



CLOUD COMPUTING

Project Report

Student Name	Mail ID	Social Security Number
Namburu Sri Harsha	srna21@student.bth.se	20001227-T235
Suvoj Reddy Kovvuri	Suko21@student.bth.se	20010306-T032
Lydia Sri Divya Dommeti	Lydo21@student.bth.se	20000915-T366

Introduction:

This is a flask application, a python based one. It is usually a simple and basic face detection web application in which user can upload an image and we will get an output detecting the faces.

Our main motive is to implement the application scalable and fault tolerant.

We also wanted to deploy in a cluster, because, we will get more power rather than that of a single standalone machine.

Steps:

- Create the application
- Dockerize it
- Push the image to AWS ECR
- Create a ECS cluster
- Define a task definition
- Open the cluster and create a service
- While creating the service and task, create and configure auto-scaling and load balancing as well
- Wait for the status changing to RUNNING state
- Copy the IP address and run on the machine

Implementation:

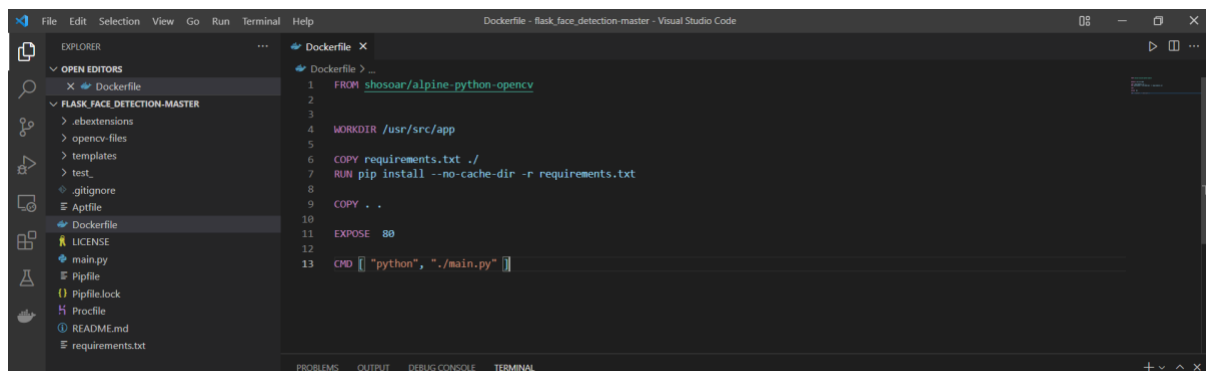
Step 1:

The application was made firstly

We then have developed a Dockerfile in which the commands to run will be there.

By building this file, we will get a customized image that contains our web facial detection flask application.

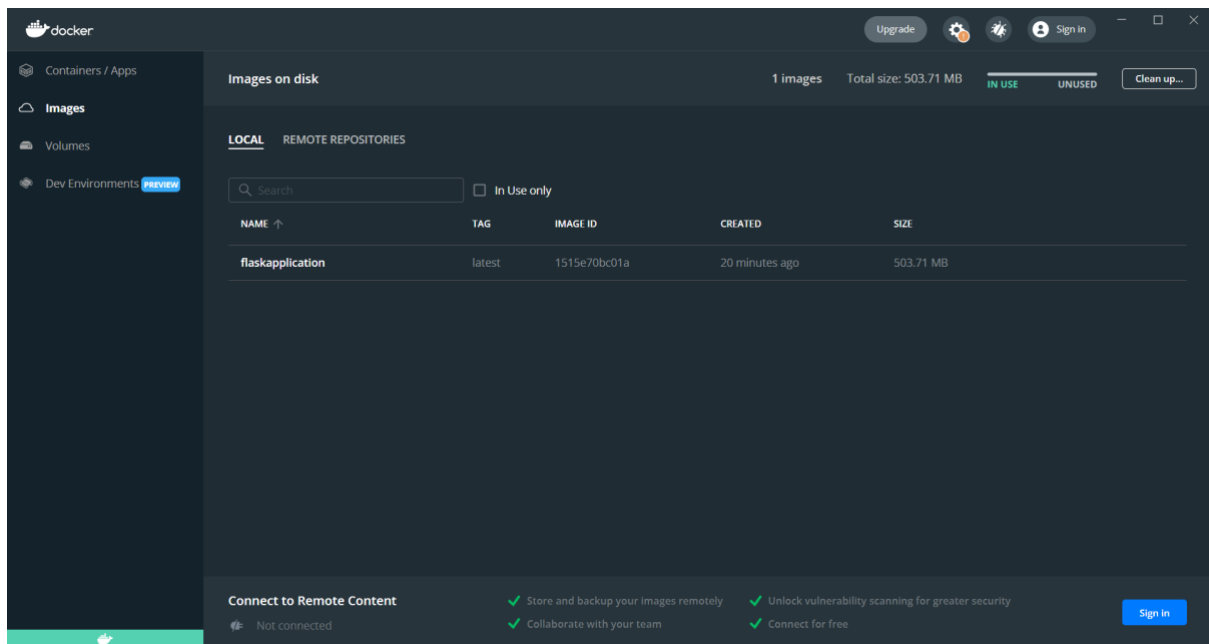
The following is the docker file that we have wrote for our flask application.

A screenshot of a Visual Studio Code editor window. The title bar reads "Dockerfile - flask_face_detection-master - Visual Studio Code". The Explorer sidebar on the left shows a file tree with folders like "FLASK_FACE_DETECTION-MASTER" and files like "Dockerfile", "LICENSE", "main.py", "Pipfile", "Pipfile.lock", "Procfile", "README.md", and "requirements.txt". The Dockerfile is open in the editor, showing the following content:

```
1 FROM shosoar/alpine-python-opencv
2
3
4 WORKDIR /usr/src/app
5
6 COPY requirements.txt ./
7 RUN pip install --no-cache-dir -r requirements.txt
8
9 COPY . .
10
11 EXPOSE 80
12
13 CMD ["python", "./main.py"]
```

Step 2:

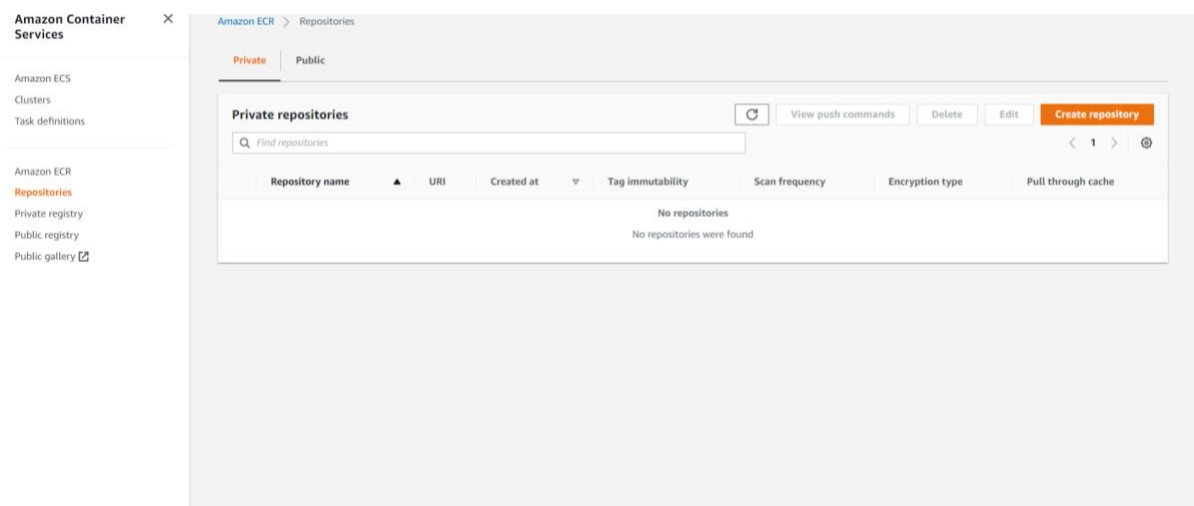
After building the application, we will get to see the image that have been build in the docker application as well



Step 3:

In the AWS account, a repository should be created in the ECR service offered by the Aws itself

The AWS provided all the things in GUI manner, so, it became easy to deploy the application;



We have created our repository in AWS as discussed above

The repository is ready.

Repository name	URI	Created at	Tag immutability	Scan frequency	Encryption type	Pull through cache
flaskapp	280041404702.dkr.ecr.ap-south-1.amazonaws.com/flaskapp	14 March 2022, 23:26:27 (UTC+05.5)	Disabled	Manual	AES-256	Inactive

Step 4:

Now the application should be pushed to the Aws repository.

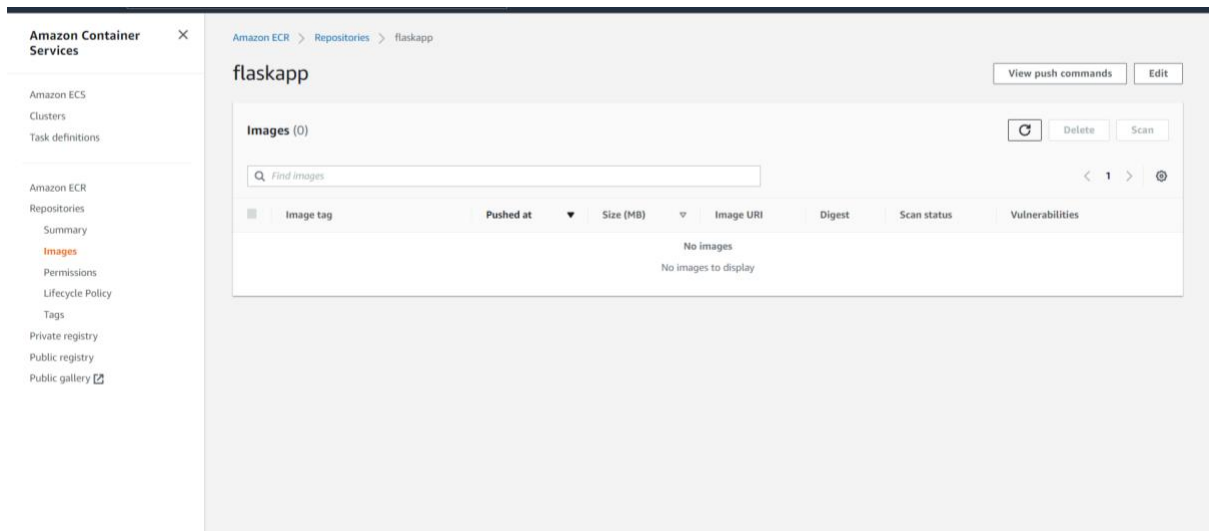
```

The push refers to repository [280041404702.dkr.ecr.ap-south-1.amazonaws.com/flaskapp]
16b6fa27fcfb: Pushing [=====] 590.6kB/96.7kB
b3713638eb49: Pushing [=====] 3.385MB/3.88MB
573a99a6cc72: Pushed
63d990b6ad8: Pushed
f5af6ce2c77: Pushing [>] 2.198MB/410.3MB
ed37abe528d2: Waiting
190791c80063: Waiting
af4997faab5b: Waiting
9f8566ee5135: Waiting

The push refers to repository [280041404702.dkr.ecr.ap-south-1.amazonaws.com/flaskapp]
16b6fa27fcfb: Pushed
b3713638eb49: Pushed
573a99a6cc72: Pushed
63d990b6ad8: Pushed
f5af6ce2c77: Pushed
ed37abe528d2: Pushed
190791c80063: Pushed
af4997faab5b: Pushed
9f8566ee5135: Pushed
latest: digest: sha256:3a793c8cae5c045a3be5495df42776400b391e1a2bbad2f83d89ac30ffa5 size: 2205
  
```

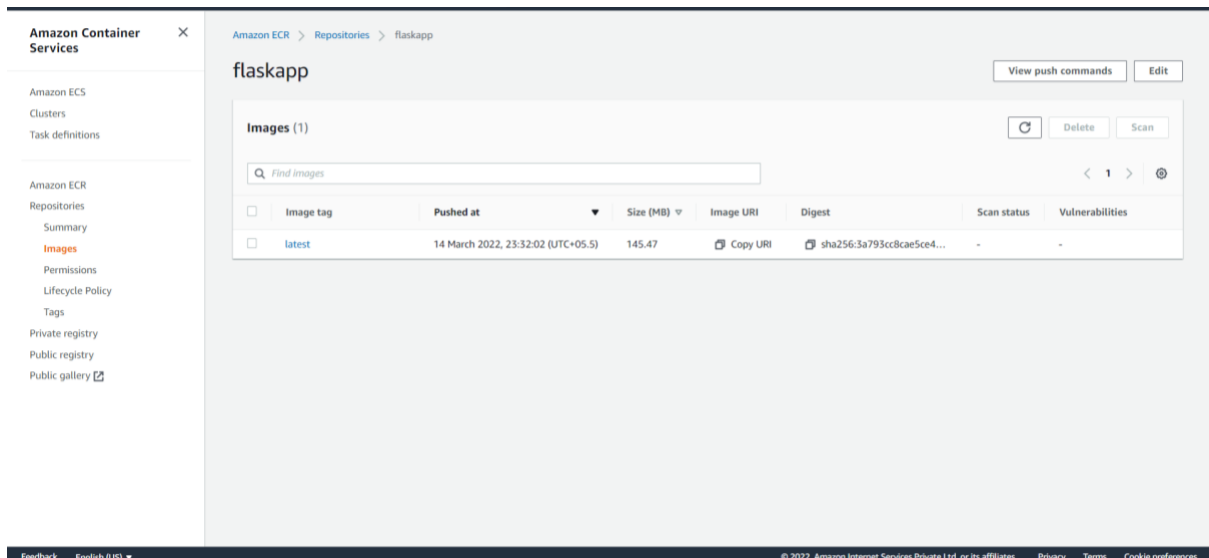
After successful push, we will get to see this

The below screenshot is the empty repository (before pushing)



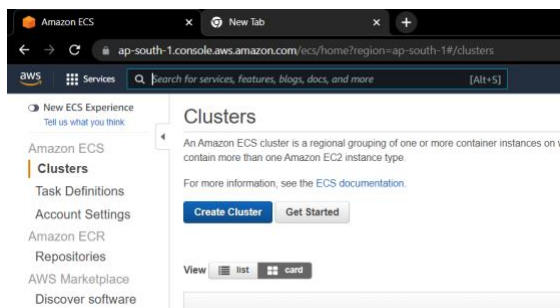
We can see our image that we have pushed from docker in the AWS repo.

This is the latest push we have done so far.

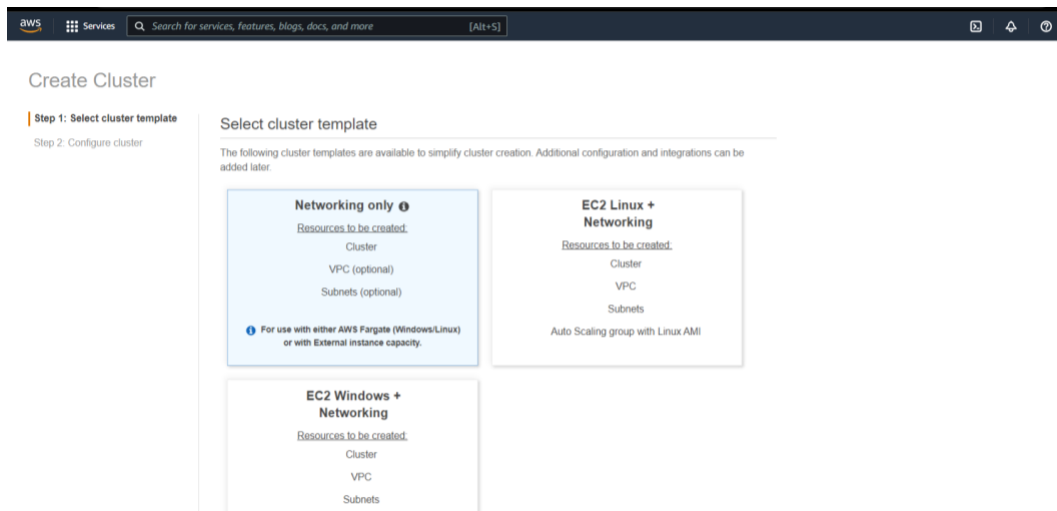


Step 5:

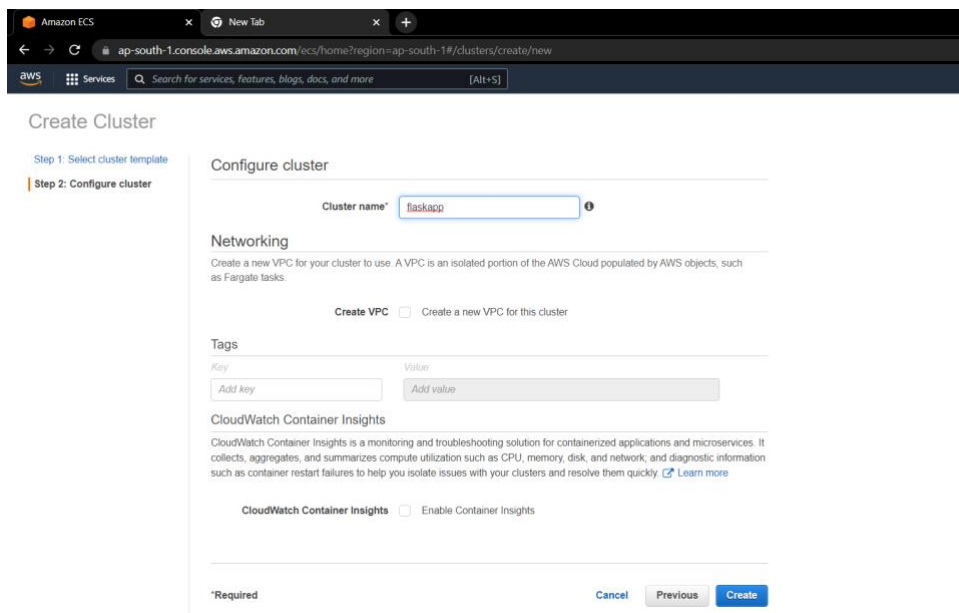
We had created an cluster using the ECS service that is beeing offered by AWS



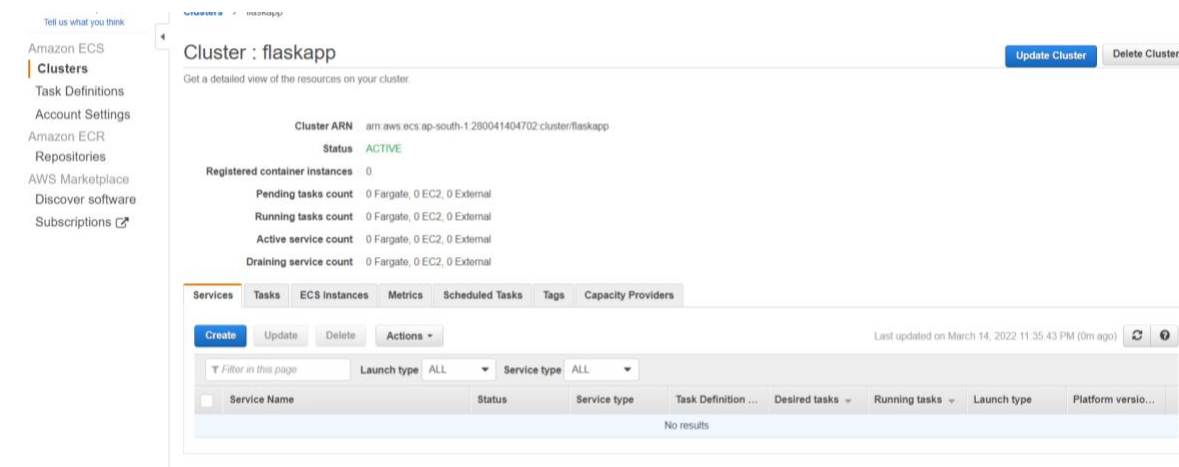
We opted for Networking only because, we will get more options to explore



After the previous step, we just have added few configurations to the cluster



At the end, we got a new cluster up and running on AWS



We had also to create a task

We opted for an auto management service that can be attached to ECS deployment, i.e., Fargate


Step 1: Select launch type compatibility

Step 2: Configure task and container definitions

Select launch type compatibility


Select which launch type you want your task definition to be compatible with based on where you want to launch your task.

FARGATE




Price based on task size
Requires network mode awsvpc
AWS-managed infrastructure, no Amazon EC2 instances to manage

EC2



Price based on resource usage
Multiple network modes available
Self-managed infrastructure using Amazon EC2 instances

EXTERNAL



Price based on instance-hours and additional charges for other AWS services used

SO, now we also started configuring the task definition

Create new task definition

Step 1: Select launch type compatibility

Step 2: Configure task and container definitions

A task definition specifies which containers are included in your task and how they interact with each other. You can also specify data volumes for your containers to use. [Learn more](#)

Task definition name* flaskapp ⓘ

Requires compatibilities* FARGATE

Task role ecsTaskExecutionRole ⓘ

Optional IAM role that tasks can use to make API requests to authorized AWS services. Create an Amazon Elastic Container Service Task Role in the IAM Console ⓘ

Network mode awsvpc ⓘ

If you choose «default», ECS will start your container using Docker's default networking mode, which is Bridge on Linux and NAT on Windows. Windows tasks support the «default» and awsvpc network modes.

Operating system family Select an operating system... ⓘ

All the important configurations needed to be done

services 🔍 Search for services, features, blogs, docs, and more [Alt+S] Mumbai anuj_77

container using Docker's default networking mode, which is Bridge on Linux and NAT on Windows. Windows tasks support the «default» and awsvpc network modes.

Operating system family Linux ⓘ

Task execution IAM role

This role is required by tasks to pull container images and publish container logs to Amazon CloudWatch on your behalf. If you do not have the ecsTaskExecutionRole already, we can create one for you.

Task execution role ecsTaskExecutionRole ⓘ

Task size

The task size allows you to specify a fixed size for your task. Task size is required for tasks using the Fargate launch type and is optional for the EC2 or External launch type. Container level memory settings are optional when task size is set. Task size is not supported for Windows containers.

Task memory (GB) 1GB ⓘ

The valid memory range for 0.5 vCPU is: 1GB - 4GB.

Task CPU (vCPU) 0.5 vCPU ⓘ

The valid CPU range for 1GB memory is: 0.25 vCPU - 0.5 vCPU.

Task memory maximum allocation for container memory reservation

Task CPU maximum allocation for containers

1024 shared of 1024 MiB

ish (US) © 2022, Amazon Internet Services Private Ltd. or its affiliates. Privacy Terms Cookie preferences

For the task to be submitted successfully, container memory reservation and CPU reservation must be set to non-zero values to prevent degradation of Windows containers.

Task memory (GB)
The valid memory range for 0.5 vCPU is: 1GB - 4GB

Task CPU (vCPU)
The valid CPU range for 1GB memory is: 0.25 vCPU - 0.5 vCPU

Task memory maximum allocation for container memory reservation

Task CPU maximum allocation for containers

Container definitions

[Add container](#)

Container Name...	Image	Hard/Soft mem...	CPU Unit...	GPU	Essential ...
No results					

Service integration

AWS App Mesh is a service mesh based on the Envoy proxy that makes it easy to monitor and control microservices. App Mesh

We need to add the container to the task definition

We should select the container that we have uploaded in the AWS container repository

We just need to add the container link to this window

amazon.com/ecs/home?region=ap-south-1#/taskDefinitions/create

Task size

The task size allows you to specify the number of vCPUs and memory for the EC2 or External Linux containers.

Task memory maximum allocation for container memory reservation

Task CPU maximum allocation for containers

Container definition

[Add container](#)

Container Name...

Service integration

AWS App Mesh is a service mesh based on the Envoy proxy that makes it easy to monitor and control microservices. App Mesh

Add container

Standard

Container name*

Image*

Private repository authentication* ☐

Memory Limits (MiB) Soft limit

[Add Hard limit](#)

Define hard and/or soft memory limits in MiB for your container. Hard and soft limits correspond to the 'memory' and 'memoryReservation' parameters, respectively, in task definitions. ECS recommends 300-500 MiB as a starting point for web applications.

Port mappings

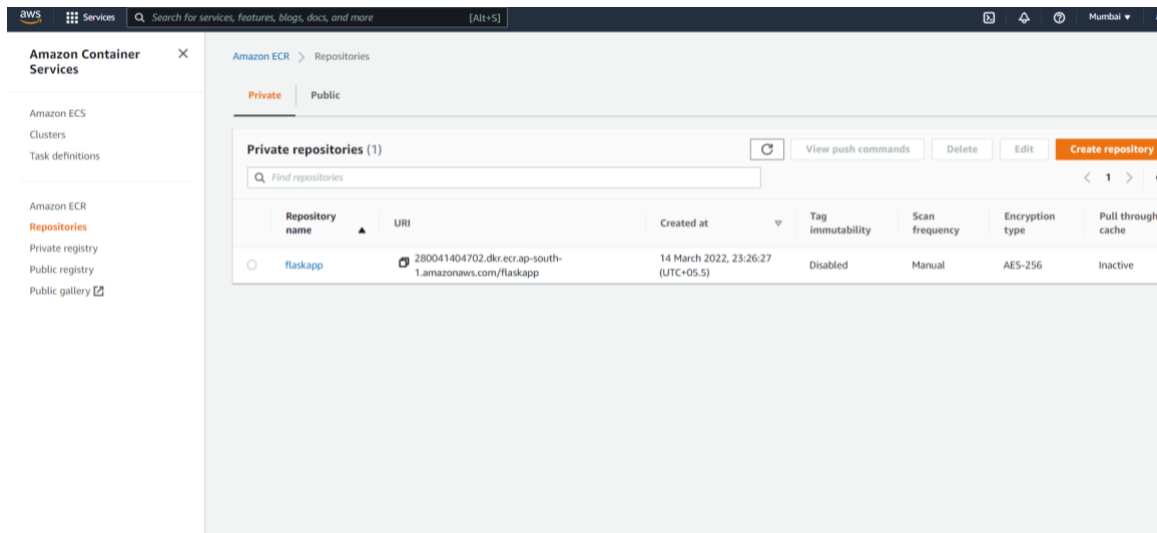
Container port Protocol

[Add port mapping](#)

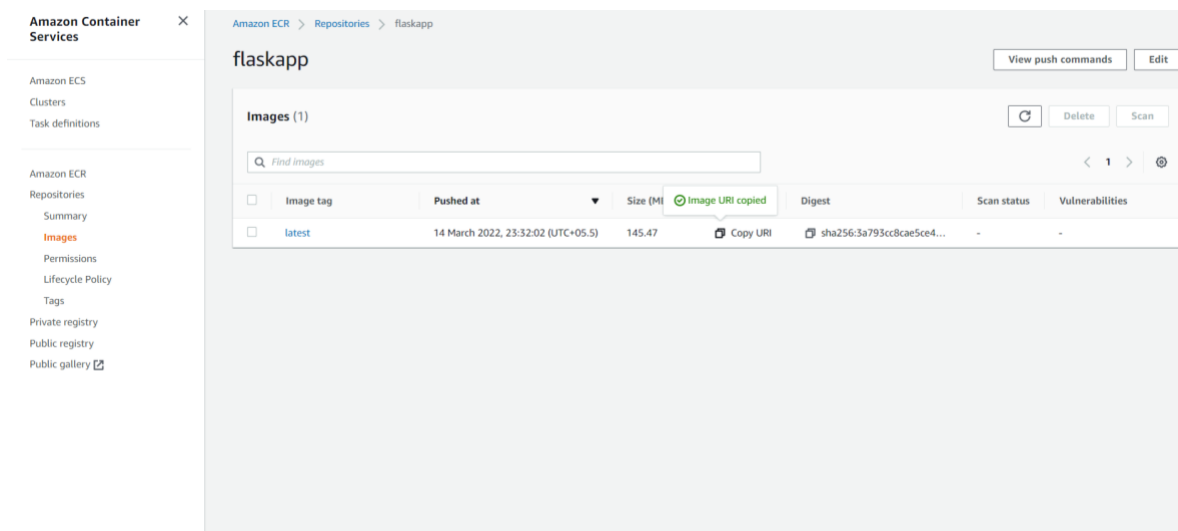
Host port mappings are not valid when the network mode for a task definition is host or awsvpc. To specify different host and container port mappings,

We can get the link from the repo.

Open the repo just by clicking on it.



Here we have the copy URI option. We just need to click on it



Paste the link in the image textbox

Do the port mappings as required

Amazon ECS

ap-south-1.console.aws.amazon.com/ecs/home?region=ap-south-1#/taskDefinitions/create

Search for services, features, blogs, docs, and more

Task size

The task size allows you to specify the EC2 or External Linux or Windows containers.

Task memory maximum

Task CPU maximum

Container definition

Add container

Container Name

Service integration

AWS App Mesh is a service mesh that standardizes how your microservices communicate.

Add container

Standard

Container name* flaskapp

Image* 280041404702.dkr.ecr.ap-south-1.amazonaws.com/flaskapp:latest

Private repository authentication* ☐

Memory Limits (MiB)

Soft limit 128

Add Hard limit

Define hard and/or soft memory limits in MiB for your container. Hard and soft limits correspond to the "memory" and "memoryReservation" parameters, respectively, in task definitions. ECS recommends 300-500 MiB as a starting point for web applications.

Port mappings

Container port 80 Protocol tcp

Add port mapping

Host port mappings are not valid when the network mode for a task definition is host or awsvpc. To specify different host and container port mappings, use the "hostPort" parameter in the "portMappings" array.

Do the other configurations if they are required

Amazon ECS

ap-south-1.console.aws.amazon.com/ecs/home?region=ap-south-1#/taskDefinitions/create

Search for services, features, blogs, docs, and more

Log configuration

The configuration details above, otherwise must be

Enable proxy configuration

Log router integration

FireLens for Amazon ECS

analysis. FireLens works choose Apply Learn more

Enable FireLens

Volumes

Use a volume configuration, and then choose a volume

Add volume

Configure via JSON

Tags

Key

Add key

*Required

Add container

Add volumes

Log configuration ☒ Auto-configure CloudWatch Logs

Log driver awslogs

Log options

Key Value

awslogs-group /ecs/flaskapp

awslogs-region ap-south-1

awslogs-stream-prefix ecs

Add key Value Add value

RESOURCE LIMITS

Limits

Limit name Soft limit Hard limit

none

Add ulimit

DOCKER LABELS

Key value pairs

Key Value

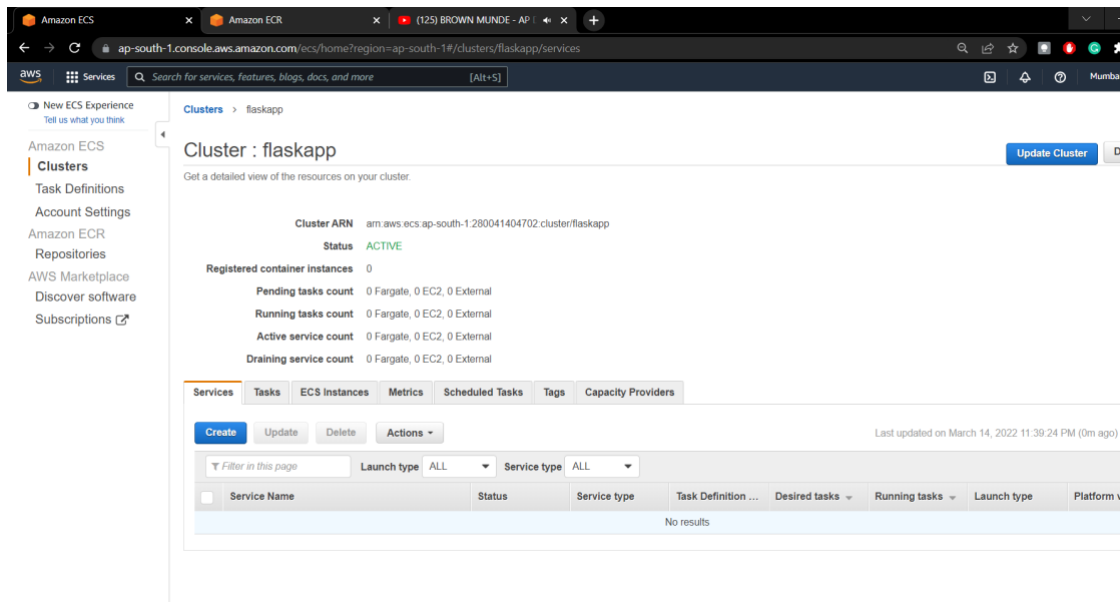
Add key Add value

The screenshot shows the 'Create' page for a new task definition in the Amazon ECS console. The browser tabs include 'Amazon ECS', 'Amazon ECR', and '(125) BROWN MUNDE - AP'. The URL is 'ap-south-1.console.aws.amazon.com/ecs/home?region=ap-south-1#/taskDefinitions/create'. The page contains several sections: 'Proxy configuration' with an 'Enable proxy configuration' checkbox; 'Log router integration' with an 'Enable FireLens integration' checkbox; 'Volumes' with an 'Add volume' button and a 'Configure via JSON' button; and 'Tags' with 'Add key' and 'Add value' input fields. At the bottom, there are 'Cancel', 'Previous', and 'Create' buttons.

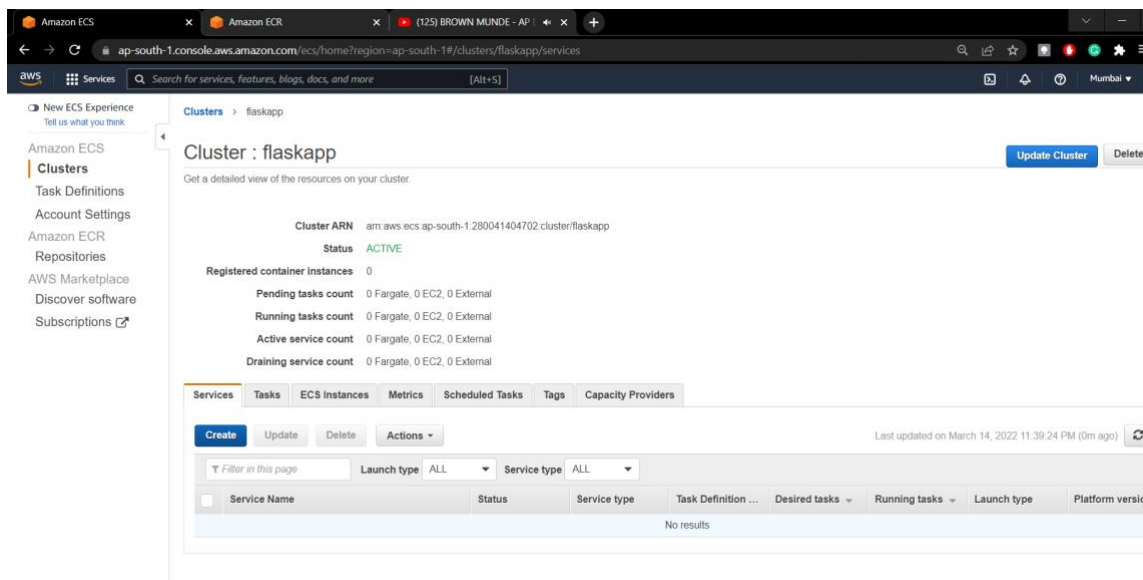
We have our Task definition ready to deploy

The screenshot shows the 'Task Definition: flaskapp:1' page in the Amazon ECS console. The left sidebar contains navigation links for 'New ECS Experience', 'Amazon ECS Clusters', 'Task Definitions', 'Account Settings', 'Amazon ECR Repositories', 'AWS Marketplace Discover software', and 'Subscriptions'. The main content area shows the task definition details for 'flaskapp'. It includes a 'Builder' tab (selected), 'JSON', and 'Tags' tabs. The 'Builder' tab shows the following configuration: 'Task definition name' is 'flaskapp', 'Task role' is 'ecsTaskExecutionRole', 'Network mode' is 'awsvpc' (with a note about Docker's default networking mode), 'Operating system family' is 'Linux', 'Compatibilities' are 'EC2, FARGATE', and 'Requires compatibilities' is 'FARGATE'. There are 'Create new revision' and 'Actions' buttons at the top.

We can now navigate to our cluster



We are now ready to create the service



Service configuration;

We have used fargate

Amazon ECS

Amazon ECR

(125) BROWN MUNDE - AP

ap-south-1.console.aws.amazon.com/ecs/home?region=ap-south-1#/clusters/flaskapp/createService

Services

Search for services, features, blogs, docs, and more

[Alt+S]

Mumbai

Configure service

A service lets you specify how many copies of your task definition to run and maintain in a cluster. You can optionally use an Elastic Load Balancing load balancer to distribute incoming traffic to containers in your service. Amazon ECS maintains that number of tasks and coordinates task scheduling with the load balancer. You can also optionally use Service Auto Scaling to adjust the number of tasks in your service.

Launch type

☒ FARGATE

☐ EC2

☐ EXTERNAL

Switch to capacity provider strategy

Operating system family

Linux

Task Definition

Family

flaskapp

Enter a value

Revision

1 (latest)

Platform version

LATEST

Cluster

flaskapp

Service name

flaskapp

Service type

REPLICA

Deployment circuit breaker

Disabled

Deployments

Choose a deployment option for the service.

Deployment type

☒ Rolling update

☐ Blue/green deployment (powered by AWS CodeDeploy)

This sets AWS CodeDeploy as the deployment controller for the service. A CodeDeploy application and deployment group are created automatically with default settings for the service. To change to the rolling update deployment type after the service has been created, you must re-create the service and select the "rolling update" deployment type.

Task tagging configuration

☒ Enable ECS managed tags

Propagate tags from

Do not propagate

Tags

Key	Value
Add key	Add value

*Required

Cancel

Next step

Create Service

Step 1: Configure service

Step 2: Configure network

Step 3: Set Auto Scaling (optional)

Step 4: Review

Do the configurations as required

We can create the load balancer

Health check grace period ⓘ

Load balancing

An Elastic Load Balancing load balancer distributes incoming traffic across the tasks running in your service. Choose an existing load balancer, or create a new one in the [Amazon EC2 console](#).

Load balancer type*

☐ None
Your service will not use a load balancer.

☐ Application Load Balancer
Allows containers to use dynamic host port mapping (multiple tasks allowed per container instance). Multiple services can use the same listener port on a single load balancer with rule-based routing and paths.

☒ Network Load Balancer
A Network Load Balancer functions at the fourth layer of the Open Systems Interconnection (OSI) model. After the load balancer receives a request, it selects a target from the target group for the default rule using a flow hash routing algorithm.

☐ Classic Load Balancer
Requires static host port mappings (only one task allowed per container instance); rule-based routing and paths are not supported.

Service IAM role Task definitions that use the awsipc network mode use the AWSServiceRoleForECS service-linked role, which is created for you automatically. [Learn more](#).

Load balancer name ⓘ

Auto configuration options are available as well

Create Service

Step 1: Configure service
Step 2: Configure network
Step 3: Set Auto Scaling (optional)
Step 4: Review

Set Auto Scaling (optional)

Automatically adjust your service's desired count up and down within a specified range in response to CloudWatch alarms. You can modify your Service Auto Scaling configuration at any time to meet the needs of your application.

Service Auto Scaling ☐ Do not adjust the service's desired count
☒ Configure Service Auto Scaling to adjust your service's desired count

Minimum number of tasks ⓘ
Automatic task scaling policies you set cannot reduce the number of tasks below this number.

Desired number of tasks ⓘ

Maximum number of tasks ⓘ
Automatic task scaling policies you set cannot increase the number of tasks above this number.

IAM role for Service Auto Scaling You are giving permission to Elastic Container Service to create and use

We have just configured the both auto scaling and load balancing options

Create security group

✔ Create security group
flaska-5808 succeeded sg-0f658f5bb5f7e4b52

Set inbound rules

✔ Set inbound rules
succeeded sg-0f658f5bb5f7e4b52

Create Service

Create service: flaskapp

✔ Service created
Service created. Tasks will start momentarily. View: flaskapp

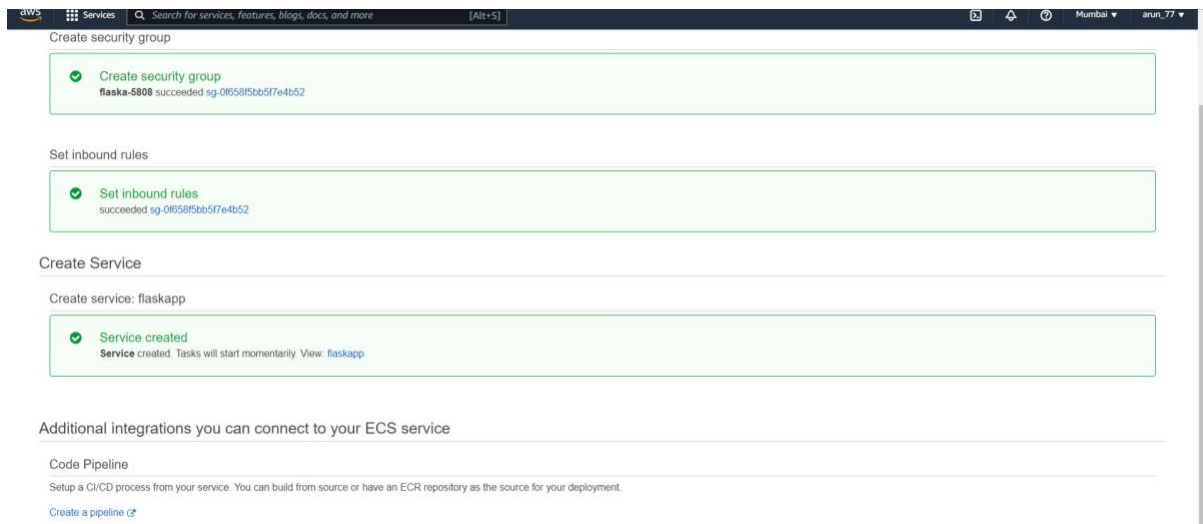
Additional integrations you can connect to your ECS service

Code Pipeline

Setup a CI/CD process from your service. You can build from source or have an ECR repository as the source for your deployment.

[Create a pipeline](#) ⓘ

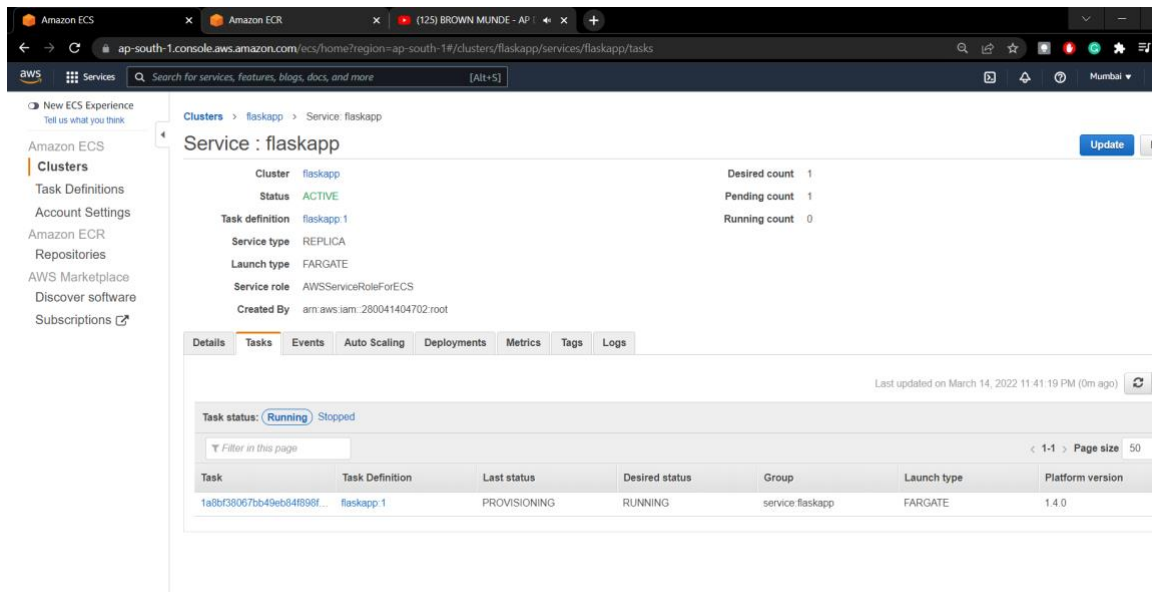
[Back](#) [View Service](#)



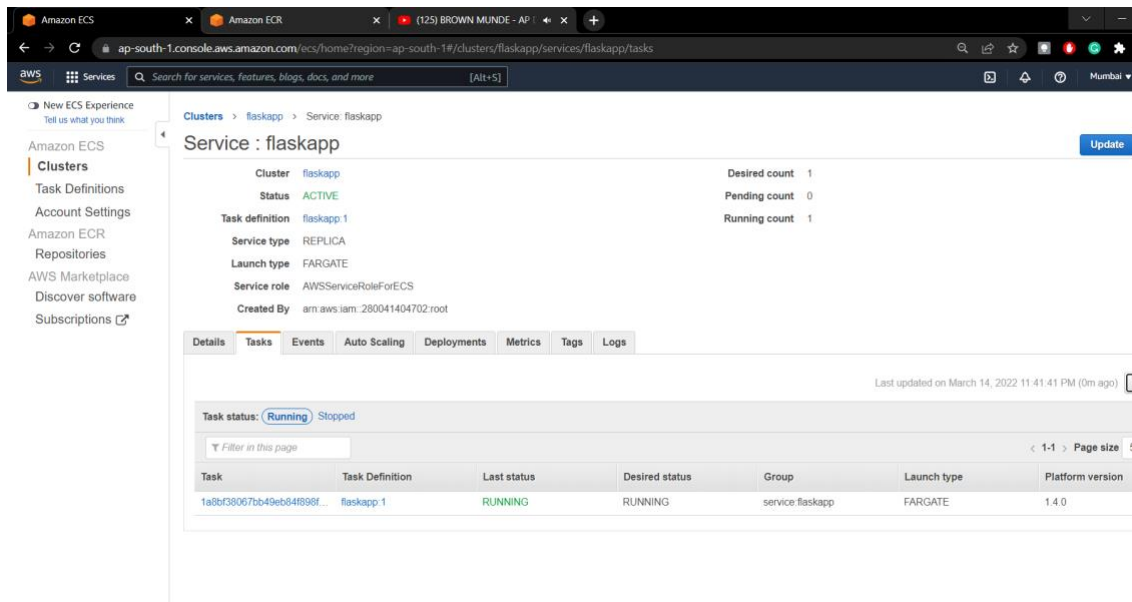
The service is up and running

We can navigate to view our service and the cluster

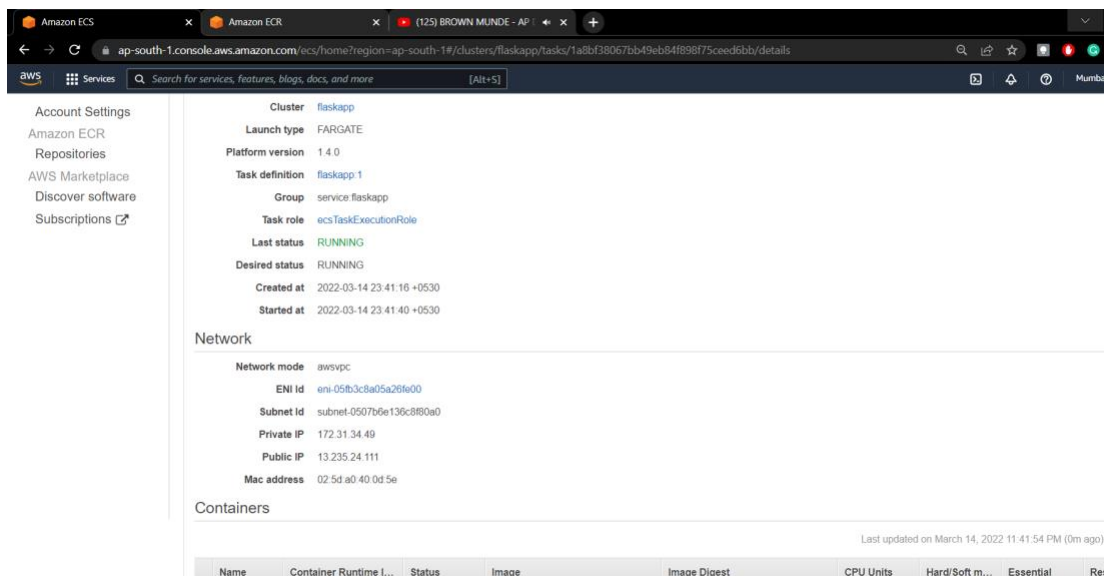
We just need a couple of minutes



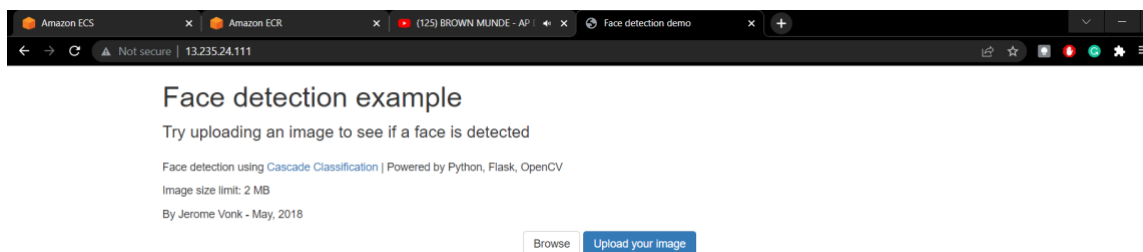
The service is ready to run and is running



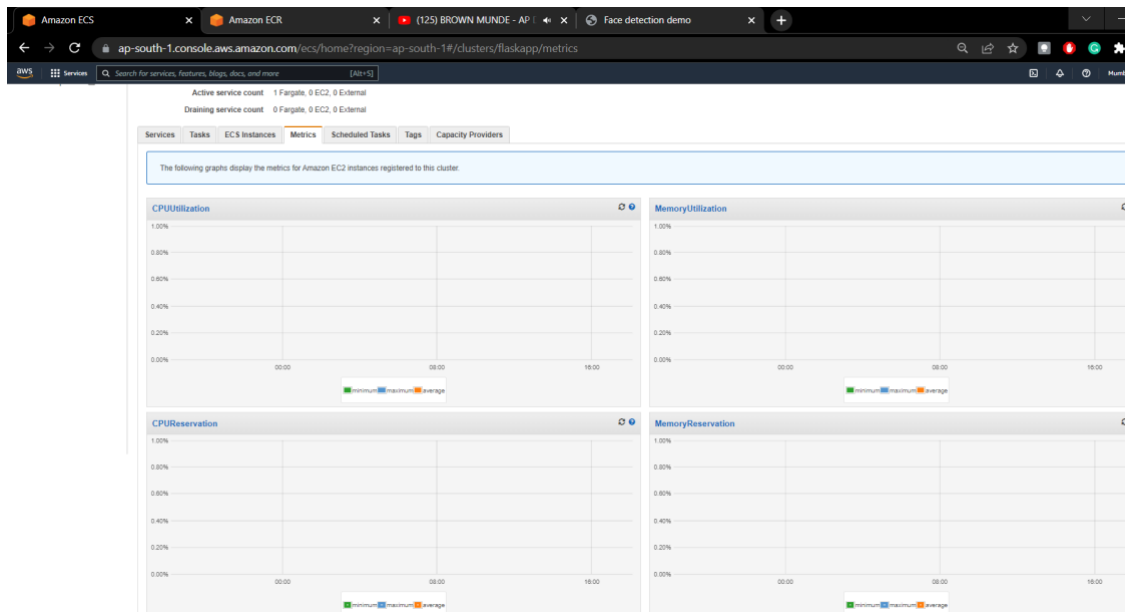
Just navigated to the service and copy the ip address (public)



Search the ip address in the browser and you will get to see our website up and running



This is the matrixes that are provided by AWS, we can see the trends.



We can also check the status, and insights of load balancing and auto scaling

We have tested on various platforms and the application was running and working seamlessly without any delay and it is very faster than the local machine because, the cluster works faster than that of a single machine.

Hence the project went successful