

Assignment I

Hadoop Exercise

19APSE4314

Tharushi Wijethunga

Department of Software Engineering

First Attempt

Step 1

Prerequisites

Docker: Ensure Docker is installed and running on your system. You can check if Docker is installed by running:

bash

Copy code

```
docker --version
```

Output

```
Microsoft Windows [Version 10.0.22631.4468]
(c) Microsoft Corporation. All rights reserved.

C:\Users\DELL>docker --version
Docker version 25.0.3, build 4deb441
```

Step 2

Step 1: Pull the Hadoop Docker Image

1. Choose a Hadoop Docker Image:

- Many Docker images are available for Hadoop single-node setups. For this lab, we'll use the `bde2020/hadoop-namenode` image from the Big Data Europe repository on Docker Hub.

Pull the Docker image:

```
docker pull bde2020/hadoop-namenode:latest
```

Then it generates error

```
C:\Users\DELL>docker pull bde2020/hadoop-namenode:latest
error during connect: this error may indicate that the docker daemon is not running: Post "http://%2F%2Fpipe%2Fdocker_engine/v1.24/images/create?fromImage=bde2020%2Fhadoop-namenode&tag=latest": open //pipe/docker_engine: The system cannot find the file specified.

C:\Users\DELL>docker pull bde2020/hadoop-namenode
Using default tag: latest
error during connect: this error may indicate that the docker daemon is not running: Post "http://%2F%2Fpipe%2Fdocker_engine/v1.20/images/create?fromImage=bde2020%2Fhadoop-namenode&tag=latest": open //pipe/docker_engine: The system cannot find the file specified.
```

Start docker

```
C:\Users\DELL>docker pull bde2020/hadoop-namenode
Using default tag: latest
latest: Pulling from bde2020/hadoop-namenode
3192219af084: Pulling fs layer
7127a1d8cced: Pulling fs layer
3192219af084: Pull complete
7127a1d8cced: Pull complete
883a89599999: Pull complete
```

```

3192219af094: Pull complete
7127ad8cced: Pull complete
883a09599980: Pull complete
77920a3e82af: Pull complete
92329e81aec4: Pull complete
f372218fec59: Pull complete
aa53513fe997: Pull complete
8b1800105b98: Pull complete
c1a04ale49c0: Pull complete
a65690a54a76: Pull complete
a29cc756c786: Pull complete
abf352b16846: Pull complete
ddd5a409e99: Pull complete
Digest: sha256:fd741188051320646cf6412435efc0919e1fb2ac5bd376c5366272fc261381e
Status: Downloaded newer image for bde2020/hadoop-namenode:latest
docker.io/bde2020/hadoop-namenode:latest

What's Next?
View a summary of image vulnerabilities and recommendations • docker scout quickview bde2020/hadoop-namenode

```

Step 3

2. Verify the Download:

List Docker images to confirm the Hadoop image is downloaded:

`docker images`

```

C:\Users\DELL>docker images
REPOSITORY          TAG         IMAGE ID      CREATED       SIZE
nginx               latest     3b25b682ea02  6 weeks ago  192MB
demos               v1        a92efe036fb2  8 months ago  187MB
<none>             <none>    ef91dd23e603  8 months ago  187MB
demo               v2        92b11f67b02b  9 months ago  187MB
redis              alpine     2b774af0fcf   10 months ago  41MB
alpine             3.12.0    8c94880f0f35  13 months ago  7.3MB
hello-world        latest     d1c90e258dcb  18 months ago  13.3kB
demo               latest     35c6c67ec39b  19 months ago  3.61MB
bde2020/hadoop-namenode latest     b638307a2119  4 years ago   1.37GB

```

Step 4

Step 2: Start the Hadoop Container

1. Run the Container:

Start a container with the necessary configurations to act as a single-node Hadoop cluster:

```

docker run -it --name hadoop-cluster -p 9870:9870 -p 8088:8088 -p
50070:50070 bde2020/hadoop-namenode:latest /bin/bash

```

```
C:\Users\DELL>docker run -it --name hadoop-cluster -p 9870:9870 -p 8080:8080 -p 50070:50070 bde2020/hadoop-namenode:latest /bin/bash
Configuring core
- Setting fs.defaultFS=hdfs://42d338f08c67:8020
Configuring hdfs
- Setting dfs.namenode.name.dir=file:///hadoop/dfs/name
Configuring yarn
Configuring httpfs
Configuring hadoop
Configuring mapred
Configuring for multihomed network
root@42d338f08c67:/#
```

Step 5

2. Start Hadoop Services:

Once inside the container's shell, start the Hadoop services:

```
start-all.sh
```

-
- This will initialize HDFS and YARN.

```
root@42d338f08c67:/# find / -name start-all.sh
/opt/hadoop-3.2.1/sbin/start-all.sh
root@42d338f08c67:/# /opt/hadoop-3.2.1/sbin/start-all.sh
Starting namenodes on [42d338f08c67]
ERROR: Attempting to operate on hdfs namenode as root
ERROR: but there is no HDFS_NAMENODE_USER defined. Aborting operation.
Starting datanodes
ERROR: Attempting to operate on hdfs datanode as root
ERROR: but there is no HDFS_DATANODE_USER defined. Aborting operation.
Starting secondary namenodes [42d338f08c67]
ERROR: Attempting to operate on hdfs secondarynamenode as root
```

From this moments onwards the above error remains same

Second Attempt

Then I have start from the beginning in the powershell

```
Windows PowerShell
Copyright (C) Microsoft Corporation. All rights reserved.

Install the latest PowerShell for new features and improvements! https://aka.ms/PSWindows

PS C:\Users\DELL> docker --version
Docker version 25.0.3, build bdebf41
PS C:\Users\DELL> docker pull bde2020/hadoop-namenode:latest
latest: Pulling from bde2020/hadoop-namenode
Digest: sha256:fd7741180851320846c6f612635efc8919e1fb2ac5bd376c5366272fc261381e
Status: Image is up to date for bde2020/hadoop-namenode:latest
docker.io/bde2020/hadoop-namenode:latest

What's Next?
View a summary of image vulnerabilities and recommendations + docker scout quickview bde2020/hadoop-namenode:latest
PS C:\Users\DELL> docker images
REPOSITORY          TAG             IMAGE ID        CREATED         SIZE
nginx                latest          3b25b682ea82   6 weeks ago    192MB
demo                 v1             a92efe036fb2   8 months ago   187MB
<none>              <none>         ef93dd23e603   8 months ago   187MB
demo                 v2             92b11f67042b   9 months ago   187MB
redis                alpine         287766fc4fcf   18 months ago  41MB
alpine               3.18.4         8ca4688f4f35   13 months ago  7.34MB
hello-world          latest          d2c94e258dcb   18 months ago  13.3kB
demo                 latest         35c6c67ec35b   19 months ago  5.61MB
bde2020/hadoop-namenode latest          b638307a2119   4 years ago    1.37GB
PS C:\Users\DELL> docker run -it --name hadoop-cluster -p 9870:9870 -p 8088:8088 -p
docker: 'run-it--name' is not a docker command.
See 'docker --help'
PS C:\Users\DELL> docker run -it --name hadoop-cluster -p 9870:9870 -p 8088:8088 -p 50070:50070 bde2020/hadoop-namenode:latest /bin/bash
docker: 'run-it--name' is not a docker command.
See 'docker --help'
PS C:\Users\DELL> docker run -it --name hadoop-cluster -p 9870:9870 -p 8088:8088 -p 50070:50070 bde2020/hadoop-namenode:
latest /bin/bash
unknown shorthand flag: '-' in --name
See 'docker run --help'
PS C:\Users\DELL> docker run -it --name hadoop-cluster -p 9870:9870 -p 8088:8088 -p 50070:50070 bde2020/hadoop-namenode:latest /bin/bash
docker: Error response from daemon: Conflict. The container name "/hadoop-cluster" is already in use by container "42d338f88c679c2457ce5276eea8453718dec3eff
a8f05d235ba9a53538cedc4". You have to remove (or rename) that container to be able to reuse that name.
See 'docker run --help'.
PS C:\Users\DELL> docker run --help
```

```
Windows PowerShell

Usage: docker run [OPTIONS] IMAGE [COMMAND] [ARG...]

Create and run a new container from an image

Aliases:
  docker container run, docker run

Options:
  --add-host list          Add a custom host-to-IP mapping
                           (host:ip)
  --annotation map         Add an annotation to the
                           container (passed through to the
                           OCI runtime) (default map[])
  -a, --attach list        Attach to STDIN, STDOUT or STDERR
  --blkio-weight uint16    Block IO (relative weight),
                           between 18 and 1000, or 0 to
                           disable (default 0)
  --blkio-weight-device list
                           Block IO weight (relative device
                           weight) (default [])
  --cap-add list           Add Linux capabilities
  --cap-drop list          Drop Linux capabilities
  --cgroup-parent string    Optional parent cgroup for the
                           container
  --cgroupns string        Cgroup namespace to use
                           (host|private)
                           'host': Run the container in
                           the Docker host's cgroup
                           namespace
                           'private': Run the container in
                           its own private cgroup namespace
                           '' : Use the cgroup
                           namespace as configured by the
                           default-cgroupns-mode
                           option on the daemon (default)
  --cidfile string         Write the container ID to the file
  --cpu-period int         Limit CPU CFS (Completely Fair
                           Scheduler) period
  --cpu-quota int          Limit CPU CFS (Completely Fair
                           Scheduler) quota
  --cpu-rt-period int      Limit CPU real-time period in
```



```
Windows PowerShell
try { service -n toryservice-3.2.1.jar:/opt/hadoop-3.2.1/share/hadoop/yarn/hadoop-yarn-server-nodemanager-3.2.1.jar:/opt/hadoop-3.2.1/share/hadoop/yarn/hadoop-yarn-server-te
sts-3.2.1.jar:/opt/hadoop-3.2.1/share/hadoop/yarn/hadoop-yarn-common-3.2.1.jar:/opt/hadoop-3.2.1/share/hadoop/yarn/hadoop-yarn-server-timeline-pluginstorage
-3.2.1.jar
STARTUP_MSG: build = https://gitbox.apache.org/repos/asf/hadoop.git -r b3cbb6467e22ea829b3888f4b7b81d97e8bf3842; compiled by 'rohithsharma' on 2019-09-1
8T15:56Z
STARTUP_MSG: java = 1.8.0_232
*****
2024-11-18 11:30:36,451 INFO namenode.NameNode: registered UNIX signal handlers for [TERM, HUP, INT]
2024-11-18 11:30:36,589 INFO namenode.NameNode: createNameNode [-Format]
Formatting using clusterid: CID-a7f2373b-9967-46a2-a88d-d1481b37fcf0
2024-11-18 11:30:37,020 INFO namenode.FSEditLog: Edit logging is async:true
2024-11-18 11:30:37,040 INFO namenode.FSNamesystem: KeyProvider: null
2024-11-18 11:30:37,043 INFO namenode.FSNamesystem: fsLock is fair: true
2024-11-18 11:30:37,044 INFO namenode.FSNamesystem: Detailed lock hold time metrics enabled: false
2024-11-18 11:30:37,051 INFO namenode.FSNamesystem: fsOwner = root (auth:SIMPLE)
2024-11-18 11:30:37,051 INFO namenode.FSNamesystem: supergroup = supergroup
2024-11-18 11:30:37,051 INFO namenode.FSNamesystem: isPermissionEnabled = true
2024-11-18 11:30:37,051 INFO namenode.FSNamesystem: HA Enabled: false
2024-11-18 11:30:37,104 INFO common.Util: dfs.datanode.fileio.profiling.sampling.percentage set to 0. Disabling file IO profiling
2024-11-18 11:30:37,125 INFO blockmanagement.DatanodeManager: dfs.block.invalidate.limit: configured=1000, counted=60, effected=1000
2024-11-18 11:30:37,125 INFO blockmanagement.DatanodeManager: dfs.namenode.datanode.registration.ip-hostname-check=true
2024-11-18 11:30:37,133 INFO blockmanagement.BlockManager: dfs.namenode.startup.delay.block.deletion.sec is set to 000:00:00:00.800
2024-11-18 11:30:37,134 INFO blockmanagement.BlockManager: The block deletion will start around 2024 Nov 18 11:38:37
2024-11-18 11:30:37,144 INFO util.GSet: Computing capacity for map BlocksMap
2024-11-18 11:30:37,144 INFO util.GSet: VM type = 64-bit
2024-11-18 11:30:37,146 INFO util.GSet: 2.0% max memory 855.5 MB = 17.1 MB
2024-11-18 11:30:37,146 INFO util.GSet: capacity = 2^21 = 2097152 entries
2024-11-18 11:30:37,155 INFO blockmanagement.BlockManager: Storage policy satisfier is disabled
2024-11-18 11:30:37,155 INFO blockmanagement.BlockManager: dfs.block.access.token.enable = false
2024-11-18 11:30:37,165 INFO Configuration.deprecation: No unit for dfs.namenode.safemode.extension(30000) assuming MILLISECONDS
2024-11-18 11:30:37,165 INFO blockmanagement.BlockManagerSafeMode: dfs.namenode.safemode.threshold-pct = 0.0000000126746033
2024-11-18 11:30:37,165 INFO blockmanagement.BlockManagerSafeMode: dfs.namenode.safemode.min.datanodes = 0
2024-11-18 11:30:37,165 INFO blockmanagement.BlockManagerSafeMode: dfs.namenode.safemode.extension = 30000
2024-11-18 11:30:37,166 INFO blockmanagement.BlockManager: defaultReplication = 3
2024-11-18 11:30:37,166 INFO blockmanagement.BlockManager: maxReplication = 512
2024-11-18 11:30:37,166 INFO blockmanagement.BlockManager: minReplication = 1
2024-11-18 11:30:37,166 INFO blockmanagement.BlockManager: maxReplicationStreams = 2
2024-11-18 11:30:37,166 INFO blockmanagement.BlockManager: redundancyRecheckInterval = 3000ms
2024-11-18 11:30:37,166 INFO blockmanagement.BlockManager: encryptDataTransfer = false
2024-11-18 11:30:37,166 INFO blockmanagement.BlockManager: maxNumBlocksToLog = 1000
2024-11-18 11:30:37,197 INFO namenode.FSDirectory: GLOBAL serial map: bits=29 maxEntries=536870911
```

```
Windows PowerShell
2024-11-18 11:30:37,197 INFO namenode.FSDirectory: GLOBAL serial map: bits=29 maxEntries=536870911
2024-11-18 11:30:37,197 INFO namenode.FSDirectory: USER serial map: bits=24 maxEntries=16777215
2024-11-18 11:30:37,197 INFO namenode.FSDirectory: GROUP serial map: bits=24 maxEntries=16777215
2024-11-18 11:30:37,197 INFO namenode.FSDirectory: XATTR serial map: bits=24 maxEntries=16777215
2024-11-18 11:30:37,212 INFO util.GSet: Computing capacity for map INodeMap
2024-11-18 11:30:37,212 INFO util.GSet: VM type = 64-bit
2024-11-18 11:30:37,213 INFO util.GSet: 1.0% max memory 855.5 MB = 8.6 MB
2024-11-18 11:30:37,213 INFO util.GSet: capacity = 2^28 = 1844576 entries
2024-11-18 11:30:37,214 INFO namenode.FSDirectory: ACLs enabled? false
2024-11-18 11:30:37,214 INFO namenode.FSDirectory: POSIX ACL inheritance enabled? true
2024-11-18 11:30:37,214 INFO namenode.FSDirectory: XAttrs enabled? true
2024-11-18 11:30:37,214 INFO namenode.NameNode: Caching file names occurring more than 18 times
2024-11-18 11:30:37,219 INFO snapshot.SnapshotManager: Loaded config captureOpenFiles: false, skipCaptureAccessTimeOnlyChange: false, snapshotDiffAllowSnapR
ootDescendant: true, maxSnapshotLimit: 65536
2024-11-18 11:30:37,221 INFO snapshot.SnapshotManager: SkipList is disabled
2024-11-18 11:30:37,227 INFO util.GSet: Computing capacity for map cachedBlocks
2024-11-18 11:30:37,227 INFO util.GSet: VM type = 64-bit
2024-11-18 11:30:37,227 INFO util.GSet: 0.25% max memory 855.5 MB = 2.1 MB
2024-11-18 11:30:37,227 INFO util.GSet: capacity = 2^18 = 262144 entries
2024-11-18 11:30:37,235 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.window.num.buckets = 10
2024-11-18 11:30:37,236 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.num.users = 10
2024-11-18 11:30:37,236 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.windows.minutes = 1,5,25
2024-11-18 11:30:37,241 INFO namenode.FSNamesystem: Retry cache on namenode is enabled
2024-11-18 11:30:37,241 INFO namenode.FSNamesystem: Retry cache will use 0.03 of total heap and retry cache entry expiry time is 600000 millis
2024-11-18 11:30:37,246 INFO util.GSet: Computing capacity for map NameNodeRetryCache
2024-11-18 11:30:37,246 INFO util.GSet: VM type = 64-bit
2024-11-18 11:30:37,247 INFO util.GSet: 0.029999999329447746% max memory 855.5 MB = 262.8 MB
2024-11-18 11:30:37,247 INFO util.GSet: capacity = 2^15 = 32768 entries
2024-11-18 11:30:37,369 INFO namenode.FSImage: Allocated new BlockPoolId: BP-1676819128-172.17.8.2-1731929437271
2024-11-18 11:30:37,338 INFO common.Storage: Storage directory /hadoop/dfs/name has been successfully formatted.
2024-11-18 11:30:37,365 INFO namenode.FSImageFormatProtobuf: Saving image file /hadoop/dfs/name/current/fsimage.ckpt_000000000000000000 using no compressio
n
2024-11-18 11:30:37,462 INFO namenode.FSImageFormatProtobuf: Image file /hadoop/dfs/name/current/fsimage.ckpt_000000000000000000 of size 399 bytes saved in
8 seconds.
2024-11-18 11:30:37,480 INFO namenode.NNStorageRetentionManager: Going to retain 1 images with txid >= 0
2024-11-18 11:30:37,488 INFO namenode.FSImage: FSImageSaver clean checkpoint: txid=0 when meet shutdown.
2024-11-18 11:30:37,488 INFO namenode.NameNode: SHUTDOWN_MSG:
/*****
SHUTDOWN_MSG: Shutting down NameNode at 04b75784f6f9/172.17.0.2
*****/
root@04b75784f6f9:/# hdfs namenode &
```



```

Windows PowerShell
2024-11-18 11:30:53.082 INFO server.AbstractConnector: Started ServerConnector@709fb071(HTTP/1.1,[http/1.1]){0.0.0.0:9876}
2024-11-18 11:30:53.082 INFO server.Server: Started #1252ms
2024-11-18 11:30:53.232 WARN namenode.FSNamesystem: Only one image storage directory (dfs.namenode.name.dir) configured. Beware of data loss due to lack of redundant storage directories!
2024-11-18 11:30:53.232 WARN namenode.FSNamesystem: Only one namespace edits storage directory (dfs.namenode.edits.dir) configured. Beware of data loss due to lack of redundant storage directories!
2024-11-18 11:30:53.281 INFO namenode.FSEditLog: Edit logging is async=true
2024-11-18 11:30:53.306 INFO namenode.FSNamesystem: KeyProvider: null
2024-11-18 11:30:53.354 INFO namenode.FSNamesystem: fslock is fair: true
2024-11-18 11:30:53.354 INFO namenode.FSNamesystem: Detailed lock hold time metrics enabled: false
2024-11-18 11:30:53.368 INFO namenode.FSNamesystem: fsOwner = root (auth:SIMPLE)
2024-11-18 11:30:53.368 INFO namenode.FSNamesystem: supergroup = supergroup
2024-11-18 11:30:53.368 INFO namenode.FSNamesystem: isPermissionEnabled = true
2024-11-18 11:30:53.368 INFO namenode.FSNamesystem: HA Enabled: false
2024-11-18 11:30:53.396 INFO common.Util: dfs.datanode.fileio.profiling.sampling.percentage set to 0. Disabling file IO profiling
2024-11-18 11:30:53.405 INFO blockmanagement.DatanodeManager: dfs.block.invalidate.limit: configured=1000, counted=00, effected=1000
2024-11-18 11:30:53.405 INFO blockmanagement.DatanodeManager: dfs.namenode.datanode.registration.ip-hostname-check=true
2024-11-18 11:30:53.408 INFO blockmanagement.DatanodeManager: dfs.namenode.startup.delay.block.deletion.sec is set to 000:00:00:00.000
2024-11-18 11:30:53.409 INFO blockmanagement.BlockManager: The block deletion will start around 2024 Nov 18 11:30:53
2024-11-18 11:30:53.411 INFO util.GSet: Computing capacity for map BlocksMap
2024-11-18 11:30:53.411 INFO util.GSet: VM type = 64-bit
2024-11-18 11:30:53.412 INFO util.GSet: 2.8% max memory 835.5 MB = 17.1 MB
2024-11-18 11:30:53.412 INFO util.GSet: capacity = 2^21 = 2097152 entries
2024-11-18 11:30:53.421 INFO blockmanagement.BlockManager: Storage policy satisfier is disabled
2024-11-18 11:30:53.421 INFO blockmanagement.BlockManager: dfs.block.access.token.enable = false
2024-11-18 11:30:53.429 INFO Configuration.deprecation: No unit for dfs.namenode.safemode.extension(30000) assuming WILLI5EC0ND5
2024-11-18 11:30:53.429 INFO blockmanagement.BlockManagerSafeMode: dfs.namenode.safemode.threshold-pct = 0.9990000128746013
2024-11-18 11:30:53.429 INFO blockmanagement.BlockManagerSafeMode: dfs.namenode.safemode.min.datanodes = 0
2024-11-18 11:30:53.429 INFO blockmanagement.BlockManagerSafeMode: dfs.namenode.safemode.extension = 30000
2024-11-18 11:30:53.430 INFO blockmanagement.BlockManager: defaultReplication = 3
2024-11-18 11:30:53.430 INFO blockmanagement.BlockManager: maxReplication = 512
2024-11-18 11:30:53.431 INFO blockmanagement.BlockManager: minReplication = 1
2024-11-18 11:30:53.431 INFO blockmanagement.BlockManager: maxReplicationStreams = 2
2024-11-18 11:30:53.431 INFO blockmanagement.BlockManager: redundancyRecheckInterval = 3000ms
2024-11-18 11:30:53.432 INFO blockmanagement.BlockManager: encryptDataTransfer = false
2024-11-18 11:30:53.433 INFO blockmanagement.BlockManager: maxNumBlocksToLog = 1000
2024-11-18 11:30:53.450 INFO namenode.FSDirectory: GLOBAL serial map: bits=29 maxEntries=536878911
2024-11-18 11:30:53.450 INFO namenode.FSDirectory: USER serial map: bits=24 maxEntries=16777215
2024-11-18 11:30:53.450 INFO namenode.FSDirectory: GROUP serial map: bits=24 maxEntries=16777215
2024-11-18 11:30:53.450 INFO namenode.FSDirectory: XATTR serial map: bits=24 maxEntries=16777215
2024-11-18 11:30:53.460 INFO util.GSet: Computing capacity for map INodeMap

```

```

Windows PowerShell
2024-11-18 11:30:54.353 INFO hdfs.StateChange: STATE* UnderReplicatedBlocks has 0 blocks
2024-11-18 11:30:54.361 INFO blockmanagement.BlockManager: Total number of blocks = 0
2024-11-18 11:30:54.361 INFO blockmanagement.BlockManager: Number of invalid blocks = 0
2024-11-18 11:30:54.362 INFO blockmanagement.BlockManager: Number of under-replicated blocks = 0
2024-11-18 11:30:54.362 INFO blockmanagement.BlockManager: Number of over-replicated blocks = 0
2024-11-18 11:30:54.362 INFO blockmanagement.BlockManager: Number of blocks being written = 0
2024-11-18 11:30:54.362 INFO hdfs.StateChange: STATE* Replication Queue initialization scan for invalid, over- and under-replicated blocks completed in 9 ms.
ec
2024-11-18 11:30:54.386 INFO ipc.Server: IPC Server listener on 8020: starting
2024-11-18 11:30:54.381 INFO ipc.Server: IPC Server Responder: starting
2024-11-18 11:30:54.388 INFO namenode.NameNode: NameNode RPC up at: 04b75784f6f9/172.17.0.2:8020
2024-11-18 11:30:54.392 INFO namenode.FSNamesystem: Starting services required for active state
2024-11-18 11:30:54.393 INFO namenode.FSDirectory: Initializing quota with 4 thread(s)
2024-11-18 11:30:54.493 INFO namenode.FSDirectory: Quota initialization completed in 11 milliseconds
name space=1
storage space=0
storage types=RAM_DISK=0, SSD=0, DISK=0, ARCHIVE=0, PROVIDED=0
2024-11-18 11:30:54.411 INFO blockmanagement.CacheReplicationMonitor: Starting CacheReplicationMonitor with interval 30000 milliseconds

root@04b75784f6f9:/# hdfs datanode &
[2] 293
root@04b75784f6f9:/# 2024-11-18 11:32:47.740 INFO datanode.DataNode: STARTUP_MSG:
/*****
STARTUP_MSG: Starting DataNode
STARTUP_MSG: host = 04b75784f6f9/172.17.0.2
STARTUP_MSG: args = []
STARTUP_MSG: version = 3.2.1
STARTUP_MSG: classpath = /etc/hadoop:/opt/hadoop-3.2.1/share/hadoop/common/lib/checker-qual-2.5.2.jar:/opt/hadoop-3.2.1/share/hadoop/common/lib/jackson-jar
xrs-1.9.13.jar:/opt/hadoop-3.2.1/share/hadoop/common/lib/commons-net-3.6.jar:/opt/hadoop-3.2.1/share/hadoop/common/lib/accessors-smart-1.2.jar:/opt/hadoop-3
.2.1/share/hadoop/common/lib/guava-27.0-jre.jar:/opt/hadoop-3.2.1/share/hadoop/common/lib/kerb-simplekdc-1.0.1.jar:/opt/hadoop-3.2.1/share/hadoop/common/lib
/curator-framework-2.13.0.jar:/opt/hadoop-3.2.1/share/hadoop/common/lib/jsr305-3.0.2.jar:/opt/hadoop-3.2.1/share/hadoop/common/lib/curator-recipes-2.13.0.jar
/opt/hadoop-3.2.1/share/hadoop/common/lib/metrics-core-2.2.0.jar:/opt/hadoop-3.2.1/share/hadoop/common/lib/kerb-crypto-1.0.1.jar:/opt/hadoop-3.2.1/share/h
adoop/common/lib/jackson-core-2.9.8.jar:/opt/hadoop-3.2.1/share/hadoop/common/lib/jack-0.1.54.jar:/opt/hadoop-3.2.1/share/hadoop/common/lib/jetty-webapp-9.3
.24.v20180605.jar:/opt/hadoop-3.2.1/share/hadoop/common/lib/gson-2.2.4.jar:/opt/hadoop-3.2.1/share/hadoop/common/lib/jetty-xml-9.3.24.v20180605.jar:/opt/had
oop-3.2.1/share/hadoop/common/lib/jackson-mapper-asl-1.9.13.jar:/opt/hadoop-3.2.1/share/hadoop/common/lib/kerb-core-1.0.1.jar:/opt/hadoop-3.2.1/share/hadoop
/common/lib/jetty-servlet-9.3.24.v20180605.jar:/opt/hadoop-3.2.1/share/hadoop/common/lib/jersey-server-1.10.jar:/opt/hadoop-3.2.1/share/hadoop/common/lib/av
ro-1.7.7.jar:/opt/hadoop-3.2.1/share/hadoop/common/lib/commons-beanutils-1.9.3.jar:/opt/hadoop-3.2.1/share/hadoop/common/lib/kerby-plex-1.0.1.jar:/opt/had
oop-3.2.1/share/hadoop/common/lib/jul-to-slf4j-1.7.25.jar:/opt/hadoop-3.2.1/share/hadoop/common/lib/kerby-util-1.0.1.jar:/opt/hadoop-3.2.1/share/hadoop/common
/lib/kerb-client-1.0.1.jar:/opt/hadoop-3.2.1/share/hadoop/common/lib/jetty-security-9.3.24.v20180605.jar:/opt/hadoop-3.2.1/share/hadoop/common/lib/woodstox-
core-5.0.3.jar:/opt/hadoop-3.2.1/share/hadoop/common/lib/kerby-asn1-1.0.1.jar:/opt/hadoop-3.2.1/share/hadoop/common/lib/kerby-protocol-1.0.1.jar:/opt/hadoop-3.2
.1/share/hadoop/common/lib/jackson-databind-2.9.8.jar:/opt/hadoop-3.2.1/share/hadoop/common/lib/kerby-asn1-1.0.1.jar:/opt/hadoop-3.2.1/share/hadoop/common/l
*****/

```

```

Windows PowerShell
2024-11-18 11:36:46,373 INFO placement.MultiNodeSortingManager: Starting NodeSortingService=MultiNodeSortingManager
2024-11-18 11:36:46,398 INFO ipc.CallQueueManager: Using callQueue: class java.util.concurrent.LinkedBlockingQueue, queueCapacity: 5000, scheduler: class or
g.apache.hadoop.ipc.DefaultRpcScheduler, ipcBackoff: false.
2024-11-18 11:36:46,404 INFO ipc.Server: Starting Socket Reader #1 for port 8031
2024-11-18 11:36:46,485 INFO pb.RpcServerFactoryPBImpl: Adding protocol org.apache.hadoop.yarn.server.api.ResourceTrackerPB to the server
2024-11-18 11:36:46,489 INFO ipc.Server: IPC Server Responder: starting
2024-11-18 11:36:46,411 INFO ipc.Server: IPC Server Listener on 8031: starting
2024-11-18 11:36:46,493 INFO util.JvmPauseMonitor: Starting JVM pause monitor
2024-11-18 11:36:46,520 INFO ipc.CallQueueManager: Using callQueue: class java.util.concurrent.LinkedBlockingQueue, queueCapacity: 5000, scheduler: class or
g.apache.hadoop.ipc.DefaultRpcScheduler, ipcBackoff: false.
2024-11-18 11:36:46,531 INFO ipc.Server: Starting Socket Reader #1 for port 8038
2024-11-18 11:36:46,540 INFO pb.RpcServerFactoryPBImpl: Adding protocol org.apache.hadoop.yarn.api.ApplicationMasterProtocolPB to the server
2024-11-18 11:36:46,541 INFO ipc.Server: IPC Server Responder: starting
2024-11-18 11:36:46,543 INFO ipc.Server: IPC Server Listener on 8038: starting
2024-11-18 11:36:46,885 INFO ipc.CallQueueManager: Using callQueue: class java.util.concurrent.LinkedBlockingQueue, queueCapacity: 5000, scheduler: class or
g.apache.hadoop.ipc.DefaultRpcScheduler, ipcBackoff: false.
2024-11-18 11:36:46,810 INFO ipc.Server: Starting Socket Reader #1 for port 8032
2024-11-18 11:36:46,814 INFO pb.RpcServerFactoryPBImpl: Adding protocol org.apache.hadoop.yarn.api.ApplicationClientProtocolPB to the server
2024-11-18 11:36:46,820 INFO ipc.Server: IPC Server Responder: starting
2024-11-18 11:36:46,824 INFO ipc.Server: IPC Server Listener on 8032: starting
2024-11-18 11:36:47,804 INFO resourceManager.ResourceManager: Transitioned to active state
2024-11-18 11:37:11,235 INFO resourceManager.ResourceTrackerService: NodeManager from node 04b75784f6f9:48639 (caPort: 48639 httpPort: 8042) registered with capabi
lity: <memory:8192, vCores:8>, assigned nodeId 04b75784f6f9:48639
2024-11-18 11:37:11,242 INFO rmNode.RMNodeImpl: 04b75784f6f9:48639 Node Transitioned from NEW to RUNNING
2024-11-18 11:37:11,256 INFO security.NMContainerTokenSecretManager: Rolling master-key for container-tokens, got key with id 174488183
2024-11-18 11:37:11,257 INFO security.NMTokenSecretManagerInNM: Rolling master-key for container-tokens, got key with id 862015861
2024-11-18 11:37:11,258 INFO nodemanager.NodeStatusUpdaterImpl: Registered with ResourceManager as 04b75784f6f9:48639 with total resource of <memory:8192, v
Cores:8>
2024-11-18 11:37:11,270 INFO capacity.CapacityScheduler: Added node 04b75784f6f9:48639 clusterResource: <memory:8192, vCores:8>
2024-11-18 11:45:49,363 INFO localizer.ResourceLocalizationService: Cache Size Before Clean: 0, Total Deleted: 0, Public Deleted: 0, Private Deleted: 0
2024-11-18 11:46:46,382 INFO scheduler.AbstractYarnScheduler: Release request cache is cleaned up
2024-11-18 11:55:49,318 INFO localizer.ResourceLocalizationService: Cache Size Before Clean: 0, Total Deleted: 0, Public Deleted: 0, Private Deleted: 0
2024-11-18 12:43:49,313 INFO localizer.ResourceLocalizationService: Cache Size Before Clean: 0, Total Deleted: 0, Public Deleted: 0, Private Deleted: 0
root@04b75784f6f9:/#

```