

# **ASSIGNMENT 03**

**Sabaragamuwa University of Sri Lanka**

**Faculty of Computing**

**Department of Software Engineering**

**SE6103 – Parallel and Distributed Systems**

Name	: Sabeeb A.I.M.
Reg. No	: 19APSE4289
Academic Period	: 03 <sup>rd</sup> Year 02 <sup>nd</sup> Semester
Degree Program	: Software Engineering
Due Date	: 06/01/2025

## Question 02

```
U:\03rd Year 02nd Semester\SE6103 - Parallel & Distributed Systems\Ass>docker-compose up -d
time="2025-01-06T15:28:26+05:30" level=warning msg="U:\03rd Year 02nd Semester\SE6103 - Parallel & Distributed Systems\Ass\docker-compose.yml: the attribute 'version' is obsolete, it will be ignored, please remove it to avoid potential confusion"
[+] Running 3/3
  ✓ Container namenode      Started          1.6s
  ✓ Container datanode      Started          2.1s
  ✓ Container historyserver Started          2.7s

U:\03rd Year 02nd Semester\SE6103 - Parallel & Distributed Systems\Ass>docker exec -it namenode hdfs dfs -mkdir -p /input

mkdir: Call From b2fd93c4ac19/172.20.0.2 to namenode:8020 failed on connection exception: java.net.ConnectException: Connection refused; For more details see: http://wiki.apache.org/hadoop/ConnectionRefused

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/
```

```
U:\03rd Year 02nd Semester\SE6103 - Parallel & Distributed Systems\Ass>docker cp "U:\03rd Year 02nd Semester\SE6103 - Parallel & Distributed Systems\Ass\sample.txt" namenode:/sample.txt
Successfully copied 2.05kB to namenode:/sample.txt
```

```
U:\03rd Year 02nd Semester\SE6103 - Parallel & Distributed Systems\Ass>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/

U:\03rd Year 02nd Semester\SE6103 - Parallel & Distributed Systems\Ass>docker exec -it namenode hdfs dfs -put /sample.txt /input
2025-01-06 10:01:28,868 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/

U:\03rd Year 02nd Semester\SE6103 - Parallel & Distributed Systems\Ass>docker exec -it namenode hdfs dfs -ls /input
OCI runtime exec failed: exec failed: unable to start container process: exec: "hdfs\u00a0dfs\u00a0ls\u00a0/input": stat hdfs dfs -ls /input: no such file or directory: unknown

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/
```

```
U:\03rd Year 02nd Semester\SE6103 - Parallel & Distributed Systems\Ass>docker exec -it namenode hadoop jar /opt/hadoop-3.2.1/share/hadoop/mapreduce/hadoop-mapreduce-examples-3.2.1.jar wordcount /input /output
2025-01-06 10:04:57,589 INFO impl.MetricsConfig: Loaded properties from hadoop-metrics2.properties
2025-01-06 10:04:57,691 INFO impl.MetricsSystemImpl: Scheduled Metric snapshot period at 10 second(s).
2025-01-06 10:04:57,692 INFO impl.MetricsSystemImpl: JobTracker metrics system started
2025-01-06 10:04:58,111 INFO input.FileInputFormat: Total input files to process : 1
2025-01-06 10:04:58,149 INFO mapreduce.JobSubmitter: number of splits:1
2025-01-06 10:04:58,440 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_local510075616_0001
2025-01-06 10:04:58,440 INFO mapreduce.JobSubmitter: Executing with tokens: []
2025-01-06 10:04:58,651 INFO mapreduce.Job: The url to track the job: http://localhost:8080/
2025-01-06 10:04:58,652 INFO mapreduce.Job: Running job: job_local510075616_0001
2025-01-06 10:04:58,656 INFO mapred.LocalJobRunner: OutputCommitter set in config null
2025-01-06 10:04:58,671 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2025-01-06 10:04:58,672 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup failures: false
2025-01-06 10:04:58,673 INFO mapred.LocalJobRunner: OutputCommitter is org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter
2025-01-06 10:04:58,731 INFO mapred.LocalJobRunner: Waiting for map tasks
2025-01-06 10:04:58,732 INFO mapred.LocalJobRunner: Starting task: attempt_local510075616_0001_m_000000_0
2025-01-06 10:04:58,767 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2025-01-06 10:04:58,768 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup failures: false
2025-01-06 10:04:58,805 INFO mapred.Task: Using ResourceCalculatorProcessTree : [ ]
2025-01-06 10:04:58,812 INFO mapred.MapTask: Processing split: hdfs://namenode:8020/input/sample.txt:0+180
2025-01-06 10:04:58,913 INFO mapred.MapTask: (EQUATOR) 0 kvi 26214396(104857584)
2025-01-06 10:04:58,913 INFO mapred.MapTask: mapreduce.task.io.sort.mb: 100
2025-01-06 10:04:58,913 INFO mapred.MapTask: soft limit at 83886080
2025-01-06 10:04:58,913 INFO mapred.MapTask: bufstart = 0; bufvoid = 104857600
2025-01-06 10:04:58,913 INFO mapred.MapTask: kvstart = 26214396; length = 6553600
2025-01-06 10:04:58,922 INFO mapred.MapTask: Map output collector class = org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2025-01-06 10:04:58,982 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2025-01-06 10:04:59,152 INFO mapred.LocalJobRunner:
2025-01-06 10:04:59,155 INFO mapred.MapTask: Starting flush of map output
2025-01-06 10:04:59,155 INFO mapred.MapTask: Spilling map output
2025-01-06 10:04:59,155 INFO mapred.MapTask: bufstart = 0; bufend = 291; bufvoid = 104857600
2025-01-06 10:04:59,155 INFO mapred.MapTask: kvstart = 26214396(104857584); kvend = 26214288(104857152); length = 109/6553600
2025-01-06 10:04:59,183 INFO mapred.MapTask: Finished spill 0
```

```

2025-01-06 10:04:59,239 INFO mapred.Task: Task 'attempt_local510075616_0001_m_000000_0' done.
2025-01-06 10:04:59,253 INFO mapred.Task: Final Counters for attempt_local510075616_0001_m_000000_0: Counters: 24
  File System Counters
    FILE: Number of bytes read=316694
    FILE: Number of bytes written=839700
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=180
    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=1
    Map output records=28
    Map output bytes=291
    Map output materialized bytes=298
    Input split bytes=102
    Combine input records=28
    Combine output records=23
    Spilled Records=23
    Failed Shuffles=0
    Merged Map outputs=0
    GC time elapsed (ms)=0
    Total committed heap usage (bytes)=273678336
  File Input Format Counters
    Bytes Read=180
2025-01-06 10:04:59,254 INFO mapred.LocalJobRunner: Finishing task: attempt_local510075616_0001_m_000000_0
2025-01-06 10:04:59,259 INFO mapred.LocalJobRunner: map task executor complete.

```

```

    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=
  Map-Reduce Framework
    Map input records=1
    Map output records=28
    Map output bytes=291
    Map output materialized bytes=298
    Input split bytes=102
    Combine input records=28
    Combine output records=23
    Reduce input groups=23
    Reduce shuffle bytes=298
    Reduce input records=23
    Reduce output records=23
    Spilled Records=46
    Shuffled Maps =1
    Failed Shuffles=0
    Merged Map outputs=1
    GC time elapsed (ms)=26
    Total committed heap usage (bytes)=548405248
  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=180
  File Output Format Counters
    Bytes Written=200

```

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/>

```

U:\03rd Year 02nd Semester\SE6103 - Parallel & Distributed Systems\Ass>docker exec -it namenode hdfs dfs -ls /output
Found 2 items
-rw-r--r-- 3 root supergroup 0 2025-01-06 10:04 /output/_SUCCESS
-rw-r--r-- 3 root supergroup 200 2025-01-06 10:04 /output/part-r-00000

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/

U:\03rd Year 02nd Semester\SE6103 - Parallel & Distributed Systems\Ass>docker exec -it namenode hdfs dfs -cat /output/part-r-00000
2025-01-06 10:09:56,957 INFO sasl.SaslDataTransferClient: SASL encryption trust check: LocalHostTrusted = false, remoteHostTrusted = false
Hadoop 2
a 2
about 1
across 1
amounts 1
an 1
and 1
computers. 1
data 1
framework 2
interesting 1
is 2
large 1
learn 1
network 1
networking 1
of 2
open-source 1
processes 1
stores 1
that 1
to 1
very 1

What's next:
  Try Docker Debug for seamless, persistent debugging tools in any container or image → docker debug namenode

```

```

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/

U:\03rd Year 02nd Semester\SE6103 - Parallel & Distributed Systems\Ass>docker-compose down
time="2025-01-06T15:40:19+05:30" level=warning msg="U:\03rd Year 02nd Semester\SE6103 - Parallel & Distributed Systems\Ass\docker-compose.yml: the attribute 'version' is obsolete, it will be ignored, please remove it to avoid potential confusion"
[*] Running 4/4
✓Container historyserver Removed 10.6s
✓Container datanode Removed 10.5s
✓Container namenode Removed 10.5s
✓Network ass_default Removed 0.3s

U:\03rd Year 02nd Semester\SE6103 - Parallel & Distributed Systems\Ass>

```