PRACTICE ENVIRONMENTS INSTALLATION GUIDE

CONTENTS

1. INTRODUCTION – PRACTICE ENVIRONMENTS FOR EPAM BIG DATA COURSE

- Practice environments are based on Docker containers, which can be downloaded and invoked on the student machines.
- To minimize resources, most topics will have their own dockers, running only the minimum required components
- For each Docker it will be stated what is the optimal resource allocation and what is the minimum resource requirements for using it.

2. PREREQUISITES FOR RUNNING THE PRACTICE ENVIRONMENT

- Docker
  - Install Docker
    - Mac: https://docs.docker.com/docker-for-mac/install/
    - Windows: https://docs.docker.com/docker-for-windows/install/
    - Linux: https://docs.docker.com/engine/install/ (Choose your Linux distribution)
  - Configure Docker
    - Open Docker settings, and configure to 8GB RAM
    - For more details see here:
      - Mac: https://docs.docker.com/docker-for-mac/
      - Windows: https://docs.docker.com/docker-for-windows/
  - Install Docker Compose (Linux only):
    - Mac & Windows: Docker compose already included along with Docker Desktop.
    - Linux: https://docs.docker.com/compose/install/
- Minimum system requirements
  - 8 GB RAM
  - 2GHz CPU
  - 10GB free space on hard disk

3. CREATE A VIRTUAL DOCKER NETWORK

- Run command in terminal to create a virtual Docker network-

  docker network create bigdata_network
- Inspect the network to find the gateway's IP address, and mark the IP address.

  (You will need this IP address for future commands):

  docker network inspect bigdata_network
- For example: (Note that you may get a different IP address)

```
C:\>docker network inspect bigdata_network
[
    {
        "Name": "bigdata_network",
        "Id": "205fb353fccbd78b3d38ad2626ad5e71740068c118281b8e6ed2e553130217be",
        "Created": "2020-09-23T10:00:38.612431565Z",
        "Scope": "local",
        "Driver": "bridge",
        "EnableIPv6": false,
        "IPAM": {
            "Driver": "default",
            "Options": {},
            "Config": [
                {
                    "Subnet": "172.18.0.0/16",
                    "Gateway": "172.18.0.1"
                }
            ]
        },
        "Internal": false,
        "Attachable": false,
        "Ingress": false,
        "ConfigFrom": {
            "Network": ""
        },
        "ConfigOnly": false,
        "Containers": {},
        "Options": {},
        "Labels": {}
    }
]
```

4. CREATING THE HBASE DOCKER CONTAINER

- To install the Docker container for the HBase practice, run the following command - `
  - Be sure to change the IP address to the IP you previously marked

docker create –name hbase –network bigdata_network -it ofrir119/hbase:1.0.0

- This can take anywhere from 2 minutes up to an hour, depending on your internet connection.
- For example: (Be sure to use the IP of your Docker network gateway)

```
C:\Users\NAYA>docker create --name hbase --network bigdata_network -it ofrir119/hbase:1.0.0
Unable to find image 'ofrir119/hbase:1.0.0' locally
1.0.0: Pulling from ofrir119/hbase
```

5. STARTING THE HBASE DOCKER CONTAINER

- Start the Docker container using the following command:

  docker start hbase

- Make sure the image is up and running:

  docker ps -a

- Sample output:

```
C:\Users\NAYA>docker ps -a
CONTAINER ID        IMAGE                  COMMAND              CREATED             STATUS              PORTS                                                                NAMES
9b62350f7faa        ofrir119/hbase:1.0.0   "/opt/hbase-server"   24 seconds ago      Up 1 second         2181/tcp, 8080/tcp, 8085/tcp, 9090/tcp, 9095/tcp, 16010/tcp   hbase
```

6. CONNECTING TO THE DOCKER ENVIRONMENT

- Open the command prompt and issue the following command:

  docker exec -it hbase /bin/bash

7. USEFUL DOCKER COMMANDS

- **Notes:**
  - All commands below are to be executed in the command prompt.
  - Assuming Docker name is “hbase”, as it is in the practice environment for Hive

VIEW DOCKER CONTAINERS AND THEIR STATUS

docker ps -a

STOPPING AND STARTING THE DOCKER CONTAINER

docker stop hbase

docker start hbase

REMOVE A PREVIOUSLY INSTALLED DOCKER CONTAINER

docker rm hbase

CHECK CONTAINER RESOURCE USAGE (RAM ETC)

docker stats

8. STOPPING THE DOCKER ENVIRONMENT

- Run command -

  docker stop hbase

9. TROUBLESHOOTING THE DOCKER ENVIRONMENT

- See Guidelines document for troubleshooting issues.