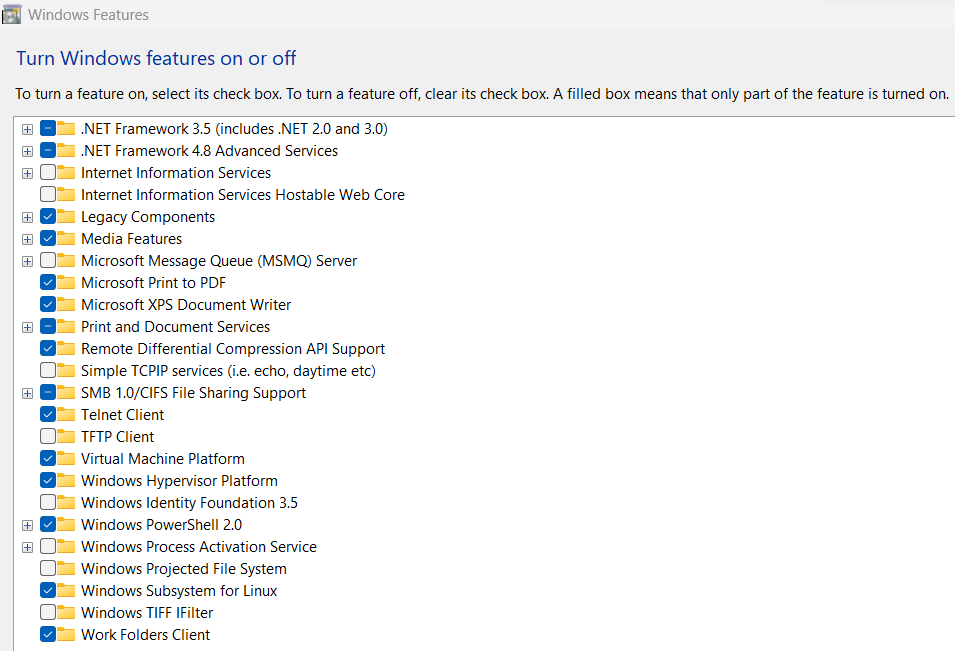
**Kadapala Rakesh Reddy -M22AI608-Assignment-1 (Fractal - 3)-** [**Algorithms for Big Data**](https://classroom.google.com/c/NjA4OTM0MDA5MDk5)

**Q>Create a Hadoop multi-node cluster using either Virtual Machines or Docker.**

**Permissions Required for windows to install Docker.**

Start🡪 Control Panel🡪Programs🡪Programs and Features🡪Turn on Windows features on or off🡪

**Enable the Following:**

* windows subsystem for Linux
* Windows Hypervisor Platform

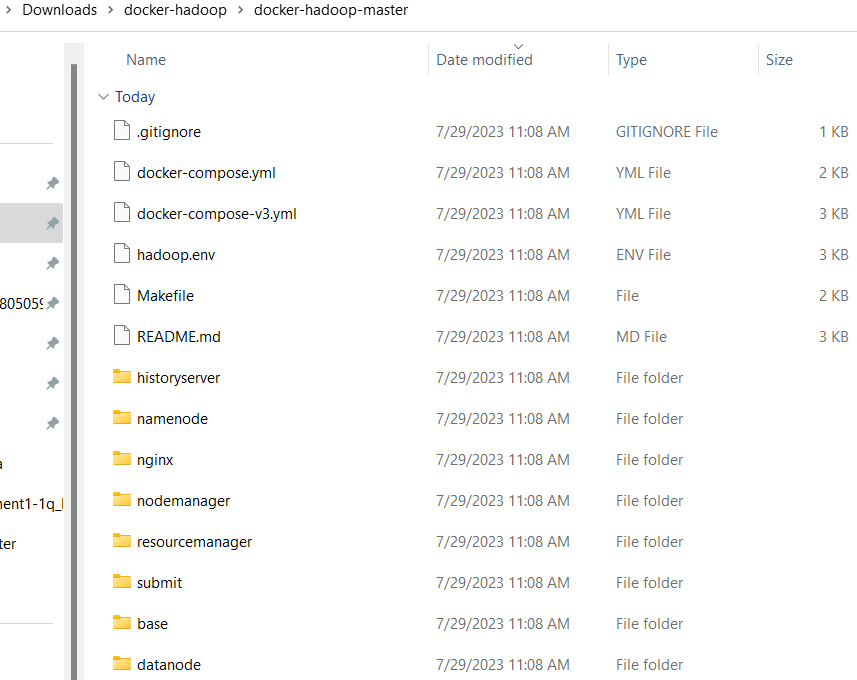
Restart the System.

Docker Installation Link 64bit: <https://docs.docker.com/desktop/install/windows-install/>

Download the Dataset from Classroom: <https://drive.google.com/file/d/1Ja6fD9ico9emdX8YNBioqOK1fD4L8g97/view?usp=sharing>

Extract the ZIP File

Open the folder –go to specific files –in file manger path ----remove path and write Powershell



**Docker-compose.yml contains 5 Data nodes:**

version: "3"

services:

namenode:

image: bde2020/hadoop-namenode:2.0.0-hadoop3.2.1-java8

container\_name: namenode

restart: always

ports:

- 9870:9870

- 9000:9000

volumes:

- hadoop\_namenode:/hadoop/dfs/name

environment:

- CLUSTER\_NAME=test

env\_file:

- ./hadoop.env

datanode01:

image: bde2020/hadoop-datanode:2.0.0-hadoop3.2.1-java8

container\_name: datanode01

restart: always

volumes:

- hadoop\_datanode01:/hadoop/dfs/data01

environment:

SERVICE\_PRECONDITION: "namenode:9870"

env\_file:

- ./hadoop.env

datanode02:

image: bde2020/hadoop-datanode:2.0.0-hadoop3.2.1-java8

container\_name: datanode02

restart: always

volumes:

- hadoop\_datanode02:/hadoop/dfs/data02

environment:

SERVICE\_PRECONDITION: "namenode:9870"

env\_file:

- ./hadoop.env

datanode03:

image: bde2020/hadoop-datanode:2.0.0-hadoop3.2.1-java8

container\_name: datanode03

restart: always

volumes:

- hadoop\_datanode03:/hadoop/dfs/data03

environment:

SERVICE\_PRECONDITION: "namenode:9870"

env\_file:

- ./hadoop.env

datanode04:

image: bde2020/hadoop-datanode:2.0.0-hadoop3.2.1-java8

container\_name: datanode04

restart: always

volumes:

- hadoop\_datanode04:/hadoop/dfs/data04

environment:

SERVICE\_PRECONDITION: "namenode:9870"

env\_file:

- ./hadoop.env

datanode05:

image: bde2020/hadoop-datanode:2.0.0-hadoop3.2.1-java8

container\_name: datanode05

restart: always

volumes:

- hadoop\_datanode05:/hadoop/dfs/data05

environment:

SERVICE\_PRECONDITION: "namenode:9870"

env\_file:

- ./hadoop.env

resourcemanager:

image: bde2020/hadoop-resourcemanager:2.0.0-hadoop3.2.1-java8

container\_name: resourcemanager

restart: always

environment:

SERVICE\_PRECONDITION: "namenode:9000 namenode:9870 datanode:9864"

env\_file:

- ./hadoop.env

nodemanager1:

image: bde2020/hadoop-nodemanager:2.0.0-hadoop3.2.1-java8

container\_name: nodemanager

restart: always

environment:

SERVICE\_PRECONDITION: "namenode:9000 namenode:9870 datanode:9864 resourcemanager:8088"

env\_file:

- ./hadoop.env

historyserver:

image: bde2020/hadoop-historyserver:2.0.0-hadoop3.2.1-java8

container\_name: historyserver

restart: always

environment:

SERVICE\_PRECONDITION: "namenode:9000 namenode:9870 datanode:9864 resourcemanager:8088"

volumes:

- hadoop\_historyserver:/hadoop/yarn/timeline

env\_file:

- ./hadoop.env

volumes:

hadoop\_namenode:

hadoop\_datanode01:

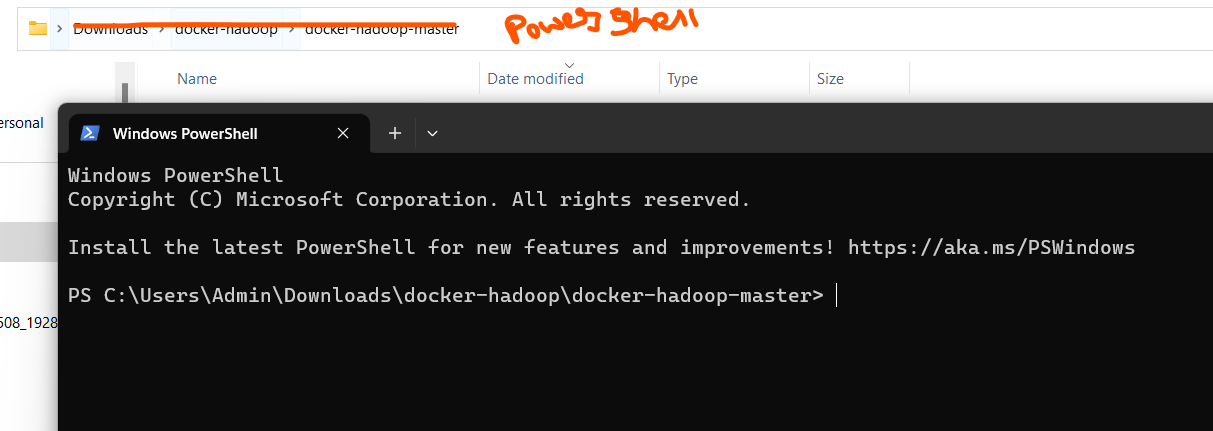
hadoop\_datanode02:

hadoop\_datanode03:

hadoop\_datanode04:

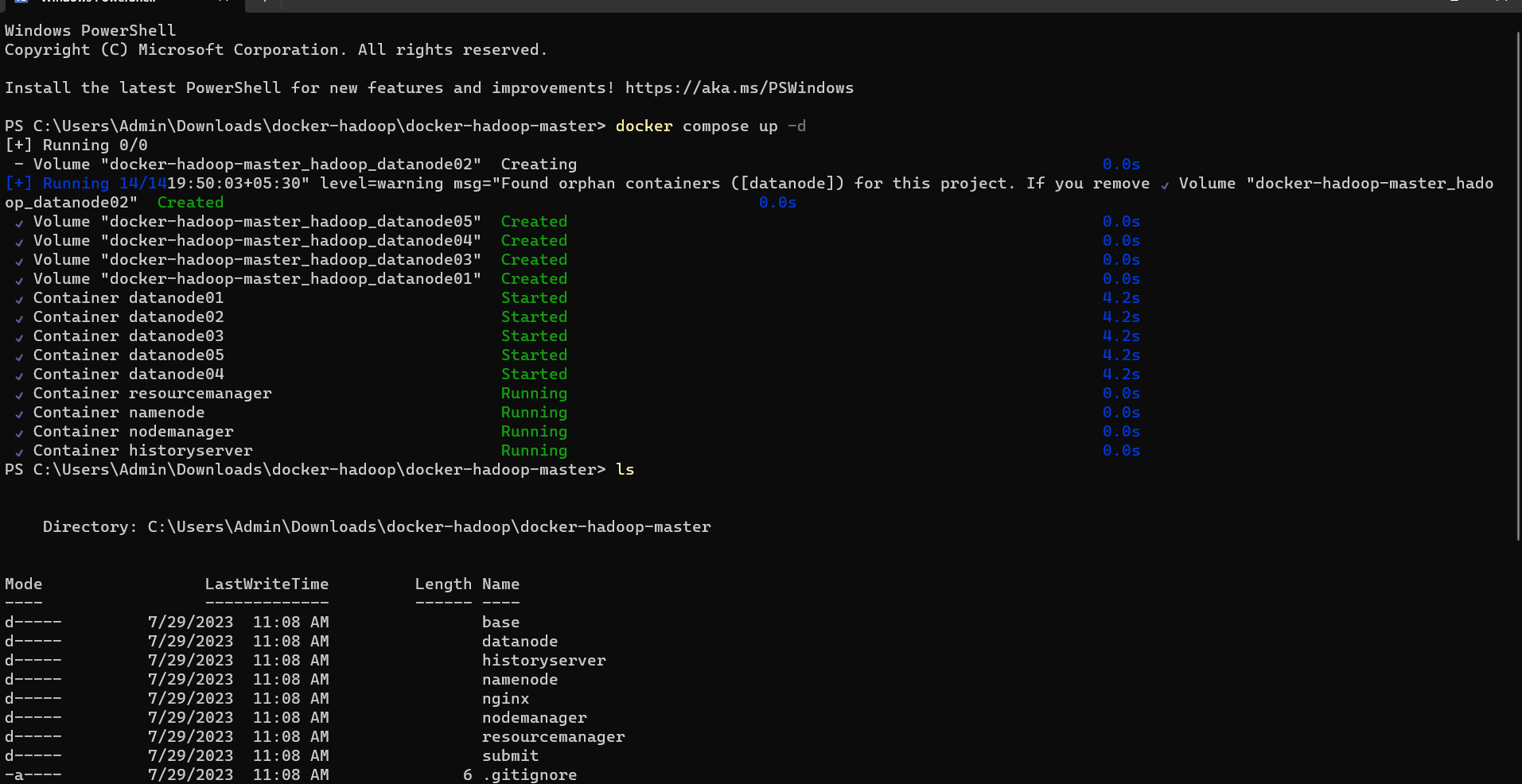
hadoop\_datanode05:

hadoop\_historyserver:

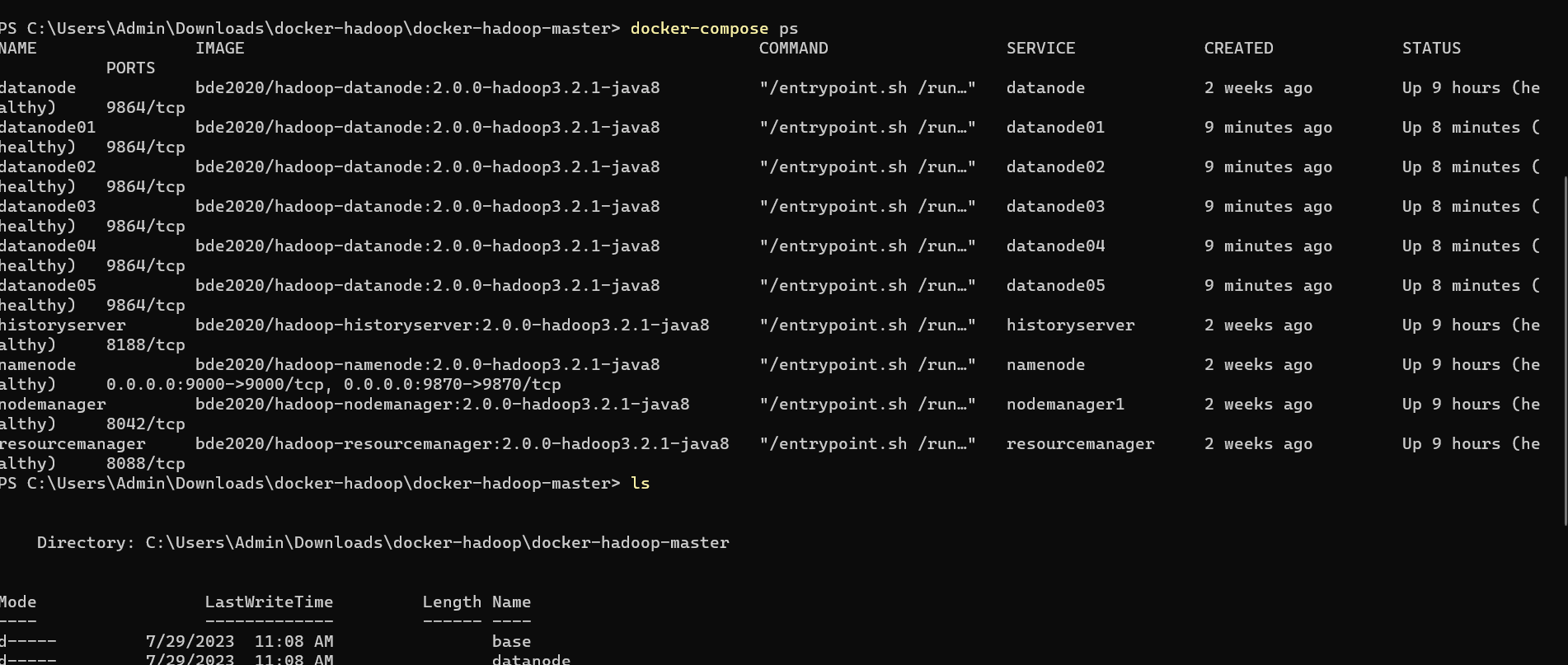


In Command Prompt type:

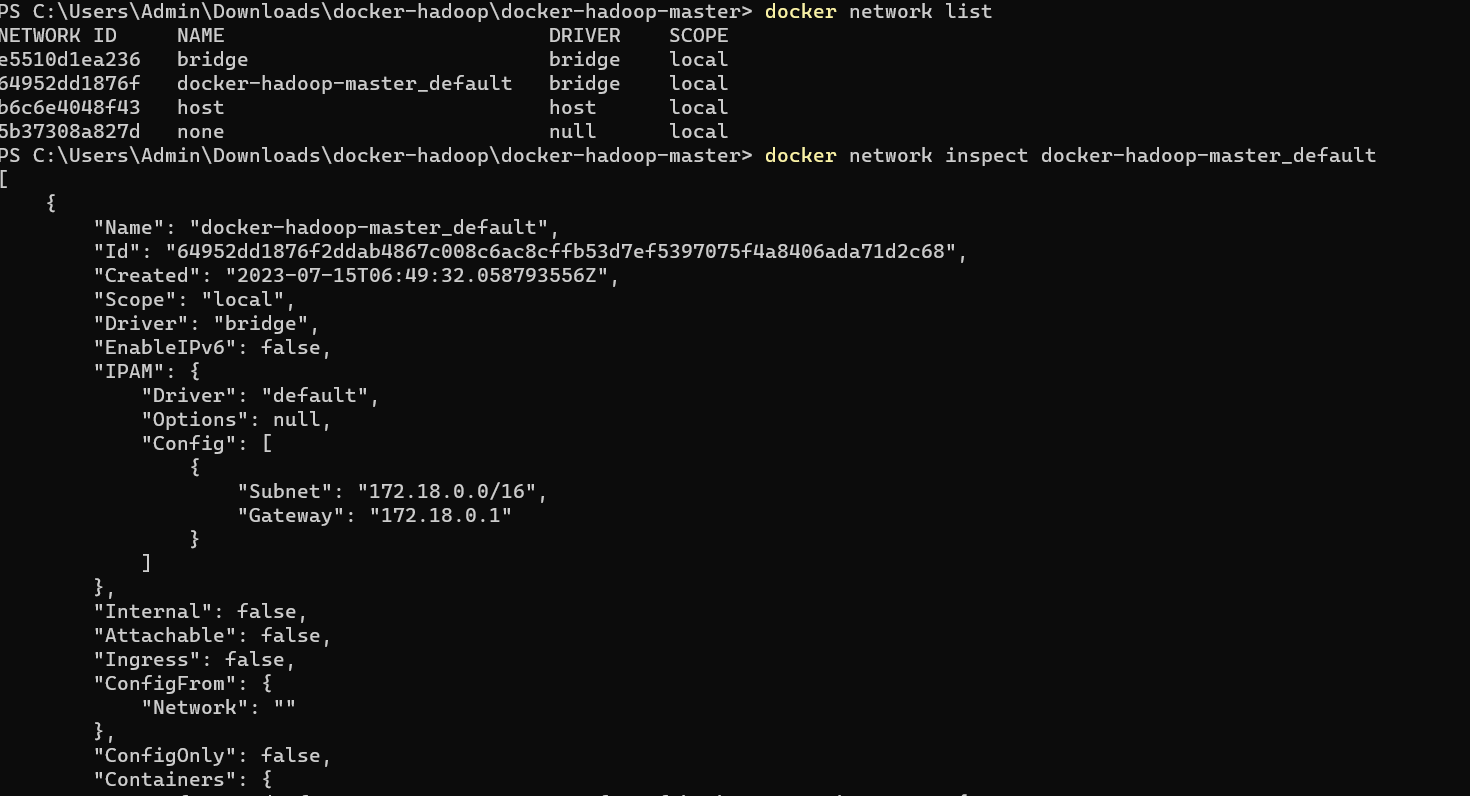
* Docker compose up -d
* ls



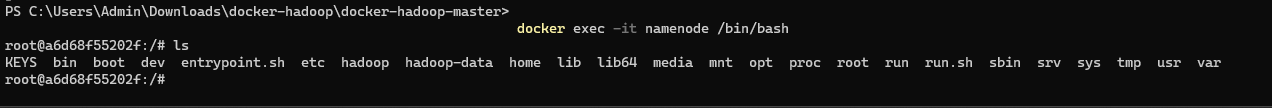
* docker-compose ps
* ls



* docker network list
* docker network inspect docker-hadoop-master\_default



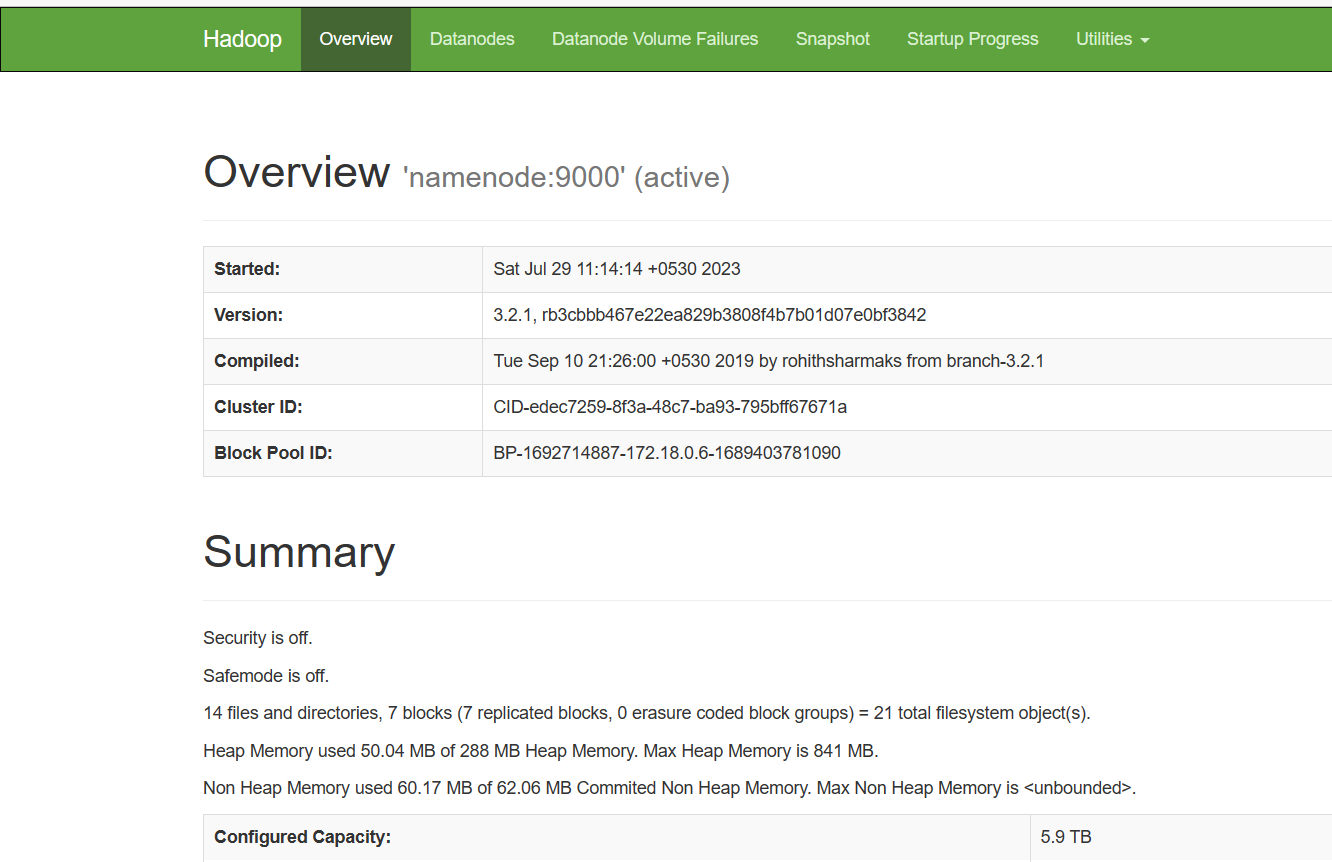
* docker exec -it namenode /bin/bash
* ls

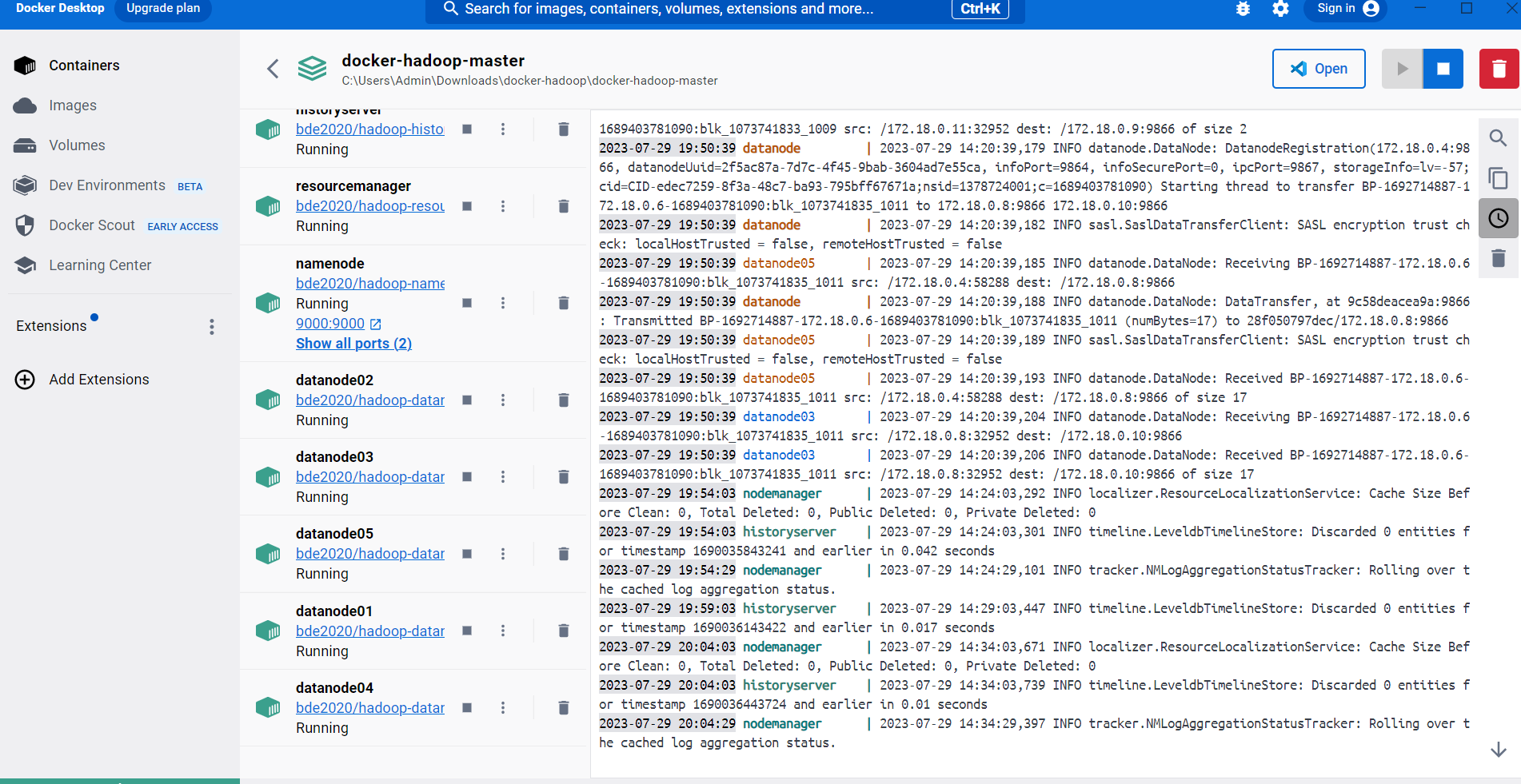


SERVICE\_PRECONDITION: "namenode:9870"

Open the URL to check nodes are running or not:

<http://localhost:9870/dfshealth.html#tab-overview>



Open Docker and Check

