Assignment 03

Sabaragamuwa University of Sri Lanka

Faculty of Computing

Department of Software Engineering

SE6103 - Parallel and Distributed Systems

Name : W.A.T.G.R.Wijethunga

Reg. No : 19APSE4314

Degree Program : Software Engineering

Academic Period : 3rd Year 2nd Semester

Question 01

Assignment 03

Parallel and Distributed Systems

Question 01

01 Virtual Machines are run a full operating system along with the application, requiring significant resources (CPU, memory, storage) Docker containers, on the other hands, share the host os kernal and run only the application, mating them more lightweight

VMs are slower to startup because they require loading a full OS, while Docker containers are much faster to start since they only need the application environment

VMs provide stronger isolation but when it comes to the Docker containers there we cannot find stronger isolation process due to kernal sharing but run in isolated process

Containers are more portable across different environments because they carry only the applications and dependencies, not the os VMs are less portable because they depend on specific os versions and configurations

02 the -d flag stands for "detached mode" which means the container runs in the background, and the terminal is freed up for other commands It's important on running services like Nginx

03 -d runs Nginx in the background
Without -d, Nginx runs in the foreground
use -d for long-running services like Nginx

04. It maps port 80 inside the container to port 8080 on the host, making the container's services accessible externally.

05 Hadoop is an open-source framework for distributed storage and processing of large datasets across clusters

06 Advantages of Apache Spark over Hadoop:

In memory processing, faster than Hadoop's disk based MapReduce.

High level API's in multiple languages

Supports batch, real-time, machine learning and graph processing

Question 2

```
PS D:\Academic\Semester 6\Docker\Assigment 3> docker-compose up -d
time="2025-01-06T14:12:25+05:30" level=warning msg="D:\\Academic\\Semester 6\\Docker\\Assigment 3\\docker-compose.yml: the attribute `ve
rsion` is obsolete, it will be ignored, please remove it to avoid potential confusion"
[+] Running 3/3
 ✓ Container historyserver  Started                                                                    6.1s
 ✓ Container namenode       Started                                                                    4.4s
 ✓ Container datanode       Started                                                                    0.6s
PS D:\Academic\Semester 6\Docker\Assigment 3>
```

```
PROBLEMS  OUTPUT  DEBUG CONSOLE  TERMINAL  PORTS  COMMENTS                                                    powershell

PS D:\Academic\Semester 6\Docker\Assigment 3> docker exec -it namenode hdfs dfs -put /sample.txt /input
2025-01-06 08:58:20,929 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = fal
se

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:\Academic\Semester 6\Docker\Assigment 3> docker exec -it namenode hdfs dfs -ls /input
-rw-r--r--   3 root supergroup        181 2025-01-06 08:58 /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:\Academic\Semester 6\Docker\Assigment 3>
```

```
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/
-rw-r--r--   3 root supergroup        181 2025-01-06 08:58 /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:\Academic\Semester 6\Docker\Assigment 3> docker cp /path/to/wordcount.jar namenode:/opt/hadoop-3.2.1/wordcount.jar
CreateFile D:\path: The system cannot find the file specified.
```

```
PROBLEMS  OUTPUT  DEBUG CONSOLE  TERMINAL  PORTS  COMMENTS                                                    powershell

2025-01-06 09:09:43,141 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = fal
se
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
    Learn more at https://docs.docker.com/go/debug-cli/
PS D:\Academic\Semester 6\Docker\Assigment 3>
```