:\LK\6th SEM\Parallel and distribution systems\labsheet>docker exec -it namenode hdfs dfs -ls /input
Found 1 items
-rw-r--r-- 3 root supergroup 676 2024-11-25 13:16 /input/sample.txt
```

```
C:\Windows\System32\cmd.exe
C:\LK\6th SEM\Parallel and distribution systems\labsheet>docker exec -it namenode hdfs dfs -ls /input
Found 1 items
-rw-r--r-- 3 root supergroup 676 2024-11-25 13:16 /input/sample.txt

What's next:
Try Docker Debug for seamless, persistent debugging tools in any container or image → docker debug namenode
Learn more at https://docs.docker.com/go/debug-cli/

C:\LK\6th SEM\Parallel and distribution systems\labsheet>docker exec -it namenode hadoop jar /opt/hadoop-3.2.1/share/hadoop/mapreduce/hadoop-mapreduce-examp
les-3.2.1.jar wordcount /input /output
2024-11-25 13:16:35,954 INFO impl.MetricsConfig: Loaded properties from hadoop-metrics2.properties
2024-11-25 13:16:35,997 INFO impl.MetricsSystemImpl: Scheduled Metric snapshot period at 10 second(s).
2024-11-25 13:16:35,997 INFO impl.MetricsSystemImpl: JobTracker metrics system started
2024-11-25 13:16:36,138 INFO input.FileInputFormat: Total input files to process : 1
2024-11-25 13:16:36,151 INFO mapreduce.JobSubmitter: number of splits:1
2024-11-25 13:16:36,226 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_local1432588286_0001
2024-11-25 13:16:36,226 INFO mapreduce.JobSubmitter: Executing with tokens: []
2024-11-25 13:16:36,282 INFO mapreduce.Job: The url to track the job: http://localhost:8080/
2024-11-25 13:16:36,282 INFO mapreduce.Job: Running job: job_local1432588286_0001
2024-11-25 13:16:36,282 INFO mapred.LocalJobRunner: OutputCommitter set in config null
2024-11-25 13:16:36,286 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2024-11-25 13:16:36,286 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup fa
ilures: false
2024-11-25 13:16:36,286 INFO mapred.LocalJobRunner: OutputCommitter is org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter
2024-11-25 13:16:36,385 INFO mapred.LocalJobRunner: Waiting for map tasks
2024-11-25 13:16:36,385 INFO mapred.LocalJobRunner: Starting task: attempt_local1432588286_0001_m_000000_0
2024-11-25 13:16:36,315 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2024-11-25 13:16:36,315 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup fa
ilures: false
2024-11-25 13:16:36,325 INFO mapred.Task: Using ResourceCalculatorProcessTree : [ ]
2024-11-25 13:16:36,327 INFO mapred.MapTask: Processing split: hdfs://namenode:8020/input/sample.txt:0+676
2024-11-25 13:16:36,355 INFO mapred.MapTask: (EQUATOR) 0 kvi 26214396(104857584)
2024-11-25 13:16:36,356 INFO mapred.MapTask: mapreduce.task.io.sort.mb: 100
2024-11-25 13:16:36,356 INFO mapred.MapTask: soft limit at 83886080
2024-11-25 13:16:36,356 INFO mapred.MapTask: bufstart = 0; bufvoid = 104857600
2024-11-25 13:16:36,359 INFO mapred.MapTask: Map output collector class = org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2024-11-25 13:16:36,378 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
2024-11-25 13:16:36,422 INFO mapred.LocalJobRunner:
2024-11-25 13:16:36,423 INFO mapred.MapTask: Starting flush of map output
2024-11-25 13:16:36,424 INFO mapred.MapTask: Spilling map output
2024-11-25 13:16:36,424 INFO mapred.MapTask: bufstart = 0; bufend = 1112; bufvoid = 104857600
2024-11-25 13:16:36,424 INFO mapred.MapTask: kvstart = 26214396(104857584); kvend = 26213956(104855824); length = 441/6553600
2024-11-25 13:16:36,432 INFO mapred.MapTask: Finished spill 0
2024-11-25 13:16:36,437 INFO mapred.Task: Task:attempt_local1432588286_0001_m_000000_0 is done. And is in the process of committing
2024-11-25 13:16:36,440 INFO mapred.LocalJobRunner: map
```

```
C:\Windows\System32\cmd.exe
2024-11-25 13:16:36,437 INFO mapred.Task: Task:attempt_local1432588286_0001_m_000000_0 is done. And is in the process of committing
2024-11-25 13:16:36,440 INFO mapred.LocalJobRunner: map
2024-11-25 13:16:36,440 INFO mapred.Task: Task 'attempt_local1432588286_0001_m_000000_0' done.
2024-11-25 13:16:36,444 INFO mapred.Task: Final Counters for attempt_local1432588286_0001_m_000000_0: Counters: 24
File System Counters
  FILE: Number of bytes read=316695
  FILE: Number of bytes written=842614
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=676
  HDFS: Number of bytes written=0
  HDFS: Number of read operations=5
  HDFS: Number of large read operations=0
  HDFS: Number of write operations=1
  HDFS: Number of bytes read erasure-coded=0
Map-Reduce Framework
  Map input records=10
  Map output records=111
  Map output bytes=1112
  Map output materialized bytes=663
  Input split bytes=102
  Combine input records=111
  Combine output records=54
  Spilled Records=54
  Failed Shuffles=0
  Merged Map outputs=0
  GC time elapsed (ms)=0
  Total committed heap usage (bytes)=284688384
File Input Format Counters
  Bytes Read=676
2024-11-25 13:16:36,444 INFO mapred.LocalJobRunner: Finishing task: attempt_local1432588286_0001_m_000000_0
2024-11-25 13:16:36,444 INFO mapred.LocalJobRunner: map task executor complete.
2024-11-25 13:16:36,446 INFO mapred.LocalJobRunner: Waiting for reduce tasks
2024-11-25 13:16:36,446 INFO mapred.LocalJobRunner: Starting task: attempt_local1432588286_0001_r_000000_0
2024-11-25 13:16:36,449 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2024-11-25 13:16:36,449 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup failures: false
2024-11-25 13:16:36,449 INFO mapred.Task: Using ResourceCalculatorProcessTree : [ ]
2024-11-25 13:16:36,451 INFO mapred.ReduceTask: Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@56af3aec
2024-11-25 13:16:36,452 WARN impl.MetricsSystemImpl: JobTracker metrics system already initialized!
2024-11-25 13:16:36,461 INFO reduce.MergeManagerImpl: MergerManager: memoryLimit=1264687488, maxSingleShuffleLimit=316171872, mergeThreshold=834693760, ioSortFactor=10, memToMemMergeOutputsThreshold=10
2024-11-25 13:16:36,462 INFO reduce.EventFetcher: attempt_local1432588286_0001_r_000000_0 Thread started: EventFetcher for fetching Map Completion Events
2024-11-25 13:16:36,473 INFO reduce.LocalFetcher: localfetcher#1 about to shuffle output of map attempt_local1432588286_0001_m_000000_0 decomp: 659 len: 663 to MEMORY
2024-11-25 13:16:36,475 INFO reduce.InMemoryMapOutput: Read 659 bytes from map-output for attempt_local1432588286_0001_m_000000_0
```

```
to MEMORY
2024-11-25 13:16:36,475 INFO reduce.InMemoryMapOutput: Read 659 bytes from map-output for attempt_local1432588286_0001_m_000000_0
2024-11-25 13:16:36,476 INFO reduce.MergeManagerImpl: closeInMemoryFile -> map-output of size: 659, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory -> 659
2024-11-25 13:16:36,477 INFO reduce.EventFetcher: EventFetcher is interrupted.. Returning
2024-11-25 13:16:36,478 INFO mapred.LocalJobRunner: 1 / 1 copied.
2024-11-25 13:16:36,478 INFO reduce.MergeManagerImpl: finalMerge called with 1 in-memory map-outputs and 0 on-disk map-outputs
2024-11-25 13:16:36,481 INFO mapred.Merger: Merging 1 sorted segments
2024-11-25 13:16:36,482 INFO mapred.Merger: Down to the last merge-pass, with 1 segments left of total size: 647 bytes
2024-11-25 13:16:36,482 INFO reduce.MergeManagerImpl: Merged 1 segments, 659 bytes to disk to satisfy reduce memory limit
2024-11-25 13:16:36,482 INFO reduce.MergeManagerImpl: Merging 1 files, 663 bytes from disk
2024-11-25 13:16:36,483 INFO reduce.MergeManagerImpl: Merging 0 segments, 0 bytes from memory into reduce
2024-11-25 13:16:36,483 INFO mapred.Merger: Merging 1 sorted segments
2024-11-25 13:16:36,483 INFO mapred.Merger: Down to the last merge-pass, with 1 segments left of total size: 647 bytes
2024-11-25 13:16:36,483 INFO mapred.LocalJobRunner: 1 / 1 copied.
2024-11-25 13:16:36,497 INFO Configuration.deprecation: mapred.skip.on is deprecated. Instead, use mapreduce.job.skiprecords
2024-11-25 13:16:36,516 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
2024-11-25 13:16:36,932 INFO mapred.Task: Task:attempt_local1432588286_0001_r_000000_0 is done. And is in the process of committing
2024-11-25 13:16:36,933 INFO mapred.LocalJobRunner: 1 / 1 copied.
2024-11-25 13:16:36,933 INFO mapred.Task: Task attempt_local1432588286_0001_r_000000_0 is allowed to commit now
2024-11-25 13:16:36,941 INFO output.FileOutputCommitter: Saved output of task 'attempt_local1432588286_0001_r_000000_0' to hdfs://namenode:8020/output
2024-11-25 13:16:36,942 INFO mapred.LocalJobRunner: reduce > reduce
2024-11-25 13:16:36,942 INFO mapred.Task: Task 'attempt_local1432588286_0001_r_000000_0' done.
2024-11-25 13:16:36,942 INFO mapred.Task: Final Counters for attempt_local1432588286_0001_r_000000_0: Counters: 30
File System Counters
  FILE: Number of bytes read=318053
  FILE: Number of bytes written=843277
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=676
  HDFS: Number of bytes written=441
  HDFS: Number of read operations=10
  HDFS: Number of large read operations=0
  HDFS: Number of write operations=3
  HDFS: Number of bytes read erasure-coded=0
Map-Reduce Framework
  Combine input records=0
  Combine output records=0
  Reduce input groups=54
  Reduce shuffle bytes=663
  Reduce input records=54
  Reduce output records=54
  Spilled Records=54
  Shuffled Maps=1
  Failed Shuffles=0
  Merged Map outputs=1
```

```
C:\Windows\System32\cmd.exe x + v

Failed Shuffles=0
Merged Map outputs=1
GC time elapsed (ms)=6
Total committed heap usage (bytes)=385875968

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

File Output Format Counters
Bytes Written=441
2024-11-25 13:16:36,942 INFO mapred.LocalJobRunner: Finishing task: attempt_local1432588286_0001_r_000000_0
2024-11-25 13:16:36,942 INFO mapred.LocalJobRunner: reduce task executor complete.
2024-11-25 13:16:37,284 INFO mapreduce.Job: Job job_local1432588286_0001 running in uber mode : false
2024-11-25 13:16:37,285 INFO mapreduce.Job: map 100% reduce 100%
2024-11-25 13:16:37,286 INFO mapreduce.Job: Job job_local1432588286_0001 completed successfully
2024-11-25 13:16:37,290 INFO mapreduce.Job: Counters: 36
File System Counters
FILE: Number of bytes read=634748
FILE: Number of bytes written=1685891
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=1352
HDFS: Number of bytes written=441
HDFS: Number of read operations=15
HDFS: Number of large read operations=0
HDFS: Number of write operations=4
HDFS: Number of bytes read erasure-coded=0
Map-Reduce Framework
Map input records=10
Map output records=111
Map output bytes=1112
Map output materialized bytes=663
Input split bytes=102
Combine input records=111
Combine output records=54
Reduce input groups=54
Reduce shuffle bytes=663
Reduce input records=54
Reduce output records=54
Spilled Records=108
Shuffled Maps =1
Failed Shuffles=0
Merged Map outputs=1

67°F
Light rain
```

```
C:\Windows\System32\cmd.exe x + v

Failed Shuffles=0
Merged Map outputs=1
GC time elapsed (ms)=6
Total committed heap usage (bytes)=670564352

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=676
File Output Format Counters
Bytes Written=441

What's next:
Try Docker Debug for seamless, persistent debugging tools in any container or image → docker debug namenode
Learn more at https://docs.docker.com/go/debug-cli/

C:\LK\6th SEM\Parallel and distribution systems\labsheet>docker exec -it namenode hdfs dfs -ls /output
'ocker' is not recognized as an internal or external command,
operable program or batch file.

C:\LK\6th SEM\Parallel and distribution systems\labsheet>docker exec -it nam
enode hdfs dfs -ls /output
Found 2 items
-rw-r--r-- 3 root supergroup 0 2024-11-25 13:16 /output/_SUCCESS
-rw-r--r-- 3 root supergroup 441 2024-11-25 13:16 /output/part-r-000000

What's next:
Try Docker Debug for seamless, persistent debugging tools in any container or image → docker debug namenode
Learn more at https://docs.docker.com/go/debug-cli/

C:\LK\6th SEM\Parallel and distribution systems\labsheet>docker exec -it namenode hdfs dfs -cat /output/part-r-000000
2024-11-25 13:17:21,420 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
afternoon 1
analysis 1
apple 3
banana 3
bike 3
bird 1
book 3
bus 2
car 4
cat 3
cheerful 2

67°F
Light rain
```

```
C:\Windows\System32\cmd.exe
cat 3
cheerful 2
code 1
dance 2
data 1
dog 4
evening 2
excited 2
fast 2
fish 2
forest 2
grape 2
hadoop 2
happy 3
hill 1
java 2
joyful 1
lake 1
mango 1
melody 2
morning 4
mountain 2
music 3
night 3
noon 1
notebook 2
orange 3
paper 3
pen 2
pencil 1
python 2
quick 2
rabbit 1
rapid 2
rhythm 2
river 3
sad 2
scala 1
scooter 1
slow 4
song 2
spark 1
steady 1
tree 3
truck 1

What's next:
```

```
C:\Windows\System32\cmd.exe
hill 1
java 2
joyful 1
lake 1
mango 1
melody 2
morning 4
mountain 2
music 3
night 3
noon 1
notebook 2
orange 3
paper 3
pen 2
pencil 1
python 2
quick 2
rabbit 1
rapid 2
rhythm 2
river 3
sad 2
scala 1
scooter 1
slow 4
song 2
spark 1
steady 1
tree 3
truck 1

What's next:
Try Docker Debug for seamless, persistent debugging tools in any container or image → docker debug namenode
Learn more at https://docs.docker.com/go/debug-cli/

C:\LK\6th SEM\Parallel and distribution systems\labsheet>docker-compose down

time="2024-11-25T18:47:36+05:30" level=warning msg="C:\\LK\\6th SEM\\Parallel and distribution systems\\labsheet\\docker-compose.yml: the attribute 'version'
[+] Running 4/4 it will be ignored, please remove it to avoid potential confusion"
✓Container historyserver Removed 10.3s
✓Container datanode Removed 10.3s
✓Container namenode Removed 10.3s
✓Network labsheet_default Removed 0.2s

C:\LK\6th SEM\Parallel and distribution systems\labsheet>
```