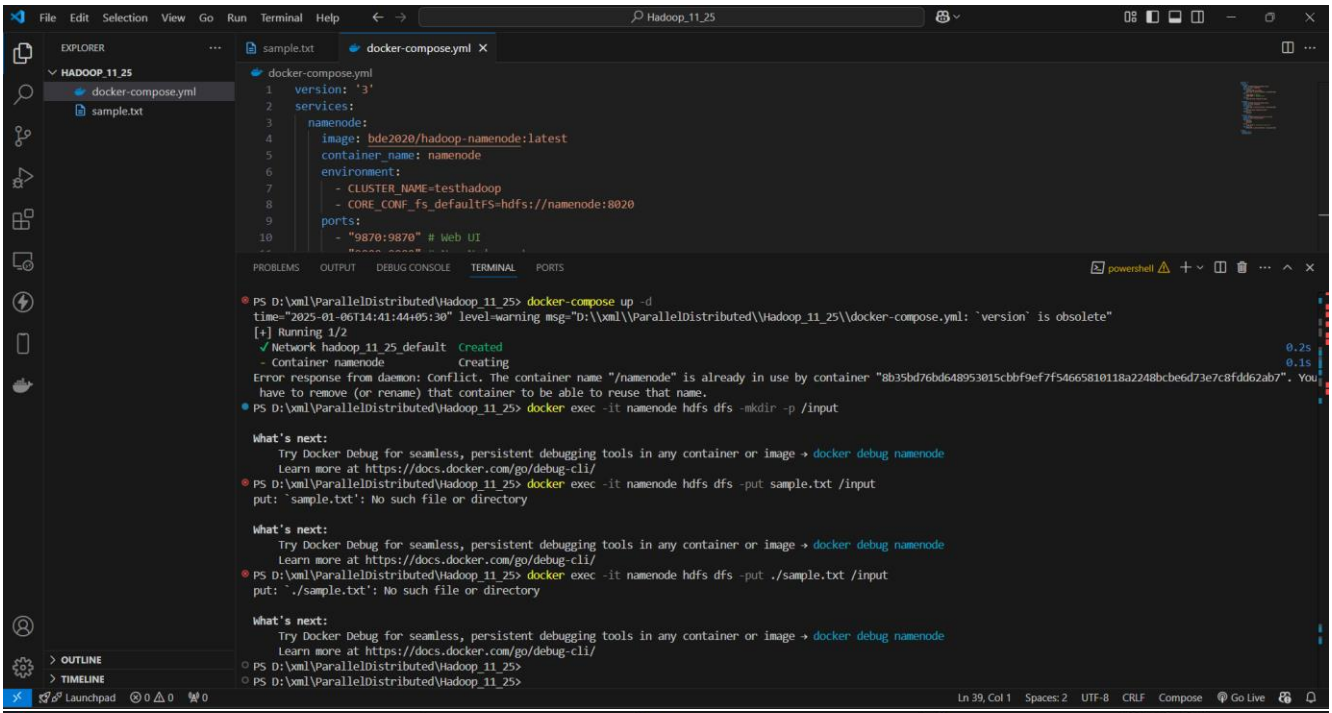
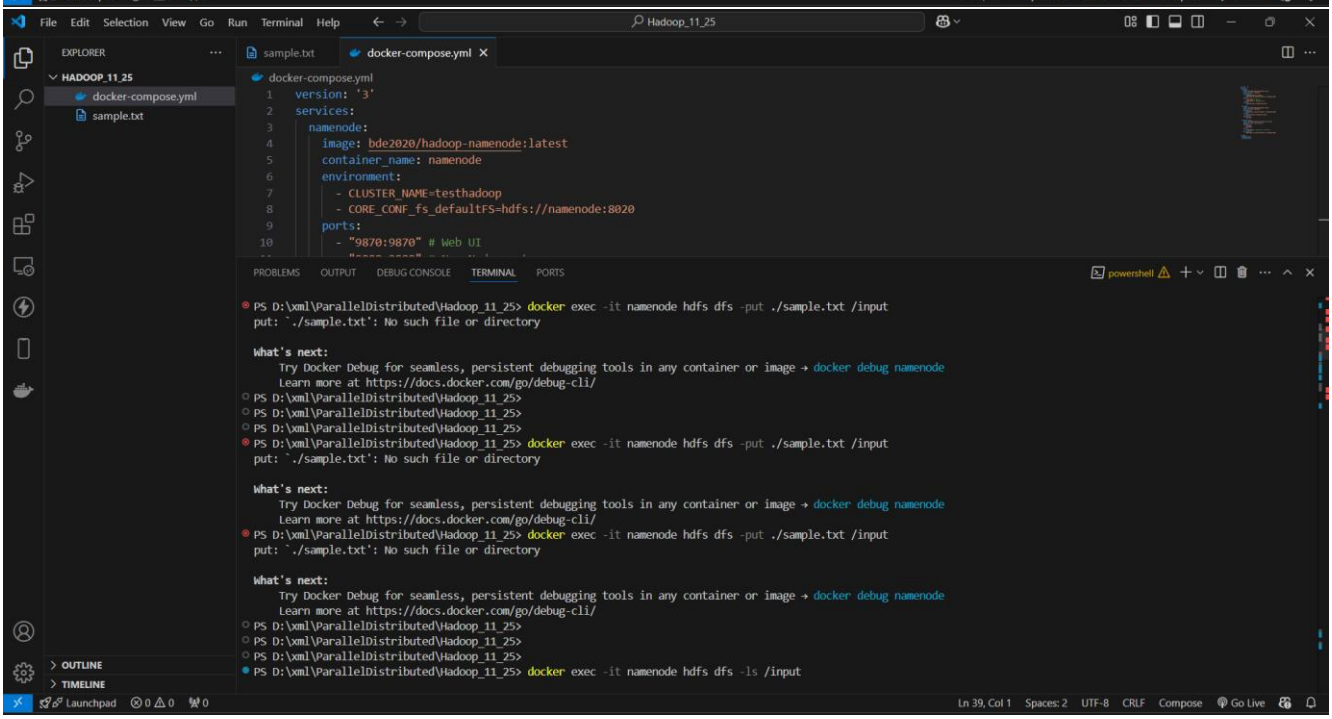
SE6103 Parallel and Distributed Systems

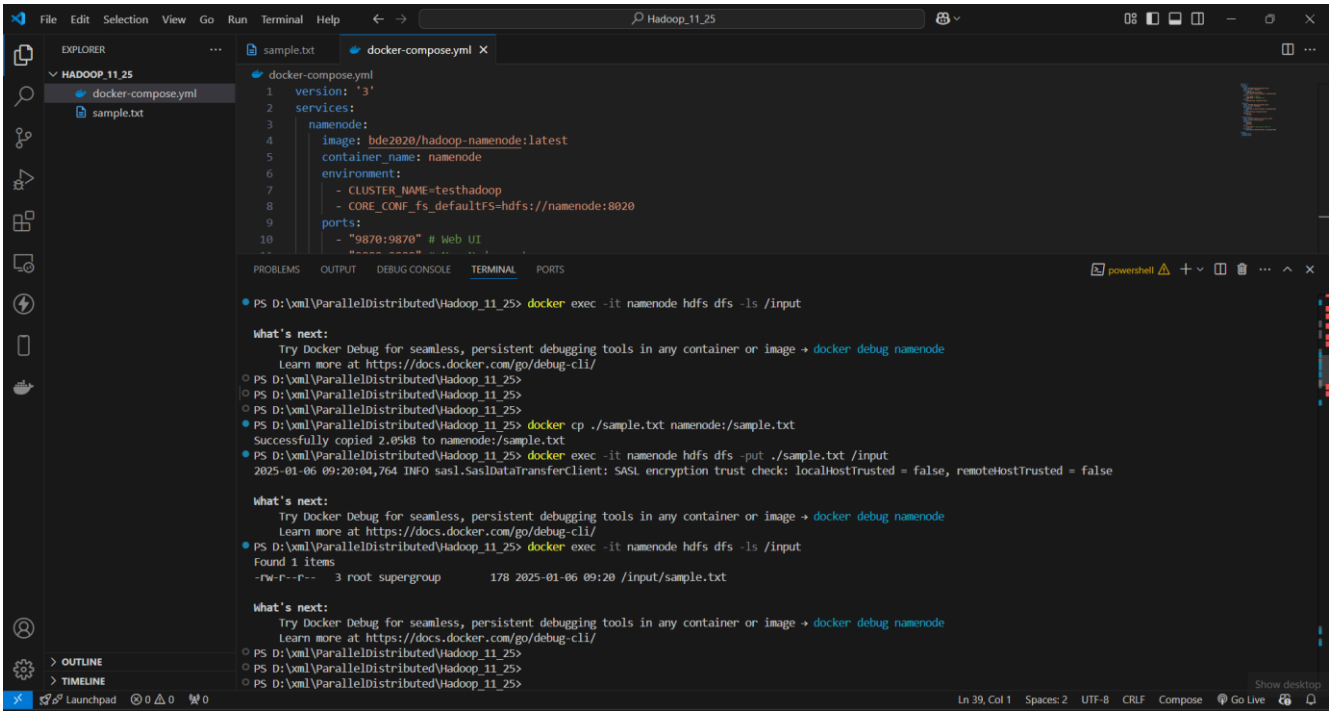
Assignment 3

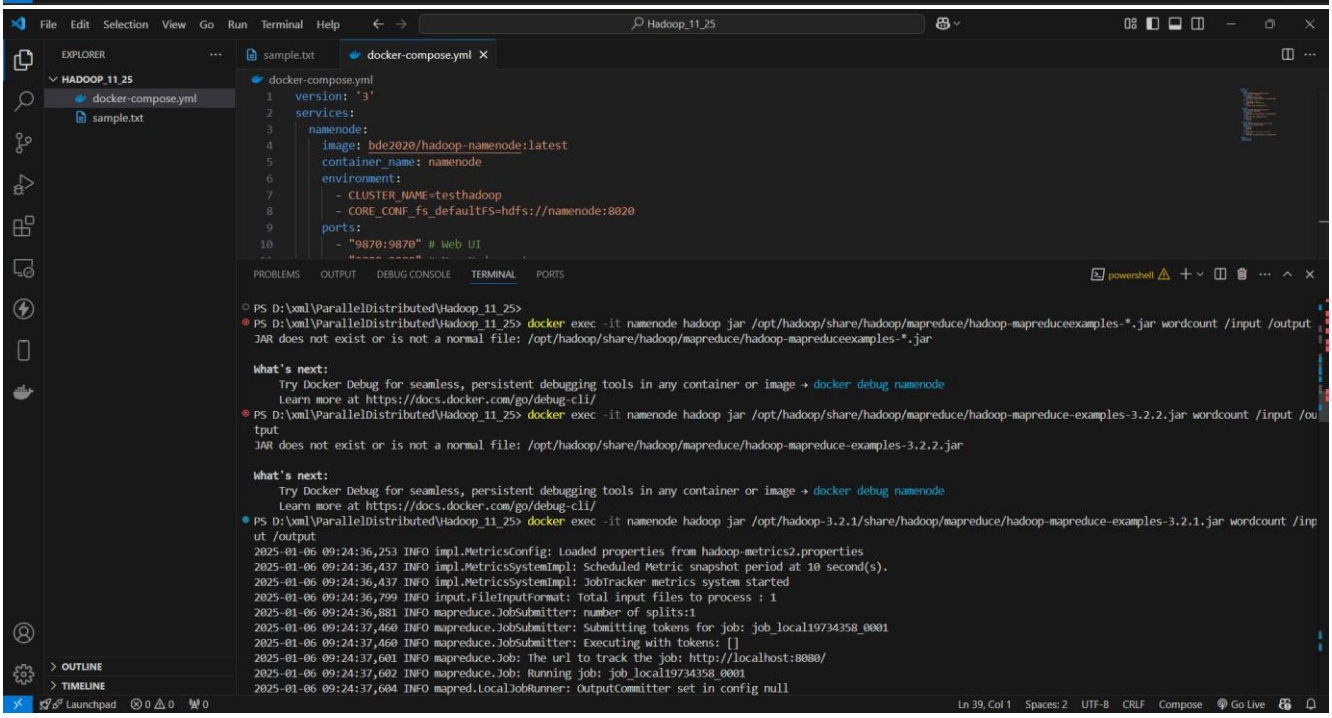
19APSE4325 – Tharindu Lakmal

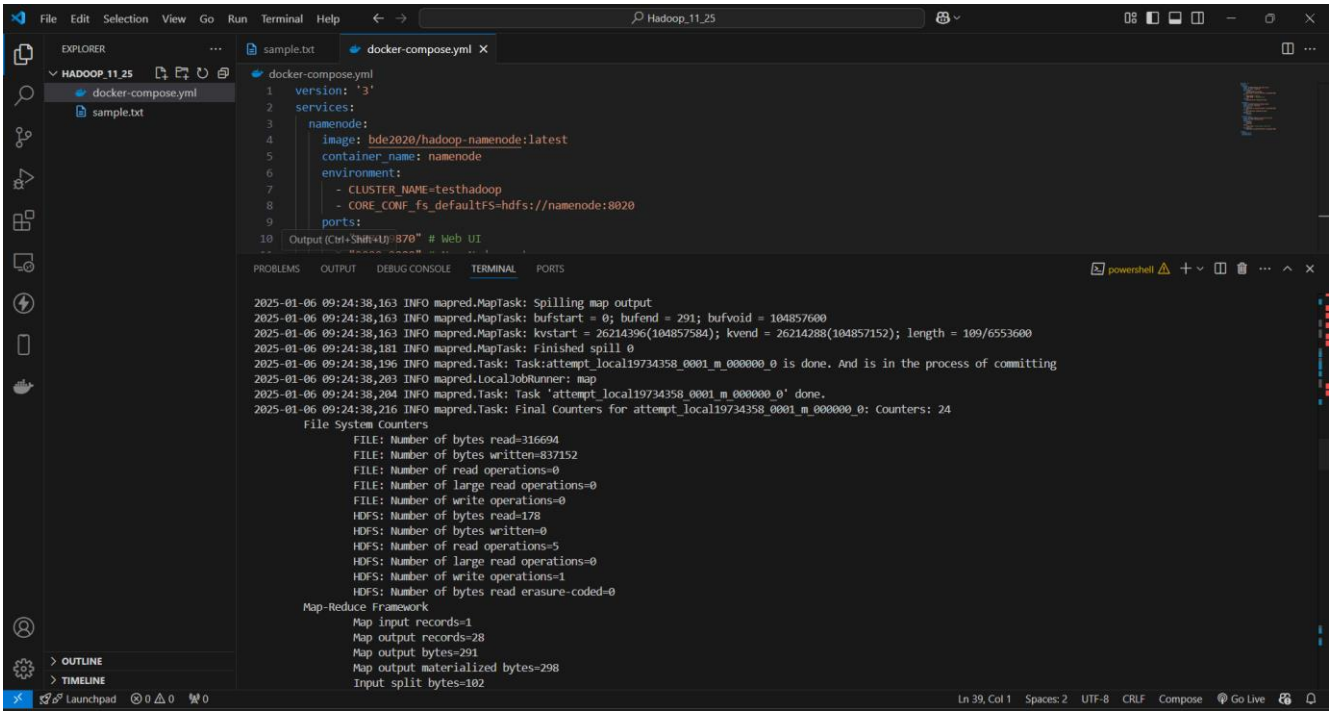
Question 2:

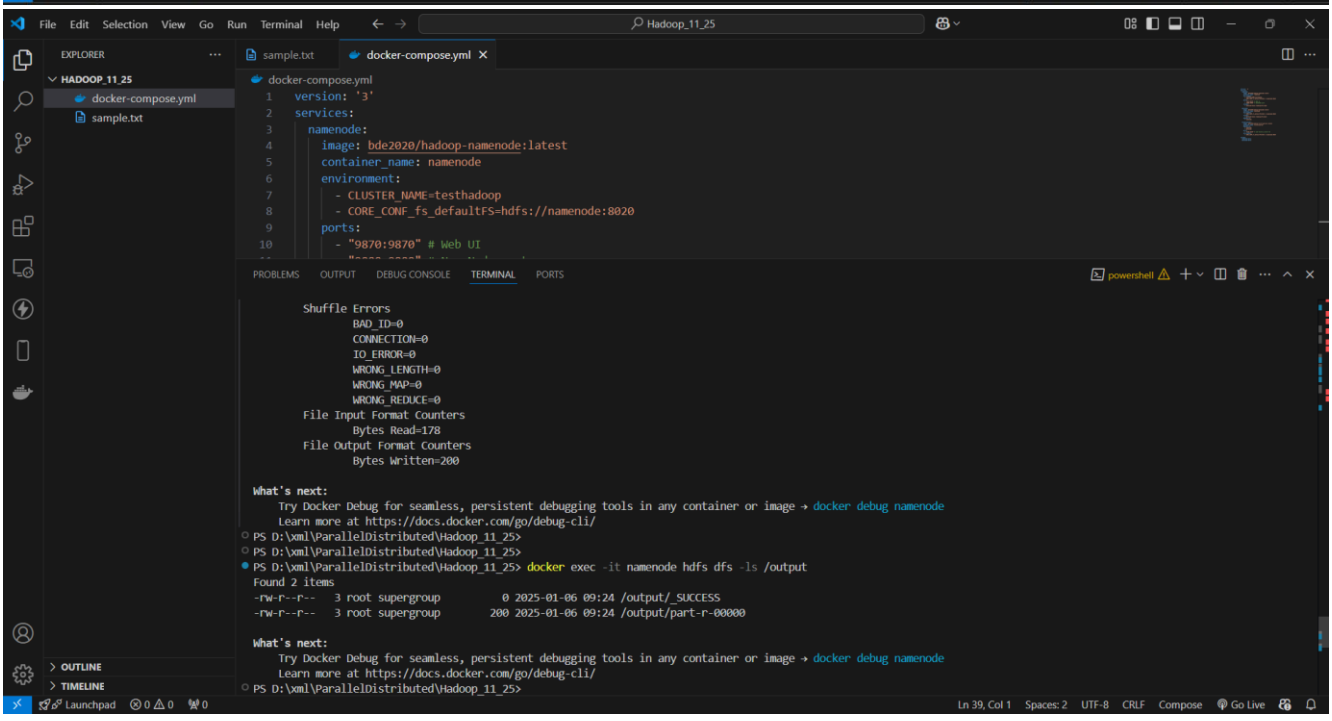


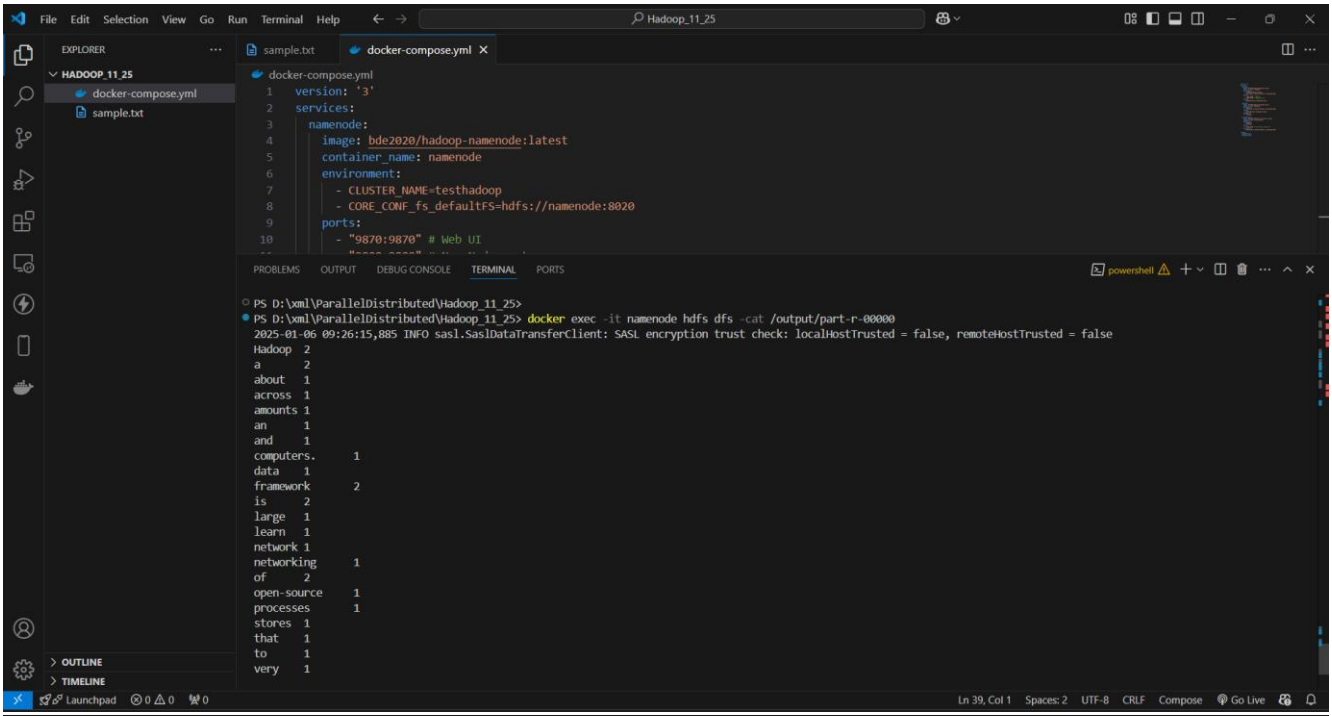


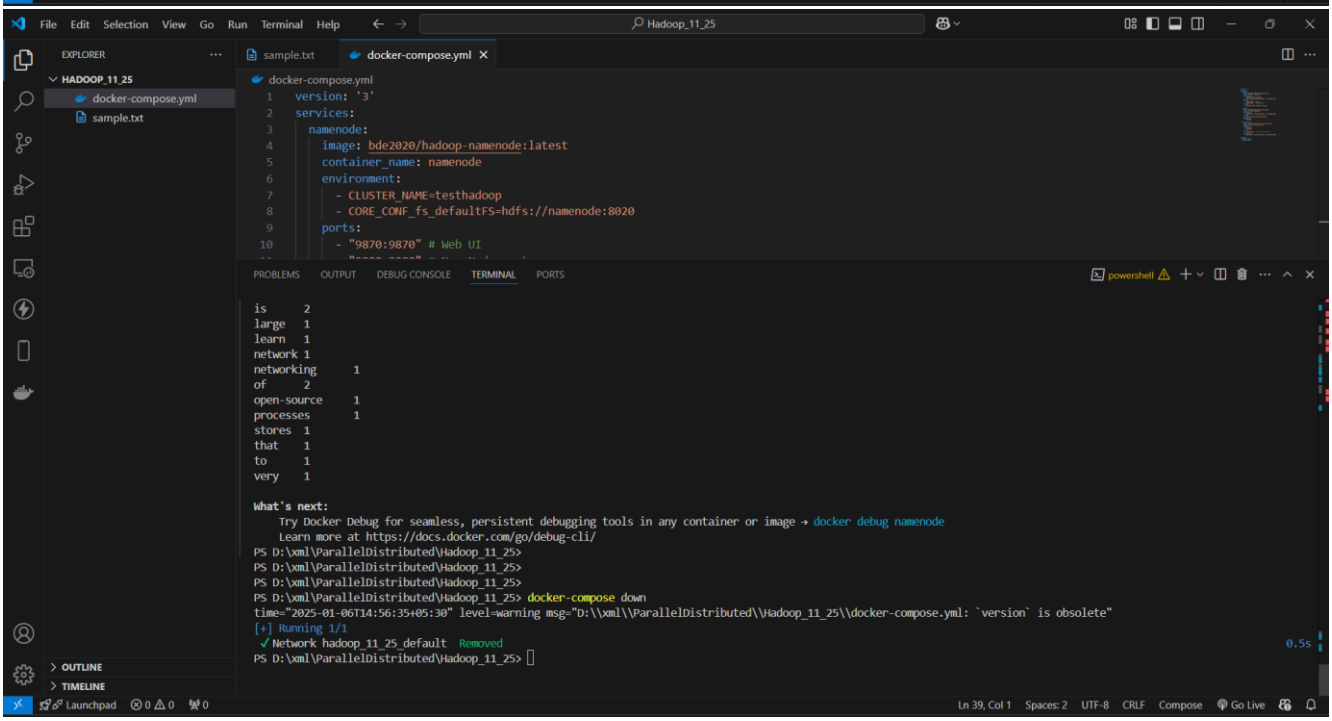

File Edit Selection View Go Run Terminal Help
Hadoop_11_25
EXPLORER
HADOOP_11_25
docker-compose.yml
sample.txt
sample.txt
docker-compose.yml
docker-compose.yml
1 version: '3'
2 services:
3 namenode:
4 image: bde2020/hadoop-namenode:latest
5 container_name: namenode
6 environment:
7 - CLUSTER_NAME=testhadoop
8 - CORE_CONF_fs_defaultFS=hdfs://namenode:8020
9 ports:
10 - "9870:9870" # Web UI
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS
powershell
PS D:\xml\ParallelDistributed\Hadoop_11_25> docker exec -it namenode hdfs dfs -ls /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/
PS D:\xml\ParallelDistributed\Hadoop_11_25>
PS D:\xml\ParallelDistributed\Hadoop_11_25>
PS D:\xml\ParallelDistributed\Hadoop_11_25>
PS D:\xml\ParallelDistributed\Hadoop_11_25> docker cp ./sample.txt namenode:/sample.txt
Successfully copied 2.05kB to namenode:/sample.txt
PS D:\xml\ParallelDistributed\Hadoop_11_25> docker exec -it namenode hdfs dfs -put ./sample.txt /input
2025-01-06 09:20:04,764 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/
PS D:\xml\ParallelDistributed\Hadoop_11_25> docker exec -it namenode hdfs dfs -ls /input
Found 1 items
-rw-r--r-- 3 root supergroup 178 2025-01-06 09:20 /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/
PS D:\xml\ParallelDistributed\Hadoop_11_25>
PS D:\xml\ParallelDistributed\Hadoop_11_25>
PS D:\xml\ParallelDistributed\Hadoop_11_25>
OUTLINE
TIMELINE
Launchpad
Ln 39, Col 1 Spaces: 2 UTF-8 CRLF Compose Go Live
Show desktop


File Edit Selection View Go Run Terminal Help
Hadoop_11_25
EXPLORER
HADOOP_11_25
docker-compose.yml
sample.txt
sample.txt
docker-compose.yml
docker-compose.yml
1 version: '3'
2 services:
3 namenode:
4 image: bde2020/hadoop-namenode:latest
5 container_name: namenode
6 environment:
7 - CLUSTER_NAME=testhadoop
8 - CORE_CONF_fs_defaultFS=hdfs://namenode:8020
9 ports:
10 - "9870:9870" # Web UI
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS
powershell
PS D:\xml\ParallelDistributed\Hadoop_11_25>
PS D:\xml\ParallelDistributed\Hadoop_11_25> docker exec -it namenode hadoop jar /opt/hadoop/share/hadoop/mapreduce/hadoop-mapreduceexamples-*.jar wordcount /input /output
JAR does not exist or is not a normal file: /opt/hadoop/share/hadoop/mapreduce/hadoop-mapreduceexamples-*.jar
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/
PS D:\xml\ParallelDistributed\Hadoop_11_25> docker exec -it namenode hadoop jar /opt/hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-3.2.2.jar wordcount /input /ou
tput
JAR does not exist or is not a normal file: /opt/hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-3.2.2.jar
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/
PS D:\xml\ParallelDistributed\Hadoop_11_25> docker exec -it namenode hadoop jar /opt/hadoop-3.2.1/share/hadoop/mapreduce/hadoop-mapreduce-examples-3.2.1.jar wordcount /inp
ut /output
2025-01-06 09:24:36,253 INFO impl.MetricsConfig: Loaded properties from hadoop-metrics2.properties
2025-01-06 09:24:36,437 INFO impl.MetricsSystemImpl: Scheduled Metric snapshot period at 10 second(s).
2025-01-06 09:24:36,437 INFO impl.MetricsSystemImpl: JobTracker metrics system started
2025-01-06 09:24:36,799 INFO input.FileInputFormat: Total input files to process : 1
2025-01-06 09:24:36,881 INFO mapreduce.JobSubmitter: number of splits:1
2025-01-06 09:24:37,460 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_local19734358_0001
2025-01-06 09:24:37,460 INFO mapreduce.JobSubmitter: Executing with tokens: []
2025-01-06 09:24:37,601 INFO mapreduce.Job: The url to track the job: http://localhost:8080/
2025-01-06 09:24:37,602 INFO mapreduce.Job: Running job: job_local19734358_0001
2025-01-06 09:24:37,604 INFO mapred.LocalJobRunner: OutputCommitter set in config null
OUTLINE
TIMELINE
Launchpad
Ln 39, Col 1 Spaces: 2 UTF-8 CRLF Compose Go Live


File Edit Selection View Go Run Terminal Help
Hadoop_11_25
EXPLORER
HADOOP_11_25
docker-compose.yml
sample.txt
sample.txt
docker-compose.yml
docker-compose.yml
1 version: '3'
2 services:
3 namenode:
4 image: bde2020/hadoop-namenode:latest
5 container_name: namenode
6 environment:
7 - CLUSTER_NAME=testhadoop
8 - CORE_CONF_fs_defaultFS=hdfs://namenode:8020
9 ports:
10 Output (Ctrl+Shift+U) 870" # Web UI
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS
powershell
2025-01-06 09:24:38,163 INFO mapred.MapTask: Spilling map output
2025-01-06 09:24:38,163 INFO mapred.MapTask: bufstart = 0; bufend = 291; bufvoid = 104857600
2025-01-06 09:24:38,163 INFO mapred.MapTask: kvstart = 26214396(104857584); kvend = 26214288(104857152); length = 109/6553600
2025-01-06 09:24:38,181 INFO mapred.MapTask: Finished spill 0
2025-01-06 09:24:38,196 INFO mapred.Task: Task:attempt_local19734358_0001_m_000000_0 is done. And is in the process of committing
2025-01-06 09:24:38,203 INFO mapred.LocalJobRunner: map
2025-01-06 09:24:38,204 INFO mapred.Task: Task 'attempt_local19734358_0001_m_000000_0' done.
2025-01-06 09:24:38,216 INFO mapred.Task: Final Counters for attempt_local19734358_0001_m_000000_0: Counters: 24
File System Counters
FILE: Number of bytes read=316694
FILE: Number of bytes written=837152
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=178
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
OUTLINE
TIMELINE
Launchpad
Ln 39, Col 1 Spaces: 2 UTF-8 CRLF Compose Go Live


File Edit Selection View Go Run Terminal Help
Hadoop_11_25
EXPLORER
HADOOP_11_25
docker-compose.yml
sample.txt
sample.txt
docker-compose.yml
docker-compose.yml
1 version: '3'
2 services:
3 namenode:
4 image: bde2020/hadoop-namenode:latest
5 container_name: namenode
6 environment:
7 - CLUSTER_NAME=testhadoop
8 - CORE_CONF_fs_defaultFS=hdfs://namenode:8020
9 ports:
10 - "9870:9870" # Web UI
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS
powershell
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=178
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/
PS D:\xml\ParallelDistributed\Hadoop_11_25>
PS D:\xml\ParallelDistributed\Hadoop_11_25>
PS D:\xml\ParallelDistributed\Hadoop_11_25> docker exec -it namenode hdfs dfs -ls /output
Found 2 items
-rw-r--r-- 3 root supergroup 0 2025-01-06 09:24 /output/_SUCCESS
-rw-r--r-- 3 root supergroup 200 2025-01-06 09:24 /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/
PS D:\xml\ParallelDistributed\Hadoop_11_25>
OUTLINE
TIMELINE
Launchpad
Ln 39, Col 1 Spaces: 2 UTF-8 CRLF Compose Go Live


File Edit Selection View Go Run Terminal Help
Hadoop_11_25
EXPLORER
HADOOP_11_25
docker-compose.yml
sample.txt
sample.txt
docker-compose.yml
docker-compose.yml
1 version: '3'
2 services:
3 namenode:
4 image: bde2020/hadoop-namenode:latest
5 container_name: namenode
6 environment:
7 - CLUSTER_NAME=testhadoop
8 - CORE_CONF_fs_defaultFS=hdfs://namenode:8020
9 ports:
10 - "9870:9870" # Web UI
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS
powershell
PS D:\xml\ParallelDistributed\Hadoop_11_25>
PS D:\xml\ParallelDistributed\Hadoop_11_25> docker exec -it namenode hdfs dfs -cat /output/part-r-00000
2025-01-06 09:26:15,885 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
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
OUTLINE
TIMELINE
Launchpad
Ln 39, Col 1 Spaces: 2 UTF-8 CRLF Compose Go Live


File Edit Selection View Go Run Terminal Help
Hadoop_11_25
EXPLORER
HADOOP_11_25
docker-compose.yml
sample.txt
sample.txt
docker-compose.yml
docker-compose.yml
1 version: '3'
2 services:
3 namenode:
4 image: bde2020/hadoop-namenode:latest
5 container_name: namenode
6 environment:
7 - CLUSTER_NAME=testhadoop
8 - CORE_CONF_fs_defaultFS=hdfs://namenode:8020
9 ports:
10 - "9870:9870" # Web UI
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS
powershell
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
Learn more at https://docs.docker.com/go/debug-cli/
PS D:\xml\ParallelDistributed\Hadoop_11_25>
PS D:\xml\ParallelDistributed\Hadoop_11_25>
PS D:\xml\ParallelDistributed\Hadoop_11_25>
PS D:\xml\ParallelDistributed\Hadoop_11_25> docker-compose down
time="2025-01-06T14:56:35+05:30" level=warning msg="D:\\xml\\ParallelDistributed\\Hadoop_11_25\\docker-compose.yml: `version` is obsolete"
[+] Running 1/1
✓ Network hadoop_11_25_default Removed
0.5s
PS D:\xml\ParallelDistributed\Hadoop_11_25>
OUTLINE
TIMELINE
Launchpad
Ln 39, Col 1 Spaces: 2 UTF-8 CRLF Compose Go Live