



DEVOPS FOR CLOUD

Project Synopsis

Developed a backend application using FastAPI, containerized with Docker, deployed on Kubernetes, and monitored with Prometheus for performance tracking.

Kaushik Paul

2024mt03072@wilp.bits-pilani.ac.in

Contents

Software and Platform Details.....	2
Project Repository.....	3
Task 1: Create the Backend Application using FastAPI	4
Task 2: Dockerize the Backend Application	9
Task 3: Run the Docker Container	13
Task 4: Deploy Docker Image to Kubernetes Cluster	15
Task 5: Configure Networking with Load Balancer in Kubernetes Cluster	20
Task 6: Configure Prometheus for Metrics Collection	23
Conclusion.....	30

Software and Platform Details

This section outlines the technologies, software, and platforms used throughout the completion of this assignment.

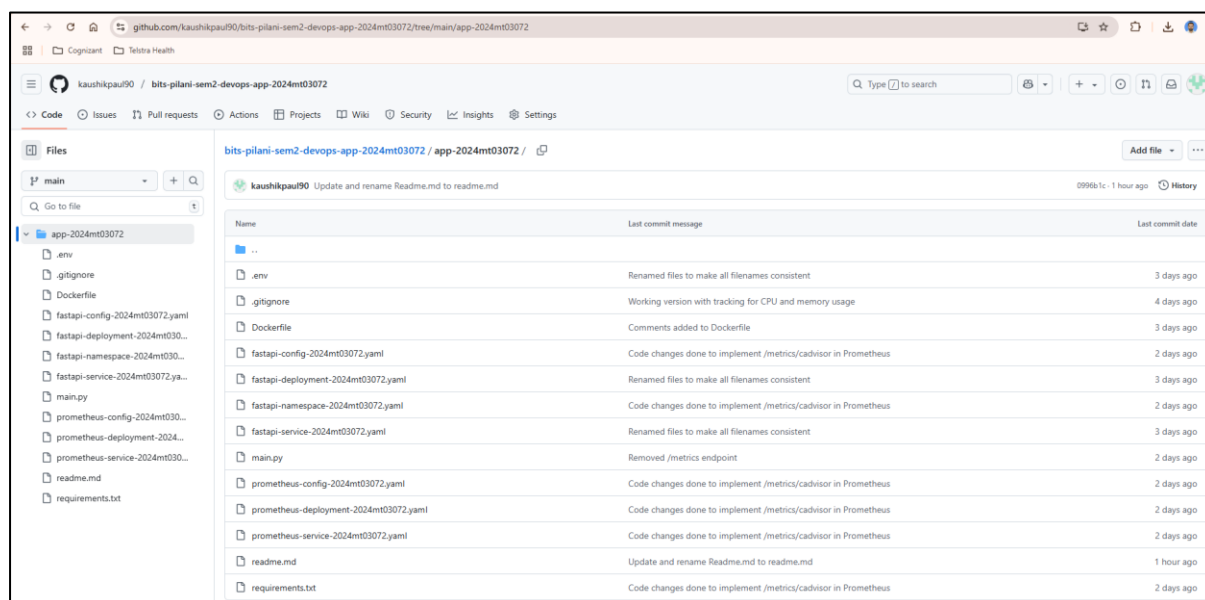
Component	Details
Operating System	Ubuntu 24.04 LTS (64-bit)
Python Version	Python 3.12.3
FastAPI Framework	FastAPI 0.115.12
ASGI Server	Uvicorn 0.34.2
Virtual Environment	venv (python -m venv .venv)
Containerization	Docker Engine (v26.1.3) installed via apt-get
Container Orchestration	Minikube v1.35.0 running Kubernetes v1.32.0
Monitoring Tool	Prometheus 3.3.1 deployed via Kubernetes YAML files
Code Editor	Visual Studio Code (VSCode) with Python and Docker extensions
Web Browser	Google Chrome
API Testing Tool	curl (command-line), browser (http://localhost:8000/get_info)
Version Control	Git + GitHub repository: link

Note: Docker Desktop was not used. Instead, Docker was installed on a native Ubuntu Linux system, and Minikube was used for local Kubernetes deployment. All functional requirements remain the same.

Project Repository

Github link → <https://github.com/kaushikpaul90/bits-pilani-sem2-devops-app-2024mt03072/tree/main/app-2024mt03072>

Screenshot



Task 1: Create the Backend Application using FastAPI

1. Project Setup

- a. Creating a new directory named **app-2024mt03072** for the Python project FastAPI.

```
bash
```

```
mkdir app-2024mt03072
```

- b. Changing the current working directory to app-2024mt03072.

```
bash
```

```
cd app-2024mt03072
```

- c. Creating a virtual environment for the project app-2024mt03072.

```
bash
```

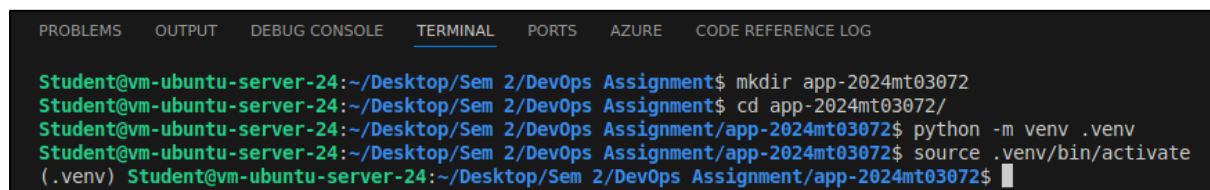
```
python -m venv .venv
```

- d. Activating the Python virtual environment named “.venv”.

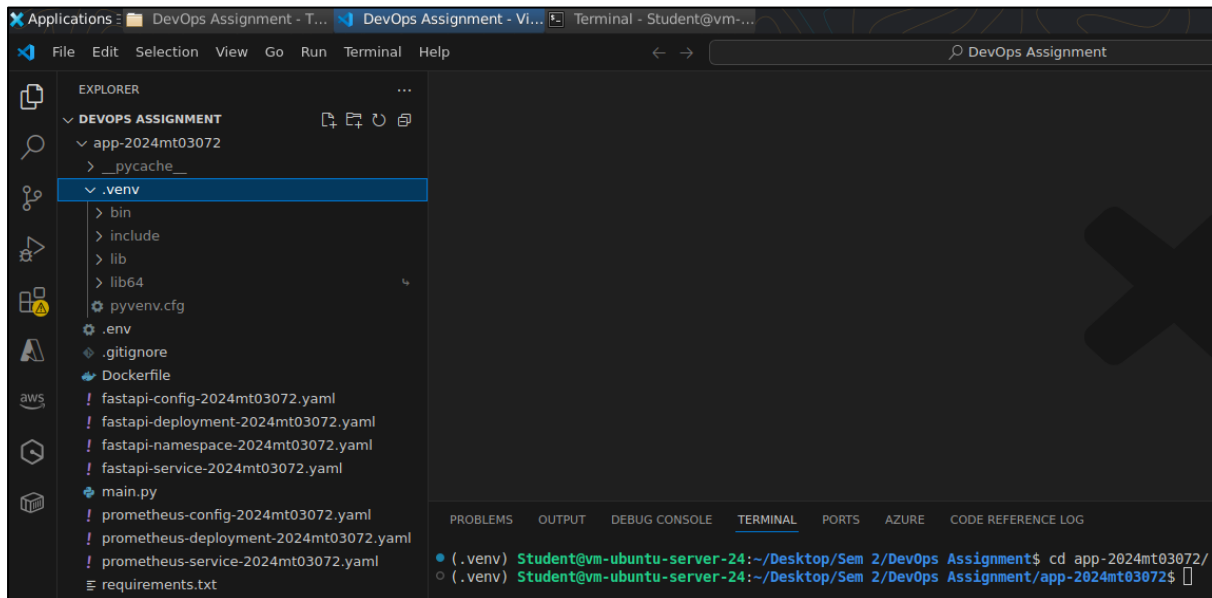
```
bash
```

```
source .venv/bin/activate
```

1.1. Screenshot



```
PROBLEMS  OUTPUT  DEBUG CONSOLE  TERMINAL  PORTS  AZURE  CODE REFERENCE LOG
Student@vm-ubuntu-server-24:~/Desktop/Sem 2/DevOps Assignment$ mkdir app-2024mt03072
Student@vm-ubuntu-server-24:~/Desktop/Sem 2/DevOps Assignment$ cd app-2024mt03072/
Student@vm-ubuntu-server-24:~/Desktop/Sem 2/DevOps Assignment/app-2024mt03072$ python -m venv .venv
Student@vm-ubuntu-server-24:~/Desktop/Sem 2/DevOps Assignment/app-2024mt03072$ source .venv/bin/activate
(.venv) Student@vm-ubuntu-server-24:~/Desktop/Sem 2/DevOps Assignment/app-2024mt03072$
```



2. Dependency Installation

- a. Adding required packages to requirements.txt file for installation.

File link → <https://github.com/kaushikpaul90/bits-pilani-sem2-devops-app-2024mt03072/blob/main/app-2024mt03072/requirements.txt>

- b. Installing the packages listed in requirements.txt using pip install.

```
bash
```

```
pip install -r requirements.txt
```

Note:

- **fastapi:** A modern, high-performance web framework for building APIs with Python 3.7+ based on standard Python type hints.
- **python-dotenv:** A tool that loads environment variables from a .env file into the Python environment.
- **pydantic-settings:** A library to manage application settings using Pydantic models with support for environment variable loading.
- **uvicorn:** A lightning-fast ASGI server for serving FastAPI and other ASGI-compatible Python web apps.
- **prometheus_client:** A Python client for exposing metrics in a format Prometheus can scrape for monitoring and alerting.

4. Creation of main.py

- Creating Python file named main.py in the project directory app-2024mt03072.

File link → <https://github.com/kaushikpaul90/bits-pilani-sem2-devops-app-2024mt03072/blob/main/app-2024mt03072/main.py>

Note:

- The `POD_NAME` attribute has been included in the returned JSON object to indicate which pod replica is processing the request.
- When running the application locally or in a containerized environment, `POD_NAME` is displayed as “unknown”.
- When the application is deployed and browsed from Minikube cluster, it shows the actual pod replica name.

4.1. Screenshot

```

1 import os # Importing os module to interact with the operating system
2 from fastapi import FastAPI # Importing FastAPI to create the web application
3 from fastapi.responses import Response # Importing Response to send custom HTTP responses
4 from contextlib import asynccontextmanager # Importing asynccontextmanager to manage application lifespan
5 from pydantic_settings import BaseSettings # Importing BaseSettings to manage application configuration
6 from prometheus_client import Counter, generate_latest, CONTENT_TYPE_LATEST # Importing Prometheus client for metrics
7
8 # Defining a Settings class to manage application settings using environment variables
9 class Settings(BaseSettings):
10     APP_VERSION: str = "1.0" # Setting the default application version
11     APP_TITLE: str = "My FastAPI App" # Setting the default application title
12
13     # Config class to specify additional configuration options for the Settings class
14     class Config:
15         env_file = ".env" # Specifying the environment file to load variables from while running locally
16
17 settings = Settings() # Creating an instance of the Settings class to access configuration values
18
19 POD_NAME = os.getenv("POD_NAME", "unknown") # Fetching the pod name from environment variables, defaulting to "unknown" if not set
20
21 # Defining Prometheus metric to count requests to the /get_info endpoint, labeled by pod
22 REQUEST_COUNT = Counter(
23     'get_info_requests_total', # Metric name
24     'Total number of requests to /get_info endpoint', # Metric description
25     ['pod'] # Label to differentiate metrics by pod
26 )
27
28 # Lifespan context manager to manage application startup and shutdown events
29 @asynccontextmanager
30 async def lifespan(app: FastAPI):
31     yield # Run the app
32
33 # Creating an instance of the FastAPI application with a custom lifespan
34 app = FastAPI(lifespan=lifespan)
35
36 # Defining a route to handle GET requests at the "/get_info" endpoint
37 @app.get("/get_info")
38 async def get_info():
39     REQUEST_COUNT.labels(pod=POD_NAME).inc() # Increment the request count metric for this pod
40     # Returning a JSON response with application version, title, and pod name
41     return {
42         "APP_VERSION": settings.APP_VERSION, # Fetching the application version from settings
43         "APP_TITLE": settings.APP_TITLE, # Fetching the application title from settings
44         "POD_NAME": POD_NAME # Including the pod name in the response
45     }

```

5. Testing FastAPI Application Locally

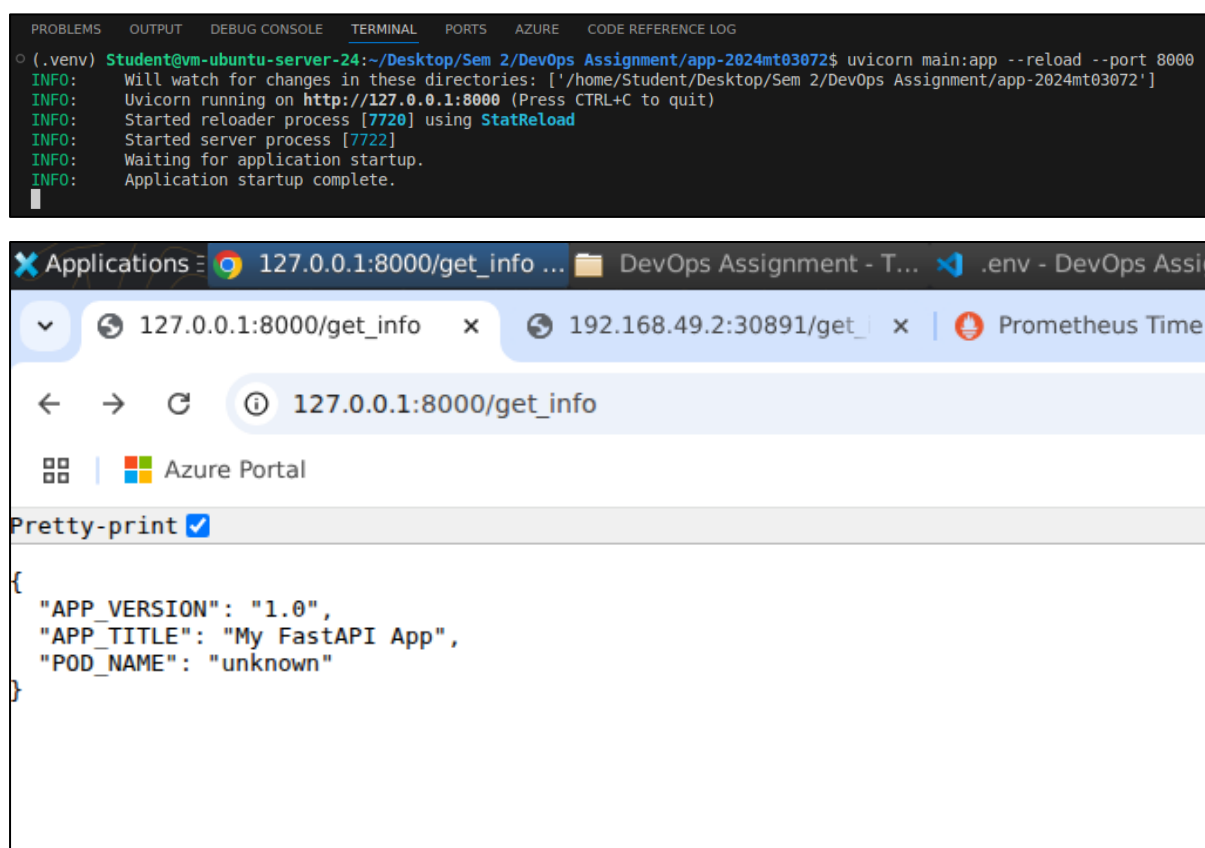
- Running the application locally on Uvicorn.

```
bash
```

```
uvicorn main:app --reload --port 8000
```

- Accessing the application via http://localhost:8000/get_info on Google Chrome.

5.1. Screenshot



6. Challenges Faced

- a. No challenges faced while creating the backend application using FastAPI.

Task 2: Dockerize the Backend Application

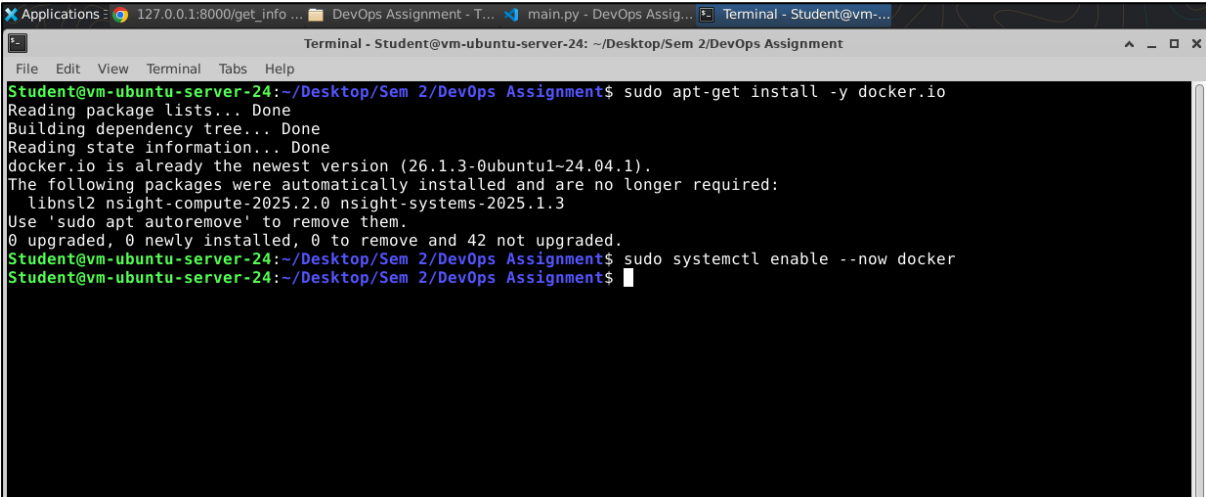
1. Docker Setup

- a. Installing and enabling Docker on Linux OS.

```
bash
```

```
sudo apt-get install -y docker.io  
sudo systemctl enable --now docker
```

1.1. Screenshot



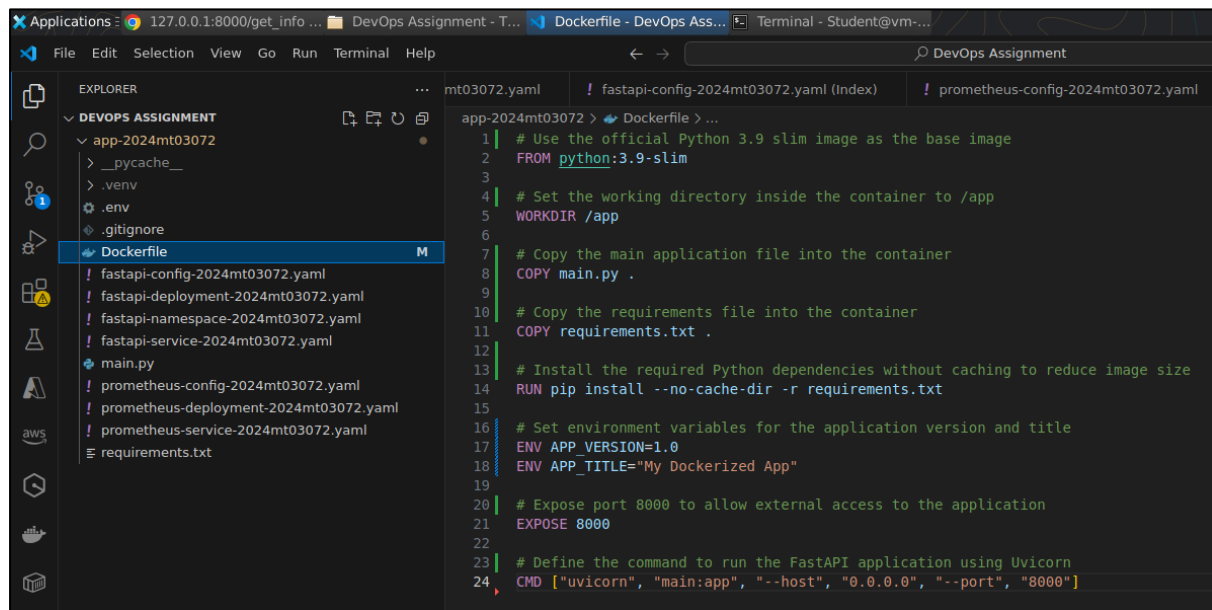
```
Terminal - Student@vm-ubuntu-server-24: ~/Desktop/Sem 2/DevOps Assignment  
Student@vm-ubuntu-server-24:~/Desktop/Sem 2/DevOps Assignment$ sudo apt-get install -y docker.io  
Reading package lists... Done  
Building dependency tree... Done  
Reading state information... Done  
docker.io is already the newest version (26.1.3-0ubuntu1~24.04.1).  
The following packages were automatically installed and are no longer required:  
  libssl2 nsight-compute-2025.2.0 nsight-systems-2025.1.3  
Use 'sudo apt autoremove' to remove them.  
0 upgraded, 0 newly installed, 0 to remove and 42 not upgraded.  
Student@vm-ubuntu-server-24:~/Desktop/Sem 2/DevOps Assignment$ sudo systemctl enable --now docker  
Student@vm-ubuntu-server-24:~/Desktop/Sem 2/DevOps Assignment$
```

2. Dockerfile Creation

- a. Creating a file named Dockerfile inside the project directory app-2024mt03072.

File link → <https://github.com/kaushikpaul90/bits-pilani-sem2-devops-app-2024mt03072/blob/main/app-2024mt03072/Dockerfile>

2.1. Screenshot



3. Docker Image Build

- a. Building the Docker image using Dockerfile. Docker image is named as **img-2024mt03072** and tag is set to **dev**.

```
bash
```

```
docker build -t img-2024mt03072:dev .
```

3.1. Screenshot

```

PROBLEMS  OUTPUT  DEBUG CONSOLE  TERMINAL  PORTS  AZURE  CODE REFERENCE LOG
(.venv) Student@vm-ubuntu-server-24:~/Desktop/Sem 2/DevOps Assignment/app-2024mt03072$ docker build -t img-2024mt03072:dev .
DEPRECATED: The legacy builder is deprecated and will be removed in a future release.
            Install the buildx component to build images with BuildKit:
            https://docs.docker.com/go/buildx/

Sending build context to Docker daemon  31.84MB
Step 1/9 : FROM python:3.9-slim
--> 9a041530811d
Step 2/9 : WORKDIR /app
--> Using cache
--> e2e89376290b
Step 3/9 : COPY main.py .
--> Using cache
--> 09939906f2eb
Step 4/9 : COPY requirements.txt .
--> Using cache
--> 743f8f1ee54c
Step 5/9 : RUN pip install --no-cache-dir -r requirements.txt
--> Using cache
--> d964acf3ae2f
Step 6/9 : ENV APP_VERSION=1.0
--> Using cache
--> 85b75356edb9
Step 7/9 : ENV APP_TITLE="My Dockerized App"
--> Using cache
--> d67cfcc181c7
Step 8/9 : EXPOSE 8000
--> Using cache
--> 0017c754318f
Step 9/9 : CMD ["uvicorn", "main:app", "--host", "0.0.0.0", "--port", "8000"]
--> Using cache
--> f1c4286f3830
Successfully built f1c4286f3830
Successfully tagged img-2024mt03072:dev
(.venv) Student@vm-ubuntu-server-24:~/Desktop/Sem 2/DevOps Assignment/app-2024mt03072$

```

4. Docker Image Verification

- a. Verifying the Docker image creation.

```
bash
```

```
docker images
```

4.1. Screenshot

```
(.venv) Student@vm-ubuntu-server-24:~/Desktop/Sem 2/DevOps Assignment/app-2024mt03072$ docker images
```

REPOSITORY	TAG	IMAGE ID	CREATED	SIZE
img-2024mt03072	dev	44294c37afe6	18 minutes ago	145MB
<none>	<none>	863d4a2be9df	2 hours ago	145MB
<none>	<none>	f4eae8d6491f	2 hours ago	147MB
<none>	<none>	f1c4286f3830	23 hours ago	147MB
<none>	<none>	02421eca7122	24 hours ago	177MB
<none>	<none>	c89f436c2cb6	27 hours ago	178MB
<none>	<none>	656d759e1f52	28 hours ago	178MB
<none>	<none>	579ddf013a12	28 hours ago	178MB
<none>	<none>	8d86c8045d29	28 hours ago	178MB
img-2024mt03072	v12	5aebbf756c8	31 hours ago	178MB
img-2024mt03072	v10	c885a16a5f45	32 hours ago	178MB
img-2024mt03072	v11	c885a16a5f45	32 hours ago	178MB
<none>	<none>	d47ce07fd16a	32 hours ago	176MB
img-2024mt03072	v9	41f6b2e8825f	2 days ago	178MB
img-2024mt03072	v8	3b966ae78b33	2 days ago	178MB
img-2024mt03072	v7	61c588922cbe	2 days ago	174MB
img-2024mt03072	v6	f93fb4c81162	2 days ago	174MB
img-2024mt03072	v5	708186ac476b	2 days ago	174MB
img-2024mt03072	v4	ecb883ad561f	2 days ago	174MB
img-2024mt03072	v3	dd44a288884c	2 days ago	174MB
<none>	<none>	05bebd61d63c	2 days ago	174MB
img-2024mt03072	v2	7f321f9a4e8e	3 days ago	173MB
<none>	<none>	44fe3d21f133	3 days ago	173MB
<none>	<none>	497ff8f1351c	3 days ago	173MB
python	3.9-slim	9a041530811d	4 weeks ago	126MB
gcr.io/k8s-minikube/kicbase	v0.0.46	e72c4cbe9b29	3 months ago	1.31GB

```
(.venv) Student@vm-ubuntu-server-24:~/Desktop/Sem 2/DevOps Assignment/app-2024mt03072$
```

5. Challenges Faced

- a. No challenges faced while dockerizing the backend application.

Task 3: Run the Docker Container

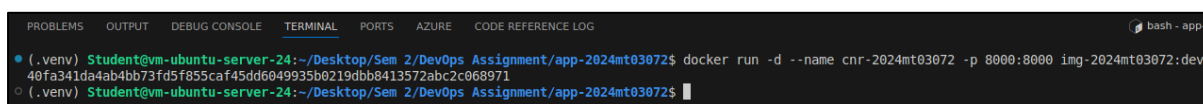
1. Docker Container Creation

- a. Creating Docker container using the built image img-2024mt03072. Container is named as **cnr-2024mt03072:dev**.

```
bash
```

```
docker run -d --name cnr-2024mt03072 -p 8000:8000 img-2024mt03072:dev
```

1.1. Screenshot



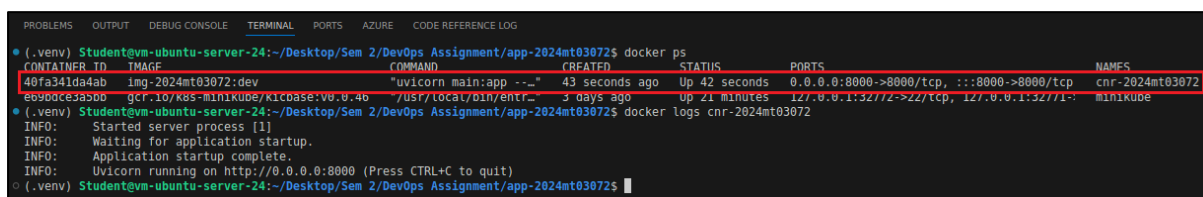
2. Docker Container Verification

- a. Verifying the created Docker container.

```
bash
```

```
docker ps
docker logs cnr-2024mt03072
```

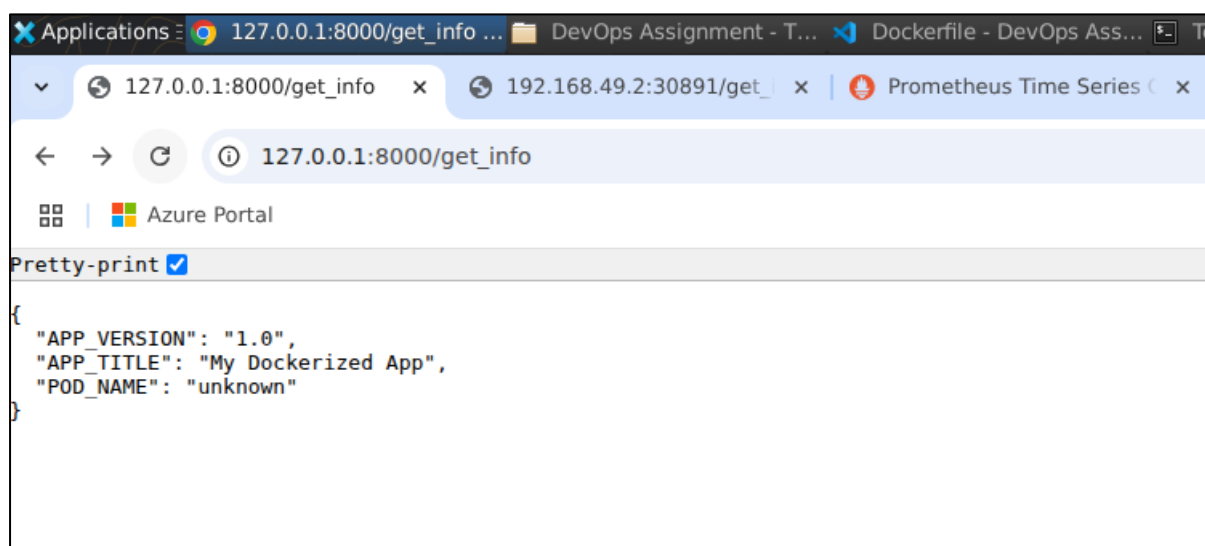
2.1. Screenshot



3. Testing the Containerized Application

- a. Accessing the containerized application via http://localhost:8000/get_info on Google Chrome.

3.1. Screenshot



4. Challenges Faced

- a. No challenges faced while running the dockerized backend application over Google Chrome.

Task 4: Deploy Docker Image to Kubernetes Cluster

1. Minikube Setup

- a. Downloading Minikube.

```
bash

curl -LO
https://storage.googleapis.com/minikube/releases/latest/minikube-
linux-amd64
```

- b. Installing Minikube.

```
bash

sudo install minikube-linux-amd64 /usr/local/bin/minikube
```

- c. Starting Minikube.

```
bash

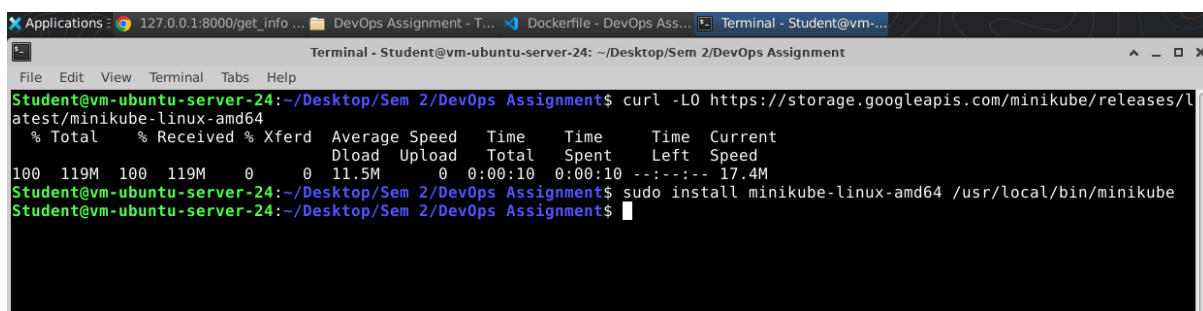
minikube start
```

- d. Verifying if the Minikube cluster is running properly.

```
bash

minikube status
```

1.1. Screenshot




```

PROBLEMS  OUTPUT  DEBUG CONSOLE  TERMINAL  PORTS  AZURE  CODE REFERENCE LOG
• (.venv) Student@vm-ubuntu-server-24:~/Desktop/Sem 2/DevOps Assignment/app-2024mt03072$ minikube start
  minikube v1.35.0 on Ubuntu 24.04
  [ ] Using the docker driver based on existing profile
  [ ] Starting "minikube" primary control-plane node in "minikube" cluster
  [ ] Pulling base image v0.0.46 ...
  [ ] Updating the running docker "minikube" container ...
  [ ] Preparing Kubernetes v1.32.0 on Docker 27.4.1 ...
    ▪ kubelet.read-only-port=10255
  [ ] Verifying Kubernetes components...
    ▪ Using image gcr.io/k8s-minikube/storage-provisioner:v5
  [ ] Enabled addons: storage-provisioner, default-storageclass
  [ ] kubectl not found. If you need it, try: 'minikube kubectl -- get pods -A'
  [ ] Done! kubectl is now configured to use "minikube" cluster and "default" namespace by default
• (.venv) Student@vm-ubuntu-server-24:~/Desktop/Sem 2/DevOps Assignment/app-2024mt03072$ minikube status
minikube
type: Control Plane
host: Running
kubelet: Running
apiserver: Running
kubeconfig: Configured
• (.venv) Student@vm-ubuntu-server-24:~/Desktop/Sem 2/DevOps Assignment/app-2024mt03072$

```

2. Loading Docker Image into Minikube

- a. Loading the image img-2024mt03072 into Minikube cluster.

```
bash
```

```
minikube image load img-2024mt03072:dev
```

2.1. Screenshot

```

PROBLEMS  OUTPUT  DEBUG CONSOLE  TERMINAL  PORTS  AZURE  CODE REFERENCE LOG
• (.venv) Student@vm-ubuntu-server-24:~/Desktop/Sem 2/DevOps Assignment/app-2024mt03072$ minikube image load img-2024mt03072:dev
• (.venv) Student@vm-ubuntu-server-24:~/Desktop/Sem 2/DevOps Assignment/app-2024mt03072$

```

3. Kubernetes Namespace Creation

- a. Creating Kubernetes Namespace **ns-fastapi-2024mt03072** using the Namespace configuration file **fastapi-namespace-2024mt03072.yaml**.

File → <https://github.com/kaushikpaul90/bits-pilani-sem2-devops-app-2024mt03072/blob/main/app-2024mt03072/fastapi-namespace-2024mt03072.yaml>.

- b. Applying the Kubernetes Namespace.

```
bash
```

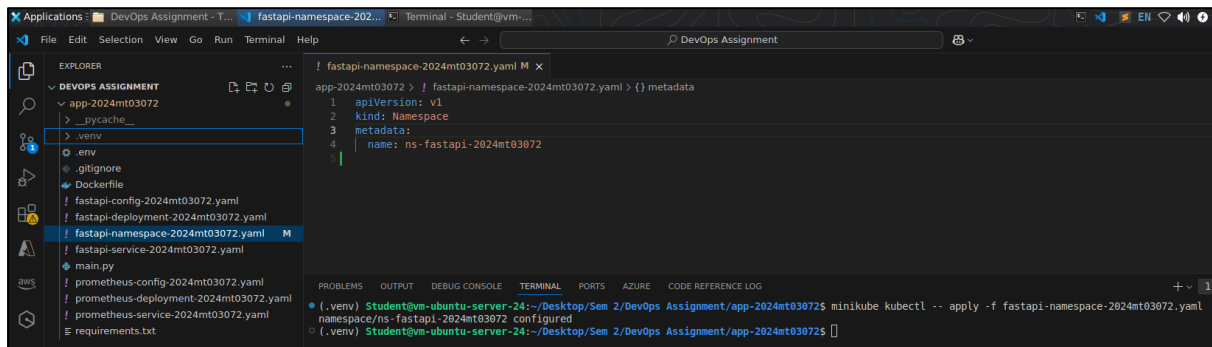
```
minikube kubectl -- apply -f fastapi-namespace-2024mt03072.yaml
```

- c. Verifying the creation of the Kubernetes Namespace ns-fastapi-2024mt03072.

```
bash
```

```
minikube kubectl -- get namespaces
```

3.1. Screenshot



4. Kubernetes ConfigMap Creation

- a. Creating Kubernetes ConfigMap **config-2024mt03072** using the ConfigMap configuration file **fastapi-config-2024mt03072.yaml**.

File → <https://github.com/kaushikpaul90/bits-pilani-sem2-devops-app-2024mt03072/blob/main/app-2024mt03072/fastapi-config-2024mt03072.yaml>.

- b. Applying the Kubernetes ConfigMap.

```
bash
```

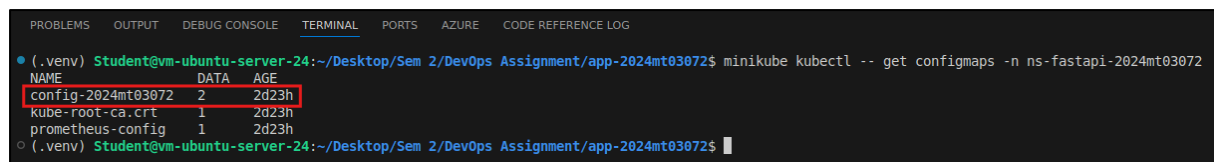
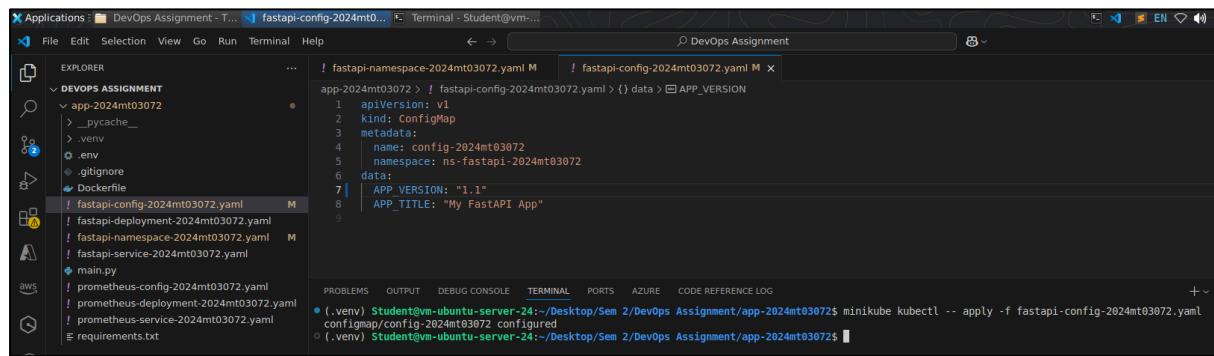
```
minikube kubectl -- apply -f fastapi-config-2024mt03072.yaml
```

- c. Verifying the creation of the Kubernetes ConfigMap config-2024mt03072.

```
bash
```

```
minikube kubectl -- get configmaps -n ns-fastapi-2024mt03072
```

4.1. Screenshot



5. Kubernetes Deployment Creation

- a. Creating Kubernetes Deployment **fastapi-deployment** using the Deployment configuration file **fastapi-deployment-2024mt03072.yaml**.

File → <https://github.com/kaushikpaul90/bits-pilani-sem2-devops-app-2024mt03072/blob/main/app-2024mt03072/fastapi-deployment-2024mt03072.yaml>.

- b. It creates 2 replicas of the FastAPI application.
- c. It reads the APP_VERSION and APP_TITLE values from Kubernetes ConfigMap named fastapi-config-2024mt03072.yaml.
- d. Applying the Kubernetes Deployment.

```
bash
```

```
minikube kubectl -- apply -f fastapi-deployment-2024mt03072.yaml
```

- e. Verifying the creation of the Kubernetes Deployment fastapi-deployment.

```
bash
```

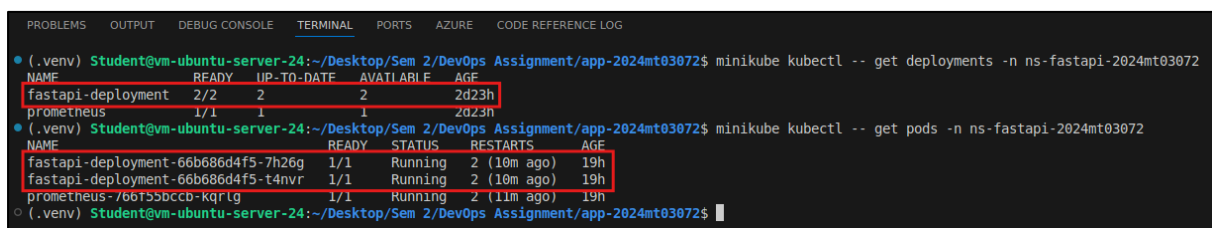
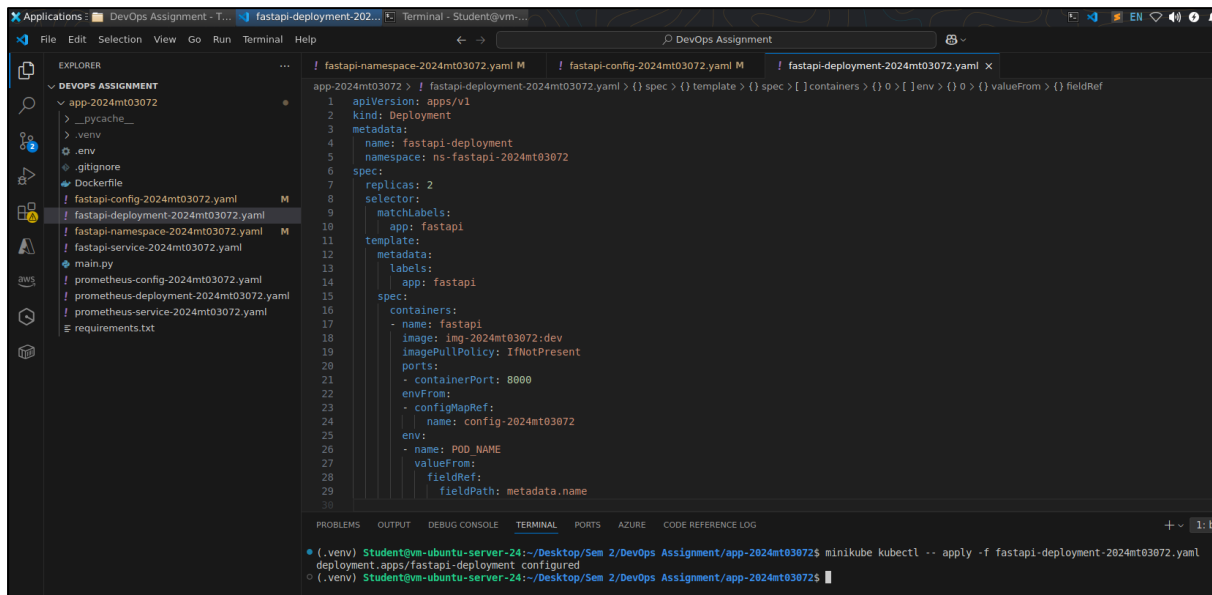
```
minikube kubectl -- get deployments -n ns-fastapi-2024mt03072
```

- f. Listing all pods running inside Kubernetes Namespace ns-fastapi-2024mt03072.

```
bash
```

```
minikube kubectl -- get pods -n ns-fastapi-2024mt03072
```

5.1. Screenshot



6. Challenges Faced

- Pods not reflecting the latest implementation:
 - Setting **imagePullPolicy** to **IfNotPresent** in fastapi-namespace-2024mt03072.yaml resolved the issue.
- Adding new image tag each time while building the image and updating the same in fastapi-namespace-2024mt03072.yaml:
 - Using fixed image tag “dev” during image build and adding it to the image name in fastapi-namespace-2024mt03072.yaml only once resolved the issue.

Task 5: Configure Networking with Load Balancer in Kubernetes Cluster

1. Kubernetes Service Creation

- Creating Kubernetes Service **fastapi-service** using the Service configuration file **fastapi-service-2024mt03072.yaml**.

File → <https://github.com/kaushikpaul90/bits-pilani-sem2-devops-app-2024mt03072/blob/main/app-2024mt03072/fastapi-service-2024mt03072.yaml>.

- It creates a Load Balancer to balance requests across both replicas.
- Applying the Kubernetes Service.

```
bash
```

```
minikube kubectl -- apply -f fastapi-service-2024mt03072.yaml
```

- Verifying the creation of the Kubernetes Service fastapi-service.

```
bash
```

```
minikube kubectl -- get services -n ns-fastapi-2024mt03072
```

1.1. Screenshot

The screenshot shows a terminal window with the following content:

```

(.venv) Student@vm-ubuntu-server-24:~/Desktop/Sem 2/DevOps Assignment/app-2024mt03072$ minikube kubectl -- apply -f fastapi-service-2024mt03072.yaml
service/fastapi-service configured
(.venv) Student@vm-ubuntu-server-24:~/Desktop/Sem 2/DevOps Assignment/app-2024mt03072$ minikube kubectl -- get services -n ns-fastapi-2024mt03072
NAME                TYPE          CLUSTER-IP      EXTERNAL-IP      PORT(S)          AGE
fastapi-service      LoadBalancer  10.100.10.202    <pending>         8000:31048/TCP   27h
prometheus           NodePort      10.98.146.20    <none>            9090:30090/TCP   2023h
  
```

The terminal output shows the successful application of the service and the resulting service configuration. The service 'fastapi-service' is of type 'LoadBalancer' and is currently in a 'pending' state for an external IP. The 'prometheus' service is also visible as a 'NodePort' service.

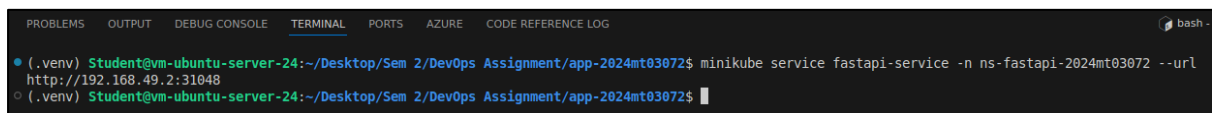
2. Obtaining Access URL

- a. Retrieving the URL to access FastAPI application from browser.

```
bash
```

```
minikube service fastapi-service -n ns-fastapi-2024mt03072 --url
```

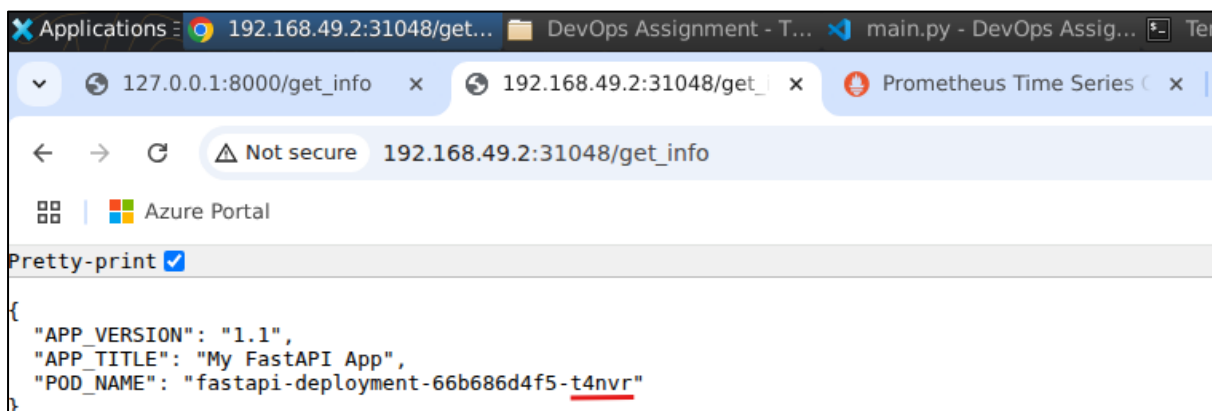
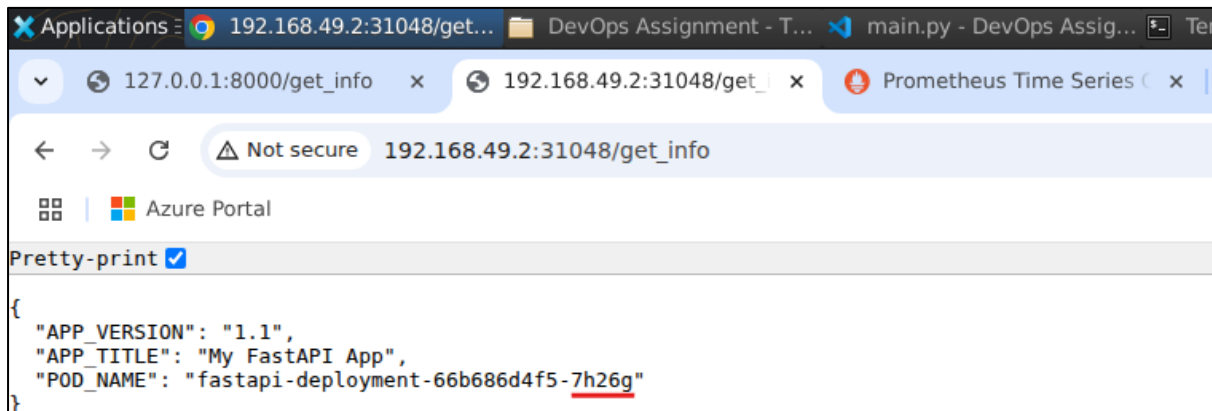
2.1. Screenshot



3. Accessing FastAPI Application

- a. Accessing the application via http://192.168.49.2:31048/get_info on Google Chrome.
- b. The fact that requests are getting routed to both the replicas is evident from the POD_NAME in the JSON object.

3.1. Screenshot



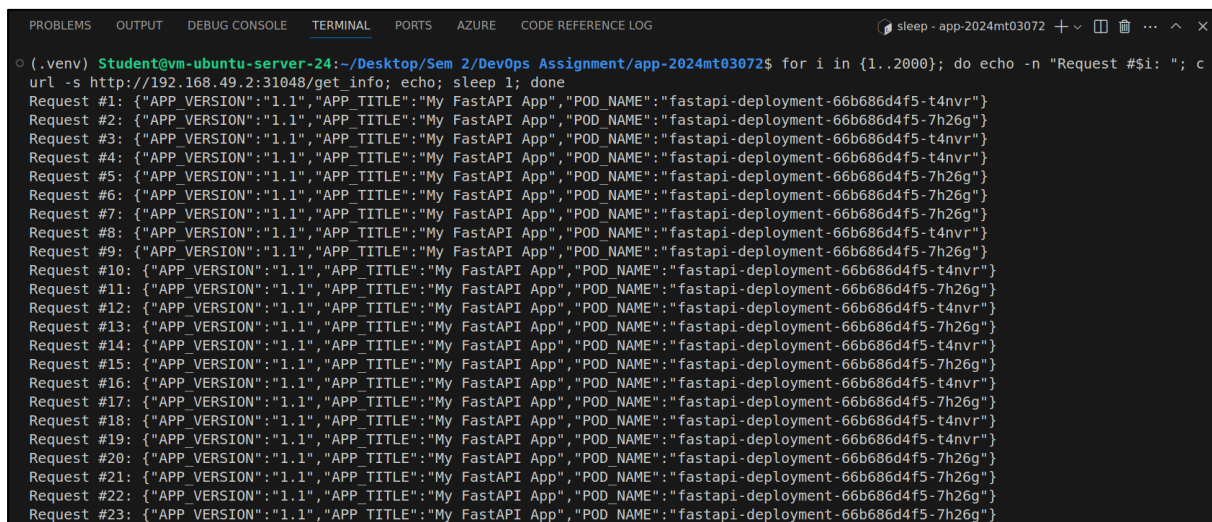
4. Load Balancer Testing

- Using cURL to make several requests to /get_info endpoint.
- The output clearly shows that requests are getting routed to both the replicas.

```
bash
```

```
for i in {1..2000}; do echo -n "Request #$i: "; curl -s  
http://192.168.49.2:31048/get_info; echo; sleep 1; done
```

4.1. Screenshot



```
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS AZURE CODE REFERENCE LOG
sleep - app-2024mt03072 + - - - - X
(.venv) Student@vm-ubuntu-server-24:~/Desktop/Sem 2/DevOps Assignment/app-2024mt03072$ for i in {1..2000}; do echo -n "Request #$i: "; c
url -s http://192.168.49.2:31048/get_info; echo; sleep 1; done
Request #1: {"APP_VERSION":"1.1","APP_TITLE":"My FastAPI App","POD_NAME":"fastapi-deployment-66b686d4f5-t4nvr"}
Request #2: {"APP_VERSION":"1.1","APP_TITLE":"My FastAPI App","POD_NAME":"fastapi-deployment-66b686d4f5-7h26g"}
Request #3: {"APP_VERSION":"1.1","APP_TITLE":"My FastAPI App","POD_NAME":"fastapi-deployment-66b686d4f5-t4nvr"}
Request #4: {"APP_VERSION":"1.1","APP_TITLE":"My FastAPI App","POD_NAME":"fastapi-deployment-66b686d4f5-t4nvr"}
Request #5: {"APP_VERSION":"1.1","APP_TITLE":"My FastAPI App","POD_NAME":"fastapi-deployment-66b686d4f5-7h26g"}
Request #6: {"APP_VERSION":"1.1","APP_TITLE":"My FastAPI App","POD_NAME":"fastapi-deployment-66b686d4f5-7h26g"}
Request #7: {"APP_VERSION":"1.1","APP_TITLE":"My FastAPI App","POD_NAME":"fastapi-deployment-66b686d4f5-7h26g"}
Request #8: {"APP_VERSION":"1.1","APP_TITLE":"My FastAPI App","POD_NAME":"fastapi-deployment-66b686d4f5-t4nvr"}
Request #9: {"APP_VERSION":"1.1","APP_TITLE":"My FastAPI App","POD_NAME":"fastapi-deployment-66b686d4f5-7h26g"}
Request #10: {"APP_VERSION":"1.1","APP_TITLE":"My FastAPI App","POD_NAME":"fastapi-deployment-66b686d4f5-t4nvr"}
Request #11: {"APP_VERSION":"1.1","APP_TITLE":"My FastAPI App","POD_NAME":"fastapi-deployment-66b686d4f5-7h26g"}
Request #12: {"APP_VERSION":"1.1","APP_TITLE":"My FastAPI App","POD_NAME":"fastapi-deployment-66b686d4f5-t4nvr"}
Request #13: {"APP_VERSION":"1.1","APP_TITLE":"My FastAPI App","POD_NAME":"fastapi-deployment-66b686d4f5-7h26g"}
Request #14: {"APP_VERSION":"1.1","APP_TITLE":"My FastAPI App","POD_NAME":"fastapi-deployment-66b686d4f5-t4nvr"}
Request #15: {"APP_VERSION":"1.1","APP_TITLE":"My FastAPI App","POD_NAME":"fastapi-deployment-66b686d4f5-7h26g"}
Request #16: {"APP_VERSION":"1.1","APP_TITLE":"My FastAPI App","POD_NAME":"fastapi-deployment-66b686d4f5-t4nvr"}
Request #17: {"APP_VERSION":"1.1","APP_TITLE":"My FastAPI App","POD_NAME":"fastapi-deployment-66b686d4f5-7h26g"}
Request #18: {"APP_VERSION":"1.1","APP_TITLE":"My FastAPI App","POD_NAME":"fastapi-deployment-66b686d4f5-t4nvr"}
Request #19: {"APP_VERSION":"1.1","APP_TITLE":"My FastAPI App","POD_NAME":"fastapi-deployment-66b686d4f5-7h26g"}
Request #20: {"APP_VERSION":"1.1","APP_TITLE":"My FastAPI App","POD_NAME":"fastapi-deployment-66b686d4f5-7h26g"}
Request #21: {"APP_VERSION":"1.1","APP_TITLE":"My FastAPI App","POD_NAME":"fastapi-deployment-66b686d4f5-7h26g"}
Request #22: {"APP_VERSION":"1.1","APP_TITLE":"My FastAPI App","POD_NAME":"fastapi-deployment-66b686d4f5-7h26g"}
Request #23: {"APP_VERSION":"1.1","APP_TITLE":"My FastAPI App","POD_NAME":"fastapi-deployment-66b686d4f5-7h26g"}
```

5. Challenges Faced

- No challenges faced while running the dockerized backend application over Google Chrome.

3. Prometheus ConfigMap Creation

- a. Creating Prometheus ConfigMap **prometheus-config** using the ConfigMap configuration file **prometheus-config-2024mt03072.yaml**.

File → <https://github.com/kaushikpaul90/bits-pilani-sem2-devops-app-2024mt03072/blob/main/app-2024mt03072/prometheus-config-2024mt03072.yaml>.

- b. Applying the Prometheus ConfigMap.

```
bash
```

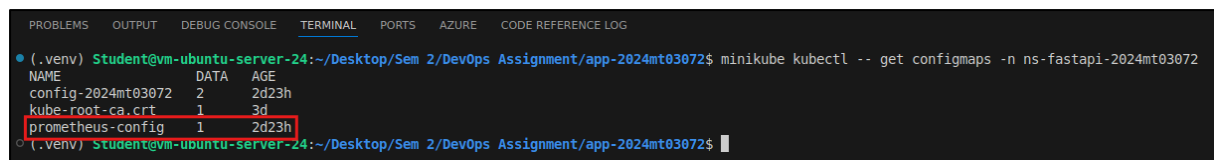
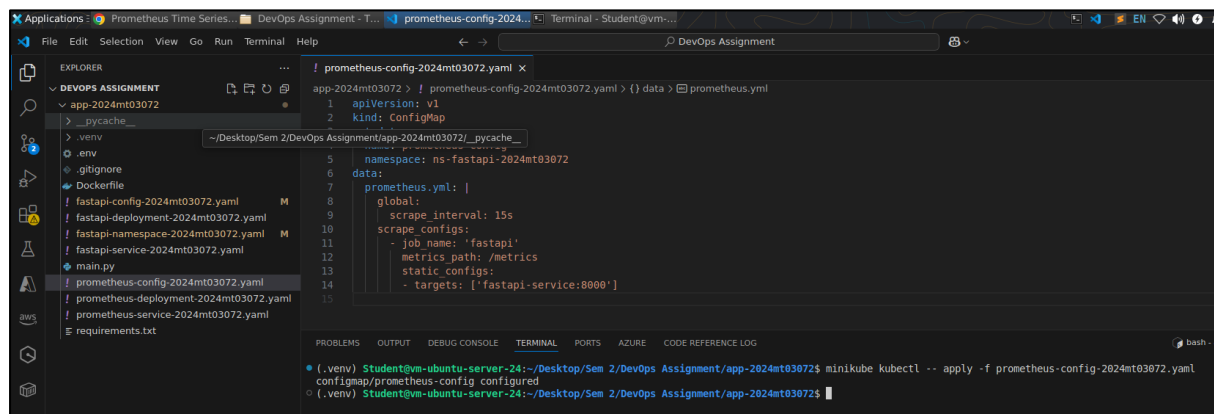
```
minikube kubectl -- apply -f prometheus-config-2024mt03072.yaml
```

- c. Verifying the creation of the Prometheus ConfigMap prometheus-config.

```
bash
```

```
minikube kubectl -- get configmaps -n ns-fastapi-2024mt03072
```

3.1. Screenshot



4. Prometheus Deployment Creation

- a. Creating Prometheus Deployment **prometheus** using the Deployment configuration file **prometheus-deployment-2024mt03072.yaml**.

File → <https://github.com/kaushikpaul90/bits-pilani-sem2-devops-app-2024mt03072/blob/main/app-2024mt03072/prometheus-deployment-2024mt03072.yaml>.

b. Applying the Prometheus Deployment.

```
bash
```

```
minikube kubectl -- apply -f prometheus-deployment-2024mt03072.yaml
```

c. Verifying the creation of the Prometheus Deployment prometheus.

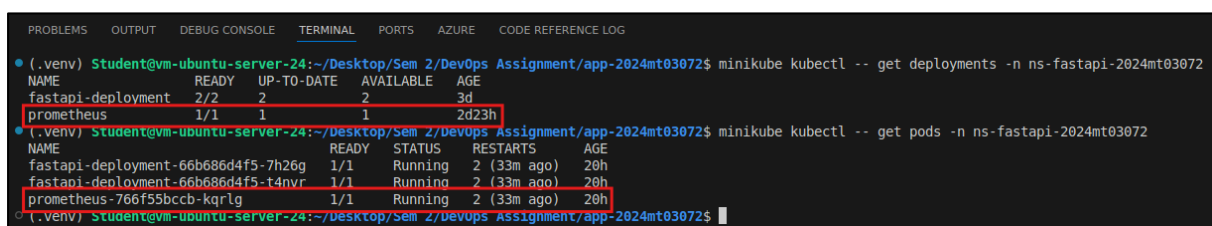
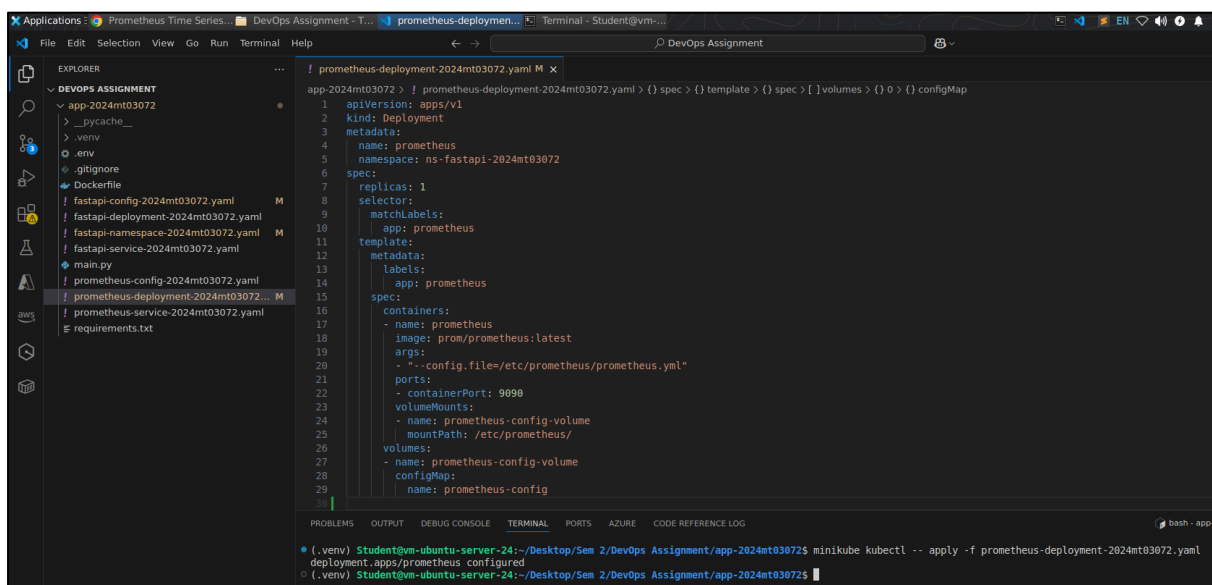
```
bash
```

```
minikube kubectl -- get deployments -n ns-fastapi-2024mt03072
```

d. Listing all pods running inside Kubernetes Namespace ns-fastapi-2024mt03072.

```
bash
```

```
minikube kubectl -- get pods -n ns-fastapi-2024mt03072
```



5. Prometheus Service Creation

- a. Creating Prometheus Service **prometheus** using the Service configuration file **prometheus-service-2024mt03072.yaml**.

File → <https://github.com/kaushikpaul90/bits-pilani-sem2-devops-app-2024mt03072/blob/main/app-2024mt03072/prometheus-service-2024mt03072.yaml>.

- b. Applying the Kubernetes Service.

```
bash
```

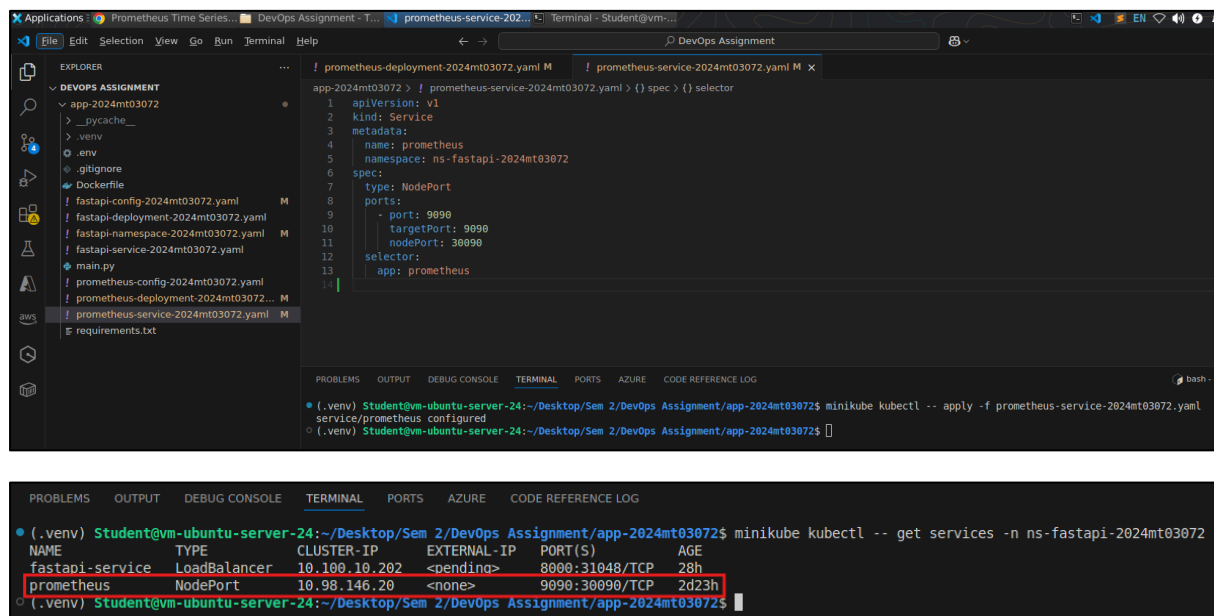
```
minikube kubectl -- apply -f prometheus-service-2024mt03072.yaml
```

- c. Verifying the creation of the Kubernetes Service prometheus.

```
bash
```

```
minikube kubectl -- get services -n ns-fastapi-2024mt03072
```

5.1. Screenshot



6. Obtaining Prometheus Access URL

- a. Retrieving the URL to access Prometheus UI from browser.

```
bash
```

```
minikube service prometheus -n ns-fastapi-2024mt03072 --url
```

Kaushik Paul

2024MT03072@WILP.BITS-PILANI.AC.IN

6.1. Screenshot

```

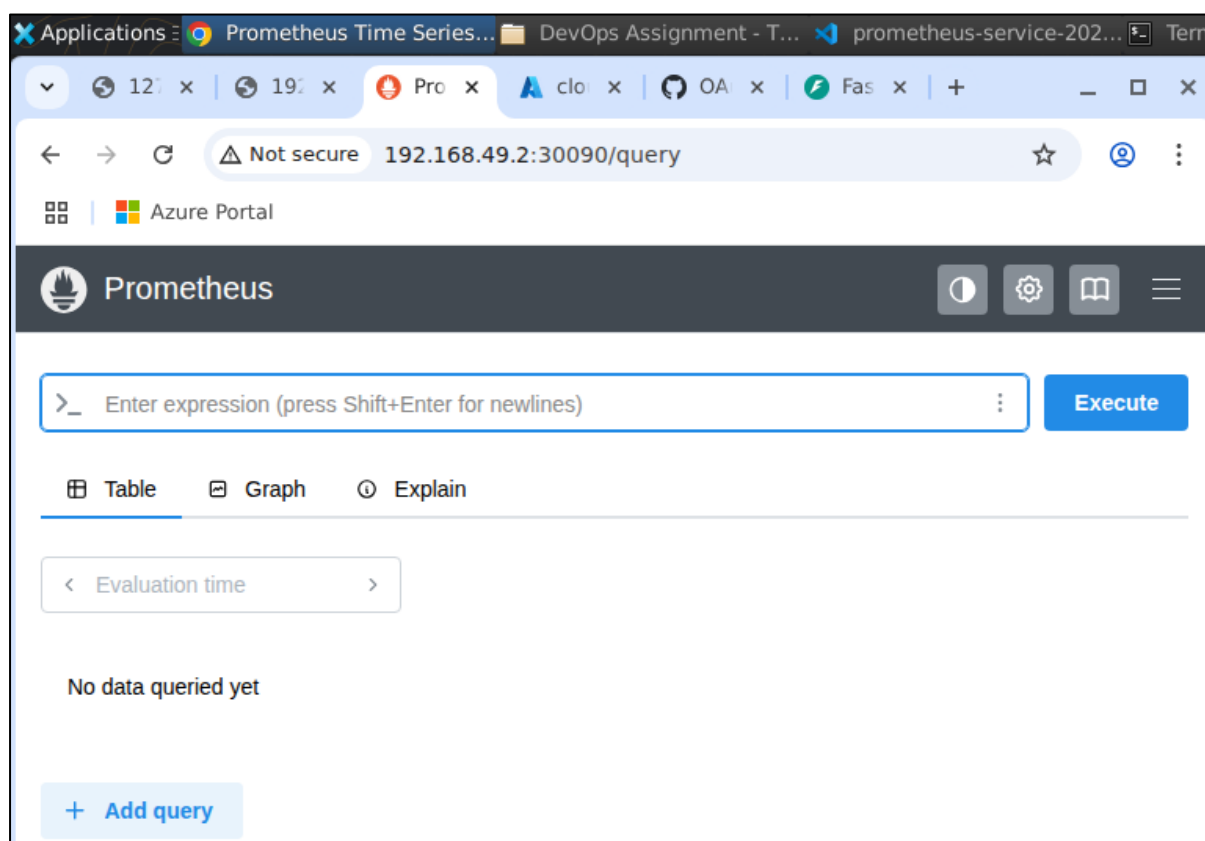
(.venv) Student@vm-ubuntu-server-24:~/Desktop/Sem 2/DevOps Assignment/app-2024mt03072$ minikube service prometheus -n ns-fastapi-2024mt03072 --url
http://192.168.49.2:30090
(.venv) Student@vm-ubuntu-server-24:~/Desktop/Sem 2/DevOps Assignment/app-2024mt03072$

```

7. Accessing Prometheus UI

- a. Accessing Prometheus UI via <http://192.168.49.2:30090> on Google Chrome.

7.1. Screenshot



8. Prometheus Metrics Collection

- a. Querying the number of requests received by the `/get_info` endpoint using metrics **`get_info_requests_total`**.
- b. Querying CPU usage for each replica using metrics **`container_cpu_usage_seconds_total`**.
 - **`container_cpu_usage_seconds_total{namespace="ns-fastapi-2024mt03072"}`** → This Prometheus query returns the total cumulative CPU

time (in seconds) consumed by each pod replica in the Kubernetes namespace ns-fastapi-2024mt03072 on the Minikube cluster.

c. Querying memory usage for each replica using metrics

container_memory_usage_bytes.

- **container_memory_usage_bytes{namespace="ns-fastapi-2024mt03072"} / (1024 * 1024)** → This Prometheus query returns the memory usage (in Mega Bytes) of each pod replica running in the Kubernetes namespace ns-fastapi-2024mt03072 on the Minikube cluster.

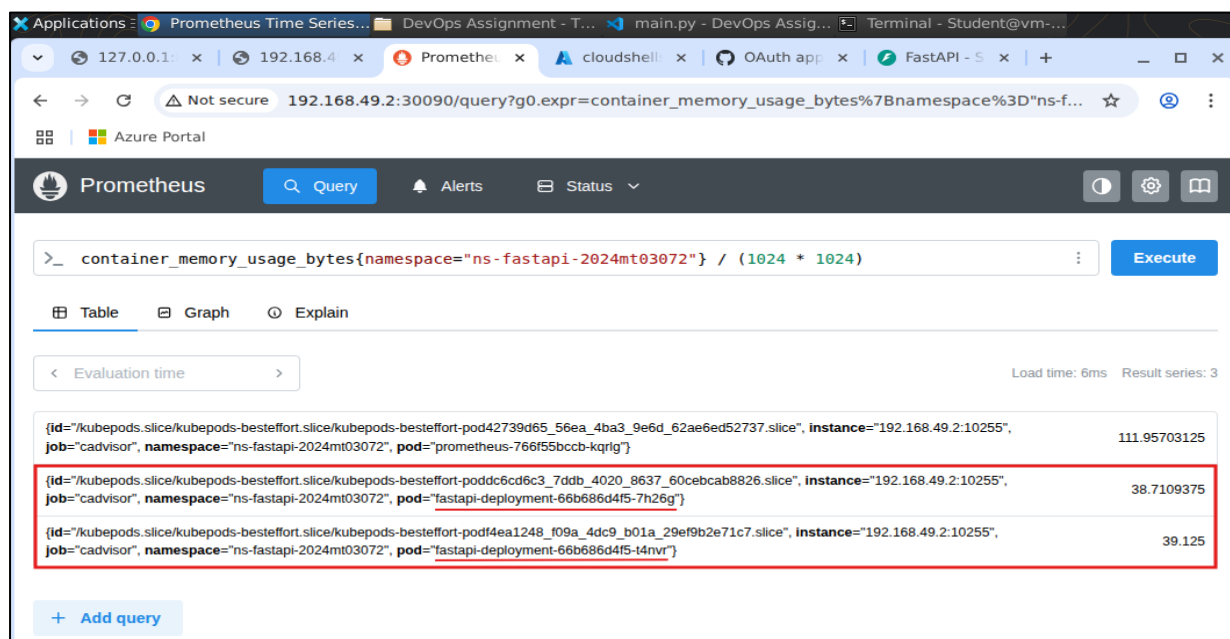
8.1. Screenshot

The screenshot shows the Prometheus web interface. The query bar contains the query: `>_ get_info_requests_total`. The 'Table' tab is selected, showing a single result with the value 1358. The query is executed successfully with a load time of 7ms and 1 result series.

get_info_requests_total(instance="fastapi-service:8000", job="fastapi")
1358

The screenshot shows the Prometheus web interface. The query bar contains the query: `>_ container_cpu_usage_seconds_total{namespace="ns-fastapi-2024mt03072"}`. The 'Table' tab is selected, showing three results. The second result is highlighted with a red box.

container_cpu_usage_seconds_total(cpu="total", id="/kubepods.slice/kubepods-besteffort.slice/kubepods-besteffort-pod42739d65_56ea_4ba3_9e6d_62ae6ed52737.slice", instance="192.168.49.2:10255", job="cadvisor", namespace="ns-fastapi-2024mt03072", pod="prometheus-766f55bccb-kqrlg")	Value
3.476243	3.476243
container_cpu_usage_seconds_total(cpu="total", id="/kubepods.slice/kubepods-besteffort.slice/kubepods-besteffort-podd6cd6c3_7ddb_4020_8637_60cebcab8826.slice", instance="192.168.49.2:10255", job="cadvisor", namespace="ns-fastapi-2024mt03072", pod="fastapi-deployment-66b686d4f5-7h26g")	12.144247
container_cpu_usage_seconds_total(cpu="total", id="/kubepods.slice/kubepods-besteffort.slice/kubepods-besteffort-podf4ea1248_f09a_4dc9_b01a_29ef9b2e71c7.slice", instance="192.168.49.2:10255", job="cadvisor", namespace="ns-fastapi-2024mt03072", pod="fastapi-deployment-66b686d4f5-t4nvr")	12.235638



9. Challenges Faced

a. Metric Discovery Complexity:

- Identifying the correct Prometheus metrics (container_cpu_usage_seconds_total, container_memory_usage_bytes) to monitor CPU/memory usage per pod.
- After lot of research and analysis, I came across the metrics container_cpu_usage_seconds_total and container_memory_usage_bytes to monitor CPU and memory usage per pod respectively.

b. Scraping Configuration:

- Even after identifying the metrics, Prometheus was not able to collect them without proper scraping configuration.
- Adding cAdvisor job to prometheus-config-2024mt03072.yaml exposed container-level metrics for Prometheus to scrape, finally resolving the issue.

Conclusion

This assignment provided a comprehensive, hands-on experience with the complete DevOps lifecycle for a cloud-native backend application. Starting from the development of a FastAPI-based microservice, the project covered containerization with Docker, orchestrated deployment using Kubernetes, and advanced monitoring with Prometheus.

Key learnings included:

- **Environment Management:** Leveraging environment variables and ConfigMaps for flexible configuration across local and cloud environments.
- **Containerization:** Building and running Docker images and containers, ensuring portability and consistency.
- **Kubernetes Orchestration:** Deploying scalable, highly available applications with replica management, namespace isolation, and service exposure for load balancing.
- **Monitoring and Observability:** Integrating Prometheus to collect and visualize application and infrastructure metrics and overcoming the challenge of identifying and scraping the correct metrics for resource usage and request tracking.
- **Troubleshooting:** Addressing common issues such as image update propagation, service exposure, and metric visibility, which are crucial skills for real-world DevOps scenarios.

Overall, this assignment not only solidified theoretical concepts but also enhanced practical skills in deploying, scaling, and monitoring modern cloud applications. The experience gained through iterative debugging and configuration will be invaluable for future projects in cloud-native DevOps and site reliability engineering.