

ASSIGNMENT 02

Sabaragamuwa University of Sri Lanka

Faculty of Computing

Department of Software Engineering

SE6103 – Parallel & Distributed Systems

Name	: A . Vithurshanan
Reg. No	: 19APSE4294
Academic Period	: 03 rd Year 02 nd Semester
Degree Program	: Software Engineering
Due Date	: 25/11/2024

Docker Compose yml file

```
PS L:\Sabragauwa.lk\6th semester\Lecture slides\SE6103 Parallel and Distributed Systems\Practical\lab-sheet-4> docker pull bde2020/hadoop-namenode:2.0.0-hadoop2.7.4-java8
2.0.0-hadoop2.7.4-java8: Pulling from bde2020/hadoop-namenode
1b5c5213415d: Download complete
937af32bfdd6: Already exists
11669b117fd9: Already exists
bad51eca7232: Already exists
16cb0e2fb479: Download complete
46210e4b5de7: Download complete
1778b563ad5e: Already exists
879a04cc05e4: Download complete
f49cf87b52c1: Download complete
869f1ffa5626: Download complete
790d93d9ab2a: Already exists
60c95a1f3ed5: Already exists
d6de864ab017: Download complete
f9d504c9f21e: Already exists
f6ecf205e109: Already exists
f77335d29f88: Already exists
Digest: sha256:54a9482c51d4e701e530f15ef2e01ca2d3a15545760c42ea6ad0e65e8196c335
Status: Downloaded newer image for bde2020/hadoop-namenode:2.0.0-hadoop2.7.4-java8
docker.io/bde2020/hadoop-namenode:2.0.0-hadoop2.7.4-java8
```

Docker Compose up -d

```
PS L:\Sabragauwa.lk\6th semester\Lecture slides\SE6103 Parallel and Distributed Systems\Practical\lab-sheet-4> docker compose up -d
[+] Running 3/3
  ✓ nodemanager Pulled                                5.0s
  ✓ 8b04488c1f10 Download complete                    1.0s
  ✓ 0c66e03c2364 Download complete                    0.5s
[+] Running 3/7
  ✓ Network lab-sheet-4_hadoop                        Created          0.1s
  ✓ Volume "lab-sheet-4_namenode_data"                Created          0.0s
  ✓ Volume "lab-sheet-4_datanode_data"                Created          0.0s
  - Container namenode                               Creating         0.1s
  - Container nodemanager                            Creating         0.1s
  - Container resourcemanager                         Creating         0.1s
  - Container datanode                               Creating         0.1s
Error response from daemon: Conflict. The container name "/resourcemanager" is already in use by container "ebc9593cf5daf424155f38c796ced5f14e5337f877247d1b385b7db16067809e". You have to remove (or rename) that container to be able to reuse that name.
PS L:\Sabragauwa.lk\6th semester\Lecture slides\SE6103 Parallel and Distributed Systems\Practical\lab-sheet-4> docker compose up -d
[+] Running 4/4
  ✓ Container namenode                               Started          1.1s
  ✓ Container datanode                               Started          1.0s
  ✓ Container nodemanager                            Running          0.0s
  ✓ Container resourcemanager                         Started          1.1s
```

Docker network create hadoop

```
PS L:\Sabragauwa.lk\6th semester\Lecture slides\SE6103 Parallel and Distributed Systems\Practical\lab-sheet-4> docker network create hadoop
436aff3fc676db086743c7f1ff2cec12fa75b9edbed433393112a724191d9697
```

Docker run namenode

```
PS L:\Sabragauwa.lk\6th semester\Lecture slides\SE6103 Parallel and Distributed Systems\Practical\lab-sheet-4> docker rm namenode
namenode
PS L:\Sabragauwa.lk\6th semester\Lecture slides\SE6103 Parallel and Distributed Systems\Practical\lab-sheet-4> docker run -d --name namenode --net hadoop -p 50070:50070 -p 8088:8088 bde2020/hadoop-namenode:2.0.0-hadoop2.7.4-java8
fcdec6038b8252b44664c50335041400a09e87d138b6d17d1ad033e9e018eeae
```

Docker run datanode

```
PS L:\Sabragauwa.lk\6th semester\Lecture slides\SE6103 Parallel and Distributed Systems\Practical\lab-sheet-4> docker rm datanode
datanode
PS L:\Sabragauwa.lk\6th semester\Lecture slides\SE6103 Parallel and Distributed Systems\Practical\lab-sheet-4> docker run -d --name datanode --net hadoop bde2020/hadoop-datanode:2.0.0-hadoop2.7.4-java8
```

Docker run resourcemanagement

```
PS L:\Sabragauwa.lk\6th semester\Lecture slides\SE6103 Parallel and Distributed Systems\Practical\lab-sheet-4> docker run -d --name resourcemanager --net hadoop bde2020/hadoop-resourcemanager:2.0.0-hadoop2.7.4-java8
Unable to find image 'bde2020/hadoop-resourcemanager:2.0.0-hadoop2.7.4-java8' locally
2.0.0-hadoop2.7.4-java8: Pulling from bde2020/hadoop-resourcemanager
e14571a9cc62: Download complete
13a12985973c: Download complete
Digest: sha256:455b866d40a241ca4f28d33abf72bbf6b474765a0a6e72758f2bebcfd2f0c9a
Status: Downloaded newer image for bde2020/hadoop-resourcemanager:2.0.0-hadoop2.7.4-java8
ebc9593cf5daf424155f38c796ced5f14e5337f877247d1b385b7db16067809e
```

Execute namenode bash

```
S L:\Sabragauwa.lk\6th semester\Lecture slides\SE6103 Parallel and Distributed Systems\Practical\lab-sheet-4> docker exec -it namenode bash
root@73f0513cba31:/# hdfs namenode -format
4/11/25 16:41:55 INFO namenode.NameNode: STARTUP_MSG:
*****
TARTUP_MSG: Starting NameNode
TARTUP_MSG:   host = 73f0513cba31/172.18.0.5
TARTUP_MSG:   args = [-format]
TARTUP_MSG:   version = 2.7.4
TARTUP_MSG:   classpath = /etc/hadoop:/opt/hadoop-2.7.4/share/hadoop/common/lib/mockito-all-1.8.5.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/guava-11.0.2.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/zookeeper-3.4.6.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/activation-1.1.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/commons-codec-1.4.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/jsr305-3.0.0.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/commons-net-3.1.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/jetty-6.1.26.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/apacheds-i18n-2.0.0-M15.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/jettison-1.1.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/protobuf-java-2.5.0.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/jetty-sslengine-6.1.26.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/xmlenc-0.52.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/commons-beanutils-core-1.8.0.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/jersey-core-1.9.jar:/opt/hadoop-2.7.4/share/hadoop/common
```

```
24/11/25 16:41:56 INFO namenode.FSNamesystem: dfs.namenode.safemode.min.datanodes = 0
24/11/25 16:41:56 INFO namenode.FSNamesystem: dfs.namenode.safemode.extension = 30000
24/11/25 16:41:56 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.window.num.buckets = 10
24/11/25 16:41:56 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.num.users = 10
24/11/25 16:41:56 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.windows.minutes = 1,5,25
24/11/25 16:41:56 INFO namenode.FSNamesystem: Retry cache on namenode is enabled
24/11/25 16:41:56 INFO namenode.FSNamesystem: Retry cache will use 0.03 of total heap and retry cache entry expiry time is 600000 millis
24/11/25 16:41:56 INFO util.GSet: Computing capacity for map NameNodeRetryCache
24/11/25 16:41:56 INFO util.GSet: VM type = 64-bit
24/11/25 16:41:56 INFO util.GSet: 0.029999999329447746% max memory 889 MB = 273.1 KB
24/11/25 16:41:56 INFO util.GSet: capacity = 2^15 = 32768 entries
Re-format filesystem in Storage Directory /hadoop/dfs/name ? (Y or N) y
24/11/25 16:42:01 INFO namenode.FSImage: Allocated new BlockPoolId: BP-1734653817-172.18.0.5-1732552921805
24/11/25 16:42:01 INFO common.Storage: Storage directory /hadoop/dfs/name has been successfully formatted.
24/11/25 16:42:01 INFO namenode.FSImageFormatProtobuf: Saving image file /hadoop/dfs/name/current/fsimage.ckpt_000000000000000000 using no compression
24/11/25 16:42:01 INFO namenode.FSImageFormatProtobuf: Image file /hadoop/dfs/name/current/fsimage.ckpt_000000000000000000 of size 321 bytes saved in 0 seconds.
24/11/25 16:42:01 INFO namenode.NNStorageRetentionManager: Going to retain 1 images with txid >= 0
24/11/25 16:42:01 INFO util.ExitUtil: Exiting with status 0
24/11/25 16:42:01 INFO namenode.NameNode: SHUTDOWN_MSG:
/*****
SHUTDOWN_MSG: Shutting down NameNode at 73f0513cba31/172.18.0.5
*****/
```

hdfs namenode

```
root@73f0513cba31:/#  
root@73f0513cba31:/# hdfs namenode  
24/11/25 16:45:39 INFO namenode.NameNode: STARTUP_MSG:  
/*****  
STARTUP_MSG: Starting NameNode  
STARTUP_MSG: host = 73f0513cba31/172.18.0.5  
STARTUP_MSG: args = []  
STARTUP_MSG: version = 2.7.4  
STARTUP_MSG: classpath = /etc/hadoop:/opt/hadoop-2.7.4/share/hadoop/common/lib/mockito-all-  
1.8.5.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/guava-11.0.2.jar:/opt/hadoop-2.7.4/share/  
hadoop/common/lib/zookeeper-3.4.6.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/activation-1.  
1.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/commons-codec-1.4.jar:/opt/hadoop-2.7.4/share  
/hadoop/common/lib/jsr305-3.0.0.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/commons-net-3.1  
.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/jetty-6.1.26.jar:/opt/hadoop-2.7.4/share/hadoo  
p/common/lib/apacheds-i18n-2.0.0-M15.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/jettison-1  
.1.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/protobuf-java-2.5.0.jar:/opt/hadoop-2.7.4/sh  
are/hadoop/common/lib/jetty-sslengine-6.1.26.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/xml  
enc-0.52.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/commons-beanutils-core-1.8.0.jar:/opt  
/hadoop-2.7.4/share/hadoop/common/lib/jersey-core-1.9.jar:/opt/hadoop-2.7.4/share/hadoop/comm  
on/lib/hamcrest-core-1.3.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/commons-collections-3.  
2.2.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/jackson-jaxrs-1.9.13.jar:/opt/hadoop-2.7.4/  
share/hadoop/common/lib/jaxb-api-2.2.2.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/curator-  
client-2.7.1.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/jersey-server-1.9.jar:/opt/hadoop-  
2.7.4/share/hadoop/common/lib/commons-httpclient-3.1.jar:/opt/hadoop-2.7.4/share/hadoop/commo  
n/lib/hadoop-annotations-2.7.4.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/commons-digester  
-1.8.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/jsp-api-2.1.jar:/opt/hadoop-2.7.4/share/ha  
dooop/common/lib/jackson-core-asl-1.9.13.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/jackson  
-xc-1.9.13.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/jaxb-impl-2.2.3-1.jar:/opt/hadoop-2.  
7.4/share/hadoop/common/lib/convlet-api-2.5.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/com
```

hdfs datanode

```
root@73f0513cba31:/# hdfs datanode
24/11/25 16:46:43 INFO datanode.DataNode: STARTUP_MSG:
/*****
STARTUP_MSG: Starting DataNode
STARTUP_MSG:   host = 73f0513cba31/172.18.0.5
STARTUP_MSG:   args = []
STARTUP_MSG:   version = 2.7.4
STARTUP_MSG:   classpath = /etc/hadoop:/opt/hadoop-2.7.4/share/hadoop/common/lib/mockito-all-
1.8.5.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/guava-11.0.2.jar:/opt/hadoop-2.7.4/share/
hadoop/common/lib/zookeeper-3.4.6.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/activation-1.
1.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/commons-codec-1.4.jar:/opt/hadoop-2.7.4/share
hadoop/common/lib/jsr305-3.0.0.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/commons-net-3.1
.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/jetty-6.1.26.jar:/opt/hadoop-2.7.4/share/hadoo
p/common/lib/apacheds-i18n-2.0.0-M15.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/jettison-1
.1.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/protobuf-java-2.5.0.jar:/opt/hadoop-2.7.4/sh
are/hadoop/common/lib/jetty-sslengine-6.1.26.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/xm
lenc-0.52.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/commons-beanutils-core-1.8.0.jar:/opt
/hadoop-2.7.4/share/hadoop/common/lib/jersey-core-1.9.jar:/opt/hadoop-2.7.4/share/hadoop/comm
on/lib/hamcrest-core-1.3.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/commons-collections-3.
2.2.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/jackson-jaxrs-1.9.13.jar:/opt/hadoop-2.7.4/
share/hadoop/common/lib/jaxb-api-2.2.2.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/curator-
client-2.7.1.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/jersey-server-1.9.jar:/opt/hadoop-
2.7.4/share/hadoop/common/lib/commons-httpclient-3.1.jar:/opt/hadoop-2.7.4/share/hadoop/commo
n/lib/hadoop-annotations-2.7.4.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/commons-digester
-1.8.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/jsp-api-2.1.jar:/opt/hadoop-2.7.4/share/ha
```


yarn resourcemanager

```
root@73f0513cba31:/# yarn resourcemanager
24/11/25 16:48:54 INFO resourcemanager.ResourceManager: STARTUP_MSG:
/*****
STARTUP_MSG: Starting ResourceManager
STARTUP_MSG:   host = 73f0513cba31/172.18.0.5
STARTUP_MSG:   args = []
STARTUP_MSG:   version = 2.7.4
STARTUP_MSG:   classpath = /etc/hadoop:/etc/hadoop:/etc/hadoop:/opt/hadoop-2.7.4/share/hadoop
/common/lib/mockito-all-1.8.5.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/guava-11.0.2.jar:
/opt/hadoop-2.7.4/share/hadoop/common/lib/zookeeper-3.4.6.jar:/opt/hadoop-2.7.4/share/hadoop/
common/lib/activation-1.1.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/commons-codec-1.4.jar
:/opt/hadoop-2.7.4/share/hadoop/common/lib/jsr305-3.0.0.jar:/opt/hadoop-2.7.4/share/hadoop/co
mmon/lib/commons-net-3.1.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/jetty-6.1.26.jar:/opt/
hadoop-2.7.4/share/hadoop/common/lib/apacheds-i18n-2.0.0-M15.jar:/opt/hadoop-2.7.4/share/hadoo
p/common/lib/jettison-1.1.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/protobuf-java-2.5.0.
jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/jetty-sslengine-6.1.26.jar:/opt/hadoop-2.7.4/sh
are/hadoop/common/lib/xmlenc-0.52.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/commons-beanu
tils-core-1.8.0.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/jersey-core-1.9.jar:/opt/hadoop
-2.7.4/share/hadoop/common/lib/hamcrest-core-1.3.jar:/opt/hadoop-2.7.4/share/hadoop/common/li
b/commons-collections-3.2.2.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/jackson-jaxrs-1.9.1
3.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/jaxb-api-2.2.2.jar:/opt/hadoop-2.7.4/share/ha
dooop/common/lib/curator-client-2.7.1.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/jersey-ser
ver-1.9.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/commons-httpclient-3.1.jar:/opt/hadoop-
2.7.4/share/hadoop/common/lib/hadoop-annotations-2.7.4.jar:/opt/hadoop-2.7.4/share/hadoop/com
mon/lib/commons-digester-1.8.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/jsp-api-2.1.jar:/o
pt/hadoop-2.7.4/share/hadoop/common/lib/jackson-core-asl-1.9.13.jar:/opt/hadoop-2.7.4/share/h
adoop/common/lib/jackson-xc-1.9.13.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/jaxb-impl-2.
2.3-1.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/servlet-api-2.5.jar:/opt/hadoop-2.7.4/sha
re/hadoop/common/lib/commons-io-2.4.jar:/opt/hadoop-2.7.4/share/hadoop/common/lib/httpcore-4.4.1.jar
```


put tweets.txt file into /input

```
root@73f0513cba31:/# hadoop fs -put L:/Sabragauwa.lk/6th semester/Lecture slides/SE6103 Parallel and Distributed Systems/Practical/lab-sheet-4/tweets.txt /input/
put: `L:/Sabragauwa.lk/6th': No such file or directory
put: `semester/Lecture': No such file or directory
put: `slides/SE6103': No such file or directory
put: `Parallel': No such file or directory
put: `and': No such file or directory
put: `Distributed': No such file or directory
put: `Systems/Practical/lab-sheet-4/tweets.txt': No such file or directory
root@73f0513cba31:/# hadoop fs -put /tweets.txt /input/
24/11/25 16:53:46 INFO hdfs.DFSClient: Exception in createBlockOutputStream
java.net.ConnectException: Connection refused
    at sun.nio.ch.SocketChannelImpl.checkConnect(Native Method)
    at sun.nio.ch.SocketChannelImpl.finishConnect(SocketChannelImpl.java:717)
    at org.apache.hadoop.net.SocketIOWithTimeout.connect(SocketIOWithTimeout.java:206)
    at org.apache.hadoop.net.NetUtils.connect(NetUtils.java:531)
    at org.apache.hadoop.hdfs.DFSOutputStream.createSocketForPipeline(DFSOutputStream.java:1702)
    at org.apache.hadoop.hdfs.DFSOutputStream$DataStreamer.createBlockOutputStream(DFSOutputStream.java:1432)
    at org.apache.hadoop.hdfs.DFSOutputStream$DataStreamer.nextBlockOutputStream(DFSOutputStream.java:1385)
    at org.apache.hadoop.hdfs.DFSOutputStream$DataStreamer.run(DFSOutputStream.java:554)
24/11/25 16:53:46 INFO hdfs.DFSClient: Abandoning BP-172920158-172.18.0.5-1732552835618:blk_1073741825_1001
24/11/25 16:53:46 INFO hdfs.DFSClient: Excluding datanode DatanodeInfoWithStorage[172.18.0.5:50010,DS-d2f06710-b4c0-4371-a691-ccbcaae743f8,DISK]
root@73f0513cba31:/# hadoop fs -ls /input/
Found 1 items
-rw-r--r--  3 root supergroup      195 2024-11-25 16:53 /input/tweets.txt
root@73f0513cba31:/#
```

make java file as jar file

```
PS L:\Sabragauwa.lk\6th semester\Lecture slides\SE6103 Parallel and Distributed Systems\Practical\lab-sheet-4> javac -classpath "D:\hadoop\hadoop-3.2.4\share\hadoop\common\*;D:\hadoop\hadoop-3.2.4\share\hadoop\common\lib\*;D:\hadoop\hadoop-3.2.4\share\hadoop\hdfs\*;D:\hadoop\hadoop-3.2.4\share\hadoop\mapreduce\*;D:\hadoop\hadoop-3.2.4\share\hadoop\yarn\*" -d . HashtagCount.java
PS L:\Sabragauwa.lk\6th semester\Lecture slides\SE6103 Parallel and Distributed Systems\Practical\lab-sheet-4> jar cf hashtagcount.jar HashtagCount*.class
PS L:\Sabragauwa.lk\6th semester\Lecture slides\SE6103 Parallel and Distributed Systems\Practical\lab-sheet-4>
```

Execute namenode bash

```
PS L:\Sabragauwa.lk\6th semester\Lecture slides\SE6103 Parallel and Distributed System
s\Practical\lab-sheet-4> docker exec -it namenode bash
root@73f0513cba31:/# hadoop jar /hashtagcount.jar HashtagCount /input /output
Exception in thread "main" java.lang.UnsupportedClassVersionError: HashtagCount has be
en compiled by a more recent version of the Java Runtime (class file version 66.0), th
is version of the Java Runtime only recognizes class file versions up to 52.0
    at java.lang.ClassLoader.defineClass1(Native Method)
    at java.lang.ClassLoader.defineClass(ClassLoader.java:763)
    at java.security.SecureClassLoader.defineClass(SecureClassLoader.java:142)
    at java.net.URLClassLoader.defineClass(URLClassLoader.java:467)
    at java.net.URLClassLoader.access$100(URLClassLoader.java:73)
    at java.net.URLClassLoader$1.run(URLClassLoader.java:368)
    at java.net.URLClassLoader$1.run(URLClassLoader.java:362)
    at java.security.AccessController.doPrivileged(Native Method)
    at java.net.URLClassLoader.findClass(URLClassLoader.java:361)
    at java.lang.ClassLoader.loadClass(ClassLoader.java:424)
```

```
PS L:\Sabragauwa.lk\6th semester\Lecture slides\SE6103 Parallel and Distributed System
s\Practical\lab-sheet-4> javac --release 8 -classpath "D:\hadoop\hadoop-3.2.4\share\ha
doo\common\*;D:\hadoop\hadoop-3.2.4\share\hadoop\common\lib\*;D:\hadoop\hadoop-3.2.4\
share\hadoop\hdfs\*;D:\hadoop\hadoop-3.2.4\share\hadoop\mapreduce\*;D:\hadoop\hadoop-3
.2.4\share\hadoop\yarn\*" -d . HashtagCount.java
warning: [options] source value 8 is obsolete and will be removed in a future release
warning: [options] target value 8 is obsolete and will be removed in a future release
warning: [options] To suppress warnings about obsolete options, use -Xlint:-options.
3 warnings
PS L:\Sabragauwa.lk\6th semester\Lecture slides\SE6103 Parallel and Distributed System
s\Practical\lab-sheet-4> jar cf hashtagcount.jar HashtagCount*.class
PS L:\Sabragauwa.lk\6th semester\Lecture slides\SE6103 Parallel and Distributed System
s\Practical\lab-sheet-4> docker cp hashtagcount.jar namenode:/hashtagcount.jar
Successfully copied 5.12kB to namenode:/hashtagcount.jar
PS L:\Sabragauwa.lk\6th semester\Lecture slides\SE6103 Parallel and Distributed System
s\Practical\lab-sheet-4> docker exec -it namenode bash
root@73f0513cba31:/#
```

Get output from /input → `hadoop jar /hashtagcount.jar HashtagCount /input /output`

```
● PS L:\Sabragauwa.lk\6th semester\Lecture slides\SE6103 Parallel and Distributed System
s\Practical\lab-sheet-4> jar cf hashtagcount.jar HashtagCount*.class
● PS L:\Sabragauwa.lk\6th semester\Lecture slides\SE6103 Parallel and Distributed System
s\Practical\lab-sheet-4> docker cp hashtagcount.jar namenode:/hashtagcount.jar
Successfully copied 5.12kB to namenode:/hashtagcount.jar
○ PS L:\Sabragauwa.lk\6th semester\Lecture slides\SE6103 Parallel and Distributed System
s\Practical\lab-sheet-4> docker exec -it namenode bash
root@73f0513cba31:/# hadoop jar /hashtagcount.jar HashtagCount /input /output
24/11/25 17:11:42 INFO Configuration.deprecation: session.id is deprecated. Instead, u
se dfs.metrics.session-id
24/11/25 17:11:42 INFO jvm.JvmMetrics: Initializing JVM Metrics with processName=JobTr
acker, sessionId=
24/11/25 17:11:42 WARN mapreduce.JobResourceUploader: Hadoop command-line option pars
ing not performed. Implement the Tool interface and execute your application with ToolR
unner to remedy this.
24/11/25 17:11:42 INFO input.FileInputFormat: Total input paths to process : 1
24/11/25 17:11:42 INFO mapreduce.JobSubmitter: number of splits:1
24/11/25 17:11:42 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_local152
0175822_0001
24/11/25 17:11:42 INFO mapreduce.Job: The url to track the job: http://localhost:8080/
24/11/25 17:11:42 INFO mapreduce.Job: Running job: job_local1520175822_0001
24/11/25 17:11:42 INFO mapred.LocalJobRunner: OutputCommitter set in config null
24/11/25 17:11:42 INFO output.FileOutputCommitter: File Output Committer Algorithm ver
sion is 1
24/11/25 17:11:42 INFO mapred.LocalJobRunner: OutputCommitter is org.apache.hadoop.map
reduce.lib.output.FileOutputCommitter
24/11/25 17:11:42 INFO mapred.LocalJobRunner: Waiting for map tasks
24/11/25 17:11:42 INFO mapred.LocalJobRunner: Starting task: attempt_local1520175822_0
001_m_000000_0
24/11/25 17:11:42 INFO output.FileOutputCommitter: File Output Committer Algorithm ver
sion is 1
24/11/25 17:11:42 INFO mapred.Task: Using ResourceCalculatorProcessTree : [ ]
24/11/25 17:11:42 INFO mapred.MapTask: Processing split: hdfs://73f0513cba31:8020/inpu
t/tweets.txt:0+195
24/11/25 17:11:42 INFO mapred.MapTask: (EQUATOR) 0 kvi 26214396(104857584)
24/11/25 17:11:42 INFO mapred.MapTask: mapreduce.task.io.sort.mb: 100
```

Output → `hadoop fs -cat /output/part-r-000000`

```
Bytes Written=135
root@73f0513cba31:/# hadoop fs -cat /output/part-r-000000
#FutureTech      1
#GoTech 1
#Love 1
#TECH 1
#Tec 1
#Tech 6
#Tech4All 1
#TechBuzz 1
#TechInspiration 1
#Tech_Life 1
#Tech_Updates 1
#tech 1
root@73f0513cba31:/#
```