ASSIGNMENT

Department of Software Engineering
Faculty of Computing
Sabaragamuwa university of Sri Lanka

Parallel and Distributed System-SE6103

Name :   Rasika Wedaarachchi
Reg. No :   19APSE4331
Academic Period : 3rd Year 2nd Semester

Setting up the Distributed Hadoop Cluster

Prepare the Docker Compose File


```yaml
    1    version: '3'
    2    services:
    3
    4      namenode:
    5        image: bde2020/hadoop-namenode:latest
    6        container_name: namenode
    7        environment:
    8          - CLUSTER_NAME=ShopSmartCluster
    9          - CORE_CONF_fs_defaultFS=hdfs://namenode:8020
   10        ports:
   11          - "9870:9870"
   12          - "9000:9000"
   13        volumes:
   14          - namenode-data:/hadoop/dfs/namenode
   15
   16      datanode:
   17        image: bde2020/hadoop-datanode:latest
   18        container_name: datanode
   19        environment:
   20          - CORE_CONF_fs_defaultFS=hdfs://namenode:8020
   21        volumes:
   22          - datanode-data:/hadoop/dfs/datanode
   23        depends_on:
   24          - namenode
   25
   26      historyserver:
   27        image: bde2020/hadoop-historyserver:latest # Replace with a valid image if necessary
   28        container_name: historyserver
   29        depends_on:
   30          - namenode
   31          - datanode
   32        ports:
```

Deploy the Cluster


```
PS D:\Docker\Docker - Practicles_02> docker-compose up -d
time="2024-11-25T14:04:25+05:30" level=warning msg="D:\\Docker\\Docker - Practicles_02\\docker-compose.yml: the attribute `version` is obs.
[+] Running 2/2
 X historyserver Error pull access denied for bde2020/yarn-historyserver, repository does not exist or may r...          4.1s
 X datanode Error        context canceled                                                                                 4.1s
Error response from daemon: pull access denied for bde2020/yarn-historyserver, repository does not exist or may require 'docker login'
PS D:\Docker\Docker - Practicles_02> docker-compose up -d
time="2024-11-25T14:06:53+05:30" level=warning msg="D:\\Docker\\Docker - Practicles_02\\docker-compose.yml: the attribute `version` is obs
olete, it will be ignored, please remove it to avoid potential confusion"
[+] Running 8/2
 √ historyserver Pulled                                                                                                   7.5s
 √ datanode Pulled                                                                                                        7.5s
[+] Running 4/4
[+] Running 4/5er-practicles_02_default        Created                                                                   0.2s
[+] Running 5/5er-practicles_02_default        Created                                                                   0.2s
 √ Network docker-practicles_02_default         Created                                                                   0.2s
[+] Running 5/6er-practicles_02_namenode-data" Created                                                                   0.0s
 √ Network docker-practicles_02_default         Created                                                                   0.2s
[+] Running 6/6er-practicles_02_namenode-data" Created                                                                   0.0s
 √ Network docker-practicles_02_default         Created                                                                   0.2s
[+] Running 5/6er-practicles_02_namenode-data" Created                                                                   0.0s
 √ Network docker-practicles_02_default         Created                                                                   0.2s
[+] Running 5/6er-practicles_02_namenode-data" Created                                                                   0.0s
 √ Network docker-practicles_02_default         Created                                                                   0.2s
[+] Running 5/6er-practicles_02_namenode-data" Created                                                                   0.0s
 √ Network docker-practicles_02_default         Created                                                                   0.2s
[+] Running 5/6er-practicles_02_namenode-data" Created                                                                   0.0s
 √ Network docker-practicles_02_default         Created                                                                   0.2s
[+] Running 5/6er-practicles_02_namenode-data" Created                                                                   0.0s
```

Verify Cluster Status

Uploading Data to HDFS

Upload Data to HDFS

Running a MapReduce Job

Run the WordCount MapReduce Job

Monitor the Job

```
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:\Docker\Docker - Practicles_02> docker exec -it namenode hdfs dfs -ls /new_output
Found 2 items
-rw-r--r--   3 root supergroup          0 2024-11-25 09:45 /new_output/_SUCCESS
-rw-r--r--   3 root supergroup        205 2024-11-25 09:45 /new_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:\Docker\Docker - Practicles_02> docker exec -it namenode hdfs dfs -cat /new_output/part-r-00000
2024-11-25 10:03:21,729 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remote
ostTrusted = false
If       1
Legend,  1
Mr.      1
Wayne.   1
a        1
an       1
and      1
become   1
can't    1
devote   1
else     1
entirely…       1
ideal,   1
if       2
just     1
make     1
man,     1
more     1
```

Clean Up the Cluster

```
PS D:\Docker\Docker - Practicles_02> docker-compose down
time="2024-11-25T15:36:38+05:30" level=warning msg="D:\\Docker\\Docker - Practicles_02\\docker-compose.yml: the attribut
e `version` is obsolete, it will be ignored, please remove it to avoid potential confusion"
[+] Running 4/4
 ✓ Container historyserver                 Removed                                                              11.7s
 ✓ Container datanode                      Removed                                                              10.5s
 ✓ Container namenode                      Removed                                                              12.0s
 ✓ Network docker-practicles_02_default    Removed                                                               0.4s
PS D:\Docker\Docker - Practicles_02>
```