Setup AWS ECS Faragte with Application Load Balancer

18 JAN 2022 Hai Tran

Step 1. Prepare a Docker script

```
FROM python:3.7-slim

COPY ./requirements.txt /app/requirements.txt

WORKDIR /app

RUN pip install -r requirements.txt

COPY . /app

EXPOSE 8081

ENTRYPOINT [ "python" ]

CMD [ "app.py" ]
```

local build

```
docker build -t flask-app .
```

local run to test

```
docker run -d -p 56733:8081 flask-app:latest
```

check containers ID runnings

```
docker ps
```

go to browser and check the web server flask is running

```
http://localhost:56733
```

push docker image to aws ecr follow this
authenticate with a profile noted

```
aws ecr get-login-password --region region | docker login --username AWS --
password-stdin aws_account_id.dkr.ecr.region.amazonaws.com
```

tag

```
sudo docker tag dc73321f7fab aws_account_id.dkr.ecr.ap-southeast-
1.amazonaws.com/flask-app:latest
```

push

```
sudo docker push aws_account_id.dkr.ecr.ap-southeast-1.amazonaws.com/flask-
app:latest
```

go to aws console and take note the image id

```
aws_account_id.dkr.ecr.ap-southeast-1.amazonaws.com/flask-app:latest
```

Step 2. Setup an ECS cluster and task definition

go to aws console ecs service and create a cluster named **FhrProcessingCluster**
create a cluster

Select cluster template

The following cluster templates are available to simplify cluster creation. Additional configuration and integrations can be added later.

Networking only ℹ

Resources to be created:

Cluster

VPC (optional)

Subnets (optional)

ℹ **For use with either AWS Fargate (Windows/Linux)
or with External instance capacity.**

**EC2 Linux +
Networking**

Resources to be created:

Cluster

VPC

Subnets

Auto Scaling group with Linux AMI

**EC2 Windows +
Networking**

Resources to be created:

Cluster

VPC

Subnets

Auto Scaling group with Windows AMI

*Required   Cancel   **Next step**

create a task definition

Task Definitions

Task definitions specify the container information for your application, such as how many containers are part of your task, what resources they will use, how they are linked together, and which host ports they will use. Learn more

Create new Task Definition   Create new revision   Actions ▾

Last updated on January 18, 2022 9:10:18 PM (0m ago)

Status: (ACTIVE) INACTIVE

▼ Filter in this page   ‹ **1-1** › **Page size** 50 ▼

☐	Task Definition	Latest revision status
☐	first-run-job-definition	ACTIVE

view task definition

ion

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

add container

Add container ✕

▼ Standard

Container name* | FhrProcessingContainer | ⓘ

Image* | 717869305038.dkr.ecr.ap-southeast-1.amazonaws.com/flask-app:latest | ⓘ

Private repository authentication* | ☐ | ⓘ

Memory Limits (MiB) | Soft limit ▼ | 2048 | ⓘ

⊕ 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.

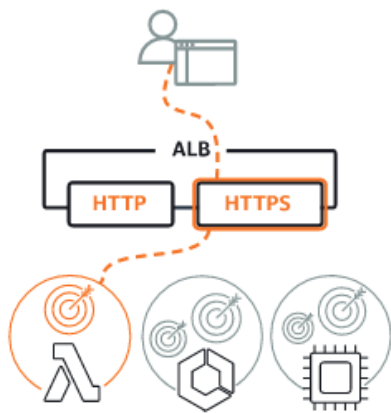**Port mappings** | *Container port* | *Protocol* | ⓘ

8081 | tcp ▼ | ✖

⊕ Add port mapping

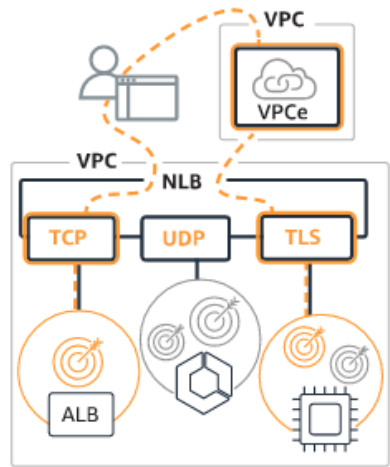create load balancer as step 4.

Load balancer types

Application Load Balancer Info

Choose an Application Load Balancer when you need a flexible feature set for your applications with HTTP and HTTPS traffic. Operating at the request level, Application Load Balancers provide advanced routing and visibility features targeted at application architectures, including microservices and containers.

Create

Network Load Balancer Info

Choose a Network Load Balancer when you need ultra-high performance, TLS offloading at scale, centralized certificate deployment, support for UDP, and static IP addresses for your applications. Operating at the connection level, Network Load Balancers are capable of handling millions of requests per second securely while maintaining ultra-low latencies.

Create

Gateway Load Balancer Info

Choose a Gateway Load Balancer when you need to deploy and manage a fleet of third-party virtual appliances that support GENEVE. These appliances enable you to improve security, compliance, and policy controls.

Create

create a service

Service IAM role   Task definitions that use the awsvpc network mode use the AWSServiceRoleForECS service-linked role, which is created for you automatically. Learn more.

Load balancer name   FhrEcsProcessingALB ▼   ⟳

Container to load balance

FhrProcessingContainer : 8081   Remove ✖

Production listener port*   80:HTTP ▼   ❶

Production listener protocol*   HTTP

Target group name   FhrEcsProcessingTar... ▼   ❶

Target group protocol   HTTP ❶

Target type   ip ❶

Path pattern   /
Evaluation order   default

Health check path   /   ❶

Additional health check options can be configured in the ELB console after you create your service.

App Mesh

To use your service with App Mesh, you must
- Ensure your task definition is configured properly. Edit your task definition if you haven't done this.
- Set up your service to use Service Discovery.

configure network for the ecs service

Configure network

VPC and security groups

VPC and security groups are configurable when your task definition uses the awsvpc network mode.

Cluster VPC*   vpc-4dd37228 (172.31.0.0/1... ▼   ⓘ

Subnets*   ⓘ

> subnet-2ca8145b ⊗
> (172.31.16.0/20) - ap-southeast-1a
> assign ipv6 on creation: Disabled

> subnet-b249dfd7 ⊗
> (172.31.0.0/20) - ap-southeast-1b
> assign ipv6 on creation: Disabled

> ▼

Security groups*   FhrPro-3349   [Edit]   ⓘ

Auto-assign public IP   ENABLED ▼   ⓘ

Health check grace period

If your service's tasks take a while to start and respond to ELB health checks, you can specify a health check grace period of up to 2,147,483,647 seconds during which the ECS service scheduler will ignore ELB health check status. This grace period can prevent the ECS service scheduler from marking tasks as unhealthy and stopping them before they have time to come up. This is only valid if your service is configured to use a load balancer.

Health check grace period   requires a load balancer.   ⓘ

Load balancing

configure container for the ecs service

Service IAM role   Task definitions that use the awsvpc network mode use the AWSServiceRoleForECS service-linked role, which is created for you automatically. Learn more.

Load balancer name   FhrEcsProcessingALB ▼   ⟳

Container to load balance

FhrProcessingContainer : 8081   Remove ✕

Production listener port*   80:HTTP ▼   ℹ

Production listener protocol*   HTTP

Target group name   FhrEcsProcessingTar... ▼   ℹ

Target group protocol   HTTP ℹ

Target type   ip ℹ

Path pattern   /

Evaluation order   default

Health check path   /   ℹ

Additional health check options can be configured in the ELB console after you create your service.

App Mesh

To use your service with App Mesh, you must
- Ensure your task definition is configured properly. Edit your task definition if you haven't done this.
- Set up your service to use Service Discovery.

important configure security for ecs service so that inbound from the ALB enabled

VPC > Security Groups > sg-02caf1288675b1b7a - FhrPro-3349 > Edit inbound rules

Edit inbound rules Info

Inbound rules control the incoming traffic that's allowed to reach the instance.

Inbound rules Info

Security group rule ID	Type Info	Protocol Info	Port range Info	Source Info	Description - optional Info	
sgr-0f02ebcf558688da6	All TCP ▼	TCP	0 - 655	Custom ▼ Q		Delete
				sg-0fd9cc095f3f5f9f6 ✕		

Add rule

Cancel   Preview changes   **Save rules**

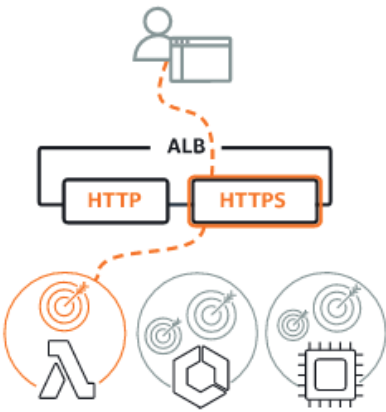## Step 3. Setup an Application Load Balancer

go to ec2 service and choose load balancer

create a load balancer

Load balancer types

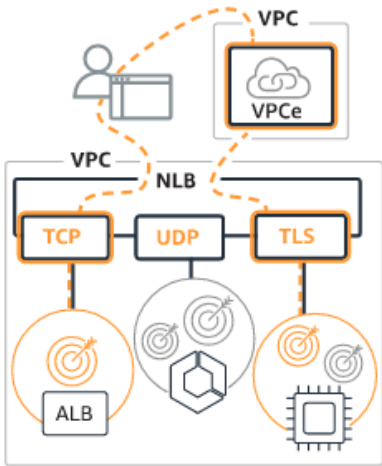**Application Load Balancer** Info

Choose an Application Load Balancer when you need a flexible feature set for your applications with HTTP and HTTPS traffic. Operating at the request level, Application Load Balancers provide advanced routing and visibility features targeted at application architectures, including microservices and containers.

Create

Network Load Balancer Info

Choose a Network Load Balancer when you need ultra-high performance, TLS offloading at scale, centralized certificate deployment, support for UDP, and static IP addresses for your applications. Operating at the connection level, Network Load Balancers are capable of handling millions of requests per second securely while maintaining ultra-low latencies.

Create

Gateway Load Balancer Info

Choose a Gateway Load Balancer when you need to deploy and manage a fleet of third-party virtual appliances that support GENEVE. These appliances enable you to improve security, compliance, and policy controls.

Create

setup security group inbound open 80 from all

Basic configuration

Load balancer name
Name must be unique within your AWS account and cannot be changed after the load balancer is created.

FhrEcsProcessingALB

A maximum of 32 alphanumeric characters including hyphens are allowed, but the name must not begin or end with a hyphen.

Scheme Info
Scheme cannot be changed after the load balancer is created.

● Internet-facing
   An internet-facing load balancer routes requests from clients over the internet to targets. Requires a public
   subnet. Learn more ⬀

○ Internal
   An internal load balancer routes requests from clients to targets using private IP addresses.

IP address type Info
Select the type of IP addresses that your subnets use.

● IPv4
   Recommended for internal load balancers.

○ Dualstack
   Includes IPv4 and IPv6 addresses.

Network mapping Info
The load balancer routes traffic to targets in the selected subnets, and in accordance with your IP address settings.

VPC Info
Select the virtual private cloud (VPC) for your targets. Only VPCs with an internet gateway are enabled for selection. The selected VPC cannot be changed after the load balancer is created. To confirm the VPC for your targets, view your target groups ⬀.

-
vpc-4dd37228
IPv4: 172.31.0.0/16   ▼   ⟳

Mappings Info
Select at least two Availability Zones and one subnet per zone. The load balancer routes traffic to targets in these Availability Zones only. Availability Zones that are not supported by the load balancer or the VPC are not available for selection. Subnets cannot be removed after the load balancer is created, but additional subnets can be added.

☑ ap-southeast-1a

Subnet

subnet-2ca8145b   ▼

IPv4 settings

forward to port 8081 of the target group ecs

setup target group with ecs fargate

Step 4. Check security group and connection