

© Germán Moltó, 2013-2024. Se prohíbe la divulgación, utilización, transmisión, distribución, reproducción y modificación total o parcial de este documento y de cualquier otro material educativo del Curso Online de Cloud Computing con Amazon Web Services por cualquier medio sin el previo y expreso consentimiento del autor, ni siquiera para ámbito académico y/o educativo. Este material es de uso estricta y exclusivamente personal.

09/10/2024

Práctica

Despliegue de Grupos Elásticos de Instancias y Gestión de Datos con Amazon Web Services usando la AWS Management Console

Práctica

Despliegue de Grupos Elásticos de Instancias y Gestión de Datos con Amazon Web Services usando la AWS Management Console

Contenido

1.	Introducción	2
2.	Resultados de Aprendizaje.....	2
3.	Conexión a la AWS Management Console.....	3
4.	Gestión de Datos en Amazon Web Services.....	4
4.1.	Usando el Servicio S3 para Almacenar Ficheros.....	4
4.2.	Usando el Servicio S3 para Servir una Web Estática	13
4.3.	Gestión de Volúmenes con EBS	19
	Gestión de Snapshots de Volúmenes EBS.....	27
4.4.	Creación y Registro de AMIs	35
5.	Balanceo de la Carga.....	41
6.	Auto-Escalado de Grupos Instancias	53
6.1.	Configuración de Auto Scaling	53
7.	Monitorización y Gestión	68
7.1.	Monitorización y Alarmas	68
8.	Conclusiones.....	76
	Información Adicional	76
	Referencias	76

1. Introducción

Amazon Web Services (AWS) [1] es un proveedor de Cloud público pionero en el campo de las tecnologías Cloud (ofreciendo servicio desde 2006). De entre todos los servicios que ofrece, en esta práctica se utilizarán principalmente los siguientes:

- Amazon Elastic Compute Cloud (EC2).
- Amazon Simple Storage Service (S3).
- Amazon Elastic Block Store (EBS)
- Amazon Elastic Load Balancing (ELB)

Esta práctica permite que el alumno realice el almacenamiento de ficheros en la nube mediante Amazon S3. También se pretende que realice el despliegue avanzado de grupos elásticos de instancias sobre la infraestructura de Cloud público de Amazon, usando EC2. Esto permitirá ofrecer una visión de la forma de trabajar desde el punto de vista del IaaS (*Infrastructure as a Service*), donde se realiza un aprovisionamiento de máquinas virtuales y de espacio de almacenamiento, que posteriormente debe gestionar de forma conveniente el usuario. Existen múltiples cuentas de usuario para los alumnos creadas mediante el servicio IAM (*Identity and Access Management*) bajo una cuenta de usuario de AWS con cargo a la tarjeta de crédito VISA del profesor. Esto permite agrupar el coste de todos los usuarios bajo una misma cuenta.

En esta práctica se utilizará fundamentalmente la AWS Management Console que permite la interacción con los principales servicios de AWS de forma gráfica a través de un navegador web. AWS también ofrece la línea de comandos para interactuar con el Cloud público mediante la herramienta AWS CLI [20]. El uso de línea de comandos facilita el *scripting* o la posibilidad de integrar la funcionalidad de AWS en aplicaciones propias del usuario. Existen a tu disposición dos versiones diferentes de esta misma práctica dependiendo de si te sientes más cómodo usando línea de comandos (AWS CLI) o un navegador web (AWS Management Console).

2. Resultados de Aprendizaje

Se espera que, una vez finalizada la práctica, el alumno sea capaz de:

- Conocer el esquema de funcionamiento a nivel de usuario de Amazon S3.
- Comprender el concepto de *bucket* así como los mecanismos de creación y destrucción de los mismos.
- Comprender el concepto de volumen EBS, así como su gestión básica.
- Conocer funcionalidad avanzada del servicio Amazon EC2.
- Entender los procedimientos de balanceo de carga ofrecidos por el servicio ELB.
- Crear grupos de autoescalado para la creación de flotas elásticas de instancias que crezcan en función de la carga de trabajo.
- Manejar con soltura la AWS Management Console para interactuar con el servicio Amazon EC2 y Amazon S3.



1. Tu profesor te habrá asignado un número. Los comandos que aparecen en el boletín hacen referencia al usuario alucloudoo. Por favor, **sustituye en los comandos que veas oo por tu número** (por ejemplo, si tu número es 06, entonces

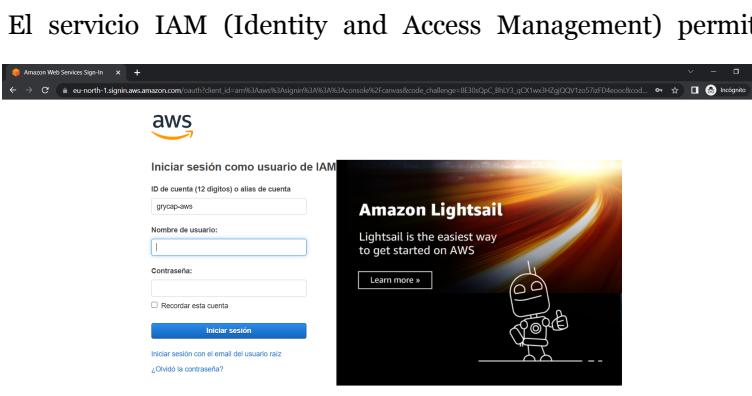
el usuario que has de utilizar es alucloudo6). Recuerda esta regla para facilitar tanto tu trabajo como el del resto de compañeros. Para facilitar tu trabajo, en los comandos se utiliza una variable de entorno llamada ID que se debe resolver a tu número.

2. En esta práctica haremos ejecuciones reales sobre un proveedor de Cloud público, que generan un coste económico. Asegúrate de **liberar apropiadamente los recursos** (terminar instancias, eliminar ficheros de S3 si ya no los vas a gastar, etc.) para **no incurrir en costes adicionales**. Si tienes alguna duda al respecto consulta con tu profesor.
3. Este boletín de prácticas está diseñado para responder a la problemática común que te puedes encontrar durante la resolución de la práctica. Si te surge algún problema durante la realización de la misma, lo más probable es que el boletín te explique a continuación una posible solución al respecto. Revisa dicha propuesta de solución antes de contactar con el instructor. Si encuentras un problema que no está explicado en el boletín contacta con el instructor proporcionando una explicación detallada del mismo para facilitar aislar el problema y encontrar una solución.
4. Por defecto trabajarás en la región us-east-1 (N. Virginia) de AWS. También tendrás acceso a la región eu-west-1 (Ireland). El acceso al resto de regiones de AWS ha sido restringido por el instructor.
5. Recuerda que hay un sistema de recompensas, definido en la Guía de Prácticas, por el que podrás aumentar la duración del curso reportando las discrepancias que encuentres en este boletín frente a posibles cambios introducidos por AWS. Cuento con tu colaboración.

3. Conexión a la AWS Management Console

En primer lugar, conéctate con las credenciales de IAM que te suministró tu instructor a la siguiente dirección.

<https://grycap-aws.signin.aws.amazon.com/console>

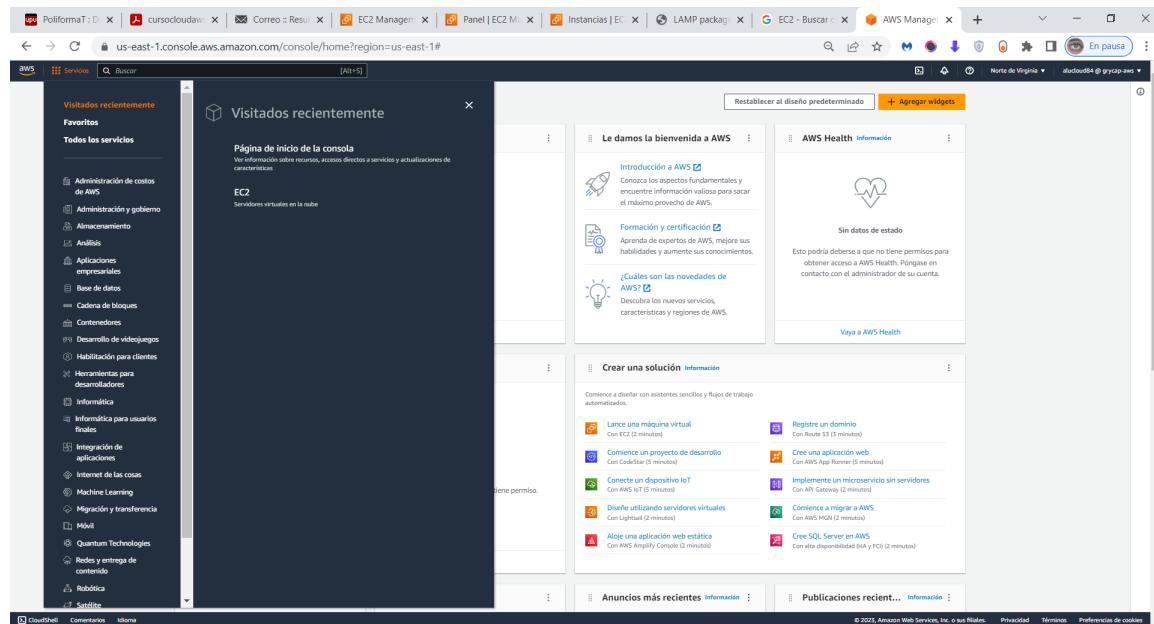


The screenshot shows the AWS sign-in interface. On the left, there's a form titled "Iniciar sesión como usuario de IAM". It has fields for "ID de cuenta (12 dígitos) o alias de cuenta" containing "grycap-aws", "Nombre de usuario" containing "grycap-aws", and "Contraseña". Below the password field is a "Recordar esta cuenta" checkbox. At the bottom of the form is a blue "Iniciar sesión" button. To the right of the form is a promotional banner for "Amazon Lightsail" with the tagline "Lightsail is the easiest way to get started on AWS". The banner features a small cartoon robot icon.

El servicio IAM (Identity and Access Management) permite la creación de múltiples cuentas vinculadas a usuario de AWS, de manera que múltiples personas pueden acceder a AWS usando las credenciales de su usuario IAM con cargo a la tarjeta de crédito del usuario AWS para el que han sido creadas.

Una vez autenticado, obtendrás acceso a la consola de administración de AWS (AWS Management Console), tal y como se muestra en la siguiente figura.

Esta herramienta web permite la administración de los principales servicios de AWS.

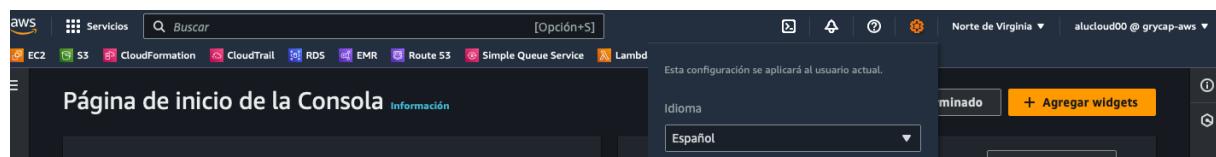


4. Gestión de Datos en Amazon Web Services

En esta parte de la práctica se trabajará con la gestión de datos en AWS. En concreto, se trabajará con Simple Storage Service (S3), que permite almacenar ficheros en la nube. Existen otros servicios relacionados con datos en AWS, como EBS, que permite construir volúmenes que pueden ser conectados a instancias de forma dinámica, al igual que si de un disco duro se tratase, para almacenar ficheros en él. También existen las bases de datos en la nube, como es el caso de Amazon SimpleDB o Relational Database Service (RDS). Todos los servicios mencionados serán abordados en las correspondientes prácticas.

4.1. Usando el Servicio S3 para Almacenar Ficheros

Amazon Simple Storage Service (S3) ofrece un sistema de almacenamiento de ficheros similar a una estructura jerárquica de directorios y ficheros como un sistema de archivos tradicional, pero almacenado en la nube. Ofrece mecanismos de replicación de ficheros para tolerancia a fallos. A continuación, se resume el proceso de interacción típico con S3. Elige desde la AWS Management Console el servicio S3. Puedes cambiar la interfaz a idioma español mediante el menú que está en la parte superior:



En ella se muestran todos los *buckets* que hay creados. Verás que ya existe un bucket creado (llamado *alucloud*). Es posible que también aparezcan algunos buckets creados temporalmente por otros alumnos durante la realización de esta misma práctica (llamados *alucloudo1*, *alucloudo2*, etc.).

The screenshot shows the 'Amazon S3' service interface. At the top, there is a note: 'Para habilitar la organización de todas las columnas de la siguiente tabla, utilice la búsqueda a fin de reducir el tamaño de la lista a 100 buckets o menos.' Below this, a table lists 'Buckets (107)'. The table has columns: Nombre, Región, Acceso, and Fecha de creación. One row is visible for the bucket 'alucloud', which is located in 'EE. UU. Este (Norte de Virginia) us-east-1', has 'Los objetos pueden ser públicos' access, and was created on '6 Oct 2014 3:29:57 PM CEST'. Action buttons at the top right include 'Copiar ARN', 'Vaciar', 'Eliminar', and 'Crear bucket'.

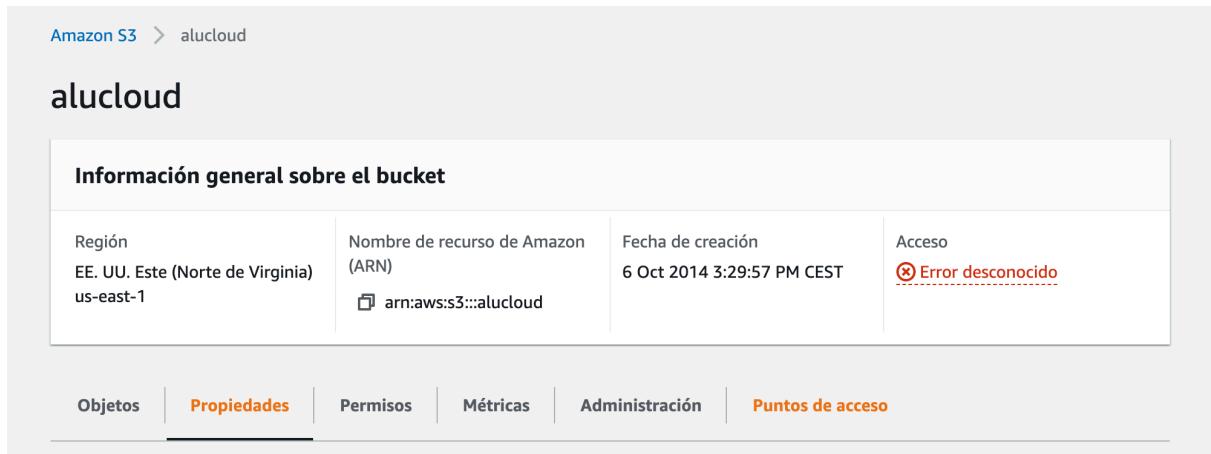
El nombre del bucket elegido (*alucloud*) debe ser único en todo Amazon S3 y el instructor ya lo creó en su momento para utilizarlo durante la realización de la práctica (de manera que ningún otro usuario de AWS cree otro bucket con el mismo nombre). Por ello, no elimines ninguno de los buckets creados en Amazon S3.

Los buckets van ligados a una región concreta. Los datos (ficheros) almacenados en dicho bucket se almacenan de forma replicada por tolerancia a fallos y alta disponibilidad, pero las réplicas siempre se mantienen dentro de la región en la que se creó el bucket. De hecho, si tratas de crear un bucket se te solicita la región en la que se almacenarán los datos y si el nombre ya existe obtienes un mensaje de error:

The screenshot shows the 'Crear bucket' (Create Bucket) wizard. The first step, 'Configuración general' (General Configuration), has a 'Nombre del bucket' field containing 'god'. A warning message says: 'Ya existe un bucket con el mismo nombre' (A bucket with this name already exists). It also specifies the region as 'EE. UU. Este (Norte de Virginia) us-east-1'. The second step, 'Configuración del bucket para Bloquear acceso público' (Bucket configuration for Blocking public access), contains several configuration options under the heading 'Bloquear todo el acceso público' (Block all public access). The 'Bloquear todo el acceso público' checkbox is checked. Other options listed include: 'Bloquear el acceso público a buckets y objetos concedido a través de nuevas listas de control de acceso (ACL)' (Block public access to buckets and objects via new ACLs), 'Bloquear el acceso público a buckets y objetos concedido a través de cualquier lista de control de acceso (ACL)' (Block public access to buckets and objects via any ACL), 'Bloquear el acceso público a buckets y objetos concedido a través de políticas de bucket y puntos de acceso públicas nuevas' (Block public access to buckets and objects via new bucket policies and access points), and 'Bloquear el acceso público u entrar cuentas a buckets y objetos concediendo a través de cualquier política de' (Block public access or allow accounts to buckets and objects by granting via any policy).

Al seleccionar un bucket concreto y pulsar la pestaña de propiedades verás que es posible obtener y definir cierta información para cada bucket (se indica entre paréntesis el nombre en inglés por si acaso estás usando la interfaz en Inglés). Únicamente se muestran las opciones principales:

- Control de versiones de buckets (*Versioning*). Permite preservar, obtener y restaurar cada versión de cada objeto almacenado en dicho bucket. Permite recuperar versiones pasadas en caso de borrados accidentales, por ejemplo.
- Etiquetas (*Tags*). Permite añadir etiquetas al bucket para luego obtener la información de coste agregada por dichas etiquetas [52].
- Cifrado predeterminado (*Default encryption*). Permite cifrar automáticamente objetos almacenados en un bucket.
- Configuraciones de agrupación por niveles inteligente. Permite que los objetos almacenados en la clase de almacenamiento Agrupación por niveles inteligente cambien al nivel Acceso a archivos o al nivel Acceso a archivos profundos. Estos niveles están optimizados para objetos a los que se accede de forma poco frecuente a lo largo de un gran período de tiempo [81].
- Registros de acceso del servidor (*Logging*) [29], para guardar un registro de acceso detallado a tus buckets de S3 (quién ha accedido a qué recurso y cuándo, en el siguiente formato [30]). Dicho registro se guarda, a su vez, en un bucket de S3. Nunca indiques como bucket destino de los logs el propio bucket, puesto que provocarás la generación de miles de ficheros de logs a lo largo del tiempo.
- Alojamiento de sitios web estáticos (*Static Website Hosting*). Soporta HTML básico, nada de PHP, por ejemplo.
- Eventos de datos de AWS CloudTrail (*Object-level logging*). Permite guardar un registro de los accesos a los ficheros, utilizando para ello el servicio CloudTrail.
- Eventos (*Events*). Permite ser notificado mediante el servicio de notificaciones Amazon SNS (Simple Notification Service) [31] ante diferentes eventos. Por ejemplo, cuando un fichero almacenado con almacenamiento basado en redundancia reducida (*Reduced Redundancy Storage* [32]) se pierde. Las alertas también pueden ser enviadas a una cola de Amazon SQS (Simple Queue Service) [50] o a una función Lambda [51] para ejecutar código de forma reactiva.
- Aceleración de Transferencia (*Transfer Acceleration*) [55]. Permite acelerar la transferencia de entrada y salida de datos del bucket (carga y descarga) hasta un 300% y únicamente se incurre en un cargo adicional si efectivamente se ha producido una mejora.
- Bloqueo de objeto [75]. Permite evitar que se eliminan o sobrescriban ciertos ficheros durante un determinado periodo de tiempo, o de forma indefinida.
- Pago por el solicitante (*Requester Pays*) [54]. Permite que el solicitante pague por la descarga de los ficheros del bucket.



The screenshot shows the 'Properties' tab of the 'alucloud' bucket in the Amazon S3 console. The top navigation bar shows 'Amazon S3 > alucloud'. The main title is 'alucloud'. Below it, the 'Información general sobre el bucket' (General information about the bucket) section displays the following details:

Región	Nombre de recurso de Amazon (ARN)	Fecha de creación	Acceso
EE. UU. Este (Norte de Virginia) us-east-1	<input type="button"/> arn:aws:s3:::alucloud	6 Oct 2014 3:29:57 PM CEST	<input checked="" type="checkbox"/> Error desconocido

Below this, there are tabs for 'Objetos', 'Propiedades' (which is selected), 'Permisos', 'Métricas', 'Administración', and 'Puntos de acceso'.

La pestaña “Permisos” te permite ver la configuración de acceso público, la lista de control de acceso y la política de bucket y la configuración de CORS [71].

La pestaña “Administración” te permite definir:

- Reglas de ciclo de vida (*Lifecycle*). Permite la creación de reglas para, por ejemplo, eliminar los ficheros de un bucket cada cierto tiempo (útil para usarlo como almacenamiento temporal) o para moverlo a Amazon Glacier [33], el servicio de archivo de bajo coste y a largo plazo, pasado un tiempo.
- Reglas de replicación (*Cross-Region Replication*) [53]. Permite replicar los ficheros subidos a dicho bucket, de forma automática y asíncrona, en otro bucket (que puede estar en otra región).
- Inventario (*Inventory*) [58]. Permite obtener una serie de ficheros y sus metadatos para tener un control detallado del contenido del mismo.

En primer lugar, crearemos una carpeta de trabajo dentro del bucket *alucloud*, con tu identificador de usuario. De esta manera, los ficheros del alumno alucloud190 estarán en la carpeta 190 del bucket *alucloud*. Es muy importante que respetas esta nomenclatura de recursos para facilitar tanto tu trabajo como el de tus compañeros. Asegúrate de tener seleccionado el bucket *alucloud* y pulsa el botón “Crear carpeta”. Introduce el identificador de alumno como nombre de carpeta (por ejemplo, 190 en la siguiente figura). Si el directorio ya está creado puedes borrarlo sin problema (incluso aunque contenga datos) y luego crear el tuyo vacío. Deja el valor por defecto, ya que no queremos cifrar el contenido de dicha carpeta.

The screenshot shows the AWS S3 'Objetos' (Objects) page with 122 items. A modal window titled 'Crear carpeta' (Create folder) is open. In the 'Nombre de la carpeta' (Folder name) field, '190' is entered. Below it, a note states: 'Si su política del bucket requiere que los objetos se cifren con una clave de cifrado específica, deberá especificar la misma clave de cifrado al crear una carpeta. De lo contrario, se producirá un error al crear la carpeta.' (If your bucket policy requires objects to be encrypted with a specific key, you must specify the same encryption key when creating a folder. Otherwise, an error will occur when creating the folder.) At the bottom right of the modal is a yellow 'Crear carpeta' (Create folder) button.

Pulsa sobre la nueva carpeta correspondiente a tu identificador para entrar dentro de ella. Debe estar vacía:

The screenshot shows the AWS S3 'Objetos' (Objects) page for the '190/' folder. It displays a message: 'No hay objetos' (No objects) and 'No tiene objetos en esta carpeta.' (This folder does not contain any objects). At the bottom right of the page is a yellow 'Cargar' (Upload) button.

A continuación, subimos algún fichero al bucket asociado a tu usuario (no importa el contenido del fichero, elige cualquier fichero de tu disco duro que tenga un tamaño razonable). Deberás seleccionar el bucket alucloud y tu carpeta correspondiente y luego pulsar sobre el botón “Cargar”.

Tras elegir el fichero (permite elegir varios ficheros al mismo tiempo), puedes ver las “Propiedades”:

Clase de almacenamiento	Diseñado para	Zonas de disponibilidad	Duración de almacenamiento
<input checked="" type="radio"/> Estándar	Datos a los que se accede con frecuencia (más de una vez al mes) con acceso en milisegundos	≥ 3	-
<input type="radio"/> Agrupación por niveles inteligente	Datos con patrones de acceso desconocidos o cambiantes	≥ 3	-
<input type="radio"/> Estándar - Acceso poco frecuente	Datos a los que se accede con poca frecuencia (una vez al mes) con acceso en milisegundos	≥ 3	30 días
<input type="radio"/> Única zona - Acceso poco frecuente	Datos recreables y a los que se accede con poca frecuencia (una vez al mes) almacenados en una única zona de disponibilidad con acceso en milisegundos	1	30 días
<input type="radio"/> Recuperación instantánea de Glacier	Datos de archivo de larga duración a los que se obtiene acceso una vez al trimestre con recuperación instantánea en milisegundos	≥ 3	90 días
<input type="radio"/> Glacier Flexible Retrieval (anteriormente Glacier)	Acceso a datos de archivo de larga duración una vez al año con recuperación de minutos a horas	≥ 3	90 días
<input type="radio"/> Glacier Deep Archive	Datos de archivo de larga duración a los que se obtiene acceso menos de una vez al año con recuperación en horas	≥ 3	180 días
<input type="radio"/> Redundancia reducida	Datos no críticos y a los que se accede con frecuencia con acceso en milisegundos (no se recomienda porque S3 Standard es más rentable)	≥ 3	-

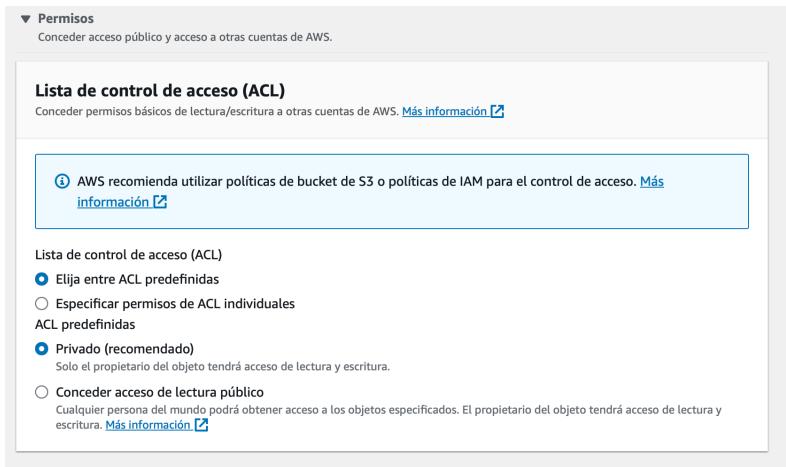
Se puede elegir la clase de almacenamiento, que puede ser, entre otras, “Estándar” (durabilidad del 99.99999999%, disponibilidad del 99.99%), “Niveles inteligentes”, para datos de larga duración con patrones de acceso variables o “Estándar-Acceso poco frecuente” (durabilidad del 99.99999999%, disponibilidad del 99%, más económico, pero con sobrecargo por acceso a datos). La opción “Única

zona – Acceso poco frecuente) permite guardar los ficheros replicados, pero solo en una Zona de Disponibilidad. Es posible reducir todavía más la durabilidad de los datos eligiendo la opción “Redundancia reducida” que permite reducir la durabilidad al 99.99%, disminuyendo el coste de almacenamiento. Por último, tienes opciones de almacenamiento a muy bajo coste, pero que requiere más tiempo de acceso mediante Glacier [33]. Tienes una tabla que resume las diferencias entre las diferentes clases de almacenamiento en Adicionalmente, se puede utilizar cifrado del lado del servidor (*server side encryption*) [34]. También es posible definir metadatos y etiquetas adicionales a los ficheros.

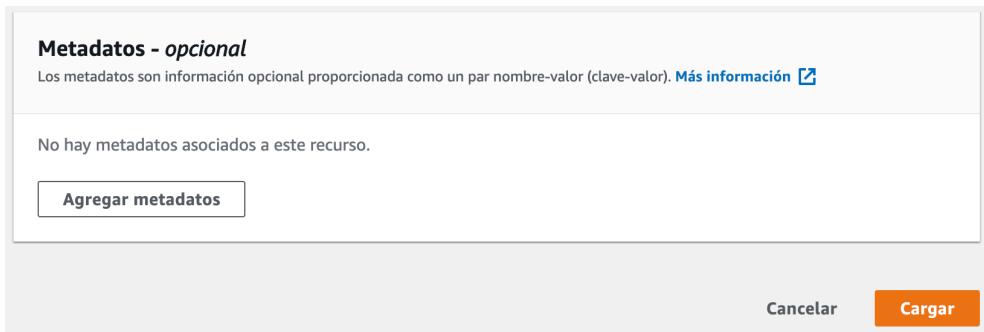
También es posible cambiar los permisos de acceso a los ficheros y permitir el acceso desde otras cuentas de AWS. Aparecen dos tipos de permisos:

- Objetos. Otorga permisos al usuario para listar, crear, sobrescribir y eliminar objetos de este bucket.
- ACL del objeto. Otorga permisos al usuario para leer o escribir en la lista de control de accesos (ACL) del bucket.

No haremos ningún cambio de permisos.



Finalmente, procederemos a subir el fichero pulsando el botón “Cargar”:



Subimos el fichero pulsando sobre el botón “Cargar”. Finalizada la subida, selecciona el fichero y verás las características del mismo:

Puedes ver que los ficheros subidos a un bucket de S3 llevan asociado una URL para su acceso desde el exterior.

Si tratas de acceder a dicha URL deberías obtener el siguiente mensaje de error:

```
-<Error>
<Code>AccessDenied</Code>
<Message>Access Denied</Message>
<RequestId>E6ABF7D5C7A2EB37</RequestId>
-<HostId>
    12xstxTqoU+BdYFOl6BfF24VftnTjF5Nw8Dr2xF8mambbqQWI/8FMQzpmBMAAn3T15YqZCTnRsHo=
</HostId>
</Error>
```

Este error es debido a que por defecto el fichero no es público. Si se desea hacerlo público (que pueda ser accedido desde cualquier punto de Internet) se tiene que modificar la lista de control de acceso (ACL) del fichero en S3 para permitir la lectura pública [21]. La forma más sencilla de hacerlo es seleccionando el fichero, pulsar sobre “Acciones de los objetos” y, posteriormente, “Hacer público”.

Si tratas de acceder ahora al fichero desde la URL (en este caso es:

https://s3.amazonaws.com/alucloud/190/cursocloudaws_logo.png

verás que ahora ya es posible acceder al fichero.

Finalmente eliminamos dicho fichero. Selecciónalo y pulsa sobre el botón “Eliminar”.

No es necesario que elimines la carpeta. La utilizarás más adelante.

4.2. Usando el Servicio S3 para Servir una Web Estática

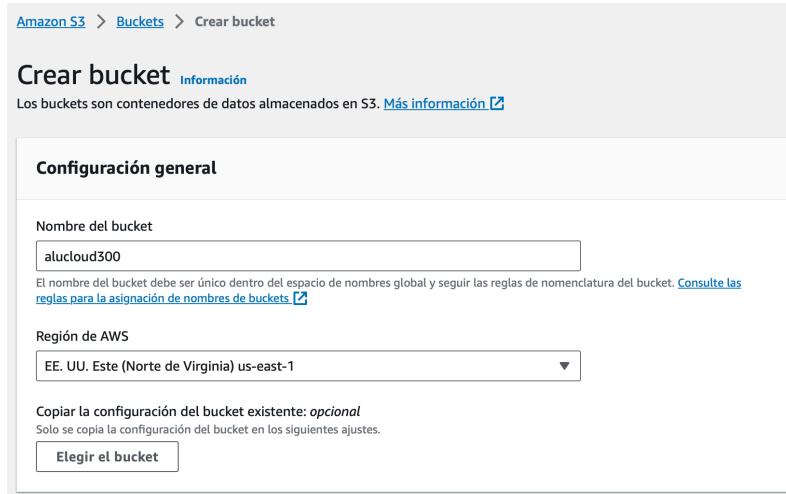
Amazon S3 incluye una característica llamada *Static Website Hosting* que permite actuar como servidor de páginas web (estáticas, sin soporte a lenguajes interpretados por el servidor, como podría ser el caso de PHP). Esto permite dejar una página web estática (junto con sus imágenes, ficheros, etc.) en un bucket de S3 y luego acceder a la web con un navegador a la siguiente URL:

alucloudoo.s3-website-us-east-1.amazonaws.com

Si quieres probar esta funcionalidad deberás utilizar temporalmente un nuevo bucket (en la región US Standard, que es North Virginia) dado que esta característica se debe activar a nivel de bucket. Te recomiendo que crees temporalmente un nuevo bucket (con la nomenclatura alucloudXX donde XX es tu identificador de usuario). Acuérdate de eliminarlo cuando acabes esta sección.

Pulsa sobre el botón “Crear bucket”:

Te saldrá un panel para indicar la configuración del bucket.



Es importante que conozcas las principales opciones. En primer lugar, es posible mantener todas las versiones de un fichero para llevar una trazabilidad de los cambios realizados [63]. Se pueden asignar etiquetas al bucket para facilitar el desglose de costes incurridos por proyectos [65]. También se cifran automáticamente los ficheros subidos al bucket [67]. No introduzcas ningún cambio en dichos paneles, salvo para indicar el nombre del bucket. Verás un panel para indicar los permisos de acceso:

Configuración de bloqueo de acceso público para este bucket

Se concede acceso público a los buckets y objetos a través de listas de control de acceso (ACL), políticas de bucket, políticas de puntos de acceso o todas las anteriores. A fin de garantizar que se bloquee el acceso público a todos sus buckets y objetos, active Bloquear todo el acceso público. Si activa esta opción, no podrá modificar los niveles de acceso. AWS recomienda activar Bloquear todo el acceso público, pero, antes de aplicar cualquiera de estos ajustes, asegúrese de que las aplicaciones funcionarán correctamente sin acceso público. Si necesita cierto nivel de acceso público a los buckets u objetos, puede personalizar la configuración individual a continuación para adaptarla a sus casos de uso de almacenamiento específicos. [Más información](#).

- Bloquear todo el acceso público**
Activar esta configuración equivale a activar las cuatro opciones que aparecen a continuación. Cada uno de los siguientes ajustes son independientes entre sí.
 - Bloquear el acceso público a buckets y objetos concedido a través de nuevas listas de control de acceso (ACL)**
S3 bloqueará los permisos de acceso público aplicados a objetos o buckets agregados recientemente, y evitará la creación de nuevas ACL de acceso público para buckets y objetos existentes. Esta configuración no cambia los permisos existentes que permiten acceso público a los recursos de S3 mediante ACL.
 - Bloquear el acceso público a buckets y objetos concedido a través de cualquier lista de control de acceso (ACL)**
S3 ignorará todas las ACL que conceden acceso público a buckets y objetos.
 - Bloquear el acceso público a buckets y objetos concedido a través de políticas de bucket y puntos de acceso públicos nuevos**
S3 bloqueará las nuevas políticas de buckets y puntos de acceso que concedan acceso público a buckets y objetos. Esta configuración no afecta a las políticas ya existentes que permiten acceso público a los recursos de S3.
 - Bloquear el acceso público y entre cuentas a buckets y objetos concedido a través de cualquier política de bucket y puntos de acceso pública**
S3 ignorará el acceso público y entre cuentas en el caso de buckets o puntos de acceso que tengan políticas que concedan acceso público a buckets y objetos.

⚠️ Desactivar el bloqueo de todo acceso público puede provocar que este bucket y los objetos que contiene se vuelvan públicos

AWS recomienda que active la opción para bloquear todo el acceso público, a menos que se requiera acceso público para casos de uso específicos y verificados, como el alojamiento de sitios web estáticos.

Reconozco que la configuración actual puede provocar que este bucket y los objetos que contiene se vuelvan públicos.

Estas opciones están activadas por defecto para impedir el acceso público a los ficheros de S3, para evitar que haya fugas de información (como por ejemplo esta que afectó al Pentágono [72]). Sin embargo, como nosotros queremos exponer una página web pública, **es importante que desactives todas las opciones anteriores**. De lo contrario, obtendrás un problema de “Acceso Denegado” cuando trates de activar el acceso público a un fichero.

Una vez revisadas las opciones pulsa sobre “Crear bucket”.

A continuación, deberás añadir una política para permitir el acceso público al bucket de S3. Para ello, una vez seleccionado el bucket, en la pestaña Permisos del bucket, baja a la parte de “Política de bucket”:

Política de bucket
La política del bucket, escrita en JSON, proporciona acceso a los objetos almacenados en el bucket. Las políticas de bucket no se aplican a los objetos que pertenecen a otras cuentas. [Más información](#)

[Editar](#) [Eliminar](#)

Amazon S3 > alucloud00

alucloud00

Información general sobre el bucket

Región EE. UU. Este (Norte de Virginia) us-east-1	Nombre de recurso de Amazon (ARN) arn:aws:s3:::alucloud00	Fecha de creación 9 Nov 2016 3:01:12 PM CET	Acceso Error desconocido
--	--	--	---

Objetos Propiedades **Permisos** Métricas Administración Puntos de acceso

Deberás introducir esta política de bucket (asegurándote de cambiar XX por tu identificador de alumno):

```
{  
    "Version": "2012-10-17",  
    "Statement": [  
        {  
            "Sid": "PublicReadGetObject",  
            "Effect": "Allow",  
            "Principal": "*",  
            "Action": ["s3:GetObject"],  
            "Resource": ["arn:aws:s3:::alucloudXX/*"]  
        }  
    ]  
}
```

El resultado se muestra en la siguiente captura de pantalla:

Política de bucket
La política del bucket, escrita en JSON, proporciona acceso a los objetos almacenados en el bucket. Las políticas de bucket no se aplican a los objetos que pertenecen a otras cuentas. [Más información](#)

[Editar](#) [Eliminar](#)

```
{  
    "Version": "2012-10-17",  
    "Statement": [  
        {  
            "Sid": "PublicReadGetObject",  
            "Effect": "Allow",  
            "Principal": "*",  
            "Action": ["s3:GetObject"],  
            "Resource": "arn:aws:s3:::alucloud00/*"  
        }  
    ]  
}
```

[Copiar](#)

Si obtienes un error de “Acceso Denegado”, entonces es que te olvidaste de desactivar las casillas que te indiqué en la página anterior. En ese caso, lo mejor es que elimines el bucket y vuelvas a crearlo correctamente siguiendo las instrucciones. Si obtienes un error de “Network Failure”, puedes darle nuevamente a “Guardar cambios” y te permitirá guardar los cambios. Esa política habilita el acceso a los ficheros de un determinado bucket para acceder a ellos desde cualquier cliente, lo que determina efectivamente un acceso público a todos los ficheros del bucket. Ten en cuenta que esta política de bucket no es necesaria si das permisos de acceso público a todos los ficheros del sitio web estático. Tienes más información sobre las políticas en [74].

Una vez configurado el bucket, creamos una sencilla web estática (con un fichero index.html que refiere a una imagen) para luego subir el fichero a tu bucket. Para tu comodidad, aquí tienes un ejemplo de web sencilla, que puedes copiar y pegar en un fichero (index.html):

```
<html>
<p> Hola desde el Cloud </p>

</html>
```

Procede a subir dicho fichero al bucket usando la propia interfaz web del servicio Amazon S3.

A continuación, activa el *website hosting*, desde la pestaña Propiedades → Alojamiento de sitios web estáticos: tal y como se muestra en la siguiente figura:



Editar alojamiento de sitios web estáticos

Alojamiento de sitios web estáticos
Utilice este bucket para alojar un sitio web o redirigir las solicitudes. [Más información](#)

Alojamiento de sitios web estáticos
 Deshabilitar
 Habilitar

Tipo de alojamiento
 Alojar un sitio web estático
Utilice el punto de enlace del bucket como dirección web. [Más información](#)
 Redirigir las solicitudes de un objeto
Redirija las solicitudes a otro bucket o dominio. [Más información](#)

Para que sus clientes puedan obtener acceso al contenido en el punto de enlace del sitio web, debe hacer que todo el contenido sea legible públicamente. Para ello, puede editar la configuración Bloquear acceso público de S3 del bucket. Para obtener más información, consulte Utilizar Bloquear acceso público de Amazon S3

Documento de índice
Especifique la página predeterminada o de inicio del sitio web.
index.html

Documento de error
Esto se devuelve cuando se produce un error.
error.html

Puedes utilizar los siguientes valores:

Documento de índice (Index Document): index.html

Documento de error (Error Document): error.html

Reglas de redirección (opcional): (Dejarlo vacío)

Esto te permitirá acceder a la web con la siguiente URL (donde XX es tu número de alumno):

<http://alucloudXX.s3-website-us-east-1.amazonaws.com>

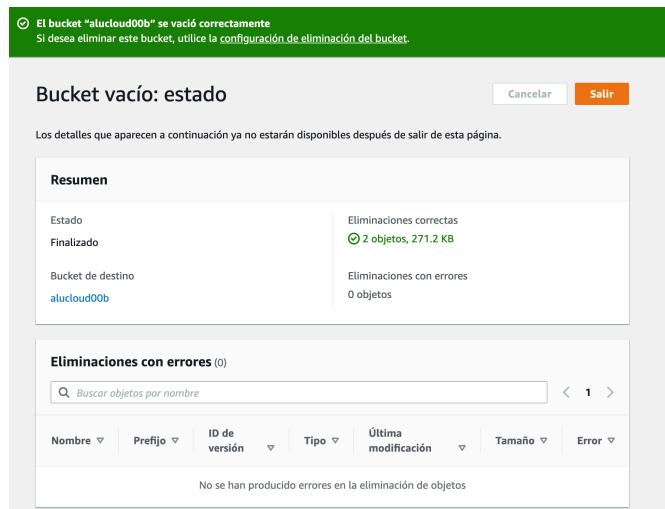
Si obtienes algún error de permiso denegado, asegúrate que el nombre del fichero HTML coincide con el indicado en la opción “Index Document” y que dichos ficheros están en la raíz del bucket, y no en alguna subcarpeta dentro del bucket. Con mínimo esfuerzo ya puedes servir una página web estática sin necesidad de desplegar una instancia en EC2, tan solo pagando por el tráfico generado. Si en lugar de mostrar la página web se muestra directamente el código HTML entonces asegúrate de que has creado correctamente la política de bucket (copia y pégala antes de modificarla). Quizá te preguntes cómo asignarle un nombre DNS concreto para acceder a dicha web con un nombre del estilo www.mi-sitio-web.com. Para ello, debes utilizar un servicio de DNS como Route 53 [69].

No olvides eliminar el bucket antes de proseguir con la práctica. Para eliminarlo, selecciona el bucket *alucloudXX* (donde XX es tu identificador de alumno) desde la consola de S3 y pulsa sobre “Delete bucket”.

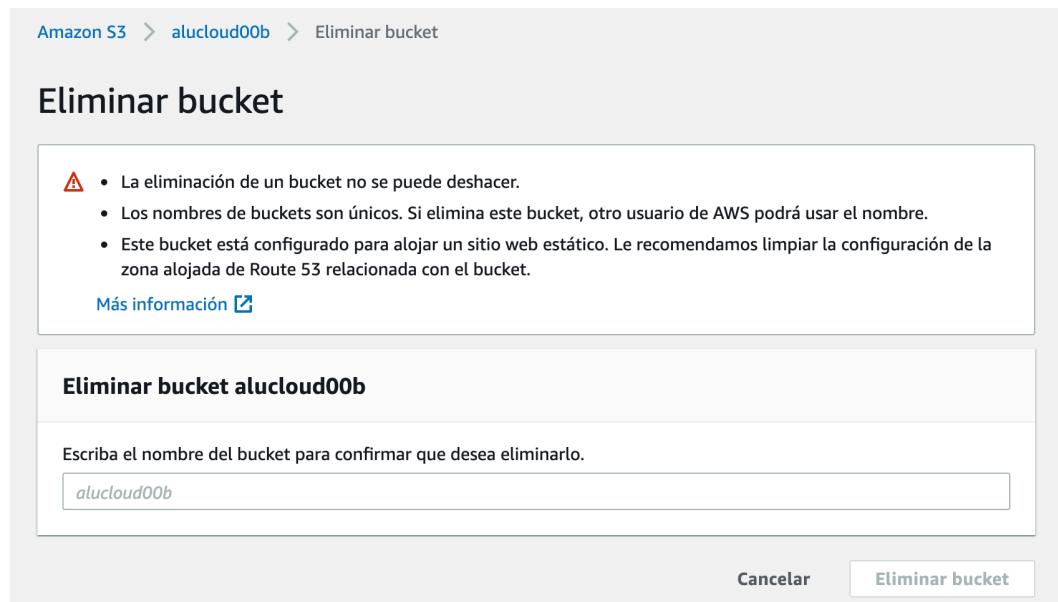
El bucket debe estar vacío antes de ser eliminado por lo que tendrás que pulsar sobre el enlace “configuración para vaciar el bucket”:

Teclea el nombre del bucket para proceder a eliminarlo y pulsa “Vaciar”:

Obtendrás una confirmación de la eliminación:



Pulsa ahora en el enlace “configuración de eliminación del bucket”, rellena el nombre del bucket y pulsa sobre “Eliminar bucket”:



Si obtienes algún error de permisos, asegúrate de que estás tratando de eliminar el bucket *aluclocloudXX* (donde XX es tu identificador de alumno) en vez del bucket *aluclocloud*, que no debes eliminar.

Amazon S3 facilita el almacenamiento de ficheros a muy bajo coste con alta disponibilidad, ofreciendo mecanismos para subir y descargar ficheros. Sin embargo, muchas aplicaciones requieren acceso a ficheros que estén directamente disponibles en la máquina en la que se estén ejecutando. Una alternativa es descargar los ficheros de S3 al sistema de almacenamiento de la máquina virtual antes de utilizarlos y volver a subirlos una vez finalizado, sin embargo, existe una forma más cómoda para gestionar este tipo de uso de ficheros: el servicio EBS.

4.3. Gestión de Volúmenes con EBS

El servicio EBS (Elastic Block Store) permite la creación dinámica de volúmenes que pueden ser conectados a instancias (máquinas virtuales en ejecución) de la misma manera que si un nuevo disco

se conectase a un equipo. Un volumen EBS solo puede ser conectado a una única instancia, de la misma manera que un disco duro solo puede estar conectado a un solo equipo.

En esta sección se procede a la creación de un volumen EBS para conectarlo de forma dinámica a una instancia en ejecución, y posteriormente formatearlo y montarlo en el sistema de archivos para usarlo como un dispositivo orientado a bloques más. Esto permite aumentar la capacidad de una instancia mediante el almacenamiento de ficheros en un volumen EBS (en lugar de en la partición raíz). También se puede utilizar para transferir datos que puedan estar en un volumen EBS a la instancia a la que se conecte.

Es posible listar los volúmenes EBS creados desde la consola de administración del servicio EC2 eligiendo la sección “Volumes” dentro de Elastic Block Store”.

Name	ID de volumen	Tipo	Tamaño	IOPS	Rendimiento	Instantánea	Creada	Zona de disponibilidad	Estado del volumen	Estado de la alarma
lamba-vpc-nat...	vol-0bd2a9f6e81b5f8a7	standard	8 GiB	-	-	snap-0b5e4e2...	2020/11/09 13:56 GMT+1	us-east-1b	En uso	Sin alarmas
-	vol-08470241bb71a56c3	gp2	30 GiB	100	-	snap-052a8f...	2021/08/18 09:01 GMT+2	us-east-1e	En uso	Sin alarmas
-	vol-07c3a397a76de8d90	gp2	10 GiB	100	-	snap-00b8679...	2023/05/04 10:24 GMT+2	us-east-1e	En uso	Sin alarmas
-	vol-09d4d9a0530945f	standard	8 GiB	-	-	snap-596eb6dc	2017/03/24 11:58 GMT+1	us-east-1a	En uso	Sin alarmas
-	vol-0da936629293b2359	gp2	20 GiB	100	-	snap-017bd4a...	2021/04/30 10:19 GMT+2	us-east-1a	En uso	Sin alarmas
-	vol-06cc7ae999a79d5e	gp2	10 GiB	100	-	snap-0b4e8fe...	2021/05/18 12:57 GMT+2	us-east-1a	En uso	Sin alarmas
-	vol-0842aa2d434aca636	gp2	8 GiB	100	-	snap-0cc1831...	2022/11/30 14:05 GMT+1	us-east-1a	En uso	Sin alarmas
-	vol-0171e54c0e91b7ec	gp2	30 GiB	100	-	snap-0c3ab6d...	2022/12/19 12:47 GMT+1	us-east-1a	En uso	Sin alarmas
-	vol-0a230b90801fc8ab8	gp2	10 GiB	100	-	snap-0973040...	2022/12/24 15:05 GMT+1	us-east-1a	En uso	Sin alarmas
-	vol-0952e289e0091dff	gp2	10 GiB	100	-	snap-0ef2202...	2023/01/08 17:12 GMT+1	us-east-1a	En uso	Sin alarmas
-	vol-0ab4ee611c6a5458	gp2	10 GiB	100	-	snap-0ef2202...	2023/01/10 22:13 GMT+1	us-east-1a	En uso	Sin alarmas
-	vol-0e08bdc1595171f9	gp2	8 GiB	100	-	snap-05b214c...	2023/02/10 14:37 GMT+1	us-east-1a	En uso	Sin alarmas
-	vol-0822c7ed2ca769736	gp2	30 GiB	100	-	snap-01beb31...	2023/05/04 14:46 GMT+2	us-east-1a	En uso	Sin alarmas
-	vol-0e7a87e37bb2e9ff8	gp2	8 GiB	100	-	snap-0da936d...	2023/05/04 18:10 GMT+2	us-east-1a	En uso	Sin alarmas

Vamos a construir un nuevo volumen de 10 GBytes vacío de tipo Magnetic (standard) en la zona de disponibilidad us-east-1a. Para ello, hay que indicar:

The screenshot shows the 'Create volume' wizard in the AWS Management Console. The steps are as follows:

- Step 1: Volume Type** (Completed): Shows 'Magnetic (standard)' selected.
- Step 2: Volume Size** (In Progress): Shows '10 GiB' selected.
- Step 3: Encryption** (In Progress): Shows the checkbox 'Encrypt this volume' is not checked.
- Step 4: Tags** (In Progress): Shows a tag 'Name: vol-00' added.
- Step 5: Review Volume** (In Progress): Shows the summary with the 'Create volume' button highlighted.

El tipo de volumen, que puede ser de tipo:

- *General Purpose SSD (GP2)*,
- *General Purpose SSD (GP3)*,
- *Provisioned IOPS SSD (IO1)* [35]
- *Provisioned IOPS SSD (IO2)*
- *Cold HDD (SC1)* [48]
- *Throughput Optimized HDD (ST1)* [48]
- *Magnetic (standard)*

Los primeros tipos proporcionan un rendimiento de E/S superior (latencias en el orden de pocos milisegundos) a los discos magnéticos al estar basados en discos duros de estado sólido (SSD – Solid State Drives). Además, la opción Provisioned IOPS es la que mejores prestaciones ofrece. Tienes información detallada sobre las características de cada tipo en [41]. Elegimos la opción más económica que es *Magnetic*. Ten en cuenta que es posible que no todas las zonas de disponibilidad soporten todos los tipos de volúmenes SSD (van añadiendo soporte poco a poco).

- El tamaño del volumen, que puede oscilar entre 1 GiB (tamaño mínimo) y 1 TiB (tamaño máximo). Elegimos 10 GiB. En realidad, existe una diferencia de tamaños entre GB y GiB [36], pero para nuestro propósito la obviaremos. Si especificas tamaños de disco superiores obtendrás un error de permisos.
- Snapshot ID (opcional), para construir un volumen EBS a partir de un snapshot (o instantánea) de otro volumen. Dejamos el campo en blanco para construir un volumen EBS vacío.
- Elegimos la zona de disponibilidad us-east-1a. Podríamos elegir cualquier otra zona de disponibilidad, pero el volumen EBS deberá estar en la misma zona de disponibilidad que la instancia a la que se trate de conectar.

- Añadimos la etiqueta Name / vol-XX, donde XX es tu identificador de alumno para encontrar fácilmente el volumen una vez creado desde la consola web de EC2.

Una vez creado el volumen obtendrás un mensaje con el identificador, que te recomiendo que apuntes para facilitar su búsqueda:

The screenshot shows a green success message at the top: "El volumen vol-042b2ccf242919ce4 se ha creado correctamente." Below it is the CloudFormation console interface with a search bar and buttons for actions and creating new stacks.

También aparecerá en el listado de volúmenes como disponible (columna Status con valor *available*). Como has especificado la etiqueta, es muy fácil saber cuál es el volumen EBS que acabas de crear. Si te olvidaste de etiquetar, puedes ordenar todos los volúmenes por fecha de creación en la consola de administración de volúmenes, elegir el último creado que esté disponible y, ahora sí, asignarle un nombre (poniéndote sobre el campo Name y pulsando sobre el lápiz).

The screenshot shows the AWS Volumes console with a single volume entry: vol-00, which was just created. It lists columns such as Name, ID de volumen, Tipo, Tamaño, IOPS, Rendimiento, Instantánea, Creada, Zona de disponibilidad, Estado del volumen, and Estado del volumen.

Un volumen EBS únicamente puede conectarse a una instancia de EC2 que haya sido desplegada en la misma zona de disponibilidad en la que ha sido creado el volumen. Por lo tanto, es muy importante que recuerdes la zona de disponibilidad en la que ha sido creado. En nuestro caso, podrás ver en la figura que el volumen ha sido creado en la zona de disponibilidad us-east-1a.

The screenshot shows the AWS Volume Details page for vol-264. It displays various volume properties like Volume ID, Size, Type, and Status. The status is shown as "Available" with "No alarms". The volume was created on "Fri Apr 08 2022 22:14:35 GMT+0200 (hora de verano de Europa central)" and is located in "us-east-1a". The volume status is "Okay" and it is "Not encrypted".

A continuación, lo conectamos dinámicamente a una máquina virtual que previamente desplegaremos en la misma zona de disponibilidad en la que fue creado el volumen EBS. Despliega una máquina virtual con las siguientes características (si no recuerdas el procedimiento, te recomiendo que revises la práctica de EC2; ten en cuenta que no es necesario que vuelvas a crear un par de claves ni un grupo de seguridad, pues puedes reaprovecharlos):

Nombre	alucloudXX-ebs (XX debe ser tu identificador de alumno)
AMI	LAMP Packaged by Bitnami
Region	Valor por defecto (us-east-1)
Tipo de instancia	t3.small
Par de claves	alucloudXX-keypair
Red	vpc-83a213fb default
Subred	La misma en la que se haya creado el volumen EBS. Por ejemplo, si está creado en us-east-1a, la subred debe ser: subnet-default-1a-public
Grupo de seguridad	Elige gs-aws-XX del desplegable (donde XX es tu identificador)

La AMI “LAMP packaged by Bitnami” (basada en EBS, con Debian, disponible para us-east-1 con acceso mediante SSH con el usuario *bitnami* [18]) la puedes buscar así:

Recuerda una última vez que es importante que la instancia se lance en la misma zona de disponibilidad dónde el volumen ha sido creado (en nuestro caso, us-east-1a.). De lo contrario, será imposible conectarlo a la instancia. Omitimos los pasos detallados de despliegue.

Una vez desplegada la instancia, si pulsas en su identificador irás directamente a la consola de administración de EC2 para ver el estado de esa instancia. Acuérdate de anotar el identificador de la instancia, ya que te hará falta más adelante.

A continuación, procederemos a conectar el volumen EBS a esa instancia. Desde el apartado de “Volúmenes” en la consola de administración de EC2, elige el volumen y con el botón derecho pulsa sobre “Asociar Volúmen”. Asegúrate de que estás cogiendo el volumen que acabas de crear (cuyo estado debe ser *available*) y no lo estás confundiendo con el volumen raíz de la instancia (cuyo estado será *in-use*).

Obtendrás un panel en el que podrás indicar a qué instancia debe conectarse el volumen EBS y en qué dispositivo se mapeará.

La elección de la ruta del dispositivo (en este caso `/dev/sdf`) responde a la sugerencia indicada en la documentación de EBS [19]. El único problema es que los kernels de Linux modernos pueden renombrar el dispositivo a `/dev/xvdf`. Además, en el caso de utilizar el hipervisor Nitro, como ocurren en los tipos de instancia modernos, el esquema de nombres cambia a `/dev/nvme0n1`, `/dev/nvme1n1`, etc. Ten en cuenta que hasta que la instancia no esté en estado *running* no será posible conectarle el volumen.

Nos conectamos a la instancia vía SSH (con usuario *bitnami*), tal y como se describió en la práctica básica de EC2. Para tu comodidad, repasamos la forma de hacerlo una última vez.

En primer lugar, averiguamos la IP pública de la instancia de EC2 filtrando por su nombre:

A continuación, una vez conectado a la instancia de prácticas (lab.cursocloudaws.net) ejecutas el siguiente comando (asegúrate de cambiar la IP por la correspondiente a tu instancia):

```
:~$ ssh -i alucloud$ID-priv.pem bitnami@44.198.58.76
```

En particular, asegúrate de que NO ejecutas los siguientes comandos en la máquina lab.cursocloudaws.net, sino en tu instancia EC2 que has desplegado previamente.

A continuación, verificamos los nombres de los dispositivos:

```
bitnami@ip-10-235-37-12:~$ df
Filesystem      1K-blocks    Used Available Use% Mounted on
udev              483588       0   483588  0% /dev
tmpfs             98484    2816   95668  3% /run
/dev/nvme0n1p1  10126600 3318476  6294988 35% /
tmpfs             492404       0   492404  0% /dev/shm
tmpfs               5120       0     5120  0% /run/lock
tmpfs             492404       0   492404  0% /sys/fs/cgroup
/dev/nvme0n1p15  126710     262   126448  1% /boot/efi
tmpfs             98480       0   98480  0% /run/user/1000

bitnami@ip-10-235-37-12:~$ lsblk
NAME      MAJ:MIN RM  SIZE RO TYPE MOUNTPOINT
nvme0n1    259:0    0   10G  0 disk
└─nvme0n1p1 259:1    0  9.9G  0 part /
└─nvme0n1p14 259:2    0    3M  0 part
└─nvme0n1p15 259:3    0 124M  0 part /boot/efi
nvme1n1    259:4    0   10G  0 disk
```

El volumen EBS ha sido conectado a /dev/nvme1n1. En el listado anterior verás que la partición raíz está montada en /, es decir, que se trata del volumen que contiene la partición de arranque, mientras que el disco /dev/nvme1n1 no está montado.

Ahora accedemos al disco con fdisk, tal y como se comenta a continuación. Para utilizar el disco para almacenar ficheros es necesario en primer lugar crear una partición y luego formatearla para que tenga un sistema de archivos. Posteriormente habrá que montar la partición en el sistema de archivos de la instancia. Procedamos paso a paso.

En primer lugar, creamos una partición de tipo Linux (código 83 en fdisk) en /dev/nvme1n1. Es posible que los valores sean diferentes a los mostrados aquí. Si no tienes claro qué valor(es)

especificar, pulsa directamente Enter porque fdisk utilizará valores por defecto que te permitirán crear una partición que se extienda a todo el tamaño del disco.

Antes de ejecutar el siguiente comando, asegúrate de que te has conectado previamente a la instancia de EC2 que tú has desplegado y que no estás ejecutando el comando en la máquina de prácticas (lab.cursocloudaws.net). De lo contrario, obtendrás el siguiente mensaje de error: Te pedirá la contraseña y, finalmente, te dirá “*“alucloudXX is not in the sudoers file. This incident will be reported.”*

```
bitnami@ip-10-235-37-12:~$ sudo fdisk /dev/nvme1n1
Welcome to fdisk (util-linux 2.27.1).
Changes will remain in memory only, until you decide to write them.
Be careful before using the write command.

Device does not contain a recognized partition table.
Created a new DOS disklabel with disk identifier 0xaec743ec.

Command (m for help): p
Disk /dev/nvme1n1: 10 GiB, 10737418240 bytes, 20971520 sectors
Disk model: Amazon Elastic Block Store
Units: sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
Disklabel type: dos
Disk identifier: 0x6f719e4b

Command (m for help): n
Command action
      p    primary (0 primary, 0 extended, 4 free)
      e    extended (container for logical partitions)
Select (default p): p
Partition number (1-4, default 1): 1
First sector (2048-20971519, default 2048): (Pulsa Enter)
Using default value 2048
Last sector, +sectors or +size{K,M,G} (2048-20971519, default 20971519): (Pulsa Enter)

Created a new partition 1 of type 'Linux' and of size 10 GiB.

Command (m for help): t
Selected partition 1
Partition type (type L to list all types): 83
Changed type of partition 'Linux' to 'Linux'.

Command (m for help): w
The partition table has been altered!
Calling ioctl() to re-read partition table.
Syncing disks.
```

A continuación, formateamos la partición con el sistema de archivos ext3 usando la herramienta mkfs.ext3

```
bitnami@ip-10-235-37-12:~$ sudo mkfs.ext3 /dev/nvme1n1p1
mke2fs 1.44.5 (15-Dec-2018)
```

```
Creating filesystem with 2621184 4k blocks and 655360 inodes
Filesystem UUID: 3aa4696f-401e-4156-b090-5dd5e31a693f
Superblock backups stored on blocks:
    32768, 98304, 163840, 229376, 294912, 819200, 884736, 1605632

Allocating group tables: done
Writing inode tables: done
Creating journal (16384 blocks): done
Writing superblocks and filesystem accounting information: done
```

Montamos el disco y ya está disponible para ser utilizado para introducir datos en él. Crearemos una carpeta y le daremos permisos para que el usuario bitnami pueda escribir.

```
bitnami@ip-10-235-37-12:~$ sudo mkdir /mnt/volebs
bitnami@ip-10-235-37-12:~$ sudo mount /dev/nvme0n1p1 /mnt/volebs
bitnami@ip-10-235-37-12:~$ df
Filesystem      1K-blocks   Used   Available  Use% Mounted on
udev             483588     0    483588   0% /dev
tmpfs            98484    2820    95664   3% /run
/dev/nvme0n1p1  10126600  3318492  6294972  35% /
tmpfs            492404     0    492404   0% /dev/shm
tmpfs            5120      0     5120   0% /run/lock
tmpfs            492404     0    492404   0% /sys/fs/cgroup
/dev/nvme0n1p15 126710     262    126448   1% /boot/efi
tmpfs            98480      0    98480   0% /run/user/1000
/dev/nvme0n1p1  10254648  23096   9707316   1% /mnt/volebs
bitnami@ip-10-235-37-12:~$ sudo mkdir /mnt/volebs/datos
bitnami@ip-10-235-37-12:~$ sudo chown bitnami:bitnami /mnt/volebs/datos
```

A partir de ese momento, se pueden guardar datos en dicho disco. Por ejemplo, creamos un fichero de texto dentro de ese volumen.

```
bitnami@ip-10-235-37-12:~$ echo "Hola" > /mnt/volebs/datos/hola.txt
```

Puedes verificar fácilmente que el fichero ha sido creado mostrando su contenido:

```
bitnami@ip-10-235-37-12:~$ cat /mnt/volebs/datos/hola.txt
Hola
```

¡Enhorabuena! Has creado un volumen EBS, lo has conectado a la instancia, le has dado formato y has almacenado algún fichero en él. Dicho volumen puede ser desconectado de la instancia y re-conectado a otra instancia de forma dinámica. Recuerda que un volumen solo puede estar conectado a una instancia. Esto permite que una aplicación que se ejecute en dicha instancia de EC2 pueda acceder a ficheros de forma local, aunque en realidad dichos ficheros están almacenados en un volumen EBS externo (conectado a la instancia) que almacena los datos de forma replicada (dentro de una misma zona de disponibilidad).

Gestión de Snapshots de Volúmenes EBS

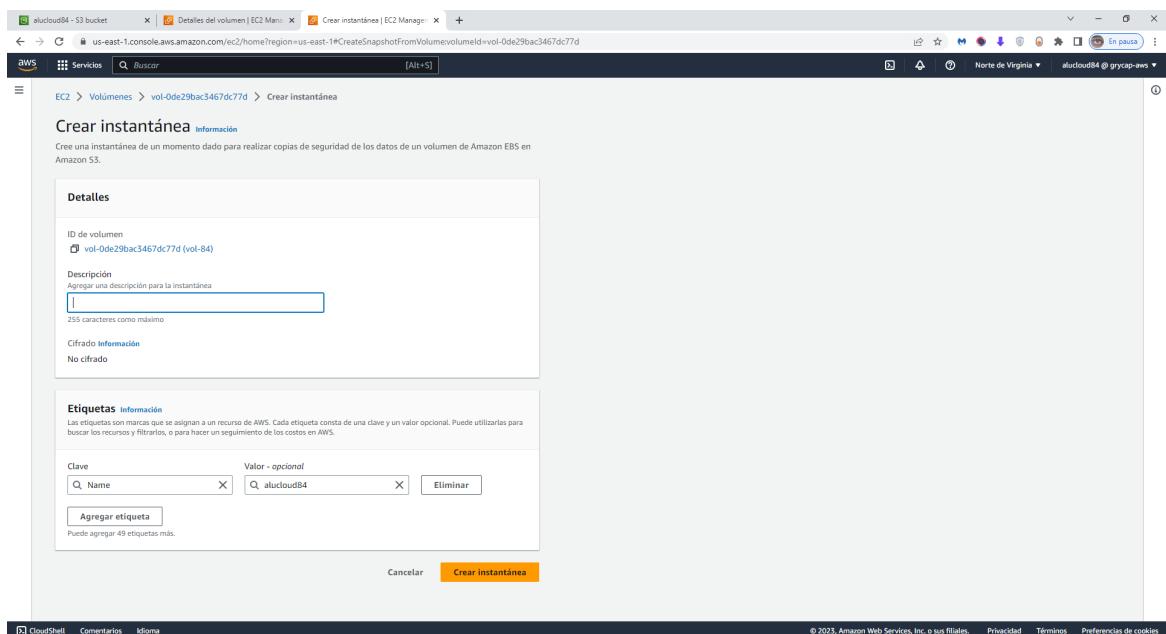
Podemos realizar una instantánea (*snapshot*) del volumen EBS para tener, por ejemplo, una copia de seguridad del volumen EBS (que se almacena en Amazon S3). Este es también el procedimiento tradicional utilizado para ampliar el tamaño de un volumen y/o moverlo a otra zona de disponibilidad (realizar un snapshot del volumen y luego crear un nuevo volumen de tamaño mayor, opcionalmente

en otra zona de disponibilidad). Para algunos tipos de instancia modernas también es posible redimensionar el tamaño del volumen en caliente (sin necesidad de hacer un snapshot). En ambos casos, hay que aprovechar el tamaño extra bien redimensionando la partición o creando una nueva partición en el espacio ampliado. Tienes más información al respecto en [59]. En esta práctica, utilizaremos la primera aproximación descrita.

Sin necesidad de desconectarlo de la instancia creamos la instantánea. Si no te has acordado de etiquetar el volumen y los listas, verás que te aparecen dos volúmenes EBS conectados a la misma instancia. Uno es el que contiene la partición de arranque del sistema, procedente de utilizar una AMI basada en EBS. Lo reconocerás porque está conectado en /dev/sda1. El otro, conectado a /dev/sdf es el volumen que tú has creado y vamos a redimensionar. Asegúrate de elegir el volumen EBS conectado a /dev/sdf antes de proseguir.

The screenshot shows the AWS Management Console with the URL us-east-1.console.aws.amazon.com/ec2/home?region=us-east-1#VolumeDetails?volumeId=vol-0de29bac3467dc77d. The left sidebar shows navigation for EC2, Volumes, and EBS. The main content displays the details for volume **vol-0de29bac3467dc77d (vol-84)**. The volume is 10 GiB, standard type, and currently in use. The Actions menu is open, and the 'Crear instantánea' (Create Snapshot) option is highlighted. Other options include Asociar volumen (Associate volume), Desasociar el volumen (Disassociate volume), and Desasociar el volumen forzadamente (Force-disassociate volume). The volume was created on May 06, 2023, at 06:39:36 GMT+0200 (hora de verano de Europa central).

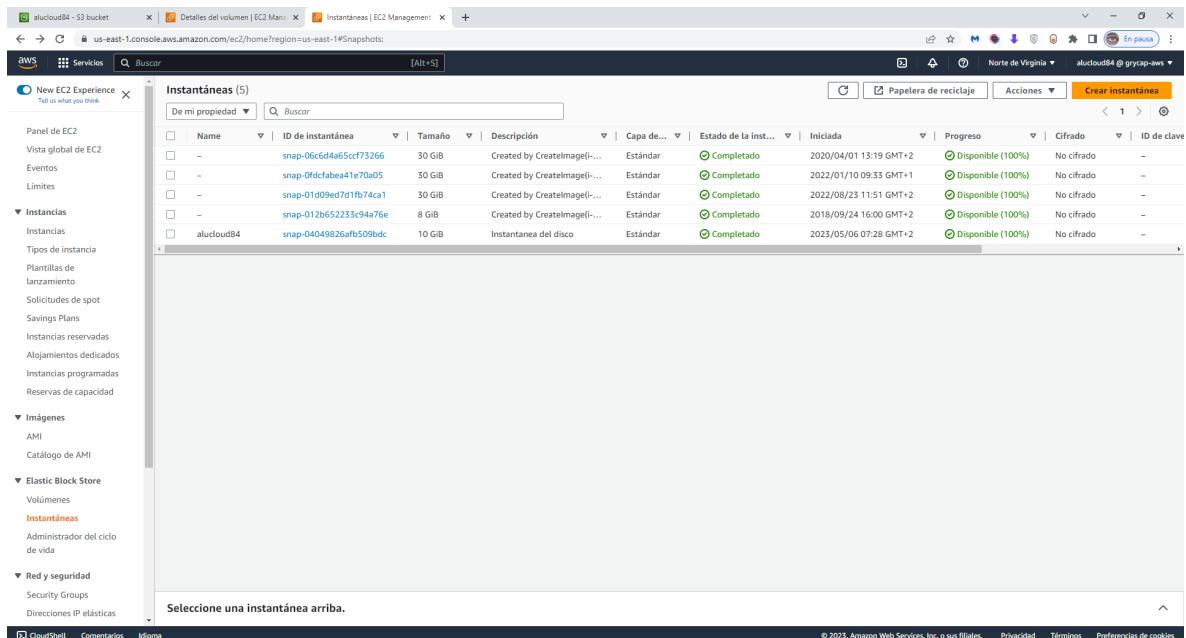
Para ello selecciona el volumen y con el botón derecho elige “Crear instantánea”.



Como siempre, te recomiendo añadir la etiqueta Name / alucloudXX, donde XX es tu identificador de usuario para facilitar la identificación del mismo.

Pulsa sobre el botón “Create Snapshot”.

Desde el panel de EC2 elige la opción “Snapshots” y verás las instantáneas creadas:



Recuerda que los *snapshots* (instantáneas) permiten realizar la copia incremental de un volumen EBS en S3. De esta manera, aunque su tamaño sea el mismo que el del volumen, en realidad tan solo se almacenan los bloques modificados dentro del volumen EBS desde el último snapshot. Es posible realizar múltiples snapshots, conforme va cambiando el contenido del volumen EBS para disponer así de un sistema efectivo de copia de seguridad incremental. En cualquier momento es posible borrar un

snapshot intermedio. Esto provocará la eliminación única y exclusivamente de aquellos datos que no sean necesarios para reconstruir cualquier otra instantánea. Por tanto, independientemente de la instantánea eliminada, cualquier instantánea activa dispone de la información necesaria para reconstruir el volumen.

Probamos a construir un nuevo volumen de tamaño más grande (quizá porque hubiéramos llegado al límite de datos de 10 GB) a partir de la instantánea. El volumen deberá ser creado en la misma zona de disponibilidad donde esté la instancia de EC2 en ejecución (en nuestro caso, us-east-1a) y tendrá 15 GB.

Selecciona el snapshot y con el botón derecho elige la opción “Crear volumen a partir de una instantánea”.

The screenshot shows the AWS EC2 Management console with the "Instantáneas" (Snapshots) page open. A specific snapshot, "alucloud84", is selected. A context menu is open over this snapshot, with the option "Crear volumen a partir de una instantánea" (Create volume from snapshot) highlighted. The main table lists five snapshots, including the selected one, with columns for Name, ID de instantánea, Tamaño (Size), Descripción (Description), Capa de... (Volume type), Estado de la inst... (State), Iniciada (Created), Progreso (Progress), Cifrado (Encrypted), and ID de clave (Volume ID). The selected snapshot has "alucloud84" in the Name column and "snap-04049826afb509bdc" in the ID de instantánea column. The progress bar indicates 100% completion.

Deberás crear un volumen EBS de tipo estándar de tamaño 15 GiB en la zona de disponibilidad us-east-1a (en general, que coincide con la zona de disponibilidad en la que desplegaste la instancia de EC2). Utiliza el tipo Magnetic y añade la etiqueta volXX-larger, para diferenciarlo del volumen anterior:

The screenshot shows the 'Create volumen' step in the AWS EBS wizard. It's a multi-step process with the current step being 'Configuración del volumen'. The configuration includes:

- ID de Instantánea:** snap-04049826af509bdc (alucloud84)
- Tipo de volumen:** Magnético (estándar)
- Tamaño (GiB):** 15 (Min: 1 GiB, máx.: 1024 GiB)
- IOPS:** No se aplica
- Rendimiento (MiB/s):** No se aplica
- Zona de disponibilidad:** us-east-1a
- Restauración rápida de instantáneas:** No habilitado para la instantánea seleccionada
- Cifrado:** Utilice el cifrado de Amazon EBS como una solución de cifrado para los recursos de EBS asociados a las instancias EC2.

Etiquetas - opcional: A section for optional tags, showing one entry: Name: vol84-larger. There is also an 'Agregar etiqueta' button.

At the bottom are 'Cancelar' and 'Crear volumen' buttons, with 'Crear volumen' being highlighted.

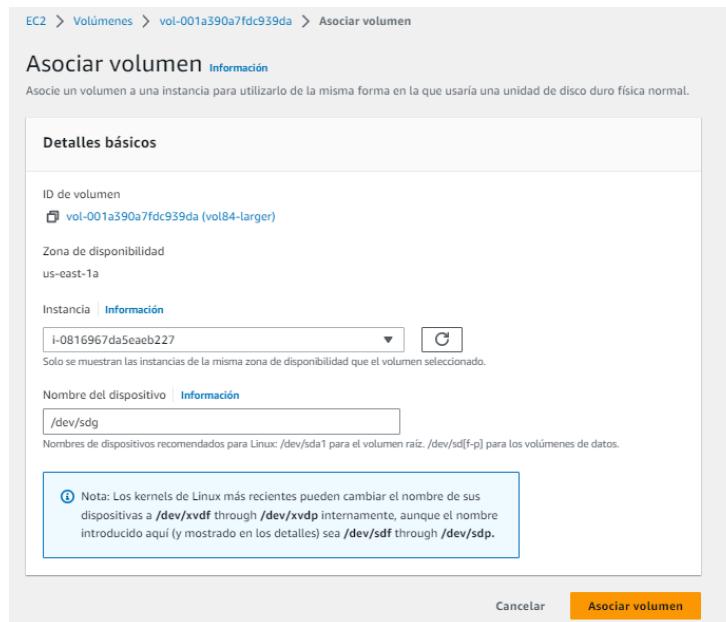
Pulsa sobre el botón “Crear volumen” para proceder a la creación del volumen EBS a partir del snapshot. Esto permitirá tener un duplicado del volumen EBS, pero de mayor tamaño. Obtendrás el identificador del volumen:

Snapshots > Create Volume

Create Volume



Conectamos el nuevo volumen a la instancia para verificar su contenido. Puedes reutilizar la instancia que ya estaba en ejecución, pero en ese caso deberás utilizar otro dispositivo diferente al /dev/sdf, que ya estaba siendo utilizado (y que además desde la instancia se renombraba a /dev/xvdf). En nuestro caso, utilizaremos la ruta de dispositivo /dev/sdg.



Si obtienes el mensaje de error “Error attaching volume: Cannot attach volumen ‘vol-XXXXXX’ with Marketplace codes as the instance ‘i-XXXXXXX’ is not in the stopped state”, es que hiciste el snapshot del volumen raíz de la instancia, en lugar del volumen adicional que añadiste, por lo que tendrás que hacerlo nuevamente de forma correcta.

Desde dentro de la instancia hay que montar la partición del volumen al sistema de archivos antes de verificar que, efectivamente, tenemos una copia de los datos del volumen EBS original.

```
bitnami@ip-10-235-37-12:~$ lsblk
nvme0n1      259:0    0   10G  0 disk
└─nvme0n1p1  259:1    0   9.9G 0 part /
  └─nvme0n1p14 259:2    0   3M  0 part
    └─nvme0n1p15 259:3    0  124M 0 part /boot/efi
nvme1n1      259:4    0   10G  0 disk
└─nvme1n1p1  259:5    0   10G 0 part /mnt/volebs
nvme2n1      259:6    0   15G  0 disk
└─nvme2n1p1  259:7    0   10G 0 part
bitnami@ip-10-235-37-12:~$ sudo mkdir /mnt/otrovolebs
bitnami@ip-10-235-37-12:~$ sudo mount /dev/nvme2n1p1 /mnt/otrovolebs
bitnami@ip-10-235-37-12:~$ ls /mnt/otrovolebs
datos  lost+found
bitnami@ip-10-235-37-12:~$ ls /mnt/otrovolebs/datos
holo.txt
bitnami@ip-10-235-37-12:~$ df
Filesystem      1K-blocks    Used Available Use% Mounted on
udev             483588       0   483588   0% /dev
tmpfs            98484    2832    95652   3% /run
/dev/nvme0n1p1  10126600 3318500  6294964  35% /
tmpfs            492404       0   492404   0% /dev/shm
tmpfs             5120       0     5120   0% /run/lock
tmpfs            492404       0   492404   0% /sys/fs/cgroup
/dev/nvme0n1p15  126710     262   126448   1% /boot/efi
tmpfs            98480       0   98480   0% /run/user/1000
/dev/nvme1n1p1  10254648 23104   9707308   1% /mnt/volebs
/dev/nvme2n1p1  10254648 23104   9707308   1% /mnt/otrovolebs
```

Si observamos el tamaño de la partición montada veremos que sigue siendo de 10 GB en lugar de 15 GB. En efecto, el nuevo volumen contiene la partición de 10 GB que había en el volumen EBS original y dispone de 5 GB en espacio por particionar (por ello los 5 GB no se ven con el comando `df`). Sería posible usar `fdisk` para reparticionar ese trozo del volumen o seguir el procedimiento indicado en [22] para redimensionar la partición ya existente. Esto implica utilizar el comando `resize2fs` para redimensionar el sistema de archivos al nuevo tamaño de la partición [14]. Tienes más información sobre cómo redimensionar una partición creada en un volumen EBS en [47].

Una vez finalizado el proceso, se deberá proceder a desmontar los volúmenes del sistema de archivos antes de desconectar el volumen de la instancia (al igual que se desmonta un disco USB antes de desconectarlo del equipo). Aunque el volumen esté montado en el sistema de archivos será posible desconectar el volumen de la instancia pero no se recomienda para evitar posibles pérdidas de datos.

```
bitnami@ip-10-235-37-12:~$ sudo umount /mnt/volebs
bitnami@ip-10-235-37-12:~$ sudo umount /mnt/otrovolebs
```

Para eliminar un volumen, es necesario desconectarlo de la instancia y posteriormente proceder a la eliminación. Para ello, desde la AWS Management Console deberás desconectar los volúmenes. Asegúrate de que eliges el volumen que está conectado a `/dev/sdf` (llamado `vol-XX`) y el volumen que está conectado a `/dev/sdg` (llamado `volXX-larger`). Por el contrario, el volumen conectado a `/dev/sda1` es el que contiene la partición raíz y no debe ser desconectado (en realidad no se te permitirá la desconexión).

En este punto puedes proceder a eliminar la instancia, los volúmenes y las instantáneas creadas. No te olvides de terminar las instancias para no generar cargos innecesarios.

Comenzamos desasociando los volúmenes y luego borrándolos (selecciona los dos y elimínalos en un solo paso):

The screenshot shows the AWS Management Console with the URL <https://us-east-1.console.aws.amazon.com/ec2/home?region=us-east-1#Volumesv3=search=vol-001a390a7fd939da;vol-84>. The left sidebar is collapsed. The main content area shows a table of volumes with columns: Name, ID del volumen, Tipo, Tamaño, IOPS, Rendimiento, Instantánea, Creada, Zona de disponibilidad, and Estado del volumen. Two volumes are selected: 'vol-84' and 'vol84-larger'. A modal window titled 'El volumen se ha desasociado correctamente.' is displayed. On the right, there's a context menu with options like 'Modificar volumen', 'Crear instantánea', 'Crear política de ciclo de vida de instantáneas', 'Eliminar volumen', etc. The status bar at the bottom indicates '© 2023, Amazon Web Services, Inc. o sus filiales. Privacidad Términos Preferencias de cookies'.

A continuación, terminamos la instancia:

The screenshot shows the AWS Management Console with the URL <https://us-east-1.console.aws.amazon.com/ec2/home?region=us-east-1#Instancesv3=search=84;instanceId=i-0816967da5eaeab227>. The left sidebar is collapsed. The main content area shows a table of instances with columns: Name, ID de la instancia, Estado de la instancia, Tipo de instancia, Comprobación de estado, Estado de la red, Zona de disponibilidad, DNS de IPv4 pública, Dirección IP, and Dirección IP privada. One instance is selected: 'alucloud84-ebs' (i-0816967da5eaeab227). A context menu on the right shows options like 'Detener instancia', 'Iniciar instancia', 'Reiniciar instancia', 'Hibernar instancia', and 'Terminar Instancia'. The status bar at the bottom indicates '© 2023, Amazon Web Services, Inc. o sus filiales. Privacidad Términos Preferencias de cookies'.

Finalmente, eliminamos el snapshot:

The screenshot shows the AWS EC2 Management Console interface. On the left, there's a sidebar with various navigation options like 'Panel de EC2', 'Instancias', 'Tipos de Instancia', 'Plantas de lanzamiento', 'Solicitudes de spot', 'Savings Plans', 'Instancias reservadas', 'Alojamientos dedicados', 'Instancias programadas', 'Reservas de capacidad', 'Imagenes', 'AMI', 'Catálogo de AMI', 'Elastic Block Store', 'Volumenes', and 'Instantáneas'. The 'Instantáneas' section is currently selected. The main area displays a table titled 'Instantáneas (1/5)' with columns: Name, ID de Instantánea, Tamaño, Descripción, Estado de la Inst., Iniciada, Progreso, and Cifrado. The last row, 'alucoud84', is selected. A detailed view of this snapshot is shown in a modal window.

Name	ID de Instantánea	Tamaño	Descripción	Estado de la Inst..	Iniciada	Progreso	Cifrado
snap-066d4a65ccf73266	snap-066d4a65ccf73266	30 GB	Created by CreateImage(...)	Estandar	2020/04/01 13:19 GMT+2	Disponible (100%)	No cifrada
snap-0fcdabaa41e70d05	snap-0fcdabaa41e70d05	30 GB	Created by CreateImage(...)	Estandar	2022/01/10 09:33 GMT+1	Disponible (100%)	No cifrada
snap-01d10edf7d1f74a1	snap-01d10edf7d1f74a1	30 GB	Created by CreateImage(...)	Estandar	2022/08/23 11:51 GMT+2	Disponible (100%)	No cifrada
snap-012b652233c94a76e	snap-012b652233c94a76e	8 GB	Created by CreateImage(...)	Estandar	2018/09/24 16:00 GMT+2	Disponible (100%)	No cifrada
alucoud84	snap-04049826afb509bdc	10 GB	Instantánea del disco	Estandar	2023/05/06 07:28 GMT+2	Disponible (100%)	No cifrada

ID de instantánea: snap-04049826afb509bdc (alucoud84)

Detalles | Permisos | Capa de almacenamiento | Etiquetas

ID de instantánea: snap-04049826afb509bdc (alucoud84)

Propietario: 97434905189

Cifrado: No cifrado

Restauración rápida de Instantáneas: Instantánea del disco

Tamaño: 10 GB

ID de volumen: vol-0db29baac3467dc77d

ID de clave de KMS: -

Progreso: Disponible (100%)

Estado de la instantánea: Completado

Última restauración: Sat May 06 2023 07:28:27 GMT+0200 (hora de verano de Europa central)

Códigos de productos: -

Alias de clave de KMS: -

ARN de clave de KMS: -



Asegúrate de que has eliminado la instancia de EC2, el volumen EBS y la instantánea antes de pasar a la siguiente sección.

Recuerda: Estás utilizando un proveedor de Cloud público y los recursos utilizados se facturan directamente a la tarjeta VISA del instructor. Si no vas a utilizar un recurso durante un tiempo, deténlo. Si no lo vas a utilizar más, terminalo.

4.4. Creación y Registro de AMIs

Hasta el momento, tan solo se han utilizado AMIs predefinidas, ya creadas por terceros, con una determinada configuración software. A menudo es necesario definir AMIs específicas con la configuración software necesaria para un determinado proyecto. Por ejemplo, imagínate que quieres desarrollar un servicio Web en Java para el tratamiento de imágenes. Para ello, sería necesario contar con una AMI que tuviera Java, un servidor de aplicaciones como Apache Tomcat y alguna librería gráfica o, por qué no, una herramienta como *ImageMagick* para manipular las imágenes.

El procedimiento más sencillo de creación de imágenes consiste en partir de una AMI que disponga de la configuración software lo más parecida posible a la que te interesa. Se debe crear una instancia a partir de la AMI sobre la que se instalará el resto de software y, finalmente, se duplicará dicha AMI y se registrará en el catálogo de AMIs de Amazon (AWS Marketplace). Esto permitirá reutilizarla en diferentes ocasiones (incluso hacerla pública para que otros usuarios puedan utilizarla, si se considera oportuno).

Otra alternativa de trabajo radica en desplegar instancias base, sin ninguna configuración especial y posteriormente desplegar dinámicamente las dependencias software necesarias para que dicha instancia asuma un determinado rol dentro de la aplicación (servidor web, servidor de base de datos,

servidor de aplicaciones, etc.). Es posible desplegar/installar software y configuraciones particulares usando herramientas tipo Ansible [23], Puppet [11], Chef [12] o AWS OpsWorks [13].

En esta parte de la práctica partiremos de una AMI de 64 bits basada en EBS con Ubuntu 22.04 (accesible vía SSH con el usuario *ubuntu*, no con el usuario *bitnami*). Estas AMIs pueden obtenerse del listado de AMIs que Ubuntu ofrece para ser desplegadas en Amazon EC2 [8] aunque concretamente ésta ha sido encontrada en el AWS MarketPlace buscando por Ubuntu 22.04 y filtrando en la barra lateral por editor “Canonical Group Limited”.

Asegúrate de que eliges la AMI ofrecida por “Canonical Group Limited”. Hay otras opciones que son AMIs de pago, porque ofrecen algún tipo de soporte por parte de una empresa externa. No tienes permisos para suscribirte y utilizar dichas AMIs.

Despliega una instancia (de tipo t3.micro, par de claves alucloud\$ID-keypair, en cualquier subred del VPC “default”, grupo de seguridad gs-aws-\$ID y con etiqueta Name / alucloudXX), y luego la configuraremos para instalarle la herramienta *ImageMagick* de manipulación de imágenes en Linux. En este punto de la práctica ya deberías saber cómo desplegar la instancia y conectarte vía SSH a la nueva instancia, por lo que omitiremos una descripción detallada. Si tienes algún problema en conectarte vía SSH a la instancia recuerda que debes conectarte con el usuario *ubuntu* (no con usuario *bitnami* ni con usuario *root*). Si necesitas un procedimiento detallado para realizar este paso, tienes a tu disposición el boletín de la práctica anterior.

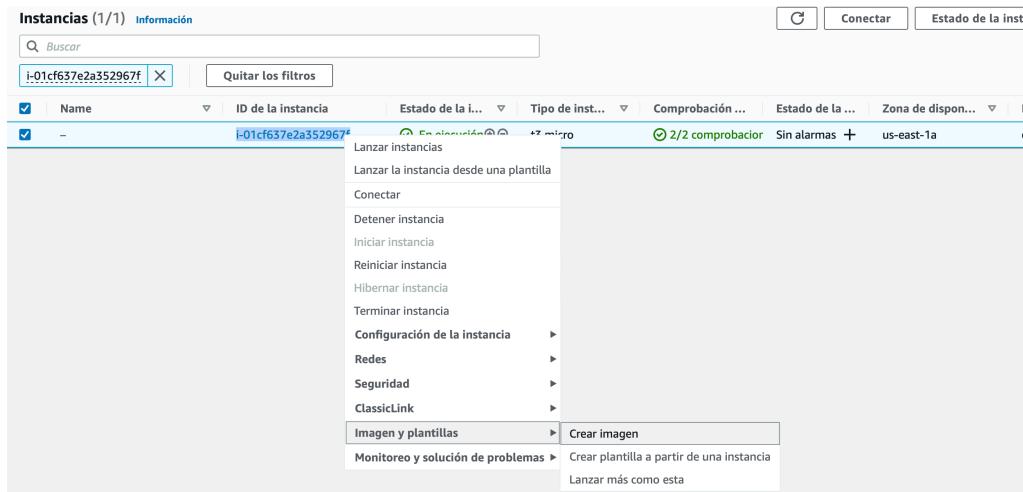
Una vez iniciada sesión por SSH con la nueva instancia (recuerda que para conectarte debes averiguar la IP pública y conectarte mediante el usuario *ubuntu*), instalamos el paquete *ImageMagick* a partir de los repositorios de software de Ubuntu.

```
ubuntu@vm:~$ sudo apt-get update && sudo apt-get install imagemagick
```

(Requiere descargarse unos 84.8 MB y la instalación no debería tardar más de 2-3 minutos)

Podríamos instalar el resto de paquetes necesarios (Apache Tomcat, etc.) usando el comando *apt-get*, aunque no lo haremos en esta práctica por simplicidad. A continuación, salimos de la instancia y creamos una nueva AMI a partir de la instancia modificada. Por defecto, Amazon EC2 detiene el

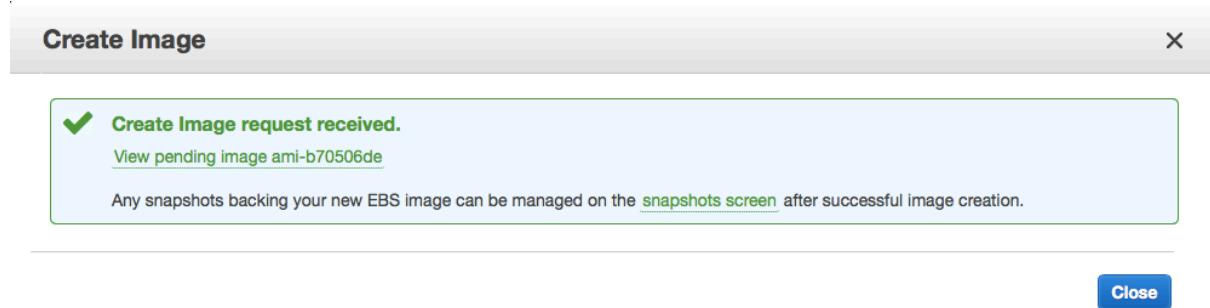
sistema operativo de la instancia, realiza un snapshot de los volúmenes conectados, crea y registra el AMI y luego reinicia el sistema operativo de la instancia, todo ello sin detener la instancia [25]. Esta operación puede realizarse en caliente (con la opción --no-reboot), aunque Amazon dice que en ese caso no hay garantía de que la nueva imagen tenga un sistema de archivos consistente [24].



Selecciona la instancia y con el botón derecho (o desde el botón de menú Actions) elige la opción “Imagen y plantillas → Crear imagen”.

Deberás indicar como mínimo el nombre de la imagen (amicursocloudaws\$ID). Opcionalmente es posible definir otros volúmenes para el AMI que se va a crear referenciando a snapshots existentes.

Esto permite al desplegar una nueva instancia de dicha AMI tener disponibles diferentes volúmenes EBS ya conectados a la instancia. Sin introducir ningún cambio, pulsa sobre “Create Image”.



Puedes obtener un listado de todas las AMIs creadas por los alumnos desde la opción AMIs de la consola de administración de EC2:

Imágenes de Amazon Machine Image (AMI) (5) Información									
De mi propiedad	Buscar	EC2 Image Builder	Acciones ▾	Lanzar instancia a partir de una imagen					
				Nombre de AMI	Origen	Propietario	Visibilidad	Estado	Fecha de creación
<input type="checkbox"/>	-	ami-028bb0f46f4e2925e	amicursocloudaws114	974349055189/amicursocloudaws114	974349055189	974349055189	Privado	Disponible	2021/12/20 16:12
<input type="checkbox"/>	-	ami-005fc91858eaaf69af	amicursocloudaws162	974349055189/amicursocloudaws162	974349055189	974349055189	Privado	Disponible	2021/12/13 13:12
<input type="checkbox"/>	-	ami-0264970bceaa1042a	cursocloudaws2021	974349055189/cursocloudaws2021	974349055189	974349055189	Privado	Disponible	2020/04/01 13:12
<input type="checkbox"/>	-	ami-05428018d638db56e	cloutrail-tracker-upload-machi...	974349055189/cloutrail-tracker-uplo...	974349055189	974349055189	Privado	Disponible	2018/09/24 16:12
<input type="checkbox"/>	-	ami-068cd861f936e9053	AMIWordPress103	974349055189/AMIWordPress103	974349055189	974349055189	Privado	Disponible	2021/12/14 12:12

Por defecto, las AMIs creadas son privadas y únicamente utilizables por su creador. No obstante, es posible hacer una AMI pública y ponerla a disposición de la comunidad de usuarios de AWS. También es posible crear una AMI con un determinado software y cobrar por su uso haciéndola disponible en el AWS Marketplace. Son las llamadas Paid AMIs [26]. Si despliegas una instancia de la AMI previamente creada, observarás que incluye el software que previamente instalaste.

Recuerda que las AMIs van ligadas a regiones concretas. Por lo tanto, si deseas utilizarla en otra región diferente, tendrás que copiarla en primer lugar. Es muy sencillo copiar una AMI a otra región usando la AWS Console, tal y como se muestra en la siguiente figura:

The screenshot shows the 'Copiar AMI' (Copy AMI) wizard. The 'Copiar Amazon Machine Image (AMI)' section includes:

- ID de AMI original:** ami-0cb1b27473454478b
- Nombre de la copia de AMI:** cursocloudaws2122
- Descripción de la copia de AMI:** [Copied ami-0cb1b27473454478b from us-east-1] cursocloudaws2122
- Región de destino:** Norte de Virginia (EE.UU. Este)
- Opciones:**
 - Copiar etiquetas
 - Cifrar instantáneas de EBS de copia de AMI
- Etiquetas:** *opcional* (radio button selected: Etiquetar imágenes e instantáneas juntas)
- No hay etiquetas asociadas a este recurso.**
- Botones:** Cancelar, Copiar AMI

Recuerda que únicamente tienes permiso para utilizar la región us-east-1 (N. Virginia) por lo que obtendrás un error si tratas de copiar la AMI a otra región diferente. Por tanto, puedes omitir este paso.

Ten en cuenta que, si creas una AMI a partir de una instancia que tenga conectados diferentes volúmenes EBS, se realizarán también snapshots de los volúmenes EBS [24].

Una vez finalizadas las pruebas con la nueva AMI creada la de-registraremos para eliminarla (apúntate el identificador de la AMI porque te hará falta):

Imágenes de Amazon Machine Image (AMI) (1/5) Información							Lanzar instancia a partir de una imagen	
De mi propiedad		Buscar		Acciones				
Name	ID de AMI	Nombre de AMI	Origen	Propietario	Visibilidad	Estado	Fecha de creación	
<input checked="" type="checkbox"/>	ami-028bb0f46f4e2923e	Lanzar instancia a partir de una imagen	189/amicursocloudaws114	974349055189	Privado	Disponible	2021/12/20 16:1	
<input type="checkbox"/>	ami-005fc91858eaa69af	Copiar AMI	189/amicursocloudaws162	974349055189	Privado	Disponible	2021/12/13 13:5	
<input type="checkbox"/>	ami-0264970bceaa1042a	Editar permisos de AMI	189/cursocloudaws2021	974349055189	Privado	Disponible	2020/04/01 13:1	
<input type="checkbox"/>	ami-05428d18d638db96e	Solicitar instancias de spot	189/cloudtrail-tracker-uplo...	974349055189	Privado	Disponible	2018/09/24 16:0	
<input type="checkbox"/>	ami-068cd861f936e9053	Administrar etiquetas	189/AMIPressT03	974349055189	Privado	Disponible	2021/12/14 12:4	
Anular registro de la AMI Cambiar descripción Registrar AMI con respaldo en el almacenamiento de la instancia								

También eliminamos la instancia a partir de la cual se creó esta nueva AMI.

<input checked="" type="checkbox"/>	–	i-01cf637e2a352967f	En ejecución	t3.micro	2/2 comprobador	Sin alarmas +	us-east-1a
<input type="checkbox"/>	vpc-nat00	i-066fd20c61cc095ba	Det	Lanzar instancias	.	Sin alarmas +	us-east-1a
<input type="checkbox"/>	rCUDA-server-nat	i-0a14070c82b5e63ba	Det	Lanzar la instancia desde una plantilla	.	Sin alarmas +	us-east-1a
<input type="checkbox"/>	test-oscar-k3s-srisco	i-045c353c9854ec4e6	Det	Conectar	.	Sin alarmas +	us-east-1a
<input type="checkbox"/>	cursocloudaws2021-lab	i-0b9628714385d9f4a	Det	Detener instancia	.	2/2 no ti +	us-east-1e
<input type="checkbox"/>	lambda-vpc-nat-instance	i-0efd8d0fe1e8da11d	Det	Iniciar instancia	.	Sin alarmas +	us-east-1b
<input type="checkbox"/>	rCUDA-server	i-08e434c6516a4a487	Det	Reiniciar instancia	.	Sin alarmas +	us-east-1e
<input type="checkbox"/>	circutor-alb-waf-test	i-02b8f5e4d3c00ba7c	Det	Hibernar instancia	.	Sin alarmas +	us-east-1d
<input type="checkbox"/>	–	i-09b70a99bae67192e	Det	Terminar instancia	.	Sin alarmas +	us-east-1e
Configuración de la instancia							

Finalmente, procederemos a eliminar el snapshot que se creó automáticamente al crear una AMI (puedes filtrar por el identificador de AMI) :

Instantáneas (1/5) Información										
	Name	ID de instantánea	Tamaño de...	Descripción	Capa de a...	Estado de la inst...	Iniciada	Progreso	Cifrado	Acciones
<input checked="" type="checkbox"/>	-	snap-06cd4a65ccf73266	30 GiB	Created by CreateImage(...)	Estándar	Completado	2020/04/01 13:19 GMT+2	Disponible (100%)	No cifrado	Detalles
<input type="checkbox"/>	-	snap-0fdcfabea41e70a05	30 GiB	Created by CreateImage(...)	Estándar	Completado	2022/01/10 09:33 GMT+1	Disponible (100%)	No cifrado	Detalles
<input type="checkbox"/>	-	snap-01d09ed7d1fb74ca1	30 GiB	Created by CreateImage(...)	Estándar	Completado	2022/08/23 11:51 GMT+2	Disponible (100%)	No cifrado	Detalles
<input type="checkbox"/>	-	snap-0e4beaa1ac9312fd6	30 GiB	Created by CreateImage(...)	Estándar	Completado	2024/01/10 14:22 GMT+1	Disponible (100%)	No cifrado	Detalles
<input type="checkbox"/>	-	snap-012b652233c94a76e	8 GiB	Created by CreateImage(...)	Estándar	Completado	2018/09/24 16:00 GMT+2	Disponible (100%)	No cifrado	Detalles

Enhorabuena. En esta sección has aprendido como crear tus propias AMIs a partir de instancias de AMIs ya existentes en el AWS MarketPlace. Esto te facilitará la creación de imágenes pre-configuradas para soportar determinados roles dentro de la arquitectura de una aplicación en la nube. El coste de las AMIs de tipo EBS depende del coste de almacenamiento del correspondiente snapshot de volumen EBS mientras que el coste de una AMI de tipo instance-store depende del coste de almacenamiento en S3 [73].



Asegúrate de que has eliminado la instancia de EC2 antes de pasar a la siguiente sección. Ten en cuenta que la siguiente parte de la práctica te llevará algún tiempo realizarla por lo que te recomiendo iniciarla cuando tengas al menos 30-60 minutos por delante, ya que deberás realizarla completamente pues, de lo contrario, algunos recursos se eliminarán transcurrido un tiempo (para evitar un consumo excesivo de recursos).

5. Balanceo de la Carga

Si un sitio web se ejecuta en una única instancia (máquina virtual) y dicho sitio se vuelve tan popular que comienza a atraer un elevado número de visitantes, es posible que la instancia (su servidor web) no sea capaz de atender todas las peticiones y se produzca algún fallo de ejecución o simplemente una denegación de servicio. En esta parte de la práctica se utiliza el servicio Elastic Load Balancing (ELB) [27] para distribuir automáticamente el tráfico entrante a diferentes instancias de EC2. Esto permite que múltiples instancias EC2 puedan atender las peticiones de visita y, de esta manera, tener una mayor garantía de disponibilidad del servicio en línea. ELB monitoriza las instancias con problemas para dirigir el tráfico sólo a las instancias sanas. La capacidad de ELB aumenta automáticamente para adaptarse a incrementos del tráfico entrante. Pero ELB únicamente utiliza una estrategia de tipo round-robin para repartir las peticiones entre instancias situadas en diferentes zonas de disponibilidad de una misma región.

Para ello en esta parte de la práctica realizaremos las siguientes acciones:

1. Subiremos una imagen (fotografía) a S3.
2. Arrancaremos una instancia a partir de una imagen que tenga un servidor web Apache (o un servidor de aplicaciones como Apache Tomcat) y modificaremos una página HTML para que incluya dicha imagen desde S3. Forzaremos su despliegue en una determinada zona de disponibilidad.
3. Arrancaremos otra instancia adicional en otra zona de disponibilidad de la misma región.
4. Crearemos un balanceador de carga con el servicio ELB.
5. Comprobaremos cómo el tráfico web del cliente se dirige a las diferentes instancias para mostrar el contenido de la web.

1. Subida de un fichero a tu bucket de S3

Elije un fichero de imagen que te guste y súbelo a tu carpeta correspondiente dentro del bucket ‘alucloud’ de S3. Puedes buscar una imagen en Google Images y descargarla. Alternativamente, en el directorio /opt/cursoaws/cloud.jpg del entorno de prácticas tienes una imagen que puedes copiar. El fichero deberá ser público para ser descargado desde fuera de AWS.

A partir de ese momento, la URL para acceder a la imagen (para el caso del usuario alucloudoo) es:
<https://s3.amazonaws.com/alucloud/oo/cloud.jpg>

Ten en cuenta que si quieres acceder a esta imagen deberás hacerla pública previamente (seleccionando el fichero y en Acciones → Hacer público a través de ACL)

2. Arrancar la instancia y modificar página web para referenciar dicha imagen.

Estamos asumiendo el uso de una AMI de Bitnami basada con LAMP (Linux, Apache MySQL y PHP), disponible en la región us-east-1 y accesible con el usuario bitnami. Puedes buscar dicha AMI desde el panel de lanzamiento de instancias buscando en el apartado “AWS Marketplace” por “bitnami lamp”.

Arrancaremos una instancia de tamaño t2.micro en la zona de disponibilidad us-east-1e (para ello tendrás que elegir la subred pública correspondiente del VPC). Si dicha zona no está disponible puedes utilizar cualquier otra zona de disponibilidad (us-east-1c, us-east-1d, us-east-1a). Recuerda indicar tu par de claves y el grupo de seguridad creado para tu usuario.

Anota el identificador de la instancia (y su zona de disponibilidad) porque te hará falta más adelante.

Posteriormente, desde dentro de la instancia crearemos una página HTML para que la sirva el servidor web Apache (`/opt/bitnami/apache2/htdocs/cloud.html`). Para ello utiliza tu editor preferido de línea

de comandos en GNU/Linux. Por ejemplo, puedes utilizar *vim* (recuerda cambiar la IP por la de tu instancia de EC2:

```
:~$ ssh -i alucloud$ID-priv.pem bitnami@107.22.37.194
bitnami@ip-10-235-37-12:~$ sudo vim /opt/bitnami/apache2/htdocs/cloud.html
<html>
<h1> Hola desde el Cloud! </h1>

</html>
```

Si nunca has utilizado *vim*, tendrás que utilizar los siguientes comandos, una vez hayas ejecutado vim para editar el nuevo fichero /opt/bitnami/apache2/htdocs/cloud.html. Pulsar *i* para pasar al modo edición, donde podrás pegar el código HTML del cuadro anterior (desde <html> hasta </html> incluído). Si quieres puedes cambiar la URL de la imagen a la que se referencia. Luego deberás guardar y salir de vim pulsando la siguiente secuencia de teclas: *Esc*, luego *:w* y luego *:q*

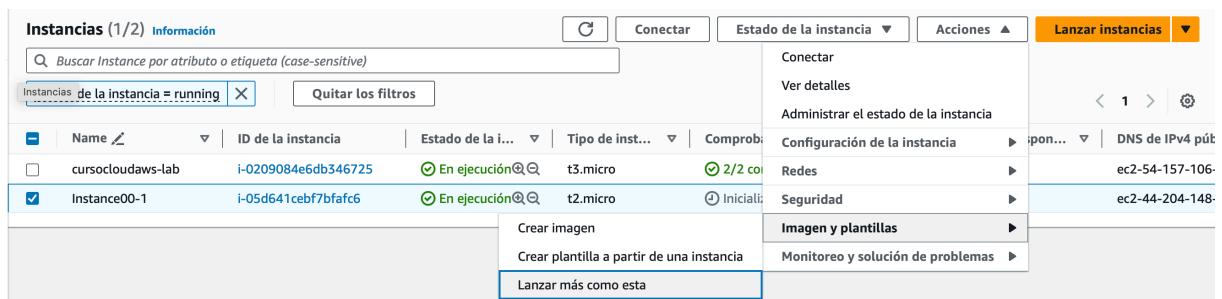
Es posible verificar que se está accediendo a la imagen almacenada en S3 usando un navegador y accediendo a <http://ec2-107-22-37-194.compute-1.amazonaws.com/cloud.html>

Si obtienes un mensaje de error indicando que la conexión ha sido rehusada, asegúrate de que no estás indicando como protocolo https en lugar de http. Recuerda que https se conectaría al puerto 443 de la instancia mientras que http se conecta al puerto 80 de la instancia.

3. Arrancar otra instancia en otra zona de disponibilidad de la misma región.

Desplegamos otra instancia de las mismas características en otra zona de disponibilidad de la misma región.

Una forma rápida es usar la opción “Acciones → Imagen y plantillas → Lanzar más como esta” una vez seleccionada la instancia de EC2 que ya está en ejecución:



Sin embargo, deberás modificar la configuración para desplegar esta nueva instancia en la región us-east-1c (editando los detalles de la instancia o seleccionando la subred pública correspondiente del VPC). Anota el identificador de la instancia (y su zona de disponibilidad) porque te hará falta más adelante.

Fíjate que en este punto hay dos instancias, una lanzada en la zona de disponibilidad us-east-1e en la que ha sido creada la página cloud.html y otra lanzada en us-east-1c en la que no existe dicha página (recuerda que los cambios en las instancias no afectan a las AMIs).

4. Creación del balanceador de carga

A continuación, se desea crear un balanceador de carga que permita distribuir el tráfico entrante a las instancias de diferentes zonas de disponibilidad, con el propósito de hacer frente a posibles fallos en

instancias y evitar así denegación de servicio. Para ello se utiliza el servicio ELB (Elastic Load Balancer), de EC2. Este servicio monitoriza las instancias y tan solo dirige tráfico a aquellas instancias sanas. A continuación, se procede a crear un balanceador de carga (*load balancer*) para que reparta el tráfico HTTP recibido en el puerto 80 del balanceador de carga entre instancias disponibles en las zonas de disponibilidad us-east-1e y us-east-1c (el tráfico se dirigirá al puerto 80 de las instancias). Le llamaremos alucloud\$ID-lb. Desde el panel de administración de EC2, elegimos el apartado “Balanceadores de carga”. Pulsamos el botón “Crear balanceador de carga”.

The screenshot shows the 'Compare and select load balancer' wizard. It has three main sections:

- Balanceador de carga de aplicaciones**: Shows a diagram where traffic from a client goes through an ALB (HTTP to HTTPS) to multiple Lambda functions.
- Balanceador de carga de red**: Shows a diagram where traffic from a client goes through an NLB (TCP, UDP, TLS) to multiple VPC endpoints (VPCe).
- Equilibrador de carga de gateway**: Shows a diagram where traffic from a client goes through a GWLB to multiple third-party devices (lock, document, firewall).

Below each section is a descriptive text block and a 'Create' button. At the bottom, there's a note about the 'Classic Load Balancer - generación anterior' and a 'Cerrar' button.

Deberás elegir la opción “Application Load Balancer” y obtendrás un panel como el que se muestra a continuación. Tienes más información sobre los balanceadores de carga a nivel de aplicación en [49]:

The screenshot shows the 'Create application load balancer' configuration wizard. It has several steps:

- Cómo funciona Elastic Load Balancing**
- Configuración básica**
 - Nombre del balanceador de carga**: 'alucloud00-lb' (highlighted in blue)
 - Espuma**: 'Exposto a Internet' (selected)
 - Interno**: Un balanceador de carga interno dirige las solicitudes de los clientes a los destinos mediante direcciones IP privadas.
 - Tipo de dirección IP**: 'IPv4' (selected)

La configuración por defecto es válida puesto que permite reenviar las peticiones recibidas en el puerto 80 del balanceador de carga al puerto 80 de las instancias. Introduce el nombre apropiado (alucloudXX-lb). Además, queremos que el balanceador reciba peticiones de cualquier punto de internet (Expuesto a Internet).

En la sección “Mapeo de red” deberás seleccionar el VPC “default” y las zonas de disponibilidad en las que has desplegado cada una de las dos instancias anteriores (en nuestro caso, us-east-1e y us-east-1c).

The screenshot shows the 'Mapping' section of the AWS Application Load Balancer configuration. At the top, it says 'Create application load balancer' and 'Map the traffic to the destinations of the selected subnets based on the IP address configuration'. Below this, under 'VPC | Info', it says 'Select the private cloud virtual network (VPC) for the destinations or create a new VPC'. It lists 'default' with 'vpc-83a213fb' and 'IPv4: 172.31.0.0/16'. A 'C' button is available to change the VPC.

Under 'Mappings | Info', it says 'Select at least one availability zone and one subnet per zone. The load balancer only directs traffic to the destinations of these availability zones. Zones of availability that are not compatible with the load balancer or the VPC are not available for selection.' It lists several availability zones:

- us-east-1a (use1-az1)
- us-east-1b (use1-az3)
- us-east-1c (use1-az4)
 - Subnet: subnet-432a9408 (selected), subnet-default-1c-public
 - Dirección IPv4: Asignado por AWS
- us-east-1d (use1-az6)
- us-east-1e (use1-az2)
 - Subnet: subnet-c0c33fef (selected), subnet-default-1e-public
 - Dirección IPv4: Asignado por AWS
- us-east-1f (use1-az5)

A los balanceadores de carga creados dentro de un VPC (más información sobre VPCs en el módulo 4) se les puede asignar grupos de seguridad, tal y como se puede hacer con las instancias de EC2 [43]. En nuestro caso, utilizaremos el grupo de seguridad asignados a las instancias de EC2 en el momento de su despliegue.

The screenshot shows the 'Groups de seguridad' (Security Groups) section in the AWS CloudFront configuration. A single security group, 'gs-aws-00' (sg-cc8b8dbe), is selected from a dropdown menu. This group is associated with a VPC and a specific subnet.

Ten en cuenta que podríamos crear un grupo de seguridad especial únicamente con el puerto 80 para el balanceador porque no tiene ningún sentido balancear las peticiones de tráfico SSH entre instancias. Sin embargo, para simplificar, reprovecharemos el grupo de seguridad gs-aws-XX.

En esta sección se puede registrar un “Agente de escucha y direccionamiento” (en inglés, *target group*) [61], que es la configuración de balanceo para las instancias de EC2 que se registrarán en el balanceador.

The screenshot shows the 'Agents de escucha y direccionamiento' (Listeners and Routing) section. It displays a configuration for an 'HTTP:80' listener. The 'Protocolo' (Protocol) is set to 'HTTP' and the 'Puerto' (Port) is set to '80'. The 'Acción predeterminada' (Default Action) is set to 'Reenviar a' (Forward to) a specific destination group. There is also a link to 'Create application load balancer' (Create application load balancer).

Pulsa en “Crear un grupo de destino”.

Indicaremos como tipo de destino instancias (de EC2). Tendrás que darle un nombre: “lb-XX-tg”. Aquí podemos dejar la configuración por defecto, que consiste en distribuir las peticiones http dirigidas al puerto 80 del balanceador al puerto 80 de las instancias de EC2 registradas en este balanceador. Como habrás visto, el balanceador también permite distribuir peticiones tanto a otras direcciones IP (para integrar recursos de cómputo de una empresa – *on-premises*) o bien hacia funciones Lambda, que es código que se ejecuta dirigido por eventos usando el servicio AWS Lambda.

Paso 1: Especificar los detalles del grupo

Especificación básica

La configuración de esta sección no se puede cambiar después de crear el grupo de destino.

Create target group un tipo de destino

- Instancias
 - Admite el balanceo de carga en instancias dentro de una VPC específica.
 - Facilita el uso de [Amazon EC2 Auto Scaling](#) para administrar y escalar la capacidad de EC2.
- Direcciones IP
 - Facilita el balanceo de carga en recursos de VPC y en las instalaciones.
 - Facilita el redireccionamiento a varias direcciones IP e interfaces de red en la misma instancia.
 - Ofrece flexibilidad con arquitecturas basadas en microservicios, lo que simplifica la comunicación entre aplicaciones.
 - Admite destinos IPv6, lo que permite la comunicación IPv6 integral y NAT de IPv4 a IPv6.
- Función Lambda
 - Facilita el redireccionamiento a una única función Lambda.
 - Accesible solo para平衡adores de carga de aplicaciones.
- Balanceador de carga de aplicaciones
 - Ofrece la flexibilidad para que un balanceador de carga de red acepte y dirija solicitudes TCP dentro de una VPC específica.
 - Facilita el uso de direcciones IP estáticas y PrivateLink con un balanceador de carga de aplicaciones.

Nombre del grupo de destino
lb-00-tg

Se permite un máximo de 32 caracteres alfanuméricos, incluidos guiones, pero el nombre no puede comenzar ni terminar por un guion.

Protocolo Puerto
HTTP : 80

1-65535

Tipo de dirección IP
Solo los destinos con el tipo de dirección IP indicado se pueden incluir en este grupo de destino.

IPv4

IPv6
Each target you register must have an assigned primary IPv6 address. This is configured on the instances default network interface (eth0). [Learn more](#)

VPC
Seleccione la VPC con las instancias que desea incluir en el grupo de destino. En esta lista solo están disponibles las VPC que admiten el tipo de dirección IP seleccionado anteriormente.

Versión del protocolo

HTTP1
Envíe solicitudes a los destinos con HTTP/1.1. Compatible cuando el protocolo de solicitud es HTTP/1.1 o HTTP/2.

HTTP2
Envíe solicitudes a los destinos con HTTP/2. Compatible cuando el protocolo de solicitud es HTTP/2 o gRPC, pero las características específicas gRPC no están disponibles.

gRPC
Envíe solicitudes a los destinos con gRPC. Compatible cuando el protocolo de solicitud es gRPC.

También es posible definir la forma de comprobar que las instancias son válidas. Por defecto, cada cinco segundos se accederá al puerto 80 de las instancias para realizar una petición http. Si se obtiene respuesta antes de 5 segundos entonces se considerará que la instancia funciona correctamente. De lo contrario, si tras dos intentos consecutivos no se recibe respuesta en el tiempo definido, se marcará como inválida y no se le reenviará ninguna petición a dicha instancia. Periódicamente se seguirá comprobando el estado de la instancia de la misma manera. Cuando hayan más de 5 peticiones que se atiendan correctamente entonces se determinará que la instancia vuelve a estar activa y entonces ya se le volverá a enviar peticiones desde el balanceador de carga.

Comprobaciones de estado
El balanceador de carga asociado envía periódicamente solicitudes, según las configuraciones que aparecen a continuación, a los destinos registrados para comprobar su estado.

Protocolo de comprobación de estado
HTTP

Ruta de comprobación de estado
Usa la ruta predeterminada de “/” para hacer ping a la raíz o especifique una ruta personalizada si lo prefiere.
/

Configuración avanzada de comprobación de estado

Puerto de comprobación de estado
El puerto predeterminado del balanceador de carga utiliza cuando realiza comprobaciones de estado en los destinos. De forma predeterminada, el puerto de comprobación de estado es el mismo que el puerto de tráfico del grupo de destino. Sin embargo, se puede especificar un puerto diferente como anulación.

Puerto de tráfico
 Sustituir

Umbral en buen estado
El número de comprobaciones de estado consecutivas necesarias antes de considerar que un destino en mal estado es correcto.
5

Umbral en mal estado
El número de errores de comprobación de estado consecutivos necesarios antes de considerar que un destino está en mal estado.
2

Tiempo de espera
La cantidad de tiempo, en segundos, durante la cual no hay respuesta significa que se produjo un error en la comprobación de estado.
5 segundos

Intervalo
Período de tiempo aproximado que transcurre entre comprobaciones de estado de un destino individual.
30 segundos

Códigos de éxito
Códigos HTTP que se deben utilizar al comprobar si se ha recibido una respuesta correcta de un destino. Puede especificar varios valores (por ejemplo, “200,202”) o un intervalo de valores (por ejemplo, “200-299”).
200

Pulsa el botón “Siguiente” para pasar a registrar las instancias que se integrarán en dicho balanceador de carga. Asegúrate de elegir únicamente las dos instancias que has creado anteriormente. En el ejemplo que se muestra en la figura se muestran dos instancias en dos zonas de disponibilidad diferentes correspondientes al usuario alucloudoo.

Registrar destinos
Se trata de la [Create target group](#) | crear un grupo de destino. Sin embargo, para asegurarse de que el balanceador de carga dirige el tráfico a este grupo de destino, debe registrar los destinos.

Instancias disponibles (2/2)

ID de instancia	Nombre	Estado	Grupos de seguridad	Zona	Dirección IPv4 privada
I-01c578260ff5730aa	Instance00-1	Ejecutando	gs-aws-00	us-east-1c	172.31.38.20
I-05d641e0bf7bfaf6	Instance00-1	Ejecutando	gs-aws-00	us-east-1e	172.31.75.158

2 seleccionados
Puertos para las instancias seleccionadas
Número para dirigir el tráfico a las instancias seleccionadas:
80
1-65535 (separar puertos múltiples con coma)
Incluir como pendiente a continuación

Revisar destinos

Destinos (0)	Eliminar todos los pendientes
Aún no se han agregado Instancias.	
Especifique las instancias anteriores o deje el grupo vacío si prefiere agregar destinos más adelante.	

0 pendientes

No te olvides de pulsar el botón “Incluir como pendiente a continuación” antes de pulsar en “Crear un grupo de destino”. De lo contrario, no añadirás las instancias al balanceador.

De vuelta a la pestaña del navegador de configuración del balanceador, refresca con el botón lateral si es necesario y selecciona el nuevo grupo de destino:

Agentes de escucha y direccionamiento [Info](#)

Un agente de escucha es un proceso que comprueba las solicitudes de conexión mediante el puerto y el protocolo que configure. Las reglas que defina para un agente de escucha determinan cómo el equilibrador de carga dirige las solicitudes a sus destinos registrados.

▼ Agente de escucha **HTTP:80** Eliminar

Protocolo	Puerto	Acción predeterminada	Info
HTTP	: 80	Reenviar a	lb-00-tg Tipo de destino: Instancia, IPv4
HTTP ▾ C			
Crear un grupo de destino			

Etiquetas del agente de escucha - *opcional*
Considere la posibilidad de agregar etiquetas al agente de escucha. Las etiquetas permiten clasificar los recursos de AWS para que pueda administrarlos con mayor facilidad.

[Agregar etiqueta de agente de escucha](#)
Puede agregar hasta 50 etiquetas más.

[Agregar agente de escucha](#)

No es necesario que pulses “Agregar agente de escucha” pues esto permite crear otro agente de escucha, escuchando en otro puerto diferente, para reenviar las peticiones a otro grupo de destino, por lo que el balanceador de carga puede ser compartido.

El balanceo de carga entre zonas de disponibilidad para los Application Load Balancers siempre es de tipo *Cross-Zone Load Balancing* [45], que permite garantizar que las peticiones de tráfico se distribuyen de forma equitativa entre las instancias de EC2 independientemente de la zona de disponibilidad en la que se encuentren. Esta opción reduce la necesidad de mantener un número de instancias equivalente en cada zona de disponibilidad. Sin embargo, AWS sigue manteniendo la recomendación de mantener un número aproximadamente equivalente de instancias en cada zona de disponibilidad para una mayor tolerancia a fallos.

Añade una etiqueta para identificar fácilmente el propietario del balanceador:

▼ Balanceador de carga etiquetas - *opcional*
Considere la posibilidad de agregar etiquetas al balanceador de carga. Las etiquetas permiten categorizar los recursos de AWS para que pueda administrarlos más fácilmente. La "Clave" es obligatoria, pero el "Valor" es opcional. Por ejemplo, puede tener Clave = production-webserver, o Clave = webserver y Valor = production.

Clave	Valor - Opcional	Eliminar
owner	alucloud00	Eliminar

[Agregar nueva etiqueta](#)
Puede agregar hasta 49 etiquetas más.

Al final de la página obtendrás un resumen de la configuración del ELB, como el que se muestra a continuación.

Resumen

Revise y confirme las configuraciones. Costo estimado [\[\]](#)

Configuración básica Editar alucloud00-lb <ul style="list-style-type: none">Exposto a InternetIPv4	Grupos de seguridad Editar gs-aws-00 sg-cc8b8dbe []	Mapeo de red Editar VPC vpc-83a213fb [] default <ul style="list-style-type: none">us-east-1c subnet-432a9408 [] subnet-default-1c-publicus-east-1e subnet-c0c33fef [] subnet-default-1e-public	Agentes de escucha y direccionamiento Editar <ul style="list-style-type: none">HTTP:80 valor predeterminado para lb-00-tg []
Servicios complementarios Editar Ninguno	Etiquetas Editar owneralucloud00		
Atributos <div style="border: 1px solid #ccc; padding: 5px;"><p>Algunos atributos predeterminados se aplicarán al balanceador de carga. Puede verlos y editarlos después de crear el balanceador de carga.</p></div>			

Pulsa sobre el botón “Crear balanceador de carga”.

Una vez creado, obtendrás un mensaje de confirmación:

Se creó correctamente el balanceador de carga: alucloud00-lb
Nota: Pueden transcurrir unos minutos hasta que el balanceador de carga esté totalmente configurado y listo para dirigir el tráfico. Los destinos también tardarán unos minutos en completar el proceso de registro y superar las comprobaciones de estado iniciales.

EC2 > Balanceadores de carga > alucloud00-lb > Crear balanceador de carga de aplicaciones

Crear balanceador de carga de aplicaciones

Pasos siguientes sugeridos

- Revise, personalice o coi Load Balancer created successfully ador de carga y de los agentes de escucha mediante las pestañas Descripción y Agentes de escucha dentro de alucloud00-lb.
- Descubra otros servicios que puede integrar con el balanceador de carga. Visite la pestaña Servicios integrados en alucloud00-lb.

[Ver el balanceador de carga](#)

Es posible obtener un listado de todos los平衡adores de carga creados con ELB desde el panel de administración de EC2, así como ver la configuración de tu balanceador:

En primer lugar, pasará a estado “Aprovisionándose”:

Balanceadores de carga (1)

Elastic Load Balancing escala automáticamente la capacidad del equilibrador de carga en respuesta a los cambios en el tráfico entrante.

Acciones	Crear balanceador de carga
<input type="checkbox"/> alucloud00-lb	Nombre: alucloud00-lb-10499524... Estado: Aprovisionándose ID de VPC: vpc-83a213fb Zonas de disponibilidad: 2 Zonas de disponibilidad Tipo: application Fecha creación: 4 de octubre 19:35 (UTC)

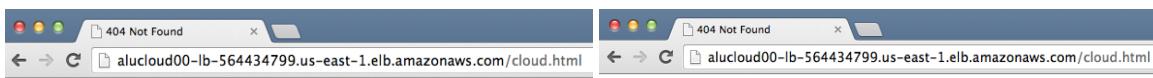
Ten en cuenta que la creación del balanceador de carga se puede demorar unos (pocos) minutos, hasta que llegue a estado “active” (no olvides refrescar la interfaz periódicamente para evitar esperas innecesarias):

Nombre	Estado	ID de VPC	Zonas de disponibilidad	Tipo	Fecha creada
alucloud00-lb	Activ	vpc-83a213fb	2 Zonas de disponibilidad	application	4 de octubre de 2023, 19:35 (UTC+02:00)

Puedes ver los detalles del balanceador de carga:

Protocolo:Puerto	Acción predeterminada	Reglas	ARN	Política de seguridad	Certificado SSL/TLS predet...	Etiquetas
HTTP:80	Reenviar al grupo de destino	1 regla	ARN	No aplicable	No applicable	0 etiquetas

Si accedes con un navegador web a la URL asociada al balanceador de carga para acceder a la nueva página creada (dicha URL podría ser, por ejemplo, <http://alucloud00-lb-1939207801.us-east-1.elb.amazonaws.com/cloud.html>) observarás que a veces obtienes el mensaje de 404 NOT Found (cuando la petición la sirve la instancia desplegada en us-east-1c) y otras obtienes la página cloud.html que muestra la imagen que se descarga de S3, cuando la petición la sirve la instancia desplegada en us-east-1e. Es posible que tengas que refrescar varias veces e incluso utilizar diferentes navegadores para verificar el efecto, puesto que la caché del propio navegador puede afectar a los resultados obtenidos. Usar el modo incógnito de Google Chrome o de Safari puede ayudar. Si no obtienes respuesta al conectarte al balanceador revisa si has asignado el grupo de seguridad que se te indicaba (ya que debe admitir conexiones entrantes al puerto 80).



Recuerda que deberás esperar unos pocos minutos hasta que el balanceador de carga esté completamente operativo. Si obtienes un error 503 es que las instancias no están registradas en el balanceador de carga. Para ello, deberás regresar a la configuración del Target Group y registrarlas correctamente.

Efectivamente, las solicitudes están siendo atendidas por diferentes instancias. Como cada instancia sirve una página diferente mediante el servidor web Apache, de ahí los diferentes resultados. En un entorno más homogéneo, donde diferentes instancias desempeñan el mismo rol, la caída de una instancia no afectaría al servicio, puesto que el balanceador de carga distribuiría las peticiones a las instancias sanas.

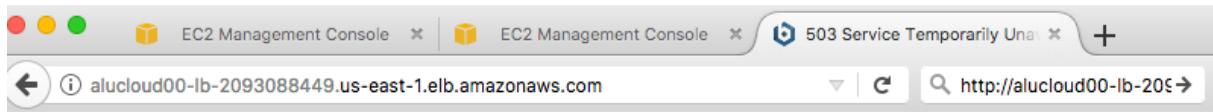
Vamos a proceder a deregistrar de forma dinámica instancias del balanceador de carga para evitar que reciban peticiones. Para ello, tendrás que ir la sección “Grupos de destino” del menú lateral izquierdo y seleccionar tu “Grupo de destino” (llamado lb-XX-tg).

The screenshot shows the AWS EC2 Groups of destination interface. On the left sidebar, under the 'Equilibrio de carga' section, 'Grupos de destino' is selected. The main pane displays a table titled 'Grupos de destino [1] Info'. The table has columns: Nombre, ARN, Puerto, Protocolo, Tipo de destino, Balanceador de carga, and ID de VPC. A single row is listed: 'lb-00-tg' with ARN 'arn:aws:elasticloadbalancing:us-east-1:1974349055189:targetgroup/lb-00-tg/418f3a14c0f7048f', Puerto '80', Protocolo 'HTTP', Tipo de destino 'Instancia', Balanceador de carga 'aludoud00-lb', and ID de VPC 'vpc-83a213fb'. There are buttons for 'Acciones' and 'Crear un grupo de destino'.

Ahora, en la pestaña “Destinos registrados” podrás deregistrar dichas instancias del balanceador de carga (seleccionándolas y pulsando el botón “Anular el registro”):

The screenshot shows the 'lb-00-tg' target group details page. At the top, it says 'lb-00-tg'. Below is a 'Detalles' section with fields: 'Tipo de destino' (Instancia), 'Protocolo : Puerto' (HTTP: 80), 'Versión del protocolo' (HTTP1), and 'VPC' (vpc-83a213fb). It also shows 'Destinos totales' (2), 'En buen estado' (2 green circles), 'En mal estado' (0 red circles), 'Sin utilizar' (0 grey circles), 'Inicial' (0 grey circles), and 'Vaciado' (0 grey circles). A note below says 'Distribución de destinos por zona de disponibilidad (AZ)'. At the bottom, there are tabs for 'Destinos' (selected), 'Monitorización', 'Comprobaciones de estado', 'Atributos', and 'Etiquetas'. The 'Destinos registrados' table shows two instances: 'i-05d641cebf7f0fa6' (Instance00-1, 80, us-east-1e, healthy) and 'i-01c578260fb3750aa' (Instance00-1, 80, us-east-1c, healthy). Buttons for 'Anular el registro' and 'Registrar destinos' are visible.

El proceso de deregistro de instancias se demora unos cuantos minutos (estado “draining”), tiempo que se dedica a finalizar las peticiones que pueda haber pendientes a los servidores web. Una vez deregistradas las instancias, si tratas de acceder a la URL del balanceador de carga obtendrás un error http 503 de servicio temporalmente no disponible (esto puede ocurrir mientras las instancias todavía están en estado “draining”):



503 Service Temporarily Unavailable

Antes de proceder a realizar la siguiente sección, asegúrate de deregistrar las instancias del balanceador de carga y terminar las instancias. Sin embargo, es importante que NO elimines el balanceador de carga, puesto que será necesario para la creación de grupos de auto-escalado.



Asegúrate de que has terminado las dos instancias desplegadas antes de pasar a la siguiente sección. La siguiente sección te llevará unos 20-30 minutos. Si crees que no te dará tiempo a terminarla de un tirón, lo mejor será que termines el balanceador ahora y vuelvas a crear uno nuevo cuando vayas terminar la práctica.

6. Auto-Escalado de Grupos Instancias

El servicio de Auto Scaling [28] permite la definición de reglas de escalado horizontal (*scale out*) para determinar cuándo desplegar más instancias en función de unas ciertas métricas de prestaciones. Esto permite, por ejemplo, desplegar un mayor número de instancias cuando la carga de trabajo (por ejemplo, medida en promedio de uso de CPU de las instancias) supera un determinado umbral. En conjunción con el servicio de Balanceo de Carga (ELB), esto permite la creación de flotas elásticas de instancias cuya capacidad puede crecer y decrecer de forma dinámica (en término del número de instancias) adaptándose así a la carga de trabajo.

6.1. Configuración de Auto Scaling

El uso de Auto Scaling requiere la creación de grupos de autoescalado (*autoscaling groups*) que agrupan un conjunto de instancias. Existen plantillas de lanzamiento (*launch templates*) (antiguamente se utilizaban en su lugar las configuraciones de lanzamiento (*launch configurations*)) que determinan qué clases de instancias deben ser lanzadas (AMI, par de claves, grupo de seguridad y tipo de instancia, por lo general). Además, también hay alarmas (*alarms*) que indican qué acción de escalado (denominada política (*policy*)) debe aplicarse. Las políticas simplemente especifican si las instancias deben ser lanzadas o terminadas.

1. Creación de la Plantilla de Lanzamiento

La plantilla de lanzamiento determina qué tipo de instancias deben ser puestas en marcha. Para este caso, se utilizará una AMI que ya se ha gastado anteriormente (Bitnami LAMP). Lo lógico sería utilizar una imagen que cumpliera un determinado rol dentro de la lógica de una aplicación. Sin embargo, para demostrar el funcionamiento del auto-escalado es suficiente.

Estas serán las características de la configuración de lanzamiento:

- AMI: El identificador de AMI va cambiando con el tiempo por lo que buscaremos una AMI concreta.
- Tipo de instancia: t3.micro
- Grupo de seguridad: El correspondiente a tu usuario.

Elije la opción “Plantillas de lanzamiento” y pulsa sobre el botón “Create launch template”:

ID de plantilla de lanzamiento...	Nombre de plantilla de lan...	Versión predet...	Última versión	Hora de creación	Creado por
It-0c73531b4465010921	Batch-It-3da7c6b-b7eb-3610...	1	1	2020-01-15T12:35:11.000Z	arn:aws:sts::974349055189:assumed-role/AWSBatchServiceRole/aws-batch
It-0c809f98c9c81c5b2	wp-launch-template-203	1	1	2023-05-27T19:49:57.000Z	arn:aws:iam::974349055189:user/aluctoud203

Introduce los siguientes valores de configuración (acuérdate de cambiar oo por tu identificador de usuario):

Crear plantilla de lanzamiento

La creación de una plantilla de lanzamiento le permite crear una configuración de instancia guardada que se puede reutilizar, compartir y lanzar más adelante. Las plantillas pueden tener varias versiones.

Nombre y descripción de la plantilla de lanzamiento

Nombre de la plantilla de lanzamiento - *obligatorio*

launch-template-00

Debe ser única para esta cuenta. Máximo de 128 caracteres. Sin espacios ni caracteres especiales, como "&", "“”, “@”.

Descripción de la versión de la plantilla

launch-template-00

Máximo de 255 caracteres

Orientación sobre Auto Scaling [Información](#)

Selecciónelo si va a utilizar esta plantilla con EC2 Auto Scaling

Proporcionar orientación que me ayude a configurar una plantilla que pueda utilizar con EC2 Auto Scaling

► Etiquetas de la plantilla

► Plantilla de origen

Para poder indicar la AMI deberás introducir “bitnami-lampstack-7” y elegir cualquiera de las AMIs que aparezcan en la pestaña “Community AMIs” (AMI de la comunidad), como se muestra a continuación. Asegúrate simplemente de elegir una cuyo tipo de virtualización sea “hvm” (ten en cuenta que los números de versión pueden haber cambiado con respecto a lo mostrado en la figura inferior):

Elegir una Amazon Machine Image (AMI)

Una AMI es una plantilla que contiene la configuración de software (sistema operativo, servidor de aplicaciones y aplicaciones) necesaria para lanzar la instancia. Puede seleccionar una AMI proporcionada por AWS o nuestra comunidad de usuarios, o bien a través de AWS Marketplace.

bitnami-lampstack-7

Crear plantilla de lanzamiento Mis AMI (56) AMI de AWS Marketplace (2056) AMI de la comunidad (46)

Acotar los resultados Borrar todos los filtros

▼ Sistema operativo ▼ Linux/UNIX □ Todos los de Linux/UNIX □ Amazon Linux □ CentOS

bitnami-lampstack-7 (46 filtrados, 46 sin filtrar)

AMI de la comunidad

Las AMI de comunidad contienen todas las AMI que son públicas, por lo tanto, cualquier persona puede publicar una AMI y se mostrará en este catálogo. Este catálogo solo contiene productos de pago. Al utilizar las AMI de comunidad, se recomienda que se asegure de que conoce y confía en el editor antes de lanzar una AMI.

bitnami-lampstack-7.4.26-16-r06-linux-debian-10-x86_64-hvm-ebs-nami-cb6af05-a3b2-4916-a3e6-bcc414f5f21

AMI ID: ami-0ad0d6cc70562c4ad

bitnami

This image may not be the latest version available and might include security vulnerabilities. Please check the latest, up-to-date, available version at <https://bitnami.com/stacks>.

Proporcionado por bitnami OwnerAlias: aws-marketplace Plataforma: Debian Arquitectura: x86_64 Propietario: 67959333241 Fecha de publicación: 2021-12-15 Tipo de dispositivo raíz: ebs Virtualización: hvm Habilitado para ENA: Sí

Seleccionar

Asegúrate de que NO estás eligiendo una AMI de la pestaña “AMI de AWS Marketplace”, pues es posible que no tengas permisos para utilizarlas.

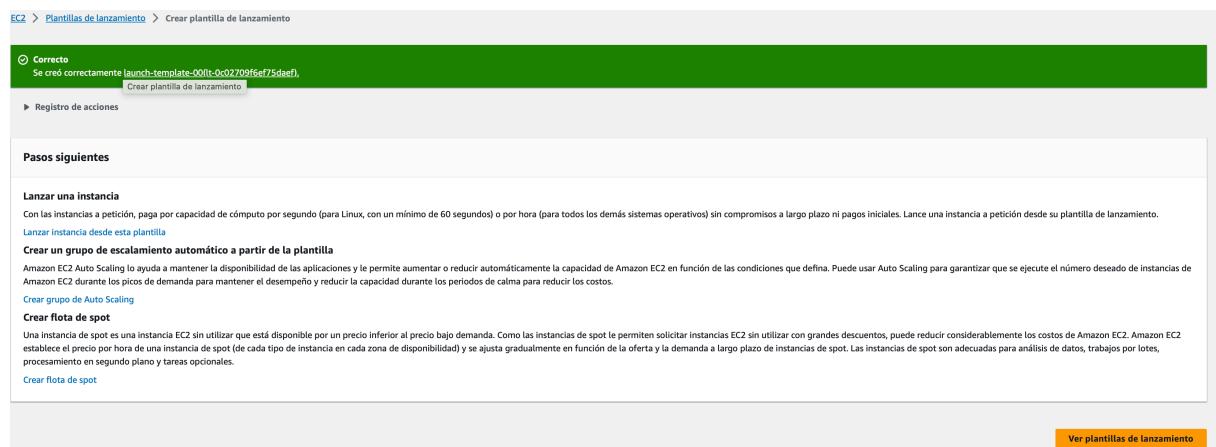
A continuación, deberás elegir el tipo de instancia (**t3.micro**), el key pair (**alucloudXX-keypair**) y la configuración de red a utilizar (VPC), indicando tu grupo de seguridad (**gs-aws-XX**, donde XX corresponde con tu identificador de usuario). No elijas ninguna subred.

The screenshot shows the 'Tipo de instancia' (Instance Type) section with 't3.micro' selected. It also shows the 'Par de claves (inicio de sesión)' (Key Pair) section with 'alucloud00-keypair' selected. In the 'Configuraciones de red' (Network Configuration) section, the 'Subred' (Subnet) dropdown is set to 'No incluir en la plantilla de lanzamiento' (Not included in launch template). The 'Firewall (grupos de seguridad)' (Security Groups) section shows 'Seleccionar un grupo de seguridad existente' (Select existing security group) selected, with 'sg-aws-00' listed. A note at the bottom says 'Configuración de red avanzada' (Advanced network configuration).

Existen aspectos de configuración avanzada, como la posibilidad cambiar el kernel utilizado para arrancar la AMI, la ejecución de un script de configuración en el momento de arranque de la instancia y la asignación de IPs públicas, entre otros. Si quisieras en dicha instancia ejecutar una aplicación que tuviera acceso a otros servicios de AWS entonces puedes seleccionar un IAM role [70] en el apartado IAM instance profile, que inyecta automáticamente en la instancia unas credenciales temporales con unos privilegios delimitados por el rol elegido.

No introduciremos ningún cambio más por lo que puedes pulsar sobre “Crear plantilla de lanzamiento”.

Obtendrás un mensaje de confirmación:



Desde esa ventana es posible la creación de un grupo de auto-escalado vinculado a dicha configuración de lanzamiento así que no la cierres y en la siguiente sección procederemos a crear el grupo de auto-escalado. De hecho, lo más probable es que directamente se inicie el wizard de creación de un grupo de auto-escalado. Si no, puedes iniciar el proceso de creación con la opción “Crear grupo de escalamiento automático”:

2. Creación del Grupo de Auto-Escalado

Un grupo de auto-escalado requiere una plantilla de lanzamiento, de ahí que en el apartado anterior se procediera a crear una. La creación del grupo de auto-escalado define:

- los tamaños mínimo y máximo del grupo (en número de instancias)
- las zonas de disponibilidad involucradas
- la configuración de lanzamiento a emplear en caso de necesitar aprovisionar instancias
- el balanceador de carga a emplear tras el que se posicionarán las nuevas instancias creadas

Al especificar un tamaño mínimo, la creación del grupo de auto-escalado provocará el despliegue inmediato de instancias. Es importante saber que las instancias ya creadas y ligadas al balanceador de carga no computan como parte del grupo de auto-escalado. Por ello, si había dos instancias ya desplegadas y se solicita un tamaño mínimo de grupo de auto-escalado de 3, se procederá al despliegue de 3 instancias y, por tanto, habrá 5 instancias tras el balanceador de carga. Ten en cuenta que el grupo de auto-escalado y el balanceador de carga son dos entidades diferentes. Únicamente las instancias desplegadas por el grupo de auto-escalado computan para las reglas de tamaño mínimo y máximo del grupo.

A continuación, continuamos con la creación de un grupo de auto-escalado de tamaño mínimo 1 instancia y como máximo 2 instancias, dentro de las zonas de disponibilidad us-east-1e y us-east-1c, ligadas al balanceador de carga ya creado anteriormente.

Para ello, pulsamos sobre la opción “Crear grupo de Auto Scaling”, si no estamos directamente ya en el wizard de creación.

No será necesario volver a especificar el nombre de la plantilla de lanzamiento. Como nombre del grupo aségñale as-group-\$ID:

Paso 1
Elija la plantilla de lanzamiento o la configuración

Paso 2
Elegir las opciones de lanzamiento de instancias

Paso 3 - opcional
Configurar las opciones avanzadas

Paso 4 - opcional
Configurar políticas de escalado y tamaño de grupo

Paso 5 - opcional
Añadir notificación

Paso 6 - opcional
Añadir etiquetas

Paso 7
Revisar

Elija la plantilla de lanzamiento o la configuración Info

Especifique una configuración de lanzamiento que contenga ajustes comunes a todas las instancias EC2 lanzadas por este grupo de Auto Scaling - Auto Scaling. Si actualmente utiliza configuraciones de lanzamiento, puede considerar la posibilidad de migrar a plantillas de lanzamiento.

Nombre

Nombre del grupo de Auto Scaling
Escriba un nombre para identificar el grupo.
as-group-00

Debe ser único para esta cuenta en la región actual y no puede superar los 255 caracteres.

Plantilla de lanzamiento Info **Cambiar a la configuración de lanzamiento**

Plantilla de lanzamiento
Elija una configuración de lanzamiento que contenga la configuración de nivel de instancia, como la imagen de máquina de Amazon (AMI), el tipo de instancia, el par de claves y los grupos de seguridad.

launch-template-00 **Crear una configuración de lanzamiento** **C**

Versión
Default (1) **C**

Crear una versión de plantilla de lanzamiento

Descripción	Plantilla de lanzamiento	Tipo de instancia
launch-template-00	launch-template-00 lt-0c02709f6ef75daef	t3.micro
AMI ID	ami-0ad0d6cc70562c4ad	Grupos de seguridad -
Nombre del par de claves	aludoud00-keypair	ID de grupos de seguridad sg-cc8b8d8e

Detalles adicionales

Almacenamiento (volumenes)	Fecha de creación Wed Oct 04 2023 19:52:21 GMT +0200 (hora de verano de Europa central)
----------------------------	---

Siguiente

Pulsa el botón “Siguiente”: Mantén la opción que indica que usaremos tipos de instancia como los indicados en la plantilla de lanzamiento. A continuación, deberás elegir el vpc “default” y las zonas de disponibilidad us-east-1e y us-east-1c (por lo que deberás elegir las subredes **públicas** del VPC “default” correspondientes):

Paso 1
Elija la plantilla de lanzamiento o la configuración

Paso 2
Elegir las opciones de lanzamiento de instancias

Paso 3 - opcional
Configurar las opciones avanzadas

Paso 4 - opcional
Configurar políticas de escalado y tamaño de grupo

Paso 5 - opcional
Añadir notificación

Paso 6 - opcional
Añadir etiquetas

Paso 7
Revisar

Elegir las opciones de lanzamiento de instancias Info

Elija el entorno de red de VPC en el que se lanzan las instancias y personalice los tipos de instancias y las opciones de compra.

Requisitos de tipo de instancias Info

Puede mantener los mismos atributos o tipos de instancias de la plantilla de lanzamiento, o bien puede optar por anular la plantilla de lanzamiento al especificar atributos de instancia diferentes o al agregar los tipos de Instancias de forma manual.

Anular la plantilla de lanzamiento

Plantilla de lanzamiento	Versión	Descripción
launch-template-00 lt-0c02709f6ef75daef	Default	launch-template-00

Tipo de instancia
t3.micro

Red Info

Para la mayoría de las aplicaciones, puede utilizar varias zonas de disponibilidad y dejar que EC2 Auto Scaling equilibre sus instancias entre las zonas. La VPC predeterminada y las subredes predeterminadas son adecuadas para comenzar rápidamente.

VPC
Elija la VPC que define la red virtual para el grupo de Auto Scaling.
vpc-83a213fb (default) **172.31.0.0/16** **C**

Crear una VPC C

Zonas de disponibilidad y subredes
Defina qué zonas de disponibilidad y subredes puede utilizar el grupo de Auto Scaling en la VPC elegida.

Seleccionar zonas de disponibilidad y subredes **C**

us-east-1c | subnet-432a9408 (subnet-default-1c-public)
172.31.32.0/20

us-east-1e | subnet-c0c33fef (subnet-default-1e-public)
172.31.64.0/20

Crear una subred C

Siguiente

Pulsa “Siguiente”. Ahora, deberás activar el balanceo de carga y seleccionar tu Grupo de Destino:

Esto permite vincular el “Grupo de destino” del balanceador de carga de ELB que creaste anteriormente con el grupo de auto-escalado. Esto permitirá que el tráfico recibido por el balanceador sea distribuido entre las instancias del grupo de auto-escalado. No es necesario que introduzcas ninguna configuración adicional.

Pulsa “Siguiente” nuevamente. Mantendremos el tamaño mínimo de grupo a 1 indicado como máximo 2, para poder demostrar la funcionalidad de auto-escalado:

Con respecto a las políticas de auto-escalado automático, está la opción:

- Políticas de escalado de seguimiento de destino (*Target Tracking Scaling Policy*) [77]. Permite ajustar el tamaño del grupo de auto-escalado para tratar de cumplir una determinada métrica, como que el promedio de uso de CPU esté situado en torno a 70%. Esto implicará desplegar instancias adicionales si el valor de dicha métrica sube por encima de dicho valor, desplegando un número de instancias adicionales en función de la discrepancia entre el valor objetivo y el valor actual de la métrica. Alternativamente, se eliminan instancias cuando dicho valor baja del valor objetivo. Esta política se puede definir en el wizard de creación del grupo de auto-escalado.

Usaremos este tipo de políticas y elegiremos un 70% como valor objetivo de la métrica “promedio de uso de CPU” de todas las instancias del grupo de auto-escalado:

Configurar políticas de escalado y tamaño de grupo - opcional

Especifique el tamaño del grupo de Auto Scaling cambiando la capacidad deseada. También puede especificar los límites de capacidad mínima y máxima. La capacidad deseada debe estar dentro del rango de límites.

Tamaño del grupo - opcional

Capacidad deseada: 1

Capacidad mínima: 1

Capacidad máxima: 2

Políticas de escalado - opcional

Elija si desea utilizar una política de escalado para cambiar el tamaño de su grupo de Auto Scaling de forma dinámica para satisfacer los cambios en la demanda. [Info](#)

Política de escalado de seguimiento de destino
Elija un resultado deseado y déjelo en manos de la política de escalado para agregar y eliminar capacidad según sea necesario para conseguir ese resultado.

Ninguno

Nombre de la política de escalado: Target Tracking Policy

Tipo de métrica [Info](#)
Métrica supervisada que determina si la utilización de recursos es demasiado baja o alta. Si utiliza métricas de EC2, considere la posibilidad de habilitar la supervisión detallada para obtener un mejor rendimiento de escalado.

Utilización promedio de la CPU

Valor de destino: 70

Preparación de la instancia [Info](#)
300 segundos

Deshabilite el escalado descendente para crear solo una política de escalado ascendente

Protección de escalado descendente de instancias - opcional

Si la opción Protección frente al escalado descendente está habilitada, las instancias recién lanzadas estarán protegidas del escalado descendente de forma predeterminada.

Habilitar la protección de escalado descendente de instancias

Siguiente

También es posible definir la política de mantenimiento de instancias [80] para controlar los requisitos de capacidad durante los eventos que causan el reemplazo de instancias, como la comprobación de salud ante una instancia que no funciona correctamente. Pero dejaremos la opción “Sin política”.

The screenshot shows the 'Politica de mantenimiento de instancias - nueva' (New Instance Maintenance Policy) configuration screen. It includes fields for 'Valor de destino' (70), 'Preparación de la instancia' (300 seconds), and a checkbox for 'Deshabilite el escalado descendente para crear solo una política de escalado ascendente' (Disable instance termination for creating an ascending scaling policy only). Below this is a detailed description of the policy's purpose and how it affects EC2 Auto Scaling.

Controle la disponibilidad y el costo durante los eventos de reemplazo

A policy of maintenance instances determines the availability of the application when EC2 Auto Scaling replaces instances. It also establishes barriers of protection that limit the capacity of the group that can be added or removed when replacing instances.

Elija un comportamiento de reemplazo en función de sus requisitos de disponibilidad

The screenshot displays four options:

- Comportamiento mixto**: Sin política. Instances will be launched in batches before all instances finish.
- Prioriza la disponibilidad**: Lance instances mas rápidamente que están listas antes de terminar el resto.
- Controla los costos**: Finaliz y lanza instances al mismo tiempo. This option permits reducing costs by launching instances in batches.
- Flexible**: Finaliz y lanza instances al mismo tiempo. This option offers more control over costs by defining minimum and maximum capacity values.

Protección de escalado descendente de instancias

This section describes the protection of instances from being terminated due to scaling events. It includes a checkbox for 'Habilitar la protección de escalado descendente de instancias' (Enable instance termination protection).

At the bottom are 'Cancelar', 'Omitir para revisar', 'Anterior', and a large orange 'Siguiente' (Next) button.

No introduciremos ningún cambio más. Pulsa en “Siguiente”.

The screenshot shows the 'Añadir notificación - opcional' (Add notification - optional) step. It includes a button to 'Añadir notificación' (Add notification) and a note about sending notifications to SNS topics when instances are launched or terminated. Below this are several optional steps:

- Paso 1**: Elija la plantilla de lanzamiento o la configuración.
- Paso 2**: Elija las opciones de lanzamiento de instancias.
- Paso 3 - opcional**: Configurar las opciones avanzadas.
- Paso 4 - opcional**: Configurar políticas de escalado y tamaño de grupo.
- Paso 5 - opcional**: Añadir notificación.
- Paso 6 - opcional**: Añadir etiquetas.
- Paso 7**: Revisar.

At the bottom are 'Cancelar', 'Omitir para revisar', 'Anterior', and a large orange 'Siguiente' (Next) button.

No añadiremos ninguna opción para las notificaciones, pues podremos ver el ciclo de vida del grupo de auto-escalado desde la consola. Pulsa Siguiente. Añadiremos una etiqueta para que podamos saber el propietario de las instancias desplegadas:

The screenshot shows the 'Add tags - optional' step of the wizard. It includes a note about adding tags to instances and their EBS volumes, a 'Tags (0)' section with a 'Add tag' button, and a note about remaining tag limits. Navigation buttons 'Cancel', 'Previous', and 'Next Step' are at the bottom.

Tras pulsar Siguiente, obtenemos un panel de revisión de la configuración del grupo de auto-escalado:

The screenshot shows the 'Review' step of the wizard, displaying the configuration details for the Auto Scaling group. It includes sections for 'Group details', 'Launch template', 'Network settings', 'Instance types', 'Advanced options', and 'Tags'. The 'Create Auto Scaling Group' button is at the bottom right.

Creamos el grupo pulsando sobre “Crear grupo de Auto Scaling” y recibiremos confirmación sobre la creación del mismo.

The screenshot shows the 'Grupos de Auto Scaling' (Auto Scaling Groups) page in the AWS Management Console. A success message at the top states: 'as-group-00, 1 Política de escalado se creó correctamente' (as-group-00, 1 scaling policy was created successfully). The main table lists one item: 'as-group-00' with a status of 'Actualización de cap.' (Scaling capacity update) set to 1, and 'Zonas de disponibilidad' (Availability zones) listed as 'us-east-1c, us-east-1e'. The table includes columns for Nombre (Name), Plantilla de lanzamiento/config... (Launch template/config...), Instanc... (Instances), Estado (State), Capacidad des... (Desired capacity), M... (Minimum), M... (Maximum), and Zonas de disponibilidad (Availability zones).

En cuanto hayas creado el grupo de auto-escalado podrás ver como automáticamente procede a desplegar una instancia (cuyas características fueron especificadas en la configuración de lanzamiento) para satisfacer el tamaño mínimo de grupo). Para ello, elige el grupo de auto-escalado y elige la pestaña “Actividad”:

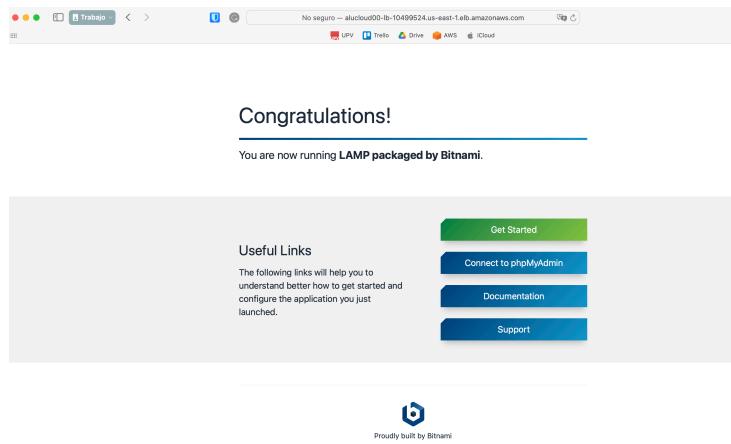
The screenshot shows the 'Actividad' (Activity) tab for the 'as-group-00' group. It displays two sections: 'Notificaciones de actividad' (Activity notifications) and 'Historial de actividad' (Activity history). Under 'Notificaciones de actividad', there is a button to 'Crear notificación' (Create notification). Under 'Historial de actividad', there is one entry: 'Launching a new EC2 instance: i-057c6028d481d2e9d' at 2023-10-04T18:05:34Z, which started in response to a difference between desired and actual capacity, increasing the capacity from 0 to 1.

3. Verificación de Funcionamiento del Grupo de Auto-Escalado

Si obtienes el nombre DNS del balanceador de la carga a través de la consola de EC2:

The screenshot shows the 'Balanceadores de carga' (Load Balancers) page in the AWS Management Console. It displays a table with one item: 'alucloud00-lb' with details such as 'Nombre' (Name), 'Estado' (State) set to 'Activo' (Active), 'ID de VPC' (VPC ID) 'vpc-83a213fb', 'Zonas de disponibilidad' (Availability zones) '2 Zonas de disponibilidad.....', 'Tipo' (Type) 'application', and 'Fecha creada' (Created date) '4 de octubre de 2023, 19:35 (UTC+02:00)'. The table includes columns for Nombre (Name), Nombre de DNS (DNS Name), Estado (State), ID de VPC (VPC ID), Zonas de disponibilidad (Availability zones), Tipo (Type), and Fecha creada (Created date).

y accedes con un navegador, verás que cualquier instancia del grupo de auto-escalado puede servir dicha petición:



El acceso al servicio proporcionado por las instancias (servir páginas web) se puede realizar accediendo a la dirección proporcionada por el balanceador de carga. Puedes averiguar la dirección del balanceador de carga empleado desde la consola de administración de EC2.

El acceso a esa URL permitirá acceder a la única instancia actualmente desplegada. Si se produce un aumento excesivo de la carga de trabajo en la(s) instancia(s) de manera que se produzcan las condiciones definidas en la alarma (promedio de uso de la CPU superior al 70%) entonces se procederá a desplegar dos nuevas instancias.

Para ello, vamos a simular *stress* en las instancias, conectándonos a la única que tenemos desplegada y ejecutando un comando que consuma CPU. En GNU/Linux, el comando *yes* muestra en un bucle infinito la letra *y* por la salida estándar. Esto provoca un aumento del uso de la CPU al 100% de su capacidad. Esta prueba equivale a simular unas condiciones de aumento de trabajo, como podría ser un aumento sobrevenido del número de peticiones HTTP a servidor, en el caso de un sitio web, o un incremento del número de peticiones a un servicio web.



Asegúrate de que ejecutas el comando *yes* en la instancia de EC2 vinculada al grupo de auto-escalado en lugar de hacerlo sobre el propio entorno de prácticas. De lo contrario, aumentarás el consumo de CPU en la instancia que aloja el entorno de prácticas con dos consecuencias. Por un lado, ralentizarás dicha instancia y notarás mayor latencia al usar el terminal y, por otro lado, no podrás ver cómo crece el grupo de auto-escalado (puesto que la instancia del grupo no estará sobrecargada). Además, tendrás que hacerlo tres veces puesto que la instancia tiene 2 vCPUs.

Deberás averiguar la IP pública de la instancia desplegada por el grupo de auto-escalado y conectarte mediante SSH:

Notificaciones de actividad (0)

Filtrar notificaciones

Enviar a

Actualmente no se especifican notificaciones

Crear notificación

Historial de actividad (1)

Filtrar historial de actividad

Estado	Descripción	Causa	Hora de inicio	Hora de finalización
Correcto	Launching a new EC2 instance: i-057c6028d481d2e9d	At 2023-10-04T18:05:34Z a user request created an AutoScalingGroup changing the desired capacity from 0 to 1. At 2023-10-04T18:05:37Z an instance was started in response to a difference between desired and actual capacity, increasing the capacity from 0 to 1.	2023 October 04, 08:05:39 PM +02:00	2023 October 04, 08:05:46 PM +02:00

Resumen de instancia de i-057c6028d481d2e9d

Detalles de la instancia

ID de la instancia: i-057c6028d481d2e9d

Dirección IPv4 pública: 3.89.96.245 [dirección abierta]

Dirección IPv4 privada: 172.31.38.157

DNS de IPv4 pública: ec2-3-89-96-245.compute-1.amazonaws.com [dirección abierta]

Estado de la instancia: En ejecución

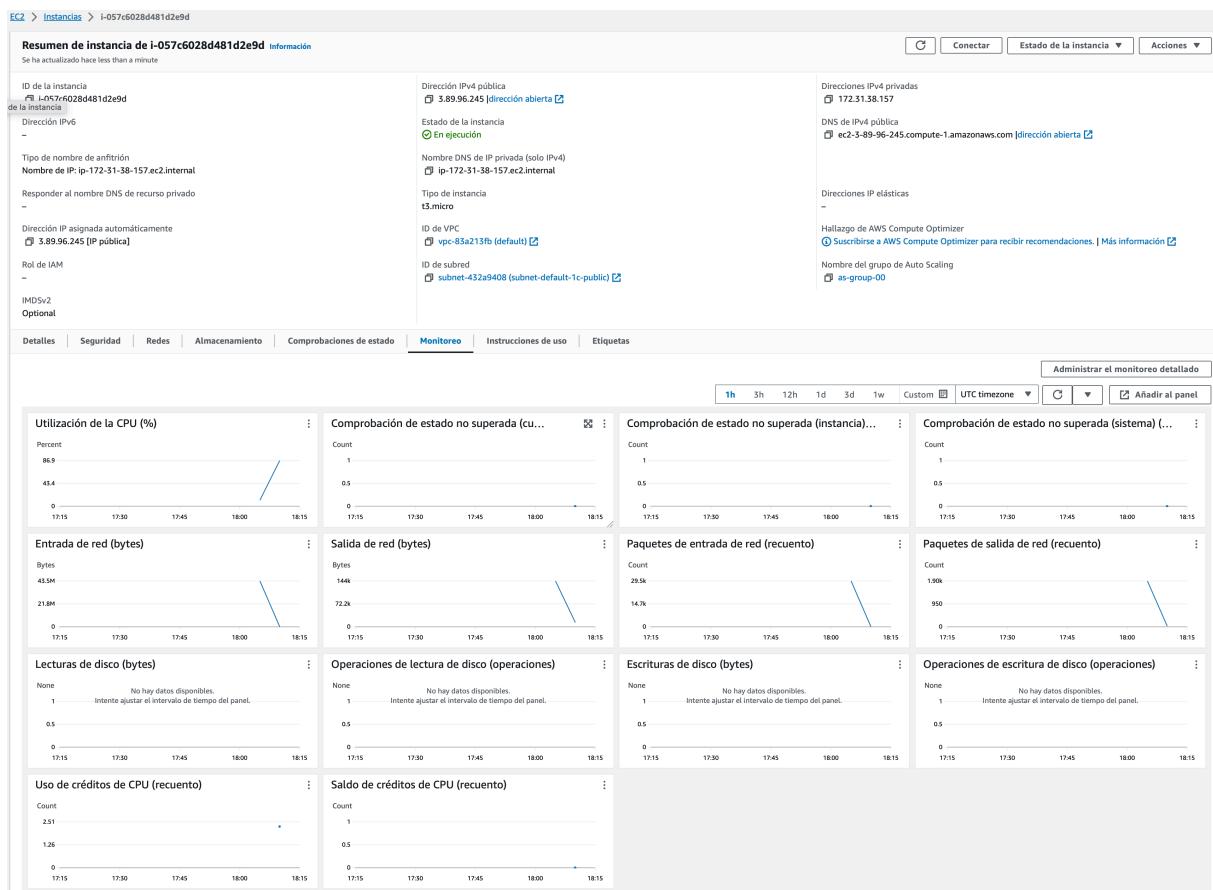
```

:~$ ssh -i alucloud$ID-priv.pem bitnami@3.89.96.245
bitnami@ip-10-235-37-12:~$ yes > /dev/null & disown
[1] 1842
bitnami@ip-10-235-37-12:~$ yes > /dev/null & disown
[1] 1843
bitnami@ip-10-235-37-12:~$ yes > /dev/null & disown
[1] 1844

```

Fíjate que ese comando provoca descartar la salida estándar del proceso (se redirige a `/dev/null`), para que no muestre nada por pantalla y además se desconecta el proceso del terminal, para que se ejecute en segundo plano. Para tu información, otra alternativa para aumentar el consumo de CPU de la instancia es utilizar el comando `stress` (`sudo apt-get install stress; stress --cpu 2 --timeout 120`).

Transcurridos unos minutos (dale un cierto margen ya que puede tardar un tiempo en recopilar las métricas de monitorización), podrás observar el aumento en el consumo de CPU tal y como se muestra en esta figura:



Te será de ayuda ir refrescando la interfaz con el siguiente botón para ver los cambios:



Al poco tiempo (unos 5-10 minutos) observarás la creación de una instancia adicional para tratar de satisfacer la métrica objetivo de que el promedio de uso de CPU esté al 70% (obviamente, en nuestro caso eso está limitado al tamaño máximo de grupo por lo que no se tendrán más de 2 instancias en marcha para este grupo de auto-escalado). Esta nueva instancia se integra automáticamente en el balanceador de carga, de manera que cualquier petición al puerto 80 del balanceador puede llegar a esa nueva instancia (o a cualquier otra del mismo grupo de auto-escalado). En efecto, el grupo de auto-escalado ajusta su capacidad de cómputo a las condiciones del entorno para tratar de satisfacer mejor la carga de trabajo.

Estado	Descripción	Causa	Hora de inicio	Hora de finalización
Esperando preparación de la instancia	Launching a new EC2 instance: i-0d2feffcd62cb3784	At 2023-10-04T18:14:39Z a monitor alarm TargetTracking-as-group-00-AlarmHigh-61e7699a-5661-4009-97c-b04c8d89894 in state ALARM triggered policy Target Tracking Policy changing the desired capacity from 1 to 2. At 2023-10-04T18:14:46Z an instance was started in response to a difference between desired and actual capacity, increasing the capacity from 1 to 2.	2023 October 04, 08:14:49 PM +02:00	
Correcto	Launching a new EC2 instance: i-0576028d481d2e9d	At 2023-10-04T18:05:34Z a user request created an AutoScalingGroup changing the desired capacity from 0 to 1. At 2023-10-04T18:05:37Z an instance was started in response to a difference between desired and actual capacity, increasing the capacity from 0 to 1.	2023 October 04, 08:05:39 PM +02:00	2023 October 04, 08:05:46 PM +02:00

Podemos ver cómo el grupo realiza el repliegue de instancias si matamos los procesos que están consumiendo la CPU (el comando yes) en la instancia inicial. Es posible hacerlo mediante un *killall*:

```
bitnami@ip-10-235-37-12:~$ killall -9 yes
```

Pasado un tiempo, dado que el promedio de uso de la CPU de las dos instancias se reduce por debajo del 70%, el grupo recupera su tamaño original (reduciendo en una instancia el tamaño del grupo) ya que la política de *scale in* consistía en eliminar una instancia del grupo de auto-escalado. La siguiente imagen resume este comportamiento.

Ten en cuenta que la eliminación de la instancia tardará unos cuantos minutos, por lo que si no quieres esperarte puedes pasar a la siguiente sección. Si vemos la actividad del grupo de auto-escalado podremos ver cómo una de las instancias se elimina del grupo de auto-escalado:

Estado	Descripción	Causa	Hora de inicio	Hora de finalización
Drenaje de conexiones en curso	Terminating EC2 instance: i-0576028d481d2e9d - Waiting For ELB Connection Draining.	At 2023-10-04T18:36:05Z a monitor alarm TargetTracking-as-group-00-AlarmLow-46a8185c-Baba-4715-b348-e54883293ac in state ALARM triggered policy Target Tracking Policy changing the desired capacity from 2 to 1. At 2023-10-04T18:36:14Z an instance was taken out of service in response to a difference between desired and actual capacity, shrinking the capacity from 2 to 1. At 2023-10-04T18:36:14Z instance i-0576028d481d2e9d was selected for termination.	2023 October 04, 08:36:14 PM +02:00	
Correcto	Launching a new EC2 instance: i-0d2feffcd62cb3784	At 2023-10-04T18:14:39Z a monitor alarm TargetTracking-as-group-00-AlarmHigh-61e7699a-5661-4009-97c-b04c8d89894 in state ALARM triggered policy Target Track Policy changing the desired capacity from 1 to 2. At 2023-10-04T18:14:46Z an instance was started in response to a difference between desired and actual capacity, increasing the capacity from 1 to 2.	2023 October 04, 08:14:49 PM +02:00	2023 October 04, 08:19:57 PM +02:00
Correcto	Launching a new EC2 instance: i-0576028d481d2e9d	At 2023-10-04T18:05:34Z a user request created an AutoScalingGroup changing the desired capacity from 0 to 1. At 2023-10-04T18:05:37Z an instance was started in response to a difference between desired and actual capacity, increasing the capacity from 0 to 1.	2023 October 04, 08:05:39 PM +02:00	2023 October 04, 08:05:46 PM +02:00

4. Finalización de Recursos

Para eliminar el grupo de auto-escalado no basta con terminar sus instancias, puesto que la configuración de tamaño mínimo de grupo a 1 provocará el despliegue de instancias adicionales para satisfacer dicha regla. Por lo tanto, hay que eliminar el grupo de auto-escalado seleccionado el grupo y pulsando el botón Acciones → Eliminar:

The screenshot shows the 'Grupos de Auto Scaling' (Auto Scaling Groups) page in the AWS Management Console. It displays one group named 'as-group-00'. The group has a launch template named 'launch-template-00' and two instances. The instance counts are 1 and 2, and the zone information is 'us-east-1c, us-east-1e'. The 'Actions' menu is open, and the 'Eliminar' (Delete) option is highlighted.

Confirma el borrado tecleando “eliminar”:

The screenshot shows the 'Eliminar grupo de Auto Scaling' (Delete Auto Scaling Group) confirmation dialog. It contains a warning message: 'El grupo de Auto Scaling contiene instancias en ejecución' (The Auto Scaling group contains running instances). It states that deleting the group will terminate all instances and that this action is不可逆 (irreversible). Below the message, it asks if the user is sure they want to delete the group. A list of groups is shown, with 'as-group-00' selected. A text input field contains the word 'eliminar'. At the bottom are 'Cancelar' (Cancel) and 'Eliminar' (Delete) buttons.

Ten en cuenta que este procedimiento se demorará unos minutos ya que se establece el tamaño deseado del grupo de auto-escalado a 0 para provocar la terminación de las instancias, antes de proceder al borrado del grupo de auto-escalado.

Posteriormente, borramos la plantilla de lanzamiento:

The screenshot shows the 'Plantillas de lanzamiento' (Launch Templates) page. It displays one template named 'launch-template-00'. The template has an ID of 'lt-0c0270f6ef75daef', version 1, and was created on '2023-10-04T17:52:21.000Z'. The 'Actions' menu is open, and the 'Eliminar plantilla' (Delete template) option is highlighted.

Al eliminar el grupo de auto-escalado se eliminan automáticamente las políticas asociadas al grupo de auto-escalado así como las alarmas asociadas a las políticas. Por tanto, no es necesario proceder a eliminarlas manualmente.

Ahora, eliminamos el balanceador de carga:

The screenshot shows the AWS Elastic Load Balancing (ELB) service page. At the top, it says 'Balanceadores de carga (1/1)'. Below that, there's a table with one row for 'alucloud00-lb'. The columns in the table are: Nombre (Name), Nombre de DNS (DNS Name), Estado (State), ID de VPC (VPC ID), Zonas de disponibilidad (Availability Zones), Tipo (Type), and Fecha creada (Created Date). The 'alucloud00-lb' entry has 'Activo' (Active) in the State column, 'vpc-83a213fb' in the VPC ID column, '2 Zonas de disponibilidad' (2 Availability Zones) in the Availability Zones column, 'application' in the Type column, and '4 de octubre de 2023, 19:35 (UTC+02:00)' in the Created Date column. On the right side of the table, there's a 'Actions' dropdown menu with options like 'Crear balanceador de carga' (Create load balancer), 'Editar el tipo de dirección IP' (Edit IP address type), 'Administrar subredes' (Manage subnets), etc. Below the table, there are tabs for 'Detalles' (Details), 'Agentes de escucha y reglas' (Listeners and rules), 'Mapeo de red' (Network mapping), 'Seguridad' (Security), 'Monitorización' (Monitoring), 'Integraciones' (Integrations), 'Atributos' (Attributes), and 'Etiquetas' (Tags).

Finalmente, eliminamos el “grupo de destino” (Target Group):

The screenshot shows the AWS Lambda service page. On the left, there's a sidebar with 'Panel de EC2', 'Grupos de destino', 'Vista global de EC2', 'Eventos', 'Console-to-Code', 'Vista previa', 'Instancias', 'Imágenes', 'Elastic Block Store', 'Red y seguridad', 'Equilibrio de carga', 'Balanceadores de carga', 'Grupos de destino' (which is selected and highlighted in blue), 'Trust Stores Nuevo', and 'Auto Scaling'. The main area is titled 'Grupos de destino (1/3) Info'. It shows a table with three rows: 'cellar-ecs' (ARN: arn:aws:elasticloadbalancing:eu-west-1:123456789012:targetgroup/cellarnode/6415e71738), 'k8s-backbone-backbone-6415e71738' (ARN: arn:aws:elasticloadbalancing:eu-west-1:123456789012:targetgroup/k8snode/6415e71738), and 'lb-00-tg' (ARN: arn:aws:elasticloadbalancing:eu-west-1:123456789012:targetgroup/lb-00-tg). The 'lb-00-tg' row has a checked checkbox next to it. On the right side of the table, there's a 'Actions' dropdown menu with options like 'Crear un grupo de destino' (Create target group), 'Eliminar' (Delete), 'Registrar destinos' (Register targets), 'Editar la configuración de comprobación de estado' (Edit state check configuration), etc. Below the table, there are tabs for 'Info', 'Listeners', 'Targets', and 'Metrics'.

Por favor, **asegúrate de haber eliminado todos los recursos** (especialmente el balanceador de carga) antes de pasar a la siguiente sección.

7. Monitorización y Gestión

A continuación, se resume alguna funcionalidad interesante que puede ser útil durante el ciclo de vida de las instancias.

7.1. Monitorización y Alarmas

Es posible activar la monitorización automática de una instancia de EC2 mediante el servicio de Amazon CloudWatch [15] y combinarlo con el uso de notificaciones del servicio Amazon SNS (Simple Notification Service) [16] para disponer de un sistema de alertas sobre el estado de la instancia de EC2.

Por ejemplo, a continuación, se describe un procedimiento para recibir un aviso por correo electrónico cuando el promedio de uso de CPU de una instancia de EC2 sea mayor o igual que el 70% durante al menos 2 períodos de 5 minutos. Esto puede permitir al usuario realizar una operación de escalado vertical (deteniendo y cambiando el tipo de instancia a una con mayor capacidad).

En primer lugar, tenemos que crear un topic de Amazon SNS para que la alarma pueda mandar un mensaje a dicho topic, lo que generará un aviso por correo electrónico. Entra en el servicio Amazon SNS y crea un topic llamado NotifyMe_alucloudX donde XX es tu identificador de alumno. Asegúrate de respetar la nomenclatura o no podrás crearlo.

Name	ARN
NotifyMe-srisco	arn:aws:sns:us-east-1:974349055189:NotifyMe-srisco
NotifyMe_CursoCloudAWS	arn:aws:sns:us-east-1:974349055189:NotifyMe_CursoCloudAWS
NotifyMe_alucloud00	arn:aws:sns:us-east-1:974349055189:NotifyMe_alucloud00
amplify_codecommit_topic	arn:aws:sns:us-east-1:974349055189:amplify_codecommit_topic
dynamodb	arn:aws:sns:us-east-1:974349055189:dynamodb

Details

Type [Info](#)
Topic type cannot be modified after topic is created

FIFO (first-in, first-out)

- Strictly-preserved message ordering
- Exactly-once message delivery
- High throughput, up to 300 publishes/second
- Subscription protocols: SQS, Lambda, HTTP, SMS, email, mobile application endpoints

Standard

- Best-effort message ordering
- At-least once message delivery
- Highest throughput in publishes/second
- Subscription protocols: SQS, Lambda, HTTP, SMS, email, mobile application endpoints

Name

Maximum 256 characters. Can include alphanumeric characters, hyphens (-) and underscores (_).

Display name - optional
To use this topic with SMS subscriptions, enter a display name. Only the first 10 characters are displayed in an SMS message. [Info](#)

Maximum 100 characters, including hyphens (-) and underscores (_).

El resto de opciones de configuración del topic puedes dejarlas en sus valores por defecto:

Si obtienes un error como el siguiente al crear el topic es que no has respetado el esquema de nombrado:

Couldnt create topic.
Error code: AuthorizationError - Error message: User: arn:aws:iam::974349055189:user/alucloud00 is not authorized to perform: SNS:CreateTopic on resource: arn:aws:sns:us-east-1:974349055189:Notify-Me-alucloud78

A continuación, deberás crear una subscripción para que cuando dicho topic reciba una notificación, recibas un correo electrónico. Elige tu topic e introduce tu correo electrónico:

Amazon SNS > Subscriptions > Create subscription

Create subscription

Details

Topic ARN
arn:aws:sns:us-east-1:974349055189:Notify

Protocol
The type of endpoint to subscribe
Email

Endpoint
An email address that can receive notifications from Amazon SNS.
gmolto@dsic.upv.es

ⓘ After your subscription is created, you must confirm it. [Info](#)

► **Subscription filter policy - optional**
This policy filters the messages that a subscriber receives. [Info](#)

► **Redrive policy (dead-letter queue) - optional**
Send undeliverable messages to a dead-letter queue. [Info](#)

[Cancel](#) [Create subscription](#)

Al crear la subscripción recibirás un e-mail de confirmación que deberás confirmar:

AWS Notifications ✉ CursoCloudAWS 19:58

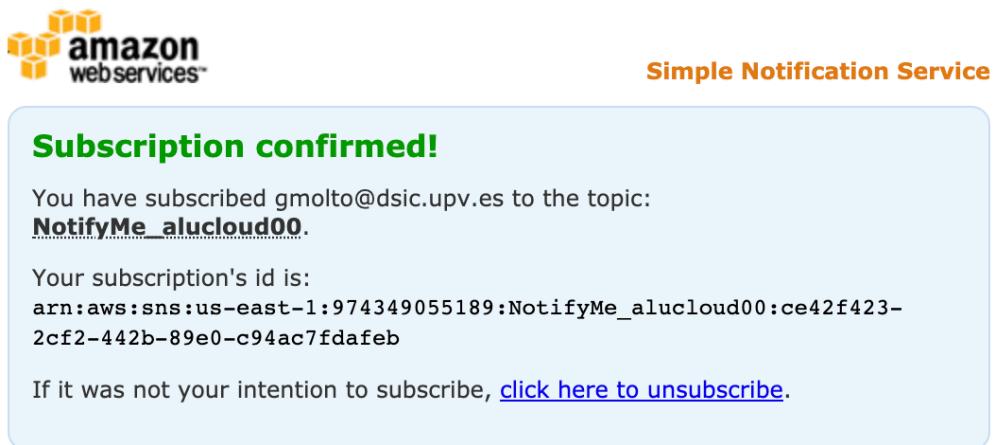
AWS Notification - Subscription Confirmation

Para: Germán Moltó Martínez

You have chosen to subscribe to the topic:
arn:aws:sns:us-east-1:974349055189:NotifyMe_alucloud00

To confirm this subscription, click or visit the link below (If this was in error no action is necessary):
[Confirm subscription](#)

Please do not reply directly to this email. If you wish to remove yourself from receiving all future SNS subscription confirmation requests please send an email to [sns-opt-out](#)



Ahora, elegimos una instancia de EC2 en ejecución y con el botón derecho, elegimos la opción “Actions → Monitor and troubleshoot → Manage CloudWatch alarms”.

Instances (1/19) Info					Connect	Instance state ▾	Actions ▾	Launch instances	▼
<input type="text"/> Filter instances									
■	Name	▼	Instance ID	Instance state ▾	Instance type	Actions	Launch instances	▼	
<input checked="" type="checkbox"/>	alocloud47 - stack		i-092f2e3f4a3dba8fe	Running ⓘ ⓘ	t2.micro	Connect	Launch instances	▼	
<input type="checkbox"/>	cursocloudaws1920		i-014d53ba18f9e1425	Running ⓘ ⓘ	t3.micro	View details	1 >	alarm status	+ Create alarm
<input type="checkbox"/>	cursocloudaws2021		i-0b9628714385d9f4a	Running ⓘ ⓘ	t3.micro	Manage instance state	alarm status	+ Create alarm	
<input type="checkbox"/>	vpc-nat00		i-066fd20c61cc095ba	Running ⓘ ⓘ	t3.micro	Instance settings	alarm status	+ Create alarm	
<input type="checkbox"/>	alucloud92		i-0ea232f9d8433b774	Running ⓘ ⓘ	t3.micro	Networking	alarm status	+ Create alarm	
<input type="checkbox"/>	-		i-09eb761b0f85bce2e	Running ⓘ ⓘ	t3.micro	Security	alarm status	+ Create alarm	
						ClassicLink	alarm status	+ Create alarm	
						Image and templates	alarm status	+ Create alarm	
						Manage detailed monitoring	alarm status	+ Create alarm	
						Manage CloudWatch alarms	Monitor and troubleshoot	alarm status	+ Create alarm

Luego, elegimos “Create a new alarm” y en “Alarm notification” introducimos el nombre del topic de SNS que habíamos creado previamente (NotifyMe_alucloudXX donde XX es tu identificador de alumno):

The screenshot shows the 'Manage CloudWatch alarms' page for an EC2 instance. It includes sections for adding a new alarm, searching for existing alarms, and configuring alarm notifications. The 'Alarm notification' section is expanded, showing a search bar with 'NotifyMe_alucloud00' selected.

Add or edit alarm Info
Indicate whether to create a new alarm or select an existing alarm to edit.

Create a new alarm
Create a new alarm for i-08a596308e7b4ea71

Edit an existing alarm
Edit an existing alarm for i-08a596308e7b4ea71

Search for alarm
Find an alarm to modify
Select an existing alarm to edit

Alarm notification Info
Configure the alarm to send notifications to an Amazon SNS topic when it is triggered.

NotifyMe_alucloud00

Alarm thresholds
Specify the metric thresholds for the alarm.

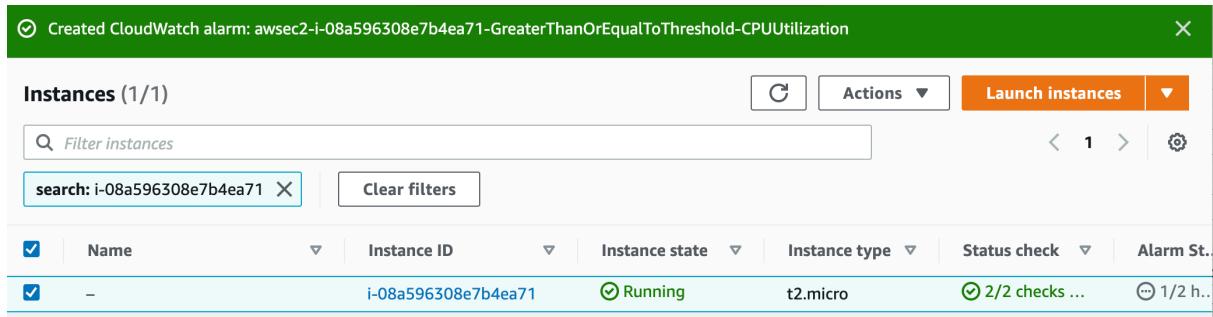
Group samples by: Average
Type of data to sample: CPU Utilization

Alarm When: >= 70 Percent

Consecutive Periods: 2 Period: 5 Minutes

Alarm name: awsec2-i-08a596308e7b4ea71-GreaterThanOrEqualToThreshold-CPUUtilization

Pulsa el botón “Create” y obtendrás un mensaje de confirmación:

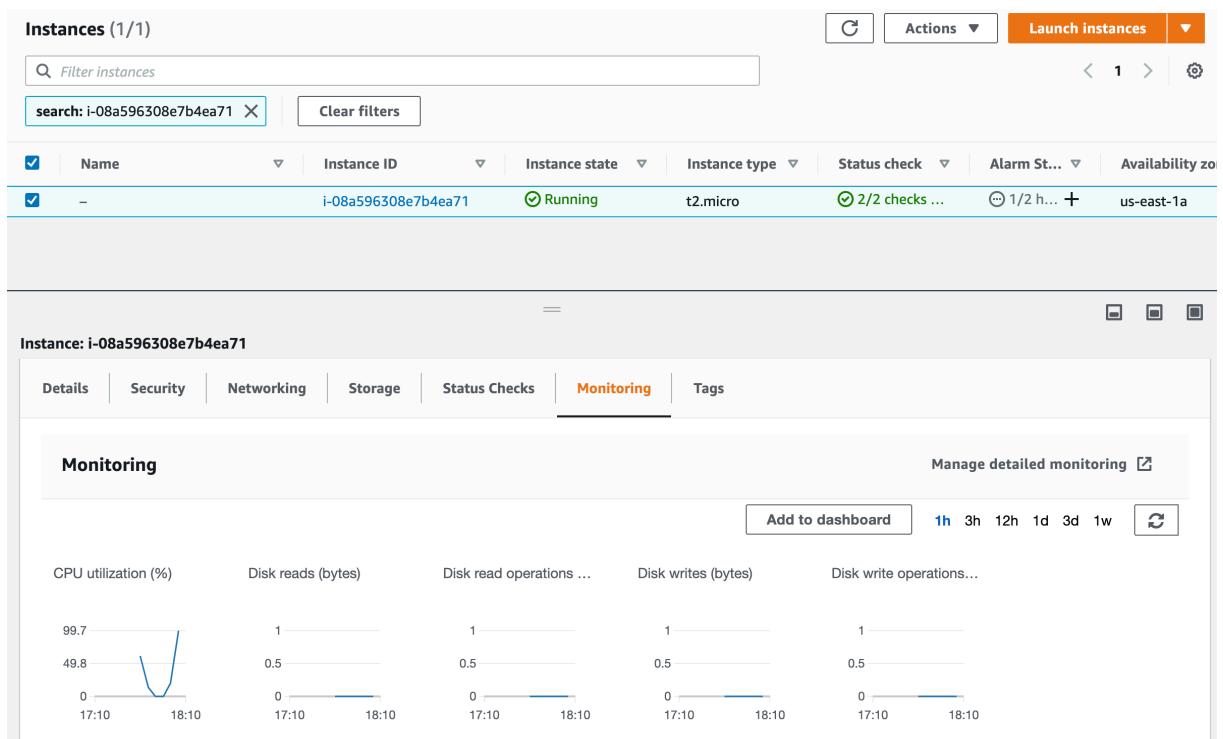


Para verificar si la alarma funciona correctamente, puedes forzar a incrementar el uso de CPU de la instancia y esperar a que la alarma se dispare. Para ello, conéctate a la instancia por SSH y ejecuta el siguiente comando:

```
(Asegúrate de estar conectado a la instancia de EC2 por SSH antes de ejecutar este comando y bajo ningún concepto lo ejecutes en la máquina de prácticas)
:~$ yes > /dev/null &
(Puedes comprobar el uso de CPU con el siguiente comando)
:~$ top
(Para salir de esta herramienta pulsa la letra q)
```

Ese comando permite incrementar el uso de CPU al 100% y permite que la instancia alcance en 10 minutos las condiciones necesarias para que se dispare la alarma.

Desde el EC2 Dashboard, pulsando sobre la instancia de EC2 y eligiendo la pestaña *Monitoring* es posible observar el incremento en consumo de CPU (gráfica *CPU Utilization (%)*):



Finalmente, es posible observar como se activa la alarma de la instancia, al cumplirse las reglas fijadas para la misma:

Instances (1)							
<input type="button" value="Actions"/> <input type="button" value="Launch instances"/>							
<input type="text" value="Filter instances"/> <input type="button" value="Clear filters"/>							
Name	Instance ID	Instance state	Instance type	Status check	Alarm Status	Avg	us-
-	i-08a596308e7b4ea71	Running	t2.micro	2/2 checks ...	1/1 in alarm	+ us-	

Y se recibe el correspondiente aviso por correo electrónico:

De: AWS Notifications <no-reply@sns.amazonaws.com>
 Asunto: ALARM: "awsec2-i-1e0af763-High-CPU-Utilization" in US - N. Virginia
 Fecha: 11 de octubre de 2013 12:31:37 GMT+02:00
 Para: Germán Moltó Martínez <gmolto@dsic.upv.es>

Ocular

You are receiving this email because your Amazon CloudWatch Alarm "awsec2-i-1e0af763-High-CPU-Utilization" in the US - N. Virginia region has entered the ALARM state, because "Threshold Crossed: 2 datapoints were greater than or equal to the threshold (70.0). The most recent datapoints: [99.868, 99.606]" at "Friday 11 October, 2013 10:31:37 UTC".

View this alarm in the AWS Management Console:
<https://console.aws.amazon.com/cloudwatch/home?region=us-east-1#s=Alarms&alarm=awsec2-i-1e0af763-High-CPU-Utilization>

Alarm Details:

- Name: awsec2-i-1e0af763-High-CPU-Utilization
- Description: Created from EC2 Console
- State Change: OK -> ALARM
- Reason for State Change: Threshold Crossed: 2 datapoints were greater than or equal to the threshold (70.0). The most recent datapoints: [99.868, 99.606].
- Timestamp: Friday 11 October, 2013 10:31:37 UTC
- AWS Account: 974349055189

Threshold:

- The alarm is in the ALARM state when the metric is GreaterThanOrEqualToThreshold 70.0 for 300 seconds.

Monitored Metric:

- MetricNamespace: AWS/EC2
- MetricName: CPUUtilization
- Dimensions: [InstanceId = i-1e0af763]
- Period: 300 seconds
- Statistic: Average
- Unit: not specified

State Change Actions:

- OK:
- ALARM: [arn:aws:sns:us-east-1:974349055189:NotifyMe_aulcloud00]
- INSUFFICIENT_DATA:

Si no recibes la notificación, asegúrate de que la alarma está configurada para avisar al topic de SNS (apartado Actions):

CloudWatch > Alarms > awsec2-i-08a596308e7b4ea71-GreaterThanOrEqualToThreshold-CPUUtilization

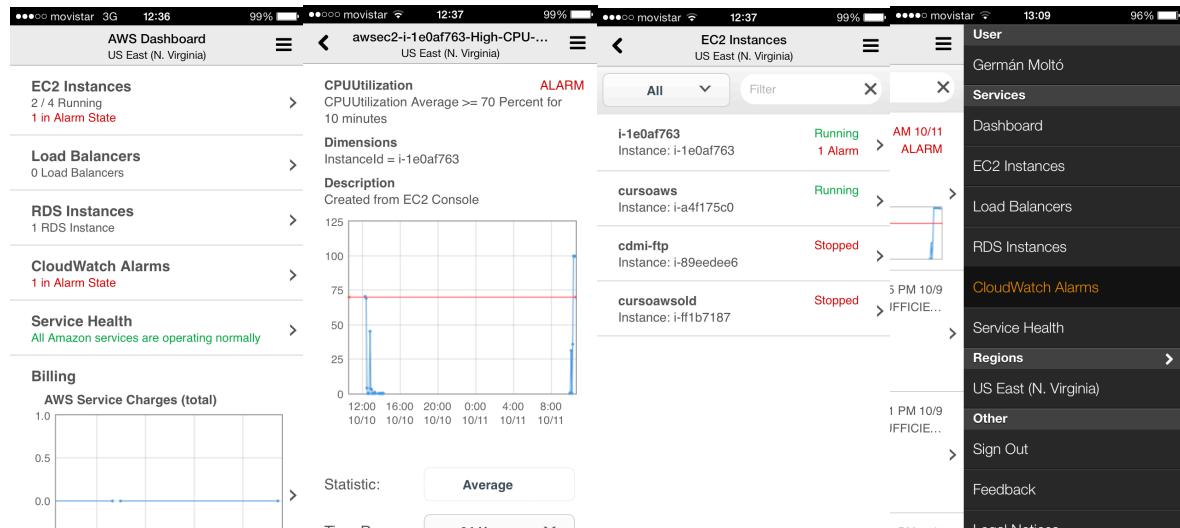
awsec2-i-08a596308e7b4ea71-GreaterThanOrEqualToThreshold-CPUUtilization

Name	Status	Metric Namespace	Datapoints to alarm
awsec2-i-08a596308e7b4ea71-GreaterThanOrEqualToThreshold-CPUUtilization	In alarm	AWS/EC2	2 out of 2
		Metric name	Missing data treatment
		CPUUtilization	Treat missing data as missing
Type	Threshold	Instanced	Percentiles with low samples evaluate
Metric alarm	CPUUtilization >= 70 for 2 datapoints within 10 minutes	i-08a596308e7b4ea71	ARN
Description	Last change	Instance name	arnawscloudwatch:us-east-1:974349055189:alarm:awsec2-i-08a596308e7b4ea71-GreaterThanOrEqualToThreshold-CPUUtilization
Created from EC2 Console	2020-09-02 18:19:12	No name specified	
		Statistic	
		Average	
		Period	
		5 minutes	

Actions

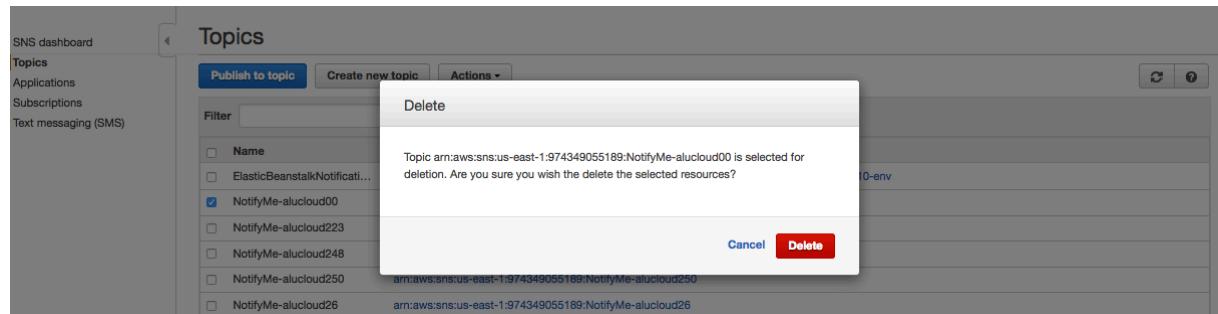
Type	Description	Config
Notification	When in alarm, send message to topic "NotifyMe_aulcloud00"	-

También es posible monitorizar y gestionar parcialmente la infraestructura desplegada en AWS con alguna aplicación móvil. Por ejemplo, a continuación, se muestran capturas de pantalla de la monitorización de la instancia usando la app para iOS/Android AWS Console [17].

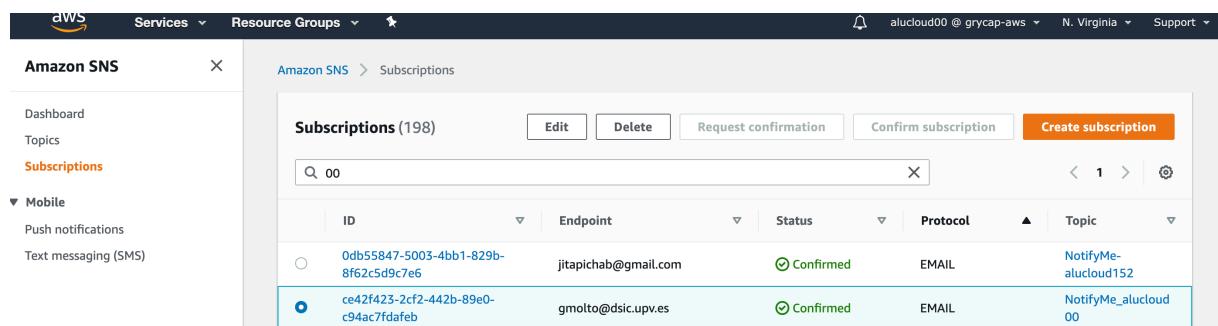


Desde tu móvil puedes detener e iniciar instancias de EC2 existentes, monitorizar alarmas de CloudWatch, gestionar instancias de Amazon RDS, etc.

Cuando termines la instancia provocará la eliminación automática de la(s) alarma(s) asociadas. Además, deberás entrar en la consola de Amazon SNS, elijas el topic NotifyMe-alucloudXX y lo elimines:



Recuerda también borrar la subscripción:





Antes de dar la práctica por finalizada, asegúrate de que has terminado/eliminado todos los recursos aprovisionados con tu cuenta. En particular revisa que hayas eliminado:

- Grupos de auto-escalado
- Configuraciones de lanzamiento
- Balanceadores de carga
- Target Group
- Topic de SNS
- Suscripción

Recuerda que puedes conectarte a la AWS Management Console para ver (y luego terminar) aquellos recursos que incluyen en su nombre tu identificador de usuario. ¡Gracias por tu colaboración!

8. Conclusiones

En esta práctica se han utilizado los servicios más relevantes de Amazon Web Services para conseguir el despliegue de infraestructuras virtuales bajo demanda en la forma de instancias. Se ha utilizado también el servicio de S3 para gestionar datos en el Cloud. Además, se han utilizado mecanismos de balanceo de carga mediante el servicio ELB. Esto ha permitido la distribución de peticiones entrantes entre diferentes zonas de disponibilidad. Finalmente, se ha utilizado Auto Scaling para la creación de grupos elásticos de instancias cuyo tamaño ha variado en función de la carga de trabajo de las instancias.

Información Adicional

Esta práctica se realiza en el marco del “Curso Online de Cloud Computing con Amazon Web Services”, ofertado por el Institut de Instrumentació per a l'Imatge Molecular de la Universitat Politècnica de València. Tienes más información sobre este curso de formación en Cloud Computing en la siguiente dirección: <http://www.grycap.upv.es/cursocloudaws>

Referencias

- [1] Amazon Web Services (AWS). <http://aws.amazon.com/es/>
- [2] Jurg van Vliet, Flavia Paganelli. “Programming Amazon EC2”, O'Reilly, 2011.
- [3] Amazon EC2 API Tools. <http://aws.amazon.com/developertools/351>
- [4] Amazon Machine Images (AMIs). <https://aws.amazon.com/amis>
- [5] Amazon Web Services Documentation. <http://aws.amazon.com/es/documentation/>
- [6] Aws: Simple Command-Line Access to Amazon EC2 and Amazon S3. <http://timkay.com/aws/>
- [7] cURL. <http://curl.haxx.se/>
- [8] <http://cloud-images.ubuntu.com/locator/ec2/>
- [9] Auto Scaling Command Line Tool. <http://aws.amazon.com/developertools/2535>
- [10] Amazon CloudWatch Command Line Tool. <http://aws.amazon.com/developertools/2534>
- [11] Puppet. <https://puppetlabs.com/>
- [12] Chef. <http://www.opscode.com/chef/>
- [13] AWS OpsWorks. <https://aws.amazon.com/es/opsworks/>
- [14] Expanding the Storage Space of a Volume
<http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/recognize-expanded-volume-linux.html>
- [15] Amazon CloudWatch. <http://aws.amazon.com/es/cloudwatch/>
- [16] Amazon Simple Notification Service (SNS). <http://aws.amazon.com/es/sns/>
- [17] AWS Management Console App. <http://aws.amazon.com/es/console/mobile/>

- [18] LAMP Stack Powered by Bitnami
https://aws.amazon.com/marketplace/pp/Bo07IN7GJA/ref=srh_res_product_title?ie=UTF8&sr=0-2&qid=1392114825842
- [19] Attaching an Amazon EBS Volume to an Instance.
<http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ebs-attaching-volume.html>
- [20] AWS CLI. <http://aws.amazon.com/es/cli/>
- [21] AWS CLI cp documentation. <http://docs.aws.amazon.com/cli/latest/reference/s3/cp.html>
- [22] Expanding a Linux Partition.
http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/storage_expand_partition.html
- [23] Ansible. <http://www.ansible.com/home>
- [24] Creating an Amazon EBS-Backed Linux AMI.
<http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/creating-an-ami-ebs.html>
- [25] EC2 create-image Does Not Fully "Stop" The Instance. <http://alestic.com/2014/02/ec2-create-image-reboot>
- [26] Paid AMIs. <http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/paid-amis.html>
- [27] Elastic Load Balancing. <http://aws.amazon.com/documentation/elasticloadbalancing/>
- [28] Auto Scaling. <http://aws.amazon.com/es/autoscaling/>
- [29] Managing Bucket Logging.
<http://docs.aws.amazon.com/AmazonS3/latest/UG/ManagingBucketLogging.html>
- [30] Server Access Log Format.
<http://docs.aws.amazon.com/AmazonS3/latest/dev/LogFormat.html>
- [31] Amazon Simple Notification Service (SNS). <http://aws.amazon.com/es/sns/>
- [32] Using Reduced Redundancy Storage.
<http://docs.aws.amazon.com/AmazonS3/latest/dev/UsingRRS.html>
- [33] Amazon Glacier. <http://aws.amazon.com/es/glacier/>
- [34] Using Server-Side Encryption.
<http://docs.aws.amazon.com/AmazonS3/latest/dev/UsingServerSideEncryption.html>
- [35] Amazon EBS Provisioned IOPS Volumes. <http://aws.amazon.com/ebs/details/>
- [36] Gibibyte. <http://es.wikipedia.org/wiki/Gibibyte>
- [37] Monitoring Your Instances with CloudWatch.
<http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/using-cloudwatch.html>
- [38] Using Your Own Linux Kernels.
<http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/UserProvidedKernels.html>
- [39] Tagging Your Amazon EC2 Resources.
http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/Using_Tags.html
- [40] Health Check.
http://docs.aws.amazon.com/AutoScaling/latest/DeveloperGuide/AS_Concepts.html#healthcheck
- [41] Amazon EBS Volume Types.
<http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/EBSVolumeTypes.html>
- [42] Scaling Policy Types. <https://docs.aws.amazon.com/AutoScaling/latest/DeveloperGuide/as-scale-based-on-demand.html#as-scaling-types>
- [43] Configure Security Groups for Your Load Balancer.
<http://docs.aws.amazon.com/ElasticLoadBalancing/latest/DeveloperGuide/elb-security-groups.html>
- [44] Create an HTTPS Load Balancer.
<http://docs.aws.amazon.com/ElasticLoadBalancing/latest/DeveloperGuide/elb-create-https-ssl-load-balancer.html>

- [45] Configure Cross-Zone Load Balancing for Your Load Balancer.
<http://docs.aws.amazon.com/ElasticLoadBalancing/latest/DeveloperGuide/enable-disable-crosszone-lb.html>
- [46] Configure Connection Draining for Your Load Balancer.
<http://docs.aws.amazon.com/ElasticLoadBalancing/latest/DeveloperGuide/config-conn-drain.html>
- [47] EC2 Can't resize volume after increasing size.
<http://stackoverflow.com/questions/11014584/ec2-cant-resize-volume-after-increasing-size>
- [48] Amazon EBS Update – New Cold Storage and Throughput Options.
<https://aws.amazon.com/es/blogs/aws/amazon-ebs-update-new-cold-storage-and-throughput-options/>
- [49] New – AWS Application Load Balancer. <https://aws.amazon.com/es/blogs/aws/new-aws-application-load-balancer/>
- [50] Amazon SQS. <https://aws.amazon.com/es/sqs/>
- [51] AWS Lambda. <https://aws.amazon.com/es/lambda/details/>
- [52] Billing and Reporting of Buckets.
<https://docs.aws.amazon.com/AmazonS3/latest/dev/BucketBilling.html>
- [53] Cross-Region Replication. <https://docs.aws.amazon.com/AmazonS3/latest/dev/crr.html>
- [54] Requester Pays Buckets.
<https://docs.aws.amazon.com/AmazonS3/latest/dev/RequesterPaysBuckets.html>
- [55] Amazon S3 Transfer Acceleration.
<https://docs.aws.amazon.com/AmazonS3/latest/dev/transfer-acceleration.html>
- [56] How do I configure storage class analysis.
<https://docs.aws.amazon.com/AmazonS3/latest/user-guide/configure-analytics-storage-class.html>
- [57] How do I configure a Request Metrics Filter?
<https://docs.aws.amazon.com/AmazonS3/latest/user-guide/configure-metrics-filter.html>
- [58] Amazon S3 Storage Inventory. <http://docs.aws.amazon.com/AmazonS3/latest/dev/storage-inventory.html>
- [59] Modifying the Size, IOPS, or Type of an EBS Volume on Linux.
<http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ebs-expand-volume.html>
- [60] Why so many weird files on my Amazon S3 account that I didn't upload?
<https://forums.aws.amazon.com/thread.jspa?threadID=71680>
- [61] Target Groups for Your Application Load Balancers.
<http://docs.aws.amazon.com/elasticloadbalancing/latest/application/load-balancer-target-groups.html>
- [62] Storage Classes. <https://docs.aws.amazon.com/AmazonS3/latest/dev/storage-class-intro.html>
- [63] Cómo habilito o suspendo el control de versiones para un bucket de S3?
https://docs.aws.amazon.com/es_es/AmazonS3/latest/user-guide/enable-versioning.html
- [64] Registro de acceso al servidor de Amazon S3.
https://docs.aws.amazon.com/es_es/AmazonS3/latest/dev/ServerLogs.html
- [65] Uso de etiquetas de buckets de S3 de asignación de costos.
https://docs.aws.amazon.com/es_es/AmazonS3/latest/dev/CostAllocTagging.html
- [66] Registro de llamadas a la API de Amazon S3 mediante AWS CloudTrail.
https://docs.aws.amazon.com/es_es/AmazonS3/latest/dev/cloudtrail-logging.html
- [67] Protección de datos mediante cifrado.
https://docs.aws.amazon.com/es_es/AmazonS3/latest/dev/UsingEncryption.html
- [68] Monitorización de métricas con Amazon CloudWatch.
https://docs.aws.amazon.com/es_es/AmazonS3/latest/dev/cloudwatch-monitoring.html

- [69] Amazon Route 53. <https://aws.amazon.com/es/route53/>
- [70] Roles de IAM. https://docs.aws.amazon.com/es_es/IAM/latest/UserGuide/id_roles.html
- [71] Cross-Origin Resource Sharing (CORS, Uso compartido de recursos entre orígenes).
https://docs.aws.amazon.com/es_es/AmazonS3/latest/dev/cors.html
- [72] Dark Cloud. Inside The Pentagon's Leaked Internet Surveillance Archive).
<https://www.upguard.com/breaches/cloud-leak-centcom>
- [73] Cost of Storing AMI. <https://stackoverflow.com/questions/18650697/cost-of-storing-ami>
- [74] Políticas y permisos:
https://docs.aws.amazon.com/es_es/IAM/latest/UserGuide/access_policies.html
- [75] Locking Objects Using Amazon S3 Object Lock.
<https://docs.aws.amazon.com/AmazonS3/latest/dev/object-lock.html>
- [76] What is AWS Global Accelerator. https://docs.aws.amazon.com/es_es/global-accelerator/latest/dg/what-is-global-accelerator.html
- [77] Políticas de escalado de seguimiento de destino para Auto Scaling de aplicaciones.
https://docs.aws.amazon.com/es_es/autoscaling/application/userguide/application-auto-scaling-target-tracking.html
- [78] Políticas de escalado por pasos para Auto Scaling de aplicaciones.
https://docs.aws.amazon.com/es_es/autoscaling/application/userguide/application-auto-scaling-step-scaling-policies.html
- [79] Amazon EBS y NVMe en instancias de Linux.
https://docs.aws.amazon.com/es_es/AWSEC2/latest/UserGuide/nvme-ebs-volumes.html
- [80] Instance maintenance policies.
https://docs.aws.amazon.com/es_es/autoscaling/ec2/userguide/ec2-auto-scaling-instance-maintenance-policy.html
- [81] Storage class for automatically optimizing data with changing or unknown access patterns.
https://docs.aws.amazon.com/AmazonS3/latest/userguide/storage-class-intro.html?icmpid=docs_amazons3_console#sc-dynamic-data-access