Despliegue Cloud

Despliegue aplicación play-products-manager en aws

Celia Herrera Ferreira cherrerafe.inf@upsa.es Marzo 2021

Descripción del sistema

En esta parte de la asignatura se va a utilizar el proveedor AWS para desplegar la aplicación que se creó en la anterior parte de la asignatura, en mi caso, la API REST play-products-manager desarrollada con el framework play.

Para ello se requerirá la utilización de herramientas como Packer y Terraform tanto para la creación y despliegue de la base de datos como del propio API REST. Se utilizarán también diferentes servicios ofrecidos por AWS, como son el servicio de bases de datos relacionales (RDS) y también el servicio EC2.

El despliegue se ha hecho en la zona de disponibilidad de París (euwest-3).

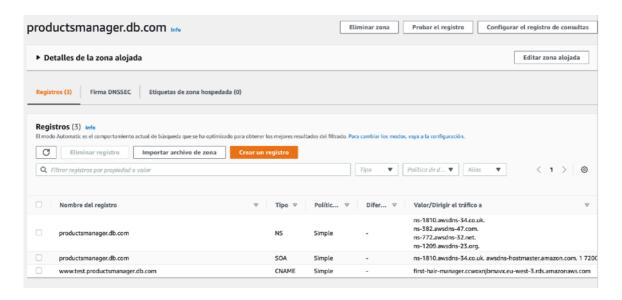
Base de datos

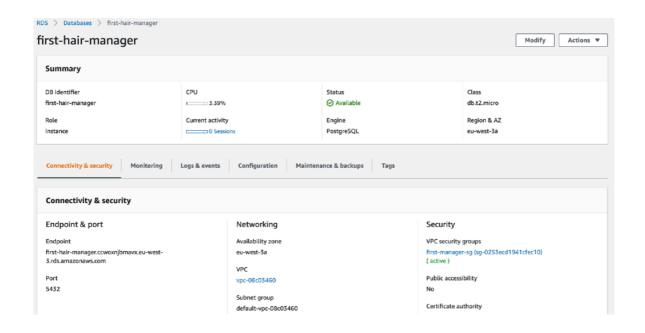
Para poder realizar el despliegue completo de la aplicación, es necesario crear una base de datos que permita la persistencia de los datos que se van a guardar y gestionar en la aplicación. Esta base de datos, que en nuestro caso será una base de datos Postgres, se creará automáticamente utilizando Terraform, para ello se hace uso del script de Terraform existente en la carpeta RDS, el archivo *main.tf*.

En este fichero se configura que la base de datos sea de tipo Postgres y la versión (en este caso la 12.5), que tenga 20gb de memoria, el tipo de la instancia que en nuestro caso utilizamos la opción gratuita de db.t2.micro, además de indicar el usuario (postgres) y la contraseña de ese usuario (celiar00t). Por otro lado se configura el puerto de entrada (5432) y también se añade un grupo de seguridad, que en este caso permite el acceso a la instancia por el puerto 5432 de conexión TCP. Además se incluye en el fichero la configuración del servicio route53 para esta instancia, en el que se indica el

nombre de la zona primaria del la Hosted Zone además del nombre que se le va a dar a la dns estática, ese record es de tipo CNAME que es el que permite el renombramiento de la dns creada por la base de datos por el que se configura en este archivo.

En él está comprendida toda la creación de la instancia de base de datos. Esta base de datos se creará con un endpoint dado de manera automática por Amazon, el cual, posteriormente habrá que indicar en el despliegue de play a donde deberá apuntar. Para evitar este cambio manual se ha configurado el servicio ofrecido por Amazon de **route53**, con este servicio se pretende crear una Hosted Zone en la que se enlaza el endpoint generado de la base de datos por una dns estática a la que siempre se apuntaría en la aplicación. Aunque se ha llegado a poder crear la Hosted Zone, no ha sido posible acceder a esa dns estática. Esto es debido a que esa dns estática, aún siendo pública, debería de estar alojada en un dominio público, el cual solo es posible utilizarlo previo pago. Debido a esto, aunque se añade el desarrollo de la creación de la Hosted Zone, no se ha terminado realizando la conexión de esta manera, y a la hora del despliegue de la aplicación habrá que cambiar la url a la que apuntar.







Posteriormente a la creación de la instancia RDS, habrá que generar la base de datos de manera manual. Para ello, habrá que acceder a la instancia creada mediante un cliente de Postgres, con el cual se creará la base de datos en ella, con el nombre de *productsmanager* y tras ello se deberá proceder a la introducción de todas las tablas de la base de datos, así como sus relaciones. Estas están contenidas en el archivo *ProductsManager.sql* contenido en la misma carpeta que el archivo *main.tf*.

Aplicación

Tras haber realizado el despliegue y configuración de la base de datos se procederá a desplegar la aplicación. Para realizar dicho despliegue lo primero que se hará será crear una AMI, es decir, una imagen de una máquina de Amazon, en nuestro caso de una instancia EC2. Esta AMI se creará de manera automática mediante la utilización de Packer, utilizando un fichero contenido en la carpeta de Aplicación, en concreto el fichero *play-products-manager.json*.

Antes de proceder a ejecutarlo habrá que editarlo cambiando el endpoint de la base de datos que se nos ha creado automáticamente en el proceso anterior.

La parte señalada es el endpoint a cambiar

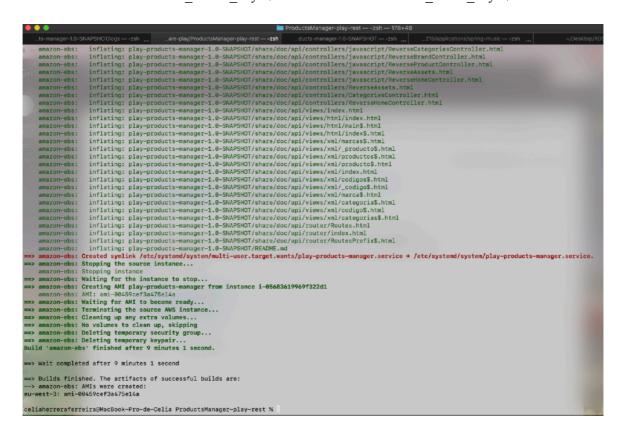
Este cambio es necesario ya que cuando se ejecute Packer este valor cambiará dentro de la configuración de de la aplicación Play para que de esta

manera se apunte correctamente a la base de datos. En mi caso he realizado el despliegue en la zona de disponibilidad de París (eu-west-3) por lo que la source_ami introducido depende de la instancia que tomé como base en esa zona de disponibilidad.

Dentro de este fichero play-products-manager.json se puede observar que está configurada la subida de un servicio (play-products-manager.service) que permitirá la ejecución de la automática de la aplicación y que se encuentra también en la carpeta, la aplicación propiamente dicha comprimida en formato zip (ya que aunque el desarrollo sea en Java, Play es un framework basado en Scala) y el script start.sh (también contenido en la carpeta) que realizará la ejecución y correcta configuración del puerto cuando el servicio (play-products-manager.service) inicie el despliegue. Se puede comprobar que tras la subida están todos los pasos necesarios descritos para realizar la correcta configuración de la imagen para que el despliegue de ella sea automático.

Para ejecutar este archivo se deberá ejecutar introduciendo las variables tanto de la clave privada de aws como de la clave secreta con el comando:

Packer build -var 'aws_access_key=\$ACCESSKEY' -var 'aws_secret_key=\$SECRETKEY'



Tras esta ejecución se procede a ejecutar el fichero de Terraform (main.tf), el cual se encuentra también contenido en la carpeta y que habrá que editar para introducir el id de la AMI que se haya creado y sustituirlo en la variable "image_id" para que la tome de referencia, que será distinta en cada ejecución.

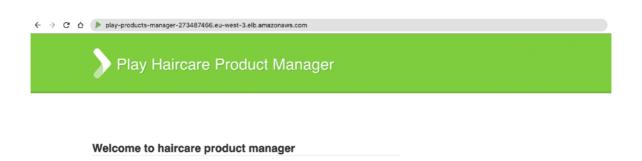
En este fichero se encuentra toda la configuración que se va a realizar para el despliegue de Cloud final de la aplicación. En él se puede observar que se va a crear un grupo de autoescalado, para tres instancias de EC2, ya que de momento se cree conveniente que existan tres instancias para soportar la capacidad deseada de uso de la aplicación, entre otras cosas por la arquitectura en la que se ha desarrollado la aplicación. En caso de que una falle, Amazon se encargará de crear otra para suplirla. Se creará también un balanceador que permita la distribución automática del tráfico a la aplicación entre las tres instancias que se crearán. También se crea un listener para ese balanceador, que en caso de que exista algún problema con el servicio, ofrecerá una pantalla de fallo con el error 404 (para nuestro caso en HTTP solamente). Se configuran también unos grupos de seguridad que permitirán la gestión de las instancias, en concreto se permitirá el acceso por el puerto 80 desde cualquier punto. Por otro lado, se creará un grupo de destino que permita al balanceador enviar las solicitudes a los destinos registrados, en nuestro caso por el puerto 80 con el protocolo HTTP. En este Target group se configura un comprobación del estado de las instancias, donde se va a esperar

durante 60 segundos la respuesta correcta de cada instancia en intervalos de 30 segundos. En el caso de que una instancia falle durante 3 veces seguidas, se creará otra para suplirla, debido a que se considera que esa instancia no está funcionando correctamente y se terminará con ella. Además, para comprobar que una instancia está funcionando correctamente, deberá de responder 3 veces consecutivas de manera correcta. De esta manera se espera obtener una tolerancia a fallos correcta de la aplicación.

En el caso de que en algún momento se decida que no es suficiente con ese número de instancias para el servicio, se podrá aumentar o disminuir gracias a este archivo.

Para ejecutar este archivo de Terraform simplemente habrá que tener introducidas de igual manera la clave privada y la clave secreta, y tras iniciar Terraform con terraform init, se podrá comprobar todos las configuraciones que se van a realizar con el comando terraform plan, y finalmente se ejecutará con el comando terraform apply.

Tras su ejecución aparecerá la dns a la que acceder para ver el despliegue realizado de la aplicación.



```
| Part | The Control | The Con
```

Otras consideraciones

El desarrollo de la práctica se encuentra tanto en los ficheros que se adjuntan en la entrega como en un repositorio de GitHub en la rama de deploy.

https://github.com/CeliaHFerreira/ProductsManager-play-rest/tree/deploy

En los archivos entregados está por una parte la carpeta RDS que contiene tanto el fichero de terraform para la creación de la base de datos como el propio archivo SQL de la base de datos.

En la carpeta Archivos despliegue se encuentran separados los archivos start.sh, play-products-manager.json, play-products-manager.service, main.tf y el proyecto para poder acceder a ellos de manera rápida, el despliegue aquí explicado si se desea replicar debe hacerse desde el fichero comprimido ProductsManager-play-rest.zip.