PRÁCTICAS DE AWS SESIÓN 5: BALANCEO DE CARGA Y AUTO ESCALADO (ELB Y EC2 AUTO SCALING)

Roberto R. Expósito (<u>roberto.rey.exposito@udc.es</u>)
Guillermo L. Taboada (<u>guillermo.lopez.taboada@udc.es</u>)

Informática coma Servizo Mestrado Universitario en Enxeñaría Informática - UDC

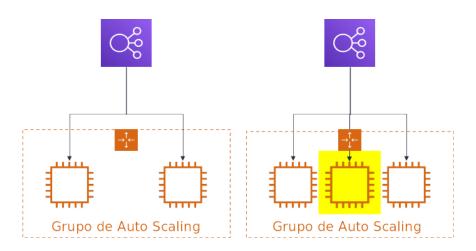


- Amazon ELB es un servicio que distribuye automáticamente el tráfico entrante de las aplicaciones o de la red entre varios destinos en una o varias zonas de disponibilidad
 - AWS escala el balanceador de carga a medida que el tráfico varía en el tiempo
 - Tipos de balanceadores de carga
 - De aplicaciones (Application Load Balancer, ALB): HTTP y HTTPS
 - Dirige el tráfico en función del contenido de la solicitud
 - De red (Network Load Balancer, NLB): tráfico TCP, UCP y TLS
 - Dirige el tráfico en función de los datos del protocolo IP
 - Clásico (generación previa, en desuso): tráfico HTTP, HTTPS, TCP y SSL
 - Balanceo entre varias instancias EC2



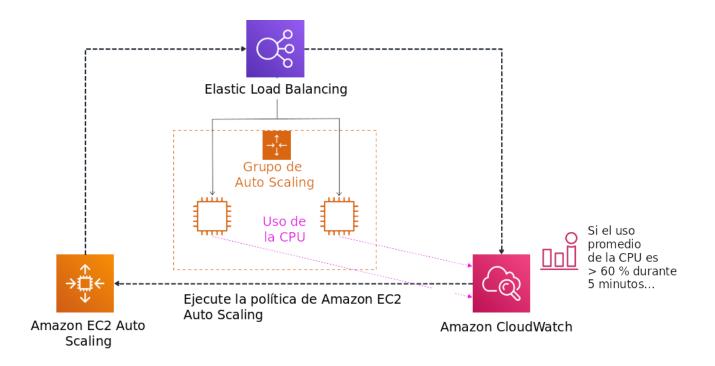


- Amazon EC2 Auto Scaling es un servicio que permite mantener la disponibilidad de las aplicaciones y aumentar o reducir automáticamente la capacidad de Amazon EC2 según las condiciones que se definan
 - Adecuado para aplicaciones con patrones de demanda estables o para aquellas cuyo uso varía cada hora, día o semana
 - Detecta instancias EC2 en mal estado y las reemplaza sin intervención
 - Soporta varias opciones de escalado: manual, programado, dinámico, bajo demanda, predictivo...
 - Grupos de Auto Scaling: colección de instancias EC2 que se tratan como una agrupación lógica a efectos de la administración y escalado automático





- EC2 Auto Scaling con escalado dinámico mediante políticas de escalado
 - Una política de escalado dinámico indica a EC2 Auto Scaling que realice un seguimiento de una métrica específica de CloudWatch (p.e., uso de CPU) y define qué acción se debe realizar cuando la alarma de CloudWatch asociada está en ALARMA
 - Las métricas que se utilizan para invocar el estado de alarma son una agregación de métricas procedentes de todas las instancias del grupo de Auto Scaling

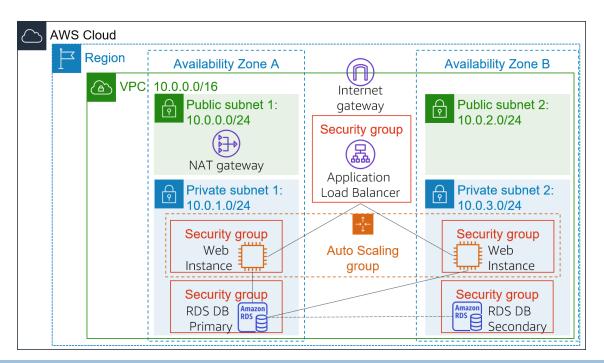




- Completar en AWS Academy el Laboratorio 6 (Escalado y balanceo de la carga de su arquitectura) del Módulo 10 (Monitoreo y escalado automático)
 - La duración estimada para completar el laboratorio es de ~40 minutos
- Prerrequisitos recomendados del módulo 10:
 - Visualizar la Sección 1 (Elastic Load Balancing) (~7 minutos)
 - Visualizar la Sección 2 (Amazon CloudWatch) (~4 minutos)
 - Visualizar la <u>Sección 3 (Amazon EC2 Auto Scaling)</u> (~9 minutos)
- Los objetivos específicos de este laboratorio son:
 - Crear una AMI a partir de una instancia EC2 en ejecución
 - Crear un balanceador de carga
 - Crear una configuración de lanzamiento y un grupo de Auto Scaling
 - □ Escalar automáticamente instancias nuevas dentro de una subred privada
 - Crear alarmas de Amazon CloudWatch y monitorear la infraestructura



- Tarea 1: crear una AMI para Auto Scaling
- Tarea 2: crear un balanceador de carga
- Tarea 3: crear una plantilla de lanzamiento y un grupo de Auto Scaling
- □ Tarea 4: verificar el funcionamiento del balanceo de carga
- Tarea 5: realizar pruebas de Auto Scaling
- Tarea 6: terminar el servidor web





 Documento PDF que incluya TODAS las capturas de pantalla mostradas en las siguientes transparencias donde se vea la información solicitada



¿Cómo completar los ejercicios de laboratorio en AWS Academy?

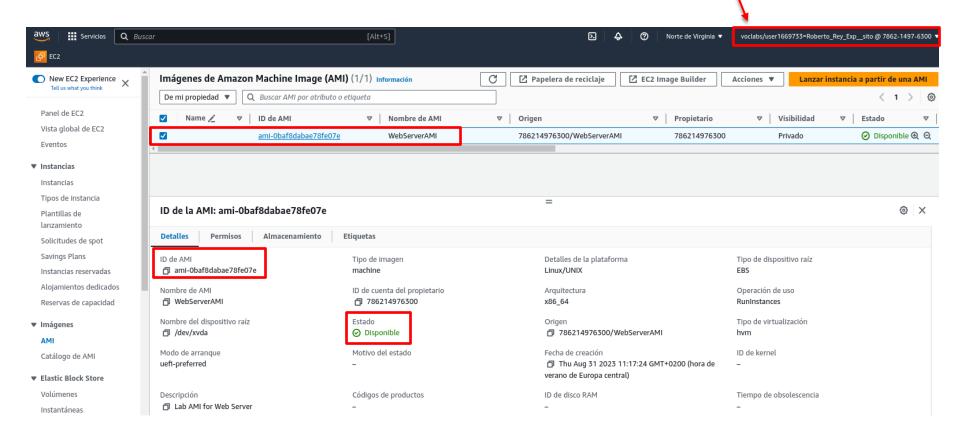
Recuerda que hay un límite de crédito para cada ejercicio de laboratorio en AWS Academy. Si alcanzas el límite, tu cuenta temporal es deshabilitada, y NO podrás iniciar el laboratorio de nuevo. Este límite es como mínimo dos veces la cantidad requerida para completar los ejercicios, pero debes usarlo de forma responsable. Recuerda finalizar el laboratorio cuando termines el ejercicio o no vayas a seguir trabajando en Academy

- ENTREGA de la práctica a través de Moodle
- NO RECORTES las capturas, DEBE verse TODA la información que sea relevante para comprobar el trabajo realizado
- Si no se siguen estas normas LA PRÁCTICA NO SE CONSIDERARÁ "APTA"



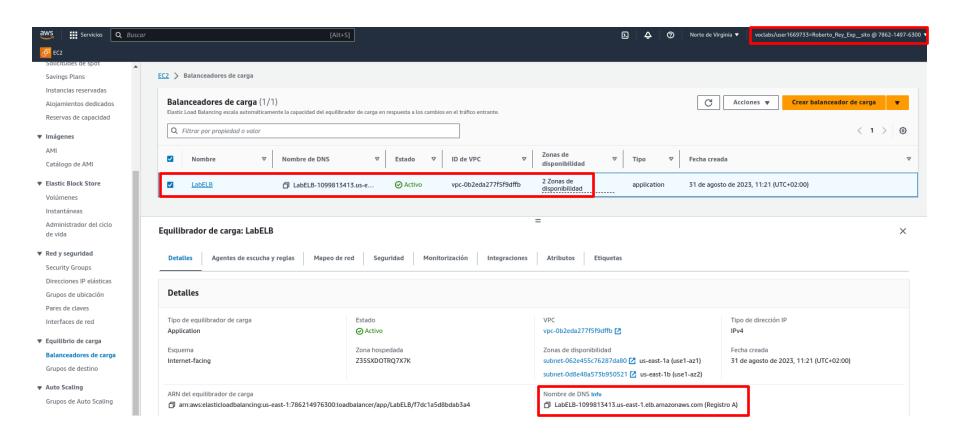


En TODAS las capturas de pantalla de la consola de AWS debes mostrar la misma información que ves en los ejemplos que se incluyen en este enunciado. Especialmente importante es que se vea SIEMPRE el ID y el nombre de usuario de tu cuenta, ya que te identifica unívocamente



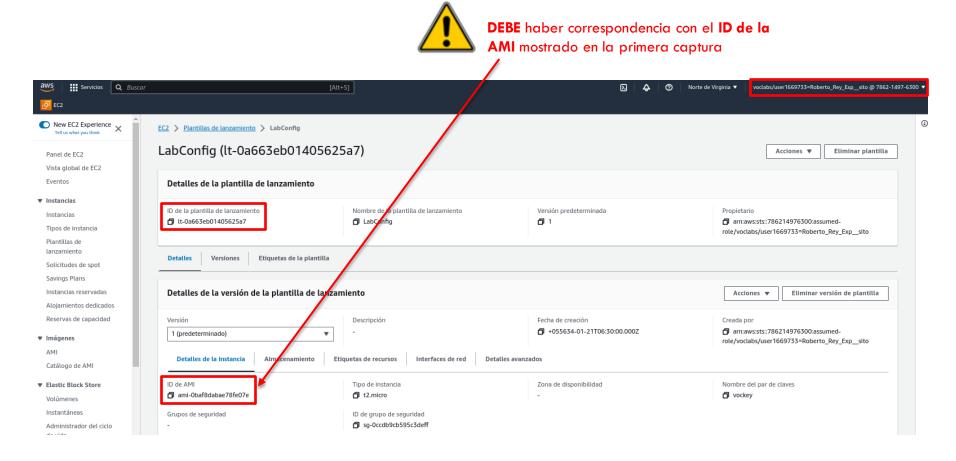
I) Detalles de la AMI (Tarea 1)





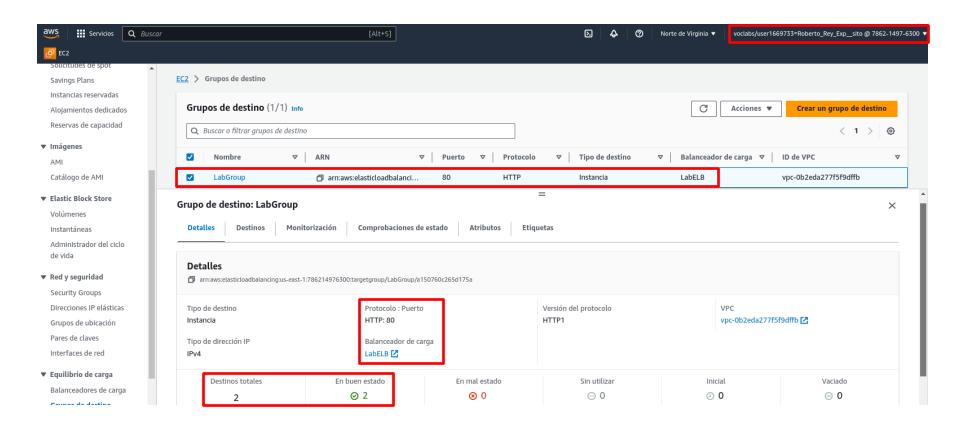
II) Detalles del balanceador de carga (Tarea 2)





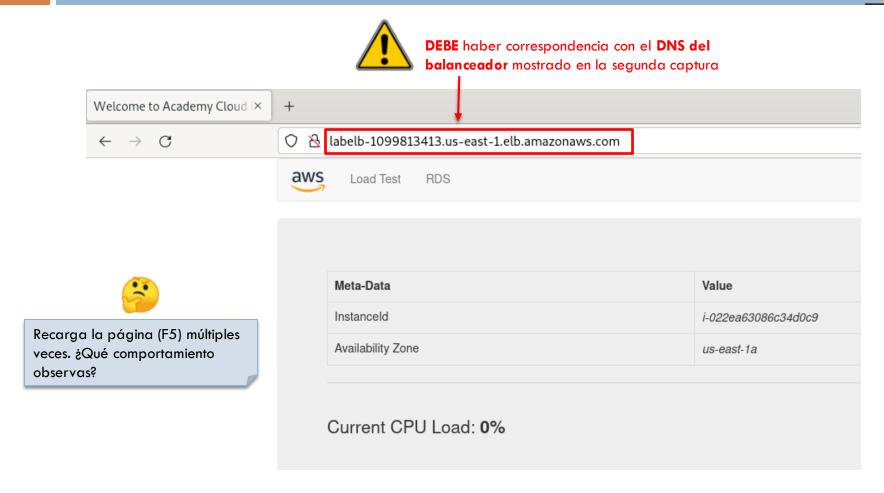
III) Detalles de la plantilla de lanzamiento (Tarea 3)





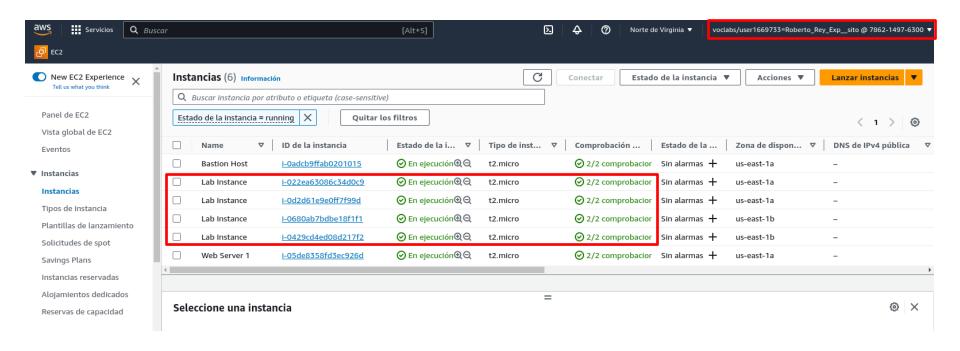
IV) Detalles del grupo de Auto Scaling
(Tarea 3)





V) Acceso a la aplicación web mediante el balanceador de carga (Tarea 4)





VI) Instancias EC2 en ejecución tras las pruebas de Auto Scaling (Tarea 5)