



Módulo 3: Back-end

Sesión 4: Despliegue en AWS

Equipo de desarrolladores:

Cristian David Ríos MSc
Daniel Escobar Grisales MSc
Nestor Rafael Calvo MSc

Coordinador del proyecto:

Prof. Dr.-Ing. Juan Rafael Orozco Arroyave



Hola!

Mi nombre es Daniel Escobar

You can find me at:

 @dangrisales



Sesión 3: Despliegue en AWS

Temas

- Creación del contenedor
 - Crear imagen localmente
 - Configurar ECR (Elastic Container Registry)
- Balanceador de carga
 - Configuración del balanceador
 - Grupos destino
- Creación de servicio
 - Definición de la tarea
 - Creación del contenedor con persistencia
 - Configuración del servicio
 - Escalamiento automático

A decorative network diagram in the top-left corner, featuring a complex web of interconnected nodes and lines, with some nodes highlighted in blue.

1.

Creación del contenedor

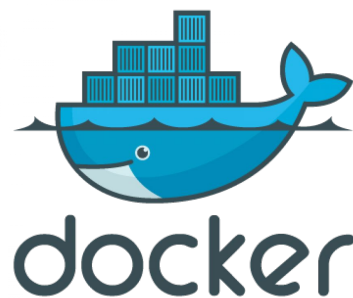
Crear imagen localmente

- ⦿ Crear .jar
 - `$/mvnw clean package`

- ⦿ Configuración del DockerFile

```
FROM openjdk:8-jdk-alpine
VOLUME /tmp
COPY *.jar app.jar
ENTRYPOINT ["java", "-jar", "app.jar"]
```

- ⦿ Creación de la imagen
 - `docker image build -t local_image .`



Configurar ECR (Elastic Container Registry)

Configuración local

Configurar credenciales de AWS

```
(base) daniel@daniel-Latitude-3400:~$ aws configure
AWS Access Key ID [*****HT6N]: ADFRGD
AWS Secret Access Key [*****r2iz]: LPJKJSKDN
Default region name [us-east-1]:
Default output format [json]:
```

Subir la imagen local al repositorio

- ⦿ ~\$aws ecr get-login-password --region us-east-1 | docker login --username AWS --password-stdin “base image URI”
- ⦿ ~\$docker tag “local_image”:latest “image URI”:latest
- ⦿ ~\$docker push “image URI”:latest

Configuración AWS

General settings

Image URI

Visibility settings Info
Choose the visibility setting for the repository.
☐ Private
Access is managed by IAM and repository policy permissions.
☐ Public
Publicly visible and accessible for image pulls.

Repository name
Provide a concise name. A developer should be able to identify the repository contents by the name.
974.dkr.ecr.us-east-1.amazonaws.com/biometria-register-api
22 out of 256 characters maximum (2 minimum). The name must start with a letter and can only contain lowercase letters, numbers, hyphens, underscores, and forward slashes.

Tag immutability Info
Enable tag immutability to prevent image tags from being overwritten by subsequent image pushes using the same tag. Disable tag immutability to allow image tags to be overwritten.
☒ Disabled

Once a repository is created, the visibility setting of the repository can't be changed.

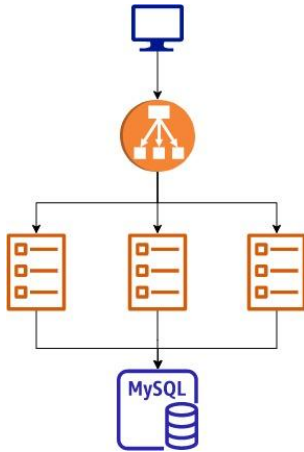
<input type="checkbox"/>	Image tag	Artifact type	Pushed at	Size (MB)	Image URI	Digest	Scan status	Vulnerabilities
<input type="checkbox"/>	latest	Image	03 de octubre de 2022, 15:24:57 (UTC+02)	111.52	Copy URI	sha256:b010f8aebd928f...	-	-
<input type="checkbox"/>	<untagged>	Image	29 de septiembre de 2022, 15:37:51 (UTC+02)	111.52	Copy URI	sha256:8cf0b2c7d6a79f...	-	-

A decorative network diagram in the top-left corner, featuring a complex web of interconnected nodes and lines. The nodes are represented by small circles, some of which are larger and have concentric circles, suggesting different levels of connectivity or importance. The lines are thin and gray, creating a mesh-like structure.

2.

Balanceador de carga

Balanceador de carga



AMIs

▼ Elastic Block Store

- Volumes
- Snapshots
- Lifecycle Manager

▼ Network & Security

- Security Groups *New*
- Elastic IPs *New*
- Placement Groups
- Key Pairs
- Network Interfaces *New*

▼ Load Balancing

- Load Balancers**
- Target Groups *New*

Create Load Balancer Actions ▼

Filter by tags and attributes or search by keyword

Name	DNS name	State	VPC ID	Availability Zones
biometria-develop-lb	biometria-develop-lb-139861...	active	vpc-43d43b3e	us-east-1f, us-east-1d
biometria-lb	biometria-lb-250897206.us-e...	active	vpc-43d43b3e	us-east-1f, us-east-1d

Select a load balancer

Select load balancer type

Elastic Load Balancing supports four types of load balancers: Application Load Balancers, Network Load Balancers, Gateway Load Balancers, and Classic Load Balancers. Choose the load balancer that best fits your needs.

[Learn more about which load balancer is right for you](#)

Application Load Balancer

HTTP
HTTPS

Create

Choose an Application Load Balancer when you need a flexible feature set for your web 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.

[Learn more >](#)

Network Load Balancer

TCP
TLS
UDP

Create

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 application. Operating at the connection level, Network Load Balancers are capable of handling millions of requests per second securely while maintaining ultra-low latencies.

[Learn more >](#)

Gateway Load Balancer

IP

Create

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.

[Learn more >](#)

Configuración del balanceador de carga

Step 1: Configure Load Balancer

Basic Configuration

To configure your load balancer, provide a name, select a scheme, specify one or more listeners, and select a network. The default listener receives HTTP traffic on port 80.

Name ⓘ biometria-lb

Scheme ⓘ ☒ internet-facing
☐ internal

IP address type ⓘ ipv4

Listeners

A listener is a process that checks for connection requests, using the protocol and port that you configured.

Load Balancer Protocol	Load Balancer Port
<input type="text" value="HTTP"/>	<input type="text" value="80"/>
<input checked="" type="text" value="HTTPS (Secure HTTP)"/>	<input type="text" value="443"/>

[Add listener](#)

Availability Zones

Specify the Availability Zones to enable for your load balancer. The load balancer routes traffic to the targets in these Availability Zones or subnets from at least two Availability Zones to increase the availability of your load balancer.

VPC ⓘ vpc-43d43b3e (172.31.0.0/16) (default)

Availability Zones ☒ us-east-1d subnet-a370acfc
☒ us-east-1f subnet-3535a03b

IPv4 address ⓘ Assigned by AWS

IPv4 address ⓘ Assigned by AWS

Step 2: Configure Security Settings

Select default certificate

AWS Certificate Manager (ACM) is the preferred tool to provision and store server certificates. If you previously stored a server listeners and certificate management.

Certificate type ⓘ ☒ Choose a certificate from ACM (recommended)
☐ Upload a certificate to ACM (recommended)
☐ Choose a certificate from IAM
☐ Upload a certificate to IAM

[Request a new certificate from ACM](#)
AWS Certificate Manager makes it easy to provision, manage, deploy, and renew SSL Certificates on the AWS

Certificate name ⓘ *.udea.edu.co (arn:aws:acm:us-east-1:346379607974:certificate:b71e93)

Select Security Policy

Security policy ⓘ ELBSecurityPolicy-2016-08

Step 3: Configure Security Groups

A security group is a set of firewall rules that control the traffic to your load balancer. On this page, you can add rules to allow specific traffic to reach your load balancer, group or select an existing one.

Assign a security group ☐ Create a new security group
☒ Select an existing security group

Security Group ID	Name	Description
<input checked="" type="checkbox"/> sg-bf20bf99	default	default VPC security group
<input type="checkbox"/> sg-00e603c7a33312b99	ec2-to-rds-efs	launch-wizard-3 created 2020-08-12T16:50:52.799-05:00
<input type="checkbox"/> sg-0355969d7c0902b79	EFS-access-sg-process	allow to connect to EFS using fargate
<input type="checkbox"/> sg-0ba7e7024bd3lef4e	efs-sg-1	Created by the LIW for EFS at 2020-11-21T14:13:37.391-05:00
<input type="checkbox"/> sg-0b9575428220db043	efs-sg-2	Created by the LIW for EFS at 2020-11-21T16:46:07.833-05:00

Configuración del balanceador de carga

Step 4: Configure Routing

Your load balancer routes requests to the targets in this target group using the protocol and port that you specify, and performs health checks on the targets using specify in this step will apply to all of the listeners configured on this load balancer; you can edit the listeners and add listeners after the load balancer is created.

Target group

Target group ⓘ

Name ⓘ

Target type

- ☐ Instance
- ☒ IP
- ☐ Lambda function

Protocol ⓘ

Port ⓘ

Protocol version ⓘ

- ☒ HTTP1
Send requests to targets using HTTP/1.1. Supported when the request protocol is HTTP/1.1 or HTTP/2.
- ☐ HTTP2

Health checks

Protocol ⓘ

Path ⓘ

▼ Advanced health check settings

Port ⓘ

- ☒ traffic port
- ☐ override

Healthy threshold ⓘ

Unhealthy threshold ⓘ

Timeout ⓘ seconds

Interval ⓘ seconds

Success codes ⓘ

Step 5: Register Targets

Register targets with your target group. If you register a target in an enabled Availability Zone, the load balancer starts routing requests to the targets as soon as the initial health checks.

prueba (target group)

Specify one or more IP addresses to register as targets

Network ⓘ	IP (Allowed ranges)	Port ⓘ	
vpc-43d43b3e (172.31.0.0/16)	<input type="text"/>	<input type="text" value="80"/>	<input type="button" value="Add to list"/>

0 total IP addresses.

Grupos destino

biometria-lb-target Delete

arn:aws:elasticloadbalancing:us-east-1: :targetgroup/biometria-lb-target/7f145c25847265d2

Basic configuration

Target type IP	Protocol : Port HTTP: 8080 Protocol version HTTP1	VPC vpc-43d43b3e	Load balancer biometria-lb
-------------------	----------------------------------------------------------------	-------------------------------------	-----------------------------------------------

Group details Targets Monitoring Tags

Health check settings Edit

Protocol HTTP	Path /
Port traffic-port	Healthy threshold 5 consecutive health check successes
Unhealthy threshold 6 consecutive health check failures	Timeout 120 seconds
Interval 300 seconds	Success codes 200

A decorative network diagram in the top-left corner, featuring a complex web of interconnected nodes and lines, with some nodes highlighted in blue.

3.

Creación del servicio Elastic Container Service (ECS)

Definición de la tarea

Task Definition Name* register-api-modify ⓘ

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. <default> is the only supported mode on Windows.

Requires compatibilities ☐ EC2
☒ FARGATE

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 launch type. Container level memory settings are optional when task size is set. Task size is not supported for Windows containers.

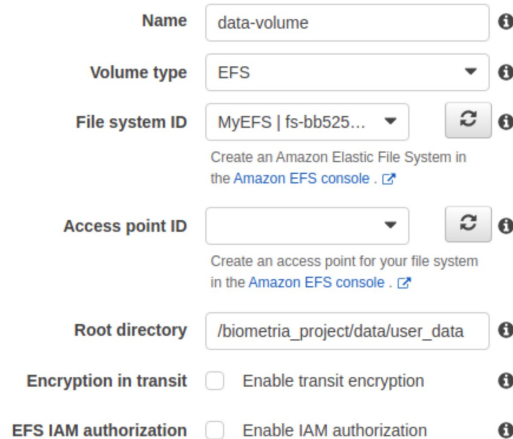
Task memory (GB) 2GB ↕

The amount of memory (in MiB) used by the task. It can be expressed as an integer using MiB, for example 1024, or as a string using GB, for example '1GB' or '1 gb'.

Task CPU (vCPU) 1 vCPU ↕

The number of CPU units used by the task. It can be expressed as an integer using CPU units, for example 1024, or as a string using vCPUs, for example '1 vCPU' or '1 vcpu'.

Creación del volumen con EFS



The screenshot displays the configuration interface for creating an Amazon Elastic File System (EFS) volume. The form includes the following fields and options:

- Name:** A text input field containing "data-volume".
- Volume type:** A dropdown menu set to "EFS".
- File system ID:** A dropdown menu showing "MyEFS | fs-bb525...". To the right is a refresh icon and an information icon. Below this field is a link: "Create an Amazon Elastic File System in the [Amazon EFS console](#) . 🔗".
- Access point ID:** An empty dropdown menu. To the right is a refresh icon and an information icon. Below this field is a link: "Create an access point for your file system in the [Amazon EFS console](#) . 🔗".
- Root directory:** A text input field containing "/biometria_project/data/user_data".
- Encryption in transit:** A checkbox labeled "Enable transit encryption", which is currently unchecked.
- EFS IAM authorization:** A checkbox labeled "Enable IAM authorization", which is currently unchecked.

Agregar contenedor

▼ Standard

Container name* register

Image*

amazon-ecr.us-east-1.amazonaws.com/biometria-register-api:latest

Image URI

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

Protocol

8080

tcp

+ Add port mapping

Host port mappings are not valid when the network mode for a task definition is host or awsvpc. To use host port mappings, choose the Bridge network mode.

► Advanced container configuration

STORAGE AND LOGGING

Read only root file system ☐

Mount points

data-volume

/user_data

Read only ☐

+ Add mount point

Volumes from

Source container

Read only

+ Add volumes

Log configuration

☐ Auto-configure CloudWatch Logs

Log driver

awslogs

Log options

Key

awslogs-group

Value

/ecs/register-api

awslogs-region

Value

us-east-1

awslogs-stream

Value

ecs

Add key

Value

Add value

Persistencia

Definición de la tarea

Container Definitions

Add container

Container Name...	Image	Hard/Soft memo...	CPU Unit...	GPU	Inference Accel...	Essential ...	
register	lkr.ecr.us-e...	--/--	0			true	✖

Volumes

Use a volume configuration to add volumes for use by the containers within a task. To add a volume, choose **Add volume**, complete the fields, and then choose **Add**. [Learn more](#)

Name

data-volume

Volume type

EFS

File system ID

fs-bb525138

Root directory

/biometria_project/data/user_data

Encryption in transit

DISABLED

EFS IAM authorization

DISABLED

Creación del servicio

Task Definition Family

Revision

Launch type ⓘ

[Switch to capacity provider strategy](#)

Platform version ⓘ

Force new deployment ☐ ⓘ

Cluster ⓘ

Service name ⓘ

Service type* ⓘ

Number of tasks ⓘ

Minimum healthy percent ⓘ

Maximum percent ⓘ

Deployment circuit breaker ⓘ

Configure network

Allowed VPC

Allowed subnets

Security groups*

Auto-assign public IP

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 ⓘ

Load balancing

Load balancing settings can only be set on service creation.

Load Balancer Name

Container Name:

Container Port:

Target Group:

*Required

[Cancel](#)

[Previous](#)

[Next step](#)

Creación del servicio

Configure network

Allowed VPC `vpc-43d43b3e`

Allowed subnets `subnet-3535a03b`

Security groups* `sg-09fe4e7c1a709dd0d`

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



Load balancing

Load balancing settings can only be set on service creation.

Load Balancer Name `biometria-lb`

Container Name: `register`

Container Port: `8080`

Target Group: `arn:aws:elasticloadbalancing:us-east-1:346379607974:targetgroup/biometria-lb-target/7f1145c25847265d2`

*Required

Cancel

Previous

Next step

Auto escalamiento

Service 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

Automatic task scaling policies

Add scaling policy

Policy name*

LowCPUPolicy

HighCPUPolicy70

LowMemoryPolicy

HighMemoryPolicy80

Políticas de auto escalamiento

- Scaling policy type**
- ☐ Target tracking
 - ☒ Step scaling

Policy name*

- Execute policy when**
- ☐ Create new Alarm
 - ☒ Use an existing Alarm

breaches the alarm threshold: CPUUtilization >= 70 for 1 out of 1 consecutive periods of 60 seconds for the metric dimensions ClusterName=processes-cluster, ServiceName=api-service

Scaling action when <= CPUUtilization <

+ Add

Cooldown period seconds between scaling actions

Revisión de los grupos destino

biometria-lb-target Delete

arn:aws:elasticloadbalancing:us-east-1:roup/biometria-lb-target/7f145c25847265d2

Basic configuration

Target type IP	Protocol : Port HTTP: 8080 Protocol version HTTP1	VPC vpc-43d43b3e ↗	Load balancer biometria-lb ↗
-------------------	----------------------------------------------------------------	---------------------------------------	-------------------------------------------------

Group details **Targets** Monitoring Tags

Registered targets (4) ↻ Deregister Register targets

<input type="checkbox"/>	IP address ▾	Port ▾	Zone ▾	Status ▾	Status details
<input type="checkbox"/>	172.31.78.69	8080	us-east-1f	✓ healthy	
<input type="checkbox"/>	172.31.70.54	8080	us-east-1f	✓ healthy	
<input type="checkbox"/>	172.31.71.150	8080	us-east-1f	✓ healthy	
<input type="checkbox"/>	172.31.66.200	8080	us-east-1f	✓ healthy	

A decorative network diagram in the top-left corner, featuring a complex web of interconnected nodes and lines, with some nodes highlighted in blue.

4.

Manos a la obra

Desplegar el back-end en AWS

Presentación del equipo

Cristian Rios

Daniel Escobar

Nestor Calvo



Gracias!

¿Alguna pregunta o comentario?

Puedes escribirme a:

daniel.esobar@udea.edu.co