

Guideline to deploy Flask App with DynamoDB on AWS ECS Farget

This document provides guidelines on how to deploy a Flask app and DynamoDB on AWS ECS Fargate.

Requirements:

- An AWS account.
- A Flask app that you want to deploy.
- A Docker file for building a Docker image of your Flask app.
- A Docker registry to store your Docker image (e.g., Amazon ECR or Docker hub)
- A DynamoDB table to store and retrieve data.
- An AWS ECS Fargate cluster to run your containers.
- An AWS ECS task definition that describes how to run your containers.
- An AWS ECS service that manages your containers and automatically scales them based on demand.
- Application Load Balancer
- A Security Group both for the Cluster and Application Load Balancer
- An AWS IAM role with the necessary permissions to access AWS resources required for your application to run. Here below are some of them:

Permission	User	Cluster
AmazonECSTaskExecutionRolePolicy	False	True
AmazonDynamoDBFullAccess	True	True
AmazonEC2ContainerRegistryFullAccess	True	True
AmazonECS_FullAccess	True	False
CloudWatchLogsFullAccess	True	True

NOTE: These are all the permissions need but not limited to them!

Here are the steps to follow to deploy a flask app and DynamoDB on AWS ECS Farget:

1. Set up your Flask app.
2. Create a Docker file.
3. Build and test the Docker image.
4. Push the Docker image to AWS ECR/Docker hub.
5. Create Security Group.

Here we need two distinct security groups. One for the cluster to accept requests from the Application Load Balancer and the other for the Application Load Balancer to accept outside http requests from anywhere.

- AWS Management Console
- Search for EC2
- Security Groups
- Create Security Group
 - Provide name for security group.
 - Select VPC
 - Add Inbound Rules

NB:

- For the Application Load Balancer security group, the inbound rule is to accept any **HTTP** request with **TCP** protocol on port **80** coming from any source **0.0.0.0/0**.
- For the Cluster security group, the inbound rule is to accept **ALL TCP** requests coming only from the **Application Load Balancer Security Group**.

6. Create a new ECS Task Definition.

Here are the steps:

- AWS Management Console.
- Search for ECS.
- Task definitions.
- Create new task definitions.
 - On Task definition configuration Section fill in:
 - Task definition family.
 - On Container Section fill in:
 - Name
 - Image URI
 - Secrets Manager ARN or name

- Port Mapping
 - Container port
 - Environment variables if any
- Next
- On Environment Section Fill in
 - App environment (select to be AWS Farget)
 - Operating system/Architecture
 - Task size
 - CPU and Memory
 - Task roles, network mode
 - Task role (select escTaskExecutionRole)
 - Task execution role (select escTaskExecutionRole)
- Next
- Create

7. Create a new ECS Cluster.

Here are the steps:

- AWS Management Console.
- Search for ECS.
- Clusters
- Create Cluster
 - On Cluster Configuration Section Fill in
 - Cluster name
 - On Networking Section
 - VPC (select default VPC)
 - Subnets (remove lambda related subnets and leave the rest 6 subnets)
 - Default namespace
 - Create

8. Launch a new ECS Service.

Here are the steps:

- AWS Management Console.
- Search ECS
- Clusters
- Select previously created cluster.
- Service Tab
- Create
 - On Environment Section Fill in
 - Compute configuration (select Launch Type)

- Launch Type
- Platform version
- On Deployment configuration Section Fill in
 - Application type (select Service)
 - Family (select previously created task definition)
 - Service name
 - Desired tasks
- On Networking Section Fill in
 - VPC (select default VPC)
 - Subnets (remove lambda related subnets and leave the rest 6 subnets)
 - Security group (select Use an existing security group)
 - Security group name (select previously created security group for the **Cluster**) and remove the default.
- On Load balancing Section Fill in
 - Load balancer type (select Application Load Balancer)
 - Application Load Balancer (select Create a new load balancer)
 - Load balancer name
 - Target group name
 - Health check grace period (mostly 20 sec)
 - Create

9. Configure Load Balancer

Here are the steps:

- AWS Management Console.
- Search EC2
- Load Balancing
- Load Balancer
 - Select the previously created load balancer.
 - Security Tab
 - Edit
 - On Security groups Section
 - Security groups (remove any other security group attached to it and give it the Application Load Balancer Security Group created previously)

10. Test

From the previous page (load balancer page) find the DNS name, copy it and check it on browser