

**Hybrid Storage and Data Migration with AWS Storage Gateway File Gateway**

Icono

Descripción generada automáticamente con confianza media

aDRIÁN gONZÁLEZ IGLESIAS

TAJAMAR

18/05/2024

Índice

[Descripción general 1](#_Toc167033393)

[Revisión de la arquitectura del proyecto, ventajas y desventajas 1](#_Toc167033394)

[Paso 1: Crear los S3 buckets primario y secundario 3](#_Toc167033395)

[Paso 2: Habilitar la replicación entre regiones de los buckets 5](#_Toc167033396)

[Paso 3: Configurar File Gateway y crear un recuro compartido de archivos NFS 8](#_Toc167033397)

[Paso 4: Montar el recurso compartido en la instancia de Linux y migrar los datos 18](#_Toc167033398)

[Paso 5: Verificar que los datos se han migrado correctamente 22](#_Toc167033399)

[Conclusión 24](#_Toc167033400)

## Descripción general

En este proyecto, utilizaremos el servicio AWS Storage Gateway File Gateway para adjuntar un sistema de archivos de red (NFS) a un almacén de datos local. Esto permitirá que los datos almacenados en el entorno local sean accesibles y gestionables a través de la nube. A continuación, replicaremos esos datos a un bucket S3 en AWS para garantizar su seguridad y disponibilidad a largo plazo. Adicionalmente, configuraremos funciones avanzadas de Amazon S3, incluyendo políticas de ciclo de vida, que automatizan la gestión de datos moviéndolos a diferentes clases de almacenamiento o eliminándolos según las necesidades. También implementaremos la replicación entre regiones para asegurar que los datos estén disponibles y protegidos incluso en caso de desastres regionales. Esta configuración optimizará tanto la gestión de datos como los costos de almacenamiento, mientras proporciona una solución escalable y segura para la infraestructura de datos.

## Revisión de la arquitectura del proyecto, ventajas y desventajas

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

**¿Qué aprenderemos en este proyecto?**

1. Configurar un File Gateway con un recurso compartido de archivos NFS y adjuntarlos a una instancia de Linux.
2. Migrar un conjunto de datos desde la instancia de Linux a un bucket S3.
3. Crear y configurar un bucket S3 prinicipal para migrar datos del servidor local a AWS.
4. Crear y configurar un bucket S3 secundario para usarlo en la replicación entre regiones.
5. Crear una política de ciclo de vida de S3 para gestionar automáticamente los datos en un bucket.

**¿Qué ventajas nos proporciona este proyecto?** [**✅**](https://www.google.com/search?sca_esv=8272a6fbd46d38f9&rlz=1C1UEAD_esES1077ES1077&sxsrf=ADLYWIIjKJSHJBXDChXJZo1ueP69ySFPCw:1716034825480&q=%C2%BFQu%C3%A9+significa+el+emoji+de+%E2%9C%85%3F&sa=X&ved=2ahUKEwilmcfxl5eGAxUd0QIHHQGHD_wQzmd6BAgeEAY)

1. Escalabilidad y flexibilidad: permite almacenamiento ilimitado en Amazon S3, adaptándose dinámicamente a las necesidades sin grandes inversiones en infraestructura.
2. Alta disponibilidad y durabilidad: asegura la disponibilidad de datos mediante replicación en múltiples ubicaciones y regiones, con una durabilidad del 99.999999999%.
3. Seguridad y cumplimiento: proporciona cifrado avanzado, control de acceso granular y cumplimiento con estándares de seguridad.
4. Gestión automatizada del ciclo de vida: reduce costos y la carga operativa al mover datos a clases de almacenamiento más económicas o eliminarlos automáticamente.
5. Reducción de costos: elimina la necesidad de hardware local y permite un modelo de costos flexible basado en el uso real.
6. Simplificación de la gestión de datos: centraliza la administración y facilita la implementación de estrategias de respaldo y recuperación de desastres.
7. Facilidad de integración: se integra perfectamente con sistemas locales a través de NFS y ofrece APIs para automatización.
8. Optimización del rendimiento: mejora el acceso a datos mediante cachés locales y optimiza las transferencias a S3 para maximizar el rendimiento.

**¿Qué desventajas nos presenta este proyecto?** ❌

1. Dependencia de la conectividad a Internet: el acceso a los datos en Amazon S3 depende de una conexión a Internet estable y rápida.
2. Costos variables: los costos pueden ser impredecibles y aumentar con grandes volúmenes de datos transferidos o almacenados.
3. Latencia: el acceso a datos en la nube puede ser más lento comparado con el acceso a datos locales, afectando aplicaciones sensibles a la latencia.
4. Complejidad de configuración inicial: la configuración de File Gateway, NFS y políticas de ciclo de vida puede ser compleja y requerir conocimientos especializados.

## Paso 1: Crear los S3 buckets primario y secundario

En el primero paso, vamos a realizar la configuración de un **primer** S3 ubicado en la región de US East (**Ohio**) us-east-2 y un **segundo** S3 ubicado en US West (**Oregon**) us-west-2, en los que activaremos el **Bucket versioning**, esto nos permitirá más adelante utilizar la replicación de tipo Cross-Region.

1. Entramos en el servicio Amazon S3 y accedemos a **Create bucket.**

Interfaz de usuario gráfica, Aplicación, Sitio web

Descripción generada automáticamente

1. Seleccionada la **región correcta**, asignamos un nombre único para reconocer el S3.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

1. Activamos el **Bucket Versioning**.

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

1. Creamos el bucket dejando los demás valores por defecto.

Forma, Rectángulo

Descripción generada automáticamente

1. Una vez creado el primer bucket, entramos ne créate bucket para crera el segundo**.**

Interfaz de usuario gráfica, Aplicación, Teams

Descripción generada automáticamente

1. Seleccionada la **región correcta**, asignamos un nombre único para reconocer el S3.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

1. Activamos el **Bucket Versioning**.

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

1. Creamos el bucket dejando los demás valores por defecto.

Forma, Rectángulo

Descripción generada automáticamente

Comprobamos que están creado los dos buckets correctamente.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

## Paso 2: Habilitar la replicación entre regiones de los buckets

El segundo paso será, habilitar la replicación entre regiones de los buckets.

Una vez activado el **bucket versioning** , vamos a crear una regla de replicación que tendrá como función **replicar** los archivos **automáticamente** de un bucket situado en **Ohio** al bucket situado en **Oregon**.

1. Accedemos al bucket situado en **Ohio** y entramos en **Create replication rule**.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

1. Asignamos un nombre a la regla y activamos el **Status** para que la regla esté activa al crearla.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

1. Aplicamos la regla a **todos** los objetos que estén ubicados en el bucket de **Ohio**.

Interfaz de usuario gráfica, Texto, Aplicación, Chat o mensaje de texto, Correo electrónico

Descripción generada automáticamente

1. Indicamos que el **destino** será el bucket situado en **Oregon.**

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. Como estamos realizando la configuración del **S3 Cross-Región**, para este paso necesitamos asignar el rol **S3-CRR-Role** que es un rol específico para el Cross-Región.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. Guardamos la configuración de la regla.

Diagrama

Descripción generada automáticamente con confianza media

1. Subimos cualquier archivo que tengamos en nuestro sistema al bucket que hemos configurado como origen de la regla, es decir, el bucket situado en la región de Ohio.

Dentro del **bucket-objects** entramos en **Upload**.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

**Add files** si es un archivo o **Add folder** si queremos subir una carpeta.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

Nosotros subimos un **.txt** creado en el momento como prueba del paso.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

Subimos el archivo pulsando en **Upload**.

Forma

Descripción generada automáticamente

1. Entramos en el bucket de **Oregon** y vemos que gracias a la configuración del **Cross-Region**, se ha **replicado automáticamente** habiendo subido solamente el archivo al bucket de **Ohio**.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

## Paso 3: Configurar File Gateway y crear un recuro compartido de archivos NFS

En el tercer paso implementaremos un dispositivo **File Gateway** como una instancia de Amazon Elastic Compute Cloud (**Amazon EC2**). Luego configuraremos un **disco de caché** utilizando un depósito de S3 para sincronizar sus archivos locales y seleccionaremos una política de **IAM** para usar. Finalmente, crearemos un recurso compartido de archivos **NFS** en File Gateway.

1. Entramos en el servicio **AWS Storage Gateway** y seleccionamos **Create Gateway**.

Interfaz de usuario gráfica, Texto

Descripción generada automáticamente

1. Asignamos un nombre para reconocer el **Gateway**.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

1. Escogemos el tipo de Gateway **S3 File Gateway**.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. Utilizaremos una **EC2** como hemos indicado antes, entramos en **Customize your settings**.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. Pinchamos en **Launch Instance** para crear la EC2 con los requisitos necesarios.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. Asignamos un **nombre único** para reconocer la EC2.

Imagen que contiene Gráfico

Descripción generada automáticamente

1. El tipo de instancia será **t2.xlarge** dado que es la que mejor se adapta a las condiciones de la EC2.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. Como estamos realizando el proyecto mediante un entorno de laboratorio se nos proporciona una **Key Pair (vockey)** específica para este paso.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

1. Utilizamos la VPC que proporciona el laboratorio **(On-Prem-VPC)** y su correspondiente **subnet**, además activamos el **Auto-assign public IP** que sirve para que la EC2 reciba automáticamente una IP pública después de su creación.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

1. En el apartado de Security groups, pinchamos en **Select existing security group** seleccionamos el grupo de seguridad **FileGatewayAccess** y además el grupo de seguridad **OnPremSshAccess.**

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

1. Generamos un nuevo volumen pinchando en **Add new volume**.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

1. Asignamos **150gb** y de tipo **gp3**.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. Dejamos los demás valore por defecto y pulsamos en **Launch instance**.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

1. Cuando se haya creado la **EC2**, copiamos en un portapapeles su **IP pública**.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

**54.162.124.5**

1. Volvemos a la anterior pestaña del Gateway, marcamos **Confirm set up Gateway** y **Next**.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. En el apartado de conexión al Gateway, escogemos la opción mediante IP address y añadimos la **IP pública** **de la EC2** copiada anteriormente.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

1. Indicamos que sea **accesible públicamente.**

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. Activamos el **Gateway**.

Interfaz de usuario gráfica, Texto, Aplicación, Chat o mensaje de texto

Descripción generada automáticamente

1. Desactivamos los **registros** y **alarmas** del **CloudWatch.**

Interfaz de usuario gráfica, Texto, Aplicación, Chat o mensaje de texto, Correo electrónico

Descripción generada automáticamente

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

1. Para finalizar la configuración pulsamos en **Configure.**

Diagrama

Descripción generada automáticamente con confianza media

1. Comprobamos que ya está funcionando nuestro Gateway.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

1. A continuación, vamos a crear un **File Share**, para ello en la parte superior izquierda entramos en File Share.

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

1. Entramos en **Create File share**.

Imagen que contiene Aplicación

Descripción generada automáticamente

1. Indicamos que utilice el **Gateway** creada en el paso anterior, que el protocolo como hemos explicado al principio será **NFS** y que el **bucket** que utilizará será el situado en la región de **Ohio**.

Interfaz de usuario gráfica, Texto, Aplicación

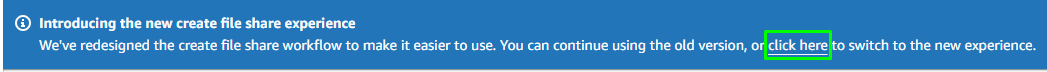
Descripción generada automáticamente

1. Entramos en **Customize configuration**.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

1. En cuanto nos redirija, pincharemos en el **cuadro azul** que nos aparece en la parte superior en **click here**, para que nos aparezca la configuración necesaria.



1. Seleccionamos el **Gateway** creado anteriormente, lo alojaremos en el **bucket de Ohio** y lo vamos a llamar **share**.

Interfaz de usuario gráfica, Texto, Aplicación

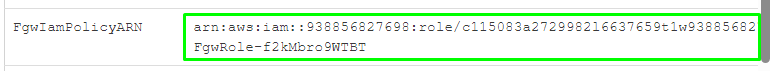
Descripción generada automáticamente

1. En la configuración del alojamiento en el S3, seleccionamos tipo Standard y marcamos las dos primeras opciones de metadatos.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

1. Antes de ir con el paso siguiente, copiamos el **ARN** del rol que vamos a utilizar, este valor nos lo proporciona el laboratorio y es un rol necesario para la configuración.



1. Seleccionamos **Use an existing IAM role** y pegamos el valor de **ARN** del rol.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

1. Creamos el File Share.

Diagrama

Descripción generada automáticamente

1. Una vez creado el File Share entramos en él y copiamos el comando que nos proporciona el File Share **On Linux**.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

Aplicación

Descripción generada automáticamente con confianza baja

**mount -t nfs -o nolock,hard 10.10.1.139:/share [MountPath]**

## Paso 4: Montar el recurso compartido en la instancia de Linux y migrar los datos

En el cuarto paso, antes de migrar los datos al NFS, primero vamos a montar el **recurso compartido** que lo que haremos será, montar el **NFS** en **un servidor Linux** y copiaremos los **datos** al **recurso compartido**.

1. Descargamos el **PPK** que nos proporciona en entorno del laboratorio y copiamos la **IP** del **servidor Linux** que también proporciona el laboratorio.

Interfaz de usuario gráfica

Descripción generada automáticamente con confianza media Interfaz de usuario gráfica, Aplicación, Word

Descripción generada automáticamente

1. Una vez descargado el PPK y copiada la IP de la instancia Linux, descargamos **Putty.**

Interfaz de usuario gráfica

Descripción generada automáticamente con confianza media

1. Entramos al **Putty**, ingresamos la **IP** de la instancia **Linux**, y escogemos el tipo de conexión **SSH**.

Interfaz de usuario gráfica, Texto, Aplicación, Chat o mensaje de texto

Descripción generada automáticamente

1. Dentro de **Connecion-SSH-Auth** ingresamos el **PPK** descargado anteriormente.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. Abrimos **conexión** y cuando nos aparezca esta ventana, pulsamos **Accept**.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. Nos logeamos con el usuario **ec2-user**.

Texto

Descripción generada automáticamente

1. Hacemos un **ls /media/data** y vemos que el servidor Linux tiene **guardado 20 archivos .png**.

Texto

Descripción generada automáticamente con confianza media

1. Creamos un directorio que se utilizará para **sincronizar** los **datos** con el **S3** mediante el comando sudo **mkdir -p /mnt/nfs/s3**.



1. Copiamos la **IP** privada de la instancia **File Gateway Appliance.**

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

**10.10.1.139**

1. Montamos el recurso compartido en la instancia Linux mediante el comando **sudo mount -t nfs -o nolock,hard <File-Gateway-appliance-private-IP-address>:/share /mnt/nfs/s3**. (Ingresamos la IP privada entre los <>)

Interfaz de usuario gráfica

Descripción generada automáticamente

1. Verificamos que se ha montado correctamente el recurso compartido mediante el comando **df -h**.

Interfaz de usuario gráfica

Descripción generada automáticamente con confianza baja

1. Copiamos los datos de la instancia Linux al S3 mediante el comando **cp -v /media/data/\*.png /mnt/nfs/s3.**

Imagen que contiene Interfaz de usuario gráfica

Descripción generada automáticamente

## Paso 5: Verificar que los datos se han migrado correctamente

Accedemos al bucket situado en **Ohio**, que es el que se ha utilizado para realizar las configuraciones necesarias del **recurso compartido** y comprobamos que se han migrado los **datos** de la instancia **Linux** correctamente.

Interfaz de usuario gráfica

Descripción generada automáticamente

Tabla

Descripción generada automáticamente

Como al principio del proyecto hemos configurado el **Croos-Region** entre **buckets**, si accedemos al bucket situado en **Oregon**, comprobamos que se ha n replicado los datos automáticamente.

Interfaz de usuario gráfica, Tabla

Descripción generada automáticamente

Pantalla de computadora con números

Descripción generada automáticamente con confianza media

## Conclusión

La implementación de este proyecto ofrece una solución integral y avanzada para la gestión de datos, combinando la infraestructura local con la potencia y flexibilidad de la nube de AWS. Al utilizar AWS Storage Gateway File Gateway para adjuntar un sistema de archivos de red (NFS) a un almacén de datos local y replicar esos datos a un bucket S3 en AWS, garantizamos la seguridad y disponibilidad de los datos a largo plazo.

Además, las funciones avanzadas de Amazon S3, como las políticas de ciclo de vida y la replicación entre regiones, automatizan y optimizan la gestión de datos, reduciendo costos operativos y mejorando la resiliencia ante desastres.

En resumen, este proyecto no solo mejora la eficiencia y la protección de los datos, sino que también proporciona una base sólida para la escalabilidad y el crecimiento futuro.