Optimización de la infraestructura



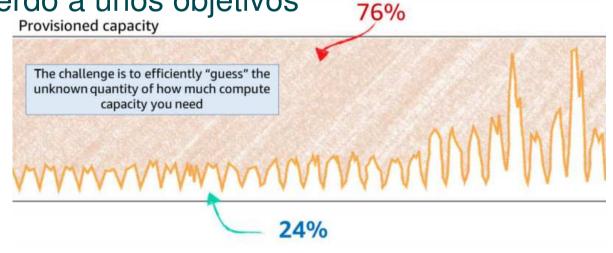


- Monitorización \*
- Elasticidad
- Escalabilidad
- Tolerancia a fallos
- Mensajería
- ServerLess/API Gateway

#### Mondragon Unibertsitatea Goi Eskola Politeknikoa

#### Conceptos básicos

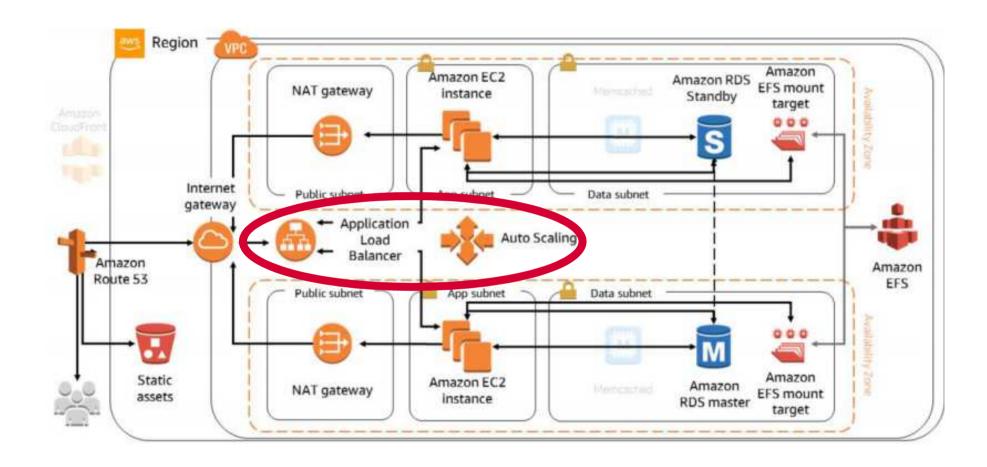
- Escalabilidad vertical
  - Aumentar capacidad de computación
- Escalabilidad horizontal
  - Aumentar números de recursos
- Elasticidad mediante escalabilidad
- Elasticidad es la habilidad del sistema para monitorizar la demanda y automáticamente incrementar/decrementar los recursos desplegados de acuerdo a unos objetivos





Politeknikoa

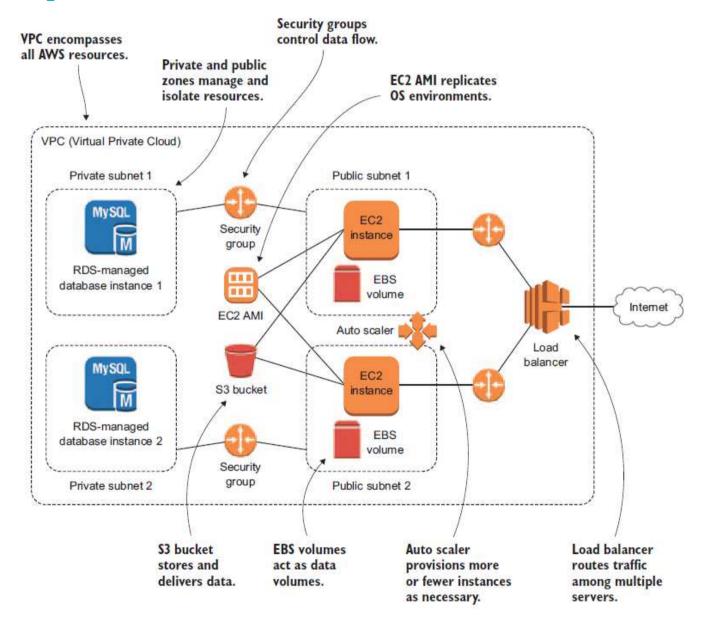
## **Conceptos básicos**





Politeknikoa

#### Conceptos básicos



3.1

## Alta disponibilidad

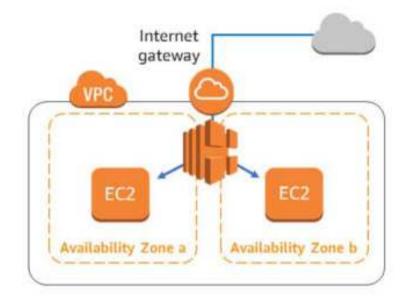


#### 3.1 Zonas de disponibilidad

Goi Eskola Politeknikoa

Start with two Availability Zones per AWS Region.

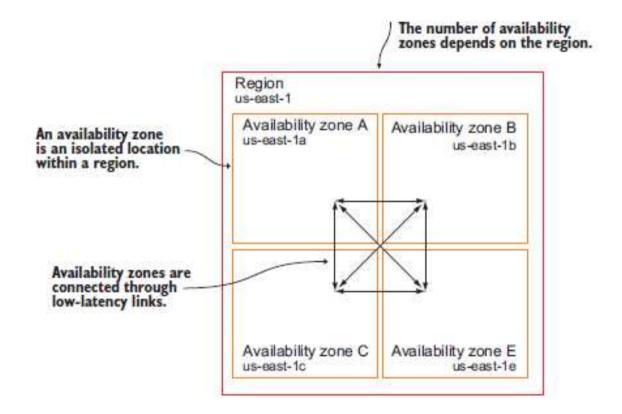
If resources in one Availability Zone are unreachable, your application shouldn't fail.





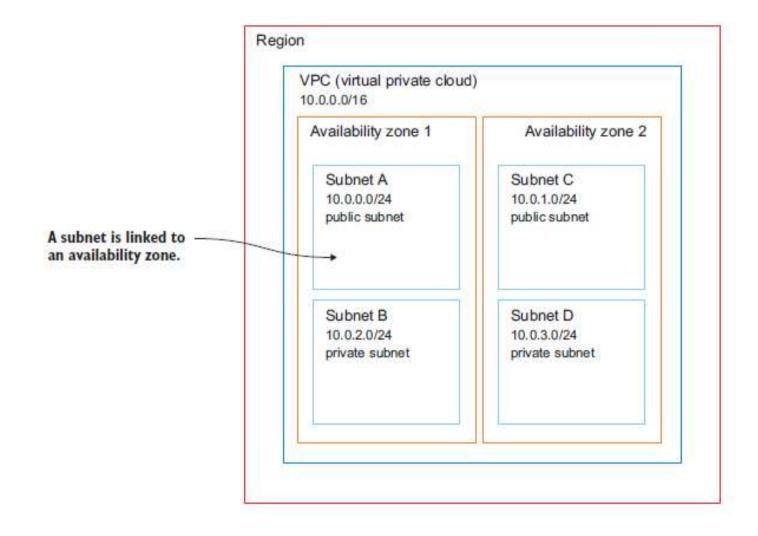


Politeknikoa





#### 3.1 Zonas de disponibilidad

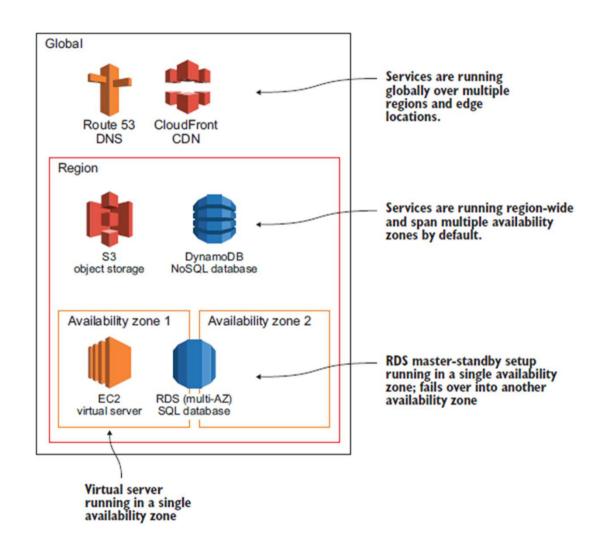


Goi Eskola Politeknikoa





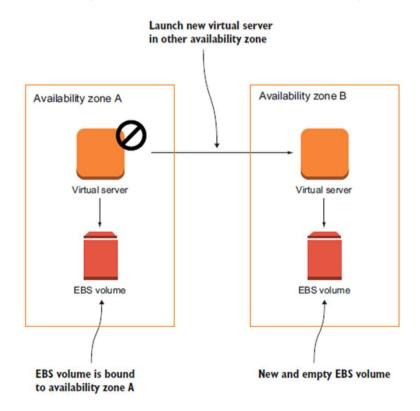
Goi Eskola Politeknikoa





#### 3.1 Zonas de disponibilidad

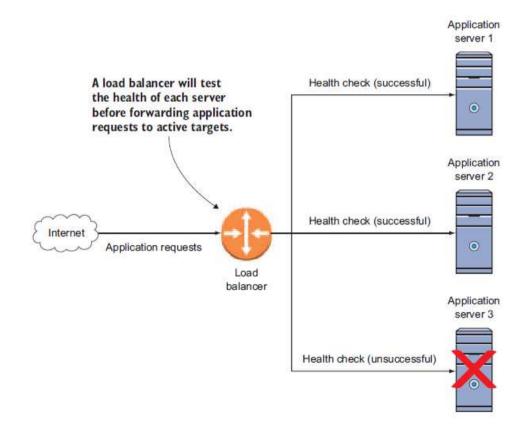
- Para reutilizar una EBS habrá que utilizar servicios que permiten ser gestionados en una región: s3, dynamoDB, RDS
- Por ejemplo snapshots de EBS se guardan en S3





Goi Eskola Politeknikoa

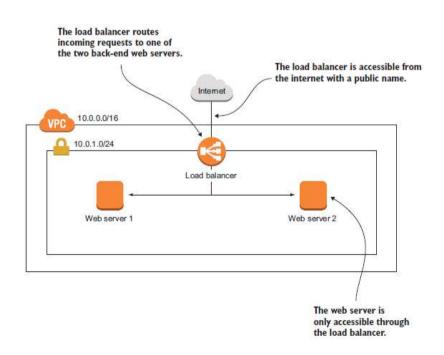
- ELB
  - Desacople síncrono mediante balanceo de carga







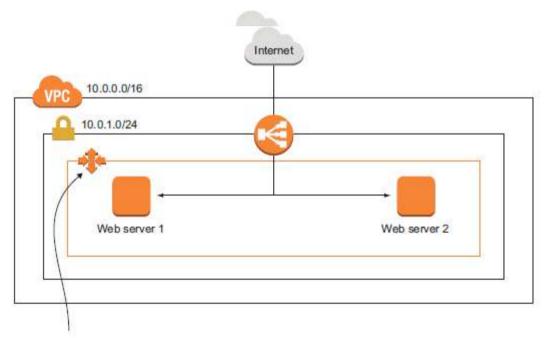
- ELB
  - El servicio de balanceo de carga de AWS es escalable y tolerante a fallos
  - Se cobra por hora de uso y trafico
    - 0.025\$/h
    - 0.0.008\$/GB
  - Tipos de balanceo
    - Trafico
    - SSL endpoint
    - Logs
    - Aplicación





Politeknikoa

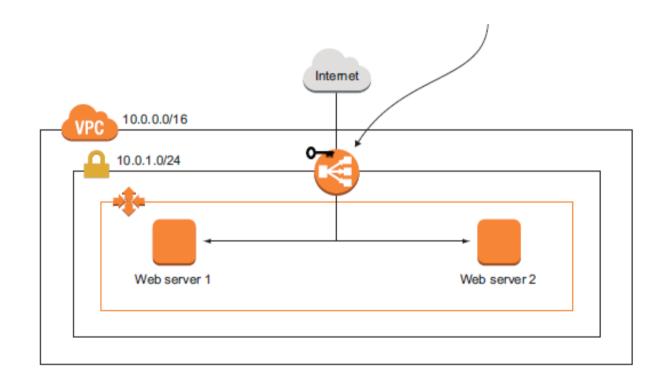
- ELB : Grupos de autoescalado
  - Tiempo de registro
  - Health-check



- The auto-scaling group observes
- i two web servers. If a new server
- is started, the auto-scaling group registers it with the ELB.

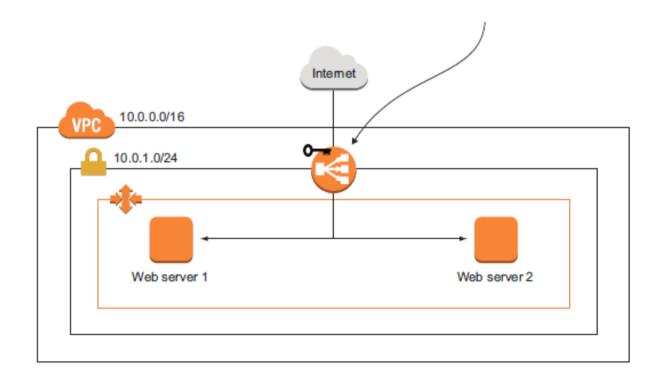


• ELB: Finalización SSL



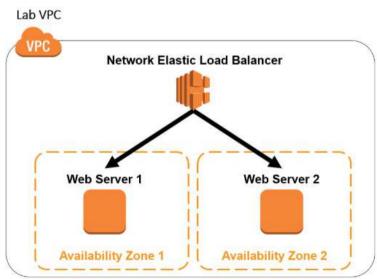


- ELB: Tolerancia a fallos
  - Balanceo entre zonas de disponibilidad
  - Deshabilitado por defecto





- Ejercicio: QWIKSLAB: Introducing to Elastic Load Balancing
  - 1. Crear un Balanceador de carga de Red
    - Agregarle el "Grupo Destino" (Target Group, instancias de redirección )
      - Las dos instancias EC2 que actúan como Servidores WEB
  - 2. Comprobar en el Target group creado la disponibilidad de las maquinas
  - 3. Validar el Balanceador de carga





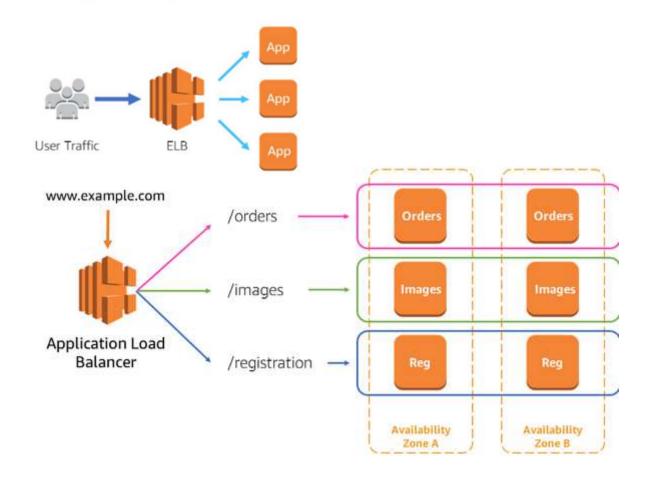
Politeknikoa

## 3.2 Balanceo de carga

#### Balanceo de aplicaciones



A managed load balancing service that distributes incoming application traffic across multiple Amazon EC instances, containers, and IP addresses.

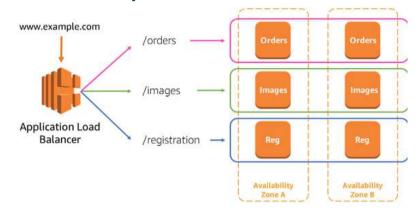




Politeknikoa

#### 3.2 Balanceo de carga

#### Balanceo de aplicaciones



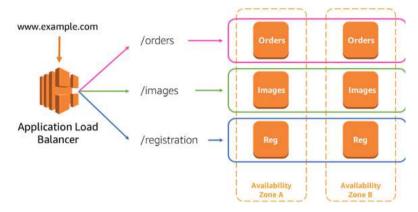
#### ELB | HTTP:80 (3 rules)

Rule limits for condition values, wildcards, and total rules.

1 am15b8c ▼	IF  ✓ Path is /bi	THEN Forward to Group2
2 arn83357 ▼	IF  ✓ Path is /orders	THEN Forward to MyTarget
last HTTP 80: default action  This rule cannot be moved or deleted	<ul><li>IF</li><li>✓ Requests otherwise not routed</li></ul>	THEN Forward to MyTarget



Balanceo de aplicaciones: Seguridad con Filtrado



Servicio WAF

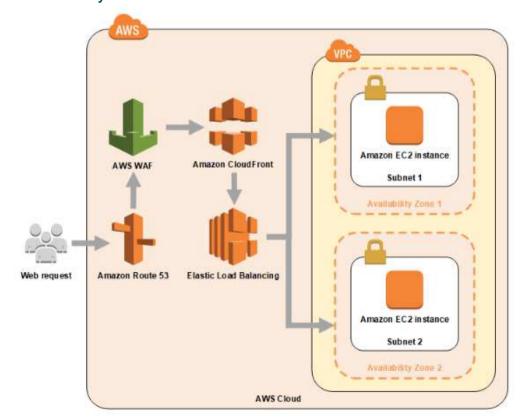


• Ej: DDoS

https://docs.aws.amazon.com/es\_es/waf/latest/developerguide/tutorials-ddos-cross-service.html

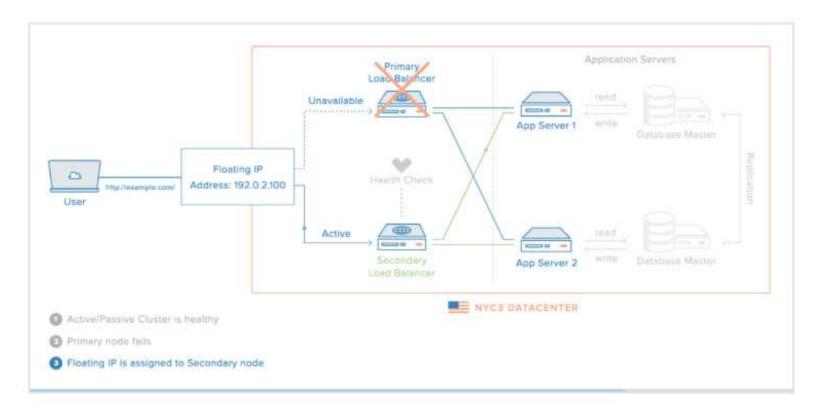


- Balanceo de aplicaciones: Seguridad con Filtrado
- Paso 1: lanzamiento de un servidor virtual mediante Amazon EC2
- Paso 2: escalado del tráfico con Elastic Load Balancing
- Paso 3: mejora del desempeño y absorción de ataques mediante Amazon CloudFront
- Paso 4: registre su nombre de dominio e implemente el servicio DNS mediante Route 53
- Paso 5: detección y filtrado de solicitudes web malintencionadas mediante AWS WAF



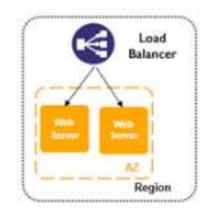


- Patrón para balanceo de carga
  - Cluster : keepalive e IP flotante (Tolerancia a fallos)
  - Cluster: Politica de DNS (Escalabilidad)





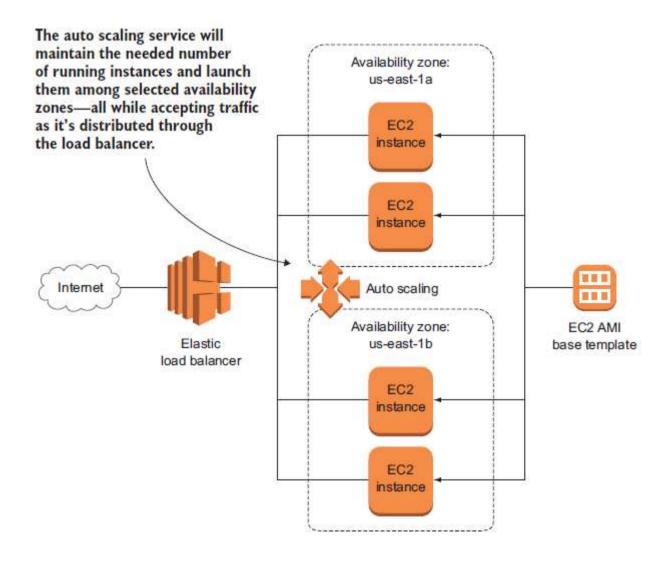
- Ejercicio: QWIKSLAB: Working with Elastic Load Balancing
  - Realizar la práctica en el classroom siguiendo los pasos de QWIKSLAB (El entorno de QWIKSLAB da error de permisos en esta práctica)
  - 1. Crear un Balanceador de carga de Applicación
    - Agregarle el "Grupo Destino" (Target Group, instancias de redirección )
      - Las dos instancias EC2 que actúan como Servidores WEB
  - 2. Comprobar en el Target group creado la disponibilidad de las maquinas
  - 3. Validar el Balanceador de carga
  - 4. Analizar opciones de reglas y de seguridad
  - 5. Comprobar métricas en CloudWatch





Goi Eskola Politeknikoa

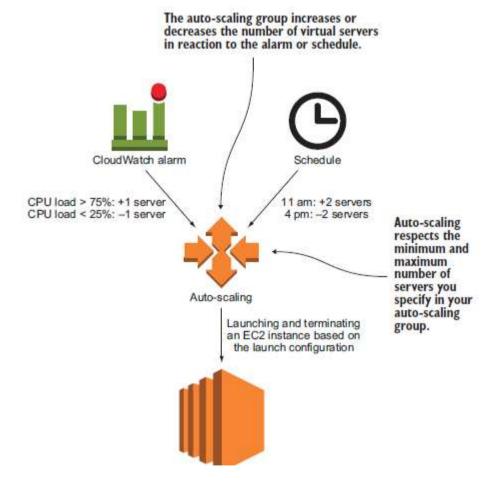
#### 3.3 Auto Escalado



#### 3.3 Auto Escalado



Goi Eskola Politeknikoa





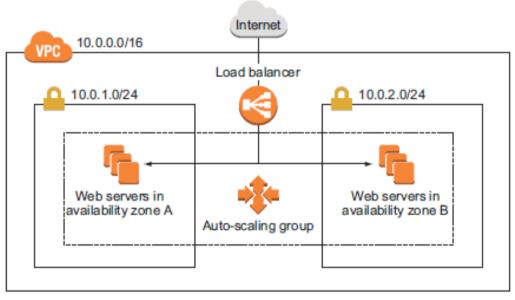
#### 3.2 AutoEscalado

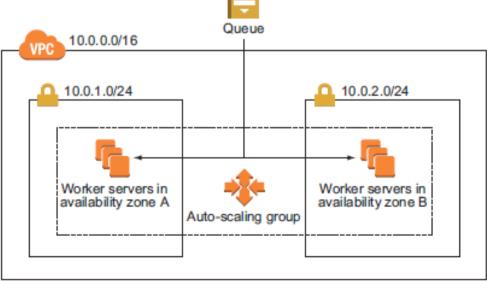
- Ejercicio: QWIKSLAB: Introduction to Amazon EC2 Auto Scaling
  - 1. Crear configuración de lanzamiento de instancias (Launch Configuration)
    - AMI Amazon Linux 2
    - Condición siempre una instancia
  - 2. Crear Grupo de autoescalado con la Configuración de lanzamiento anterior



Goi Eskola Politeknikoa

#### 3.4 Tolerancia a fallos

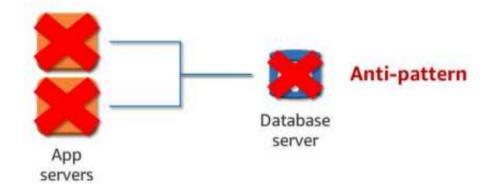


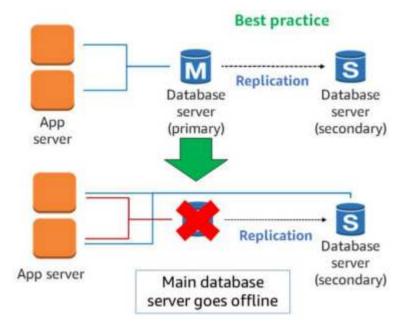






Goi Eskola Politeknikoa

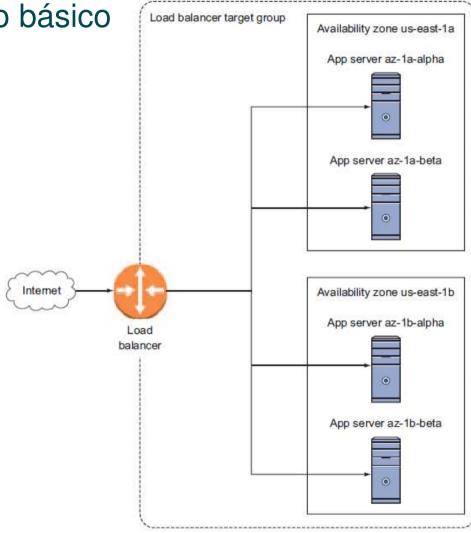






## 3.5 Alta disponibilidad

Práctica: ELB: Ejercicio básico

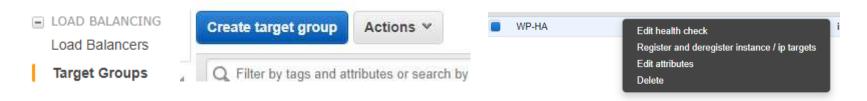




- Práctica : ELB
  - Crear instancia Wordpress con CDN S3 y base de datos RDS
  - Punto de partida:
    - Instancia EC2 con Wordpress
      - Base de datos MYSQL en RDS
      - Plugin CDN y S3
    - VPC con subredes
  - Utilizar el classroom Serverless o el Starter Account (El autescalado como mucho es para 3)
  - Nota: Para facilitar la práctica podéis realizarla solamente con un simple Servidor WEB y una simple página PHP

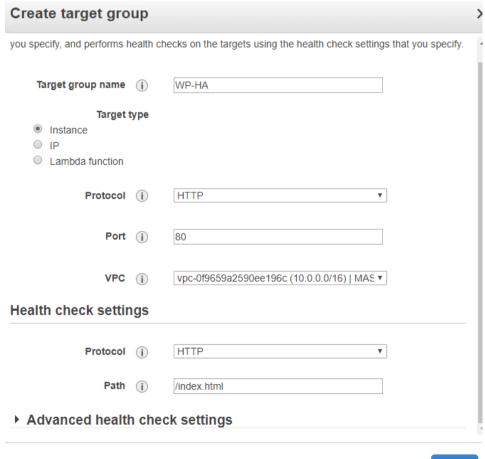


- Práctica : ELB
  - Primer caso crear subredes en dos zonas de disponibilidad
    - 2 publicas
  - Crear 4 instancias de wordpress que acceden a la misma base de datos
    - Inicialmente serán publicas
  - En EC2 Crear un "Grupo Objetivo" hacia instancias y para HTTP
  - Registrar las instancias en el grupo objetivo
  - Crear el Balanceador de carga
    - Asociarle un grupo de seguridad
    - Asociarle el grupo objetivo
  - Verificar/Testear la arquitectura





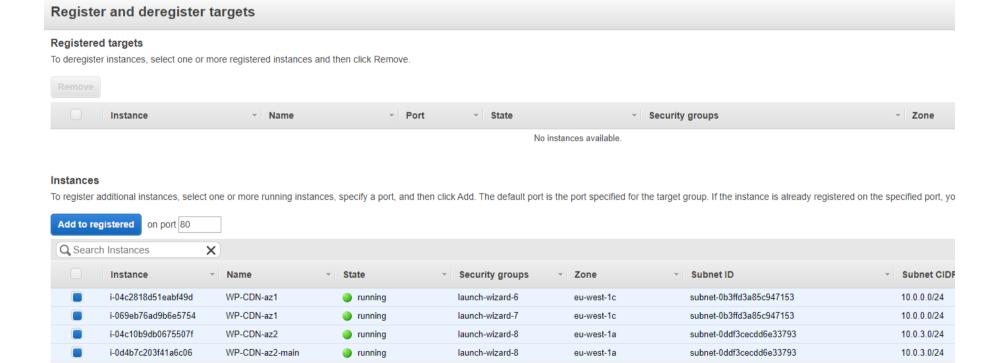
- Práctica : ELB
  - En EC2 Crear un "Grupo Objetivo" hacia instancias y para HTTP







- Práctica : ELB
- Registrar las instancias en el grupo de seguridad





- Práctica : ELB
  - Crear balanceador de carga

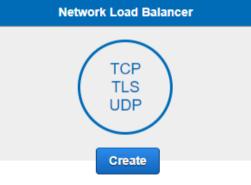
#### Select load balancer type

Elastic Load Balancing supports three types of load balancers: Application Load Balancers, Network Load Balancers (new), and Classic Load Balancers. Choose the load balancer type that meets 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

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 >



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 >

## PREVIOUS GENERATION for HTTP, HTTPS, and TCP Create

Choose a Classic Load Balancer when you have an existing application running in the EC2-Classic network.

Learn more >



Unibertsitatea

Goi Eskola Politeknikoa

- Práctica: ELB
  - Crear balanceador de carga

Step 1: Configure Load Balancer tacing load balancer in the selected network with a listener that receives HTTP traffic on port 80.



#### 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
[HTTP ▼]	80
HTTPS (Secure HTTP) ▼	443
Add listener	

#### **Availability Zones**

Specify the Availability Zones to enable for your load balancer. The load balancer routes traffic to the targets in the one subnet per Availability Zone. You must specify subnets from at least two Availability Zones to increase the ava





- Práctica : ELB
  - Crear balanceador de carga : Configuración de seguridad
    - Para HTTPS AWS puede gestionar los certificados mediante el servicio ACM
      - Crear certificado (Requiere tener un dominio)
    - Elegir un grupo de seguridad ya creado o crear uno nuevo
      - Solo dar acceso HTTP

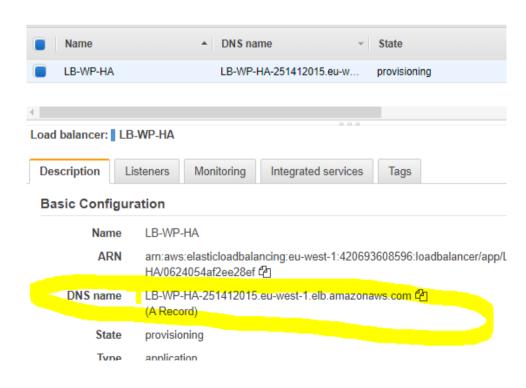


- Práctica : ELB
  - Configurar el enrutamiento al "Grupo Objetivo" creado anteriormente





- Práctica : ELB
  - Probar configuración
    - DNS público del balanceador de carga
    - Crear la página index.php para que de la IP del servidor



```
<!DOCTYPE html>
<html>
<body>
<h1>Ongi etorri orria</h1>
</php

echo "Ongi etorri orria. Load balancer froga";
$localIP = getHostByName(getHostName());
echo "Zerbiztararen IP:$localIP";

?>
</body>
</html>
```

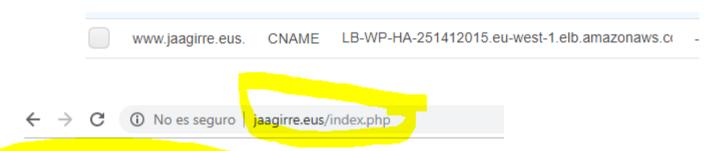


Práctica : ELBProbar configuración





- Práctica : ELB
- Si se tiene un dominio crear un CNAME del dominio que apunte al LOAD\_BALANCER



### Ongi etorri orria

Ongi etorri orria. Load balancer frogaZerbiztararen IP:10.0.0.95



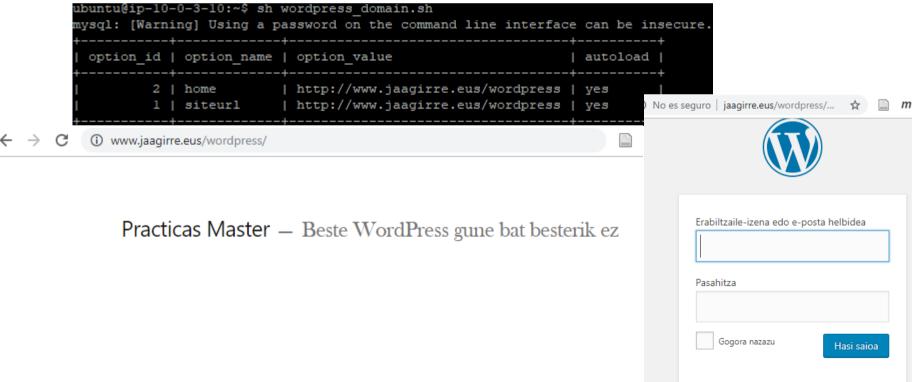
### ngi etorri orria

gi etorri orria. Load balancer frogaZerbiztararen IP:10.0.3.10



- Práctica : ELB
  - Configurar la base de datos de Wordpress con el dominio

URL=www.jaagirre.eus mysql -uwpuser -ppassword -hdatabase.endpoint.amazonaws.com --execute "SET @url='\${URL}/wordpress'; source wordpress\_domain.sql;";





Goi Eskola Politeknikoa

- Práctica : ELB
  - Pruebas de rendimiento y tolerancia a fallos
    - JMETER comparar poniendo 1,2,3 y 4 servidores
    - Entregar documento con resultados de rendimiento



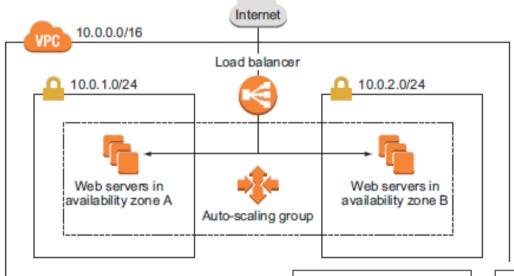
- Práctica : ELB & Autoescalado
  - Porque pagar 4 instancias constantemente
  - Mejor adaptarse a las necesidades

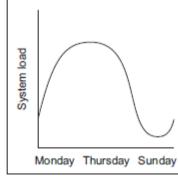
System load

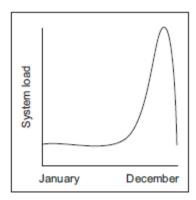
6am

12am

6pm









- Práctica : ELB & Autoescalado
- Crear configuración de "Launch"
   rear configuración de lanzamiento

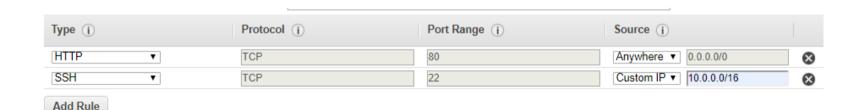


Nombre	<b>①</b>	WP-HA
Opción de compra	1	Request Spot Instances
Rol de IAM	<b>①</b>	Ninguno ▼
Monitorización	1	☑ Habilitar monitorización detallada de CloudWatch  Más información
Detalles avanzados		
ID de kernel	<b>(F)</b>	Usar valor predeterminado ▼
ID de disco de RAM	1	Usar valor predeterminado ▼
Datos de usuario	1	
		#cloud-boothook  #!/bin/bash  exec > >(tee /var/log/user-data.log logger -t user-data ) 2>&1  sysctl kernel.hostname=MACC_PRACTICA_1  echo 'export URL=\$(curl http://169.254.169.254/latest/meta-
Tipo de dirección IP	(1)	<ul> <li>Asigne una dirección IP pública solamente a las instancias lanzadas en la subred y la VPC predeterminadas. (predeterminado)</li> </ul>
		<ul> <li>Asigne una dirección IP pública a cada instancia.</li> <li>No asigne una dirección IP pública a ninguna de las instancias.</li> </ul>



- Práctica : ELB & Autoescalado
  - Crear configuración de "Launch"
  - Tipo de maquina
    - AMI
    - Dirección privada
    - Grupo de seguridad
      - Dar acceso vía BASTION





Polite

- Práctica: ELB & Autoescalado
  - Crear "Grupo de autoescalado"

 AUTO SCALING
 Configuraciones de lanzamiento

Grupos de Auto Scaling

### Crear grupo de Auto Scaling

Nota: Para crear grupos de Auto Scaling en una región diferente, seleccione su región en la barra de navegación.

### Beneficios de Auto Scaling

### Aprovisionamiento automatizado



Mantenga su grupo de Auto Scaling en equilibrio y en buen estado, ya sea que necesite una instancia o 1000.

Más información

### Capacidad ajustable



Mantenga un tamaño de grupo fijo o ajústelo dinámicamente en función de las métricas de Amazon CloudWatch

Más información

#### Compatibilidad con plantillas de lanzamiento



Aprovisione instancias de manera fácil mediante el uso de plantillas de lanzamiento EC2.

Más información



Goi Eskola Politeknikoa



- Práctica : ELB & Autoescalado
  - Crear "Grupo de autoescalado"

Configuraciones de lanzamiento Grupos de Auto Scaling

AUTO SCALING

Seleccionar la "Configuracion de lanzamiento" previa

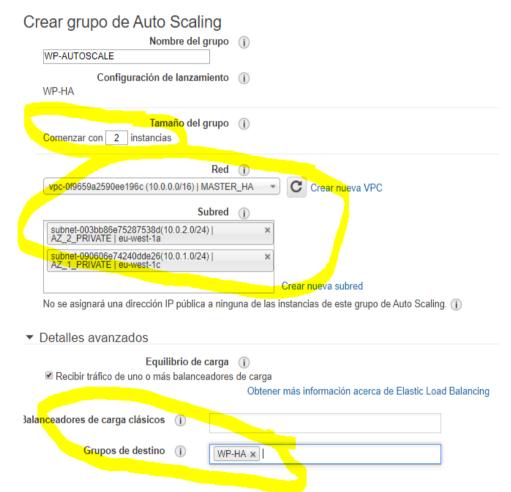




Goi Eskola Politeknikoa

- Práctica : ELB & Autoescalado
  - Crear "Grupo de autoescalado"

Configurar con balanceo de carga agregando las instancias del grupo de autoescalado al grupo objetivo de balanceador





Politeknikoa

3.6 PRACTICA Alta disponibilidad

- Práctica : ELB & Autoescalado
  - Crear "Grupo de autoescalado"

 AUTO SCALING
 Configuraciones de lanzamiento
 Grupos de Auto

Scaling

Política de escalado ate Alarm Utilice políticas de escalado para ajustar la capacidad de este grupo Escale entre 2 y 4 instancias. Estos serán el tamaño mínimo y máximo de su grupo. an use CloudWatch alarms to be notified automatically whenever metric data reaches a level you defir an alarm, first choose whom to notify and then define when the notification should be sent. Send a notification to: AUTOESCALADO-CPU-WP cancel CPU Ut Aumentar tamaño de grupo With these recipients: jaagibas@gmail.com Nombre: WP-AUTOESCALADO-AUMEI ▼ of CPU Utilization Whenever: Average Ejecutar la política awsec2-WP-AUTOSCALE-CPU-Utilization Editar Eliminar cuando: supera el umbral de alarma: CPUUtilization >= 70 durante 300 segundos ls: >= ▼ 70 Percent para las dimensiones de métricas AutoScalingGroupName = WP-AUTOSCALE For at least: 1 consecutive period(s) of 5 Minutes ▼ <= CPUUtilization < + Realizar la acción: Añadir ▼ | 1 Instancias ▼ cuando 70 Añadir paso (i) Name of alarm: awsec2-WP-AUTOSCALE-CPU-Utilization W/F Las instancias segundos para prepararse después de cada paso necesitan: Crear una política de escalado sencillo (1) reate Alarm Reducir tamaño de grupo u can use CloudWatch alarms to be notified automatically whenever metric data reaches a l∉ Nombre: WP-AUTOESCALADO-REDUC edit an alarm, first choose whom to notify and then define when the notification should be sent. awsec2-WP-AUTOSCALE-High-CPU-Utilization Editar Eliminar Ejecutar la política ■ Send a notification to: AUTOESCALADO-CPU-WP (jaagibas@c ▼ create topic) cuando: supera el umbral de alarma: CPUUtilization <= 70 durante 300 segundos para las dimensiones de métricas AutoScalingGroupName = WP-AUTOSCALE ▼ of CPU Utilization Whenever: Average Realizar la acción: Fliminar Instancias ▼ cuando 70 >= CPUUtilization > ls: <= ▼ 70 Percent Añadir paso (i) For at least: 1 consecutive period(s) of 5 Minutes ▼ Crear una política de escalado sencillo (1) Name of alarm: awsec2-WP-AUTOSCALE-High-CPU-Utilization Previous Análisis Cancel



Goi Eskola Politeknikoa

AUTO SCALING

lanzamiento

Scaling

Configuraciones de

Grupos de Auto

- Práctica: ELB & Autoescalado
  - Crear "Grupo de autoescalado"
    - Notificaiones

notificación

### Crear grupo de Auto Scaling

Configure el grupo de Auto Scaling de modo que envíe notificaciones a un punto de enlace especifica ocurran determinados eventos. Entre estos eventos, se incluyen el lanzamiento correcto, el lanzamien las instancias

Si ha creado un tema nuevo, revise su buzón de correo electrónico para encontrar un mensaje de con suscripción. Las notificaciones solo se pueden enviar a direcciones confirmadas.

Enviar una notificación	Autoescalado	utilizar tem
a:		
Con estos	jaagibas@gmail.com	
destinatarios:		
Cuando las instancias:	✓ lanzar	
	✓ finalizar	
	✓ no se puede lanzar	
	✓ no se puede terminar	
Añadir		



Unibertsitatea

Goi Eskola Politeknikoa

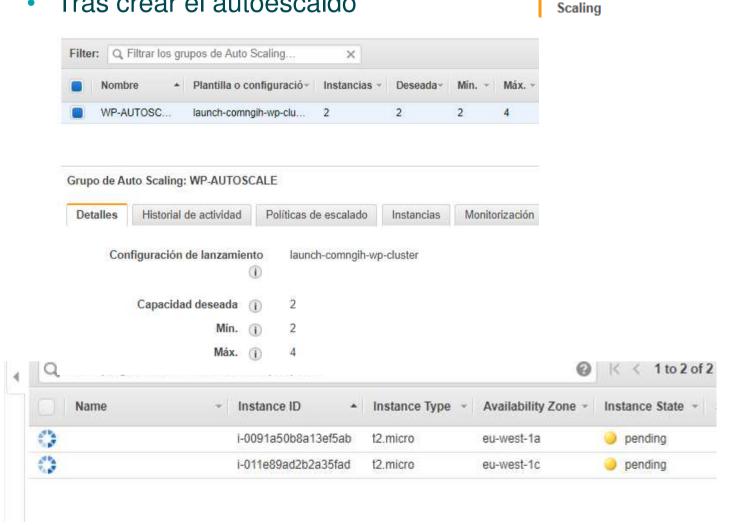
AUTO SCALING

lanzamiento

Configuraciones de

Grupos de Auto

- Práctica: ELB & Autoescalado
  - Tras crear el autoescaldo





- Mondragon Unibertsitatea Goi Eskola
- Goi Eskola Politeknikoa

- Práctica : ELB: Autoescalado
  - Tras crear el autoescalado



### Ongi etorri orria

Ongi etorri orria. Load balancer frogaZerbiztararen IP:10.0.2.136



AUTO SCALING

lanzamiento

Scaling

Configuraciones de

Grupos de Auto

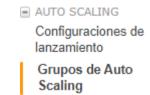
### )ngi etorri orria

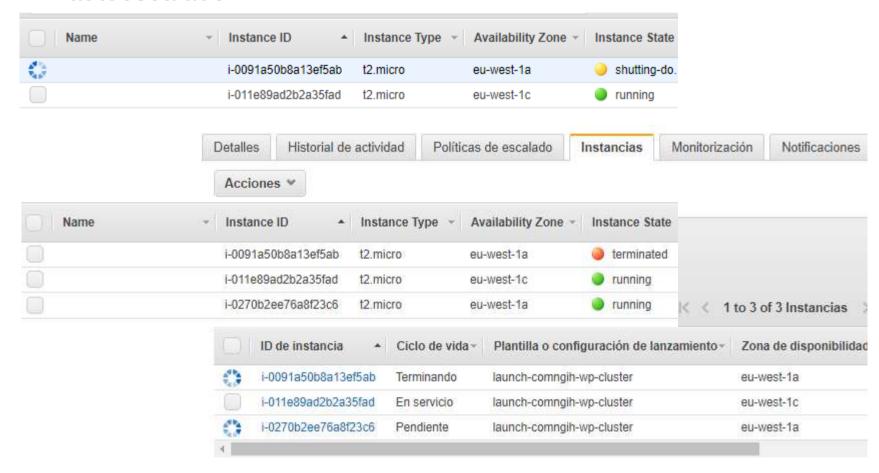
ngi etorri orria. Load balancer frogaZerbiztararen IP:10.0.1.98



- Mondragon Unibertsitatea
- Goi Eskola Politeknikoa

- Práctica : ELB & Autoescalado
  - Crear una nueva alarma por parada de servidor
  - Aumentar la carga mediante JMETER para analizar autoescalado





Goi Eskola Politeknikoa

AUTO SCALING

Scaling

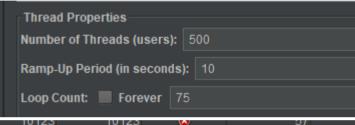
Configuraciones de lanzamiento

Grupos de Auto

### 3.6 PRACTICA Alta disponibilidad

- Práctica : ELB: Autoescalado
- Configuración:

October, 2019 12:08:55 UTC".



14.00.44.199 Hilleau Glou	HIT KEYUESI	10123	10123	<b>©</b>	37	300	129
14:08:44.454 Thread Grou	HTTP Request	10126	10126	⊗	57	308	129
14:08:34.304 Thread Grou	HTTP Request	63165	63165	⊗	57	308	129
14:08:34.314 Thread Grou	HTTP Request	63156	63156	⊗	58	308	129
14:09:37.470 Thread Grou	HTTP Request	10276	10276	⊗	191	308	129





You are receiving this email because your Amazon CloudWatch Alarm "awsec2-WP-AUTOSCALE-High-CPU-Utilization" in the EU (Ireland) region has entered the ALARM state, because "Threshold Crossed: 1 datapoint [56.503352783180524 (01/10/19 12:04:00)] was less than or equal to the threshold (70.0)." at "Tuesday 01 October, 2019 12:14:06 UTC".

▶	En curso	Terminating EC2 instance: i-0270b2ee76a8f23c6	2019 October 1 14:15:10 UTC+2	
<b>&gt;</b>	Correcto	Launching a new EC2 instance: i-08e4081e1326649d4	2019 October 1 14:09:11 UTC+2	2019 Oc

A <sup>CD</sup> A	: 0270b2cc76c0f22c6	t2 miore	au waat da	abutting do
o <sup>d b</sup> a <b>V</b> apo	i-0270b2ee76a8f23c6	t2.micro	eu-west-1a	oshutting-do
	i-08e4081e1326649	t2.micro	eu-west-1c	running



Goi Eskola Politeknikoa

AUTO SCALING

Configuraciones de lanzamiento

Grupos de Auto Scaling

Práctica : ELB: Autoescalado

Configuración:

#### **AUMENTAR-SERVERS**

Tipo de política: Escalado por pasos

Ejecutar la política

awsec2-WP-AUTOSCALE-CPU-Utilization

cuando:

supera el umbral de alarma: CPUUtilization >= 50 durante 60 segundos

para las dimensiones de métricas AutoScalingGroupName = WP-AUTOSCALE

Realizar la acción: Añadir 1 Instancias cuando 50 <= CPUUtilization < +infinito

Las instancias necesitan:

100 segundos para prepararse después de cada paso

#### reducir-servers

Tipo de política: Escalado por pasos

Ejecutar la política

awsec2-WP-AUTOSCALE-High-CPU-Utilization

cuando:

supera el umbral de alarma: CPUUtilization < 50 durante 300 segundos

para las dimensiones de métricas AutoScalingGroupName = WP-AUTOSCALE

Realizar la acción: Eliminar 1 Instancias cuando 50 >= CPUUtilization > -infinito

Thread Properties

Number of Threads (users): 500

Ramp-Up Period (in seconds): 10

Loop Count: Forever 125

		Loop Count. Tolevel	123
<b>▲▼</b>			
Basic Advanced			
Web Server			
Protocol [http]: http	Server Name or IP:	www.jaagirre.eus	
UTTD Dominet			



Goi Eskola Politeknikoa

### 3.6 PRACTICA Alta disponibilidad

- Práctica : ELB: Autoescalado
- Configuración:



Configuraciones de lanzamiento

Grupos de Auto Scaling

