**Practice Environments Installation Guide**

CONTENTS

[Introduction: Practice Environments for the EPAM Big Data Course](#_Toc126760992)

[Prerequisites for Running the Practice Environment](#_Toc126760993)

[Creating a Virtual Docker Network 2](#_Toc126760994)

[Creating the Kafka Docker Container 3](#_Toc126760995)

[Starting the Kafka Docker Container 4](#_Toc126760996)

[Connecting to the Docker Environment 4](#_Toc126760997)

[Useful Docker Commands 4](#_Toc126760998)

[Stopping and Starting the Docker Container 4](#_Toc126760999)

[Remove a Previously Installed Docker Container 5](#_Toc126761000)

[Check Container Resource Usage (e.g., RAM) 5](#_Toc126761001)

[Stopping the Docker Environment 5](#_Toc126761002)

[Creating the MySQL Docker Container 5](#_Toc126761003)

[Creating the Kafkadrop Docker Container (optional) 5](#_Toc126761004)

[Troubleshooting the Docker Environment 5](#_Toc126761005)

INTRODUCTION: PRACTICE ENVIRONMENTS FOR THE 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.

PREREQUISITES FOR RUNNING THE PRACTICE ENVIRONMENT

* Docker
* Install Docker
  + - Mac[: https://docs.docker.com/docker-for-mac/install/](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/](https://docs.docker.com/docker-for-mac/)
    - Windows[: https://docs.docker.com/docker-for-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

CREATING 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)



CREATING THE KAFKA DOCKER CONTAINER

To install the Docker container for the Kafka practice, run the following command - `

Be sure to change the IP address to the IP you previously marked

**docker create --name kafka --network bigdata\_network -p 2181:2181 -p 9092:9092 -p 8083:8083 -e ADVERTISED\_HOST=*<network gateway ip>* -it ofrir119/kafka:2.4.0**

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)



STARTING THE KAFKA DOCKER CONTAINER

* Start the Docker container using the following command:

**docker start kafka**

* Make sure the image is up and running:

**docker ps -a**

Sample output:



CONNECTING TO THE DOCKER ENVIRONMENT

Open the command prompt and issue the following command:

Be sure to include the “-l” flag to load scripts

**docker exec -it kafka /bin/bash**

USEFUL DOCKER COMMANDS

Notes:

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

VIEW DOCKER CONTAINERS AND THEIR STATUS

Run the command:

**docker ps -a**

STOPPING AND STARTING THE DOCKER CONTAINER

Run the commands:

**docker stop kafka**

**docker start kafka**

REMOVE A PREVIOUSLY INSTALLED DOCKER CONTAINER

Run the command:

**docker rm kafka**

CHECK CONTAINER RESOURCE USAGE (E.G., RAM)

Run the command:

**docker stats**

STOPPING THE DOCKER ENVIRONMENT

Run the command:

**docker stop kafka**

CREATING THE MYSQL DOCKER CONTAINER

* Install this only before performing the MySQL related practice.
* To install the MySQL Docker container, run the following command - `
* Be sure to change the IP address to the IP you previously marked

**docker run -d --name mysql --network bigdata\_network -p 3306:3306 -e MYSQL\_ROOT\_PASSWORD=root mysql**

CREATING THE KAFDROP DOCKER CONTAINER (OPTIONAL)

Install this only before performing the Kafdrop related practice. It is optional but recommended. Still, this docker is provided for students who want to install this GUI tool for Kafka. Install this only if you intend to use it.

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

**docker create --name kafdrop --network bigdata\_network -p 9000:9000 -e KAFKA\_BROKERCONNECT=172:9092 -it obsidiandynamics/kafdrop**

TROUBLESHOOTING THE DOCKER ENVIRONMENT

See the Guidelines document for troubleshooting issues.