Jonathan A nderson

AWS

Certified Solutions Architect

Associate

The ultimate guide for the

SAA-C03

Imagen que contiene lego, juguete

Descripción generada automáticamenteE X A M

## Static Website Hosting

Un caso de uso común para el almacenamiento de Amazon S3 implica el alojamiento de sitios web estáticos. Muchos sitios web, especialmente los micrositios, no requieren los servicios de un servidor web completo. Un sitio web estático significa que todas las páginas del sitio web contienen solo contenido estático y no requieren ningún procesamiento en el servidor, como PHP, ASP.NET o JSP. (Tenga en cuenta que esto no significa que el sitio web no pueda ser interactivo y dinámico; Se puede lograr con secuencias de comandos del lado del cliente, como JavaScript incrustado en páginas web HTML estáticas.) Hay varios beneficios para los sitios web estáticos: son muy rápidos, altamente escalables y son más seguros que un sitio web dinámico tradicional. El alojamiento de un sitio web estático en Amazon S3 también puede aprovechar la seguridad, la durabilidad, la disponibilidad y la escalabilidad de Amazon S3.

Dado que cada objeto de Amazon S3 tiene una URL, convertir un bucket en un sitio web es relativamente fácil. Para alojar un sitio web estático, debe configurar un bucket para el alojamiento de sitios web y, a continuación, cargar el contenido estático del sitio web en el bucket.

Para configurar un bucket de Amazon S3 para alojar sitios web estáticos:

1. Cree un bucket con el mismo nombre que el nombre de host del sitio web deseado.
2. Descargue los archivos estáticos en el bucket.
3. Hacer que todos los archivos sean públicos (legibles en todo el mundo).
4. Habilite el alojamiento de sitios web estáticos para el bucket. Esto incluye especificar un documento de índice y un documento de error.
5. El sitio web ahora está disponible en la URL del sitio web de S3: <bucket-name>

.s3-website- <AWS-region>.amazonaws.com.

1. Cree un nombre DNS descriptivo en su propio dominio para el sitio web mediante CNAME DNS o un alias de Amazon Route 53 que coincida con la URL del sitio web de Amazon S3.
2. El sitio web ahora está disponible bajo el nombre de dominio de su sitio web.

## Amazon S3 Advanced Functions

Además, para lo básico, hay algunas características avanzadas de Amazon S3 con las que también debe familiarizarse.

## Prefixes and Separators

Aunque Amazon S3 utiliza una estructura plana en un bucket, admite el uso de parámetros de prefijo y separador al enumerar nombres de clave. Esta funcionalidad le permite organizar, examinar y recuperar los objetos en un bucket de manera jerárquica. Normalmente, debe utilizar una barra diagonal (/) o una barra diagonal inversa (\) como separador y, a continuación, utilizar nombres de clave con separadores incrustados para emular una jerarquía de archivos y carpetas en el espacio de nombres de claves. un objeto plano de un compartimento.

Por ejemplo, desea almacenar una serie de registros de servidor por nombre de servidor (como server42), pero organizados por año y mes, de la siguiente manera:

Logs / January 2016 / Server42.logs Logs / February / 2016 / server42.log Logs / 2016 / March / 42.log server

La API de REST, los SDK de encapsulación, la AWS CLI y la consola de administración de Amazon admiten el uso de separadores y prefijos. Esta característica le permite organizar nuevos datos de forma lógica y administrar fácilmente la carpeta jerárquica y la estructura de archivos de los datos existentes descargados o respaldados de los sistemas de archivos tradicionales. Cuando se utilizan con políticas de bucket de IAM o Amazon S3, los prefijos y delimitadores también pueden ayudar a crear el equivalente de "subcarpetas" o "directorios de inicio" de usuarios en un solo bucket, restringiendo o compartiendo acceso a esas "subcarpetas" (definidas por prefijos) según sea necesario.

Utilice separadores de objetos y prefijos para organizar jerárquicamente los objetos de los buckets de Amazon S3, pero tenga en cuenta que Amazon S3 no es realmente un sistema de archivos.

**Storage Classes**

Amazon S3 proporciona una serie de clases de almacenamiento que son suficientes para varios casos de uso.

Amazon S3 Standard proporciona alta confiabilidad, alta disponibilidad, baja latencia y almacenamiento de objetos de alta calidad para uso general. Como ofrece baja latencia en el primer byte y alto rendimiento, el estándar es preferible para el almacenamiento a corto o largo plazo de datos de uso frecuente. Para la mayoría de los casos de uso general, es mejor comenzar con Amazon S3 Standard.

El acceso estándar poco frecuente (Standard-IA) de Amazon S3 proporciona la misma longevidad, baja latencia y alto rendimiento que Amazon S3 Standard, pero está diseñado para datos de larga duración y datos a los que se tiene menos acceso. Standard-IA tiene un costo de almacenamiento por GB mensual más bajo que el Standard, pero el modelo de precios también incluye un tamaño mínimo de objeto (128 KB), la duración mínima (30 días) y costo de recuperación por GB, por lo que es más adecuado para acceder raramente a los datos almacenados durante más de 30 días.

El almacenamiento de redundancia reducida (RRS) de Amazon S3 proporciona una confiabilidad ligeramente menor (4 nuevos) a un costo reducido que Normal o Normal-IA. Es más adecuado para datos derivados, como miniaturas de imágenes, que se pueden replicar fácilmente.

Por último, para los datos que no requieren acceso en tiempo real, como archivos y copias de seguridad a largo plazo, la clase de almacenamiento de Amazon Glacier proporciona una nube fiable, robusta y altamente rentable. almacenamiento. Para mantener los costos bajos, Amazon Glacier está optimizado para datos de uso mínimo con un tiempo de recuperación de varias horas. Para restaurar un objeto de Amazon Glacier, emita un comando de restauración con una de las API de Amazon S3; de tres a cinco horas después, el objeto Amazon Glacier se copia en Amazon S3 RRS. Tenga en cuenta que la opción de restauración simplemente hará una copia en Amazon S3 RRS; el objeto de datos original permanece en Amazon Glacier hasta que se elimina explícitamente. Además, es importante tener en cuenta que Amazon Glacier le permite recuperar hasta el 5 % de los datos de Amazon S3 almacenados en Amazon Glacier para gratis todos los meses. En el caso de las restauraciones fuera de la asignación diaria de comidas, se han puesto a disposición los gastos de entrega.

Amazon Glacier no solo se utiliza como nivel de almacenamiento en Amazon S3, sino que también es un servicio de almacenamiento independiente con una API independiente y funciones únicas. Sin embargo, cuando utiliza Amazon Glacier como clase de almacenamiento de Amazon S3, aún puede comunicarse con los datos a través de las API de Amazon S3. Consulte la sección Amazon Glacier para obtener más información.

Establecer una política de recuperación de datos para limitar la recuperación para el nivel gratuito o al máximo límite de GB por hora para evitar o minimizar los costos de recuperación de Amazon Glacier.

**Object Lifecycle Management**

La administración del ciclo de vida de los objetos de Amazon S3 es aproximadamente equivalente al almacenamiento automatizado agotador en las infraestructuras de almacenamiento de TI tradicionales. En muchos casos, los datos tienen un ciclo de vida natural, comenzando con datos "calientes" (a los que se accede con frecuencia), pasando a datos "calientes" (a los que se accede con menos frecuencia) a medida que envejecen y terminando con datos "calientes". Luego se mueve a los datos de Cold " (copia de seguridad o archivo a largo plazo) antes de una posible eliminación.

Por ejemplo, muchos documentos comerciales a menudo se ven a medida que se crean y apenas se ven con el tiempo. Sin embargo, en muchos casos, las reglas de cumplimiento requieren que los registros comerciales se archiven y se mantengan accesibles durante años. Del mismo modo, los estudios muestran que las copias de seguridad de archivos, sistemas operativos y bases de datos se abren con mayor frecuencia dentro de los primeros días de su creación, generalmente para eliminar un error accidental. error. Después de una semana o dos, estas copias de seguridad seguirán siendo un activo crítico, pero es mucho menos probable que sean accesibles para la recuperación. Las leyes de cumplimiento requieren que una variedad de copias se conserve durante muchos años en muchos casos.

Con el uso de reglas de configuración del ciclo de vida de Amazon S3, puede reducir los costos de almacenamiento transfiriendo automáticamente datos de una clase de almacenamiento a otra o incluso eliminando automáticamente los datos después de un período de tiempo. Por ejemplo, las reglas del ciclo de vida de los datos de copia de seguridad pueden ser:

1. Guardar inicialmente los datos de copia de seguridad en Amazon S3 Standard.
2. Cambio a Amazon Standard-IA después de 30 días.
3. Después de 90 días, comience la transición al glaciar Amazonas.
4. Retirar después de 3 años.

Las configuraciones del ciclo de vida están vinculadas al bucket y se pueden aplicar a todos los objetos del bucket o solo a los objetos especificados con un prefijo.

## Encryption

Se recomienda cifrar todos los datos confidenciales contenidos en Amazon S3, tanto en vuelo como en reposo.

Los puntos de enlace de la API de Capa de sockets estables (SSL) de Amazon S3 se pueden utilizar para cifrar los datos de Amazon S3 sobre la marcha. Esto implica que cuando están en tránsito a través de HTTPS, todos los datos enviados hacia y desde Amazon S3 se cifran.

Puede utilizar varias variantes de cifrado del lado del servidor (SSE) para cifrar los datos de Amazon S3 en reposo. Amazon S3 cifra los datos a nivel de objeto a medida que los escribe en las unidades de sus centros de datos y los descifra automáticamente a medida que accede a ellos. Todos los SSE ejecutados por Amazon S3 y AWS Key Management Service (Amazon KMS) utilizan el estándar de cifrado avanzado de 256 bits (AES). También puede utilizar el cifrado del lado del cliente para cifrar los datos de Amazon S3 cuando están en reposo, donde los datos se cifran para el cliente antes de enviarlos a Amazon S3.

## SSE-S3 (AWS managed keys)

Se trata de una solución de cifrado de "casilla de verificación" totalmente integrada en la que AWS administra la administración de claves y la protección de claves para Amazon S3. Con una clave única, cada objeto se cifra. A continuación, una clave maestra independiente cifra la propia clave de objeto real. Al menos una vez al mes, se lanza una nueva clave maestra, con AWS ejecutando las claves. En hosts seguros, los datos cifrados, las claves de cifrado y las claves maestras se almacenan por separado, lo que mejora aún más la seguridad.

**SSE-KMS (AWS KMS keys)**

Es una solución totalmente integrada en la que Amazon administra y protege sus claves para Amazon S3, pero puede administrar las claves. SSE-KMS ofrece varias ventajas adicionales sobre SSE-S3. Se requieren diferentes permisos para obtener acceso a la clave maestra mediante SSE-KMS, lo que ayuda a evitar el acceso no autorizado a los elementos almacenados de Amazon S3 y también agrega una capa adicional de control. AWS KMS también ofrece auditoría, por lo que está familiarizado con quien utiliza la clave para obtener acceso a los objetos y la hora de acceso a los objetos. AWS KMS también le permite ver todos los intentos fallidos de obtener acceso a los datos de los usuarios a los que no se les permitió descifrar la información.

**SSE-C (keys provided by the customer)**

Esto se usa si desea mantener sus propias claves de cifrado, pero no desea administrar o implementar su propia biblioteca de cifrado del lado del cliente. Con SSE-C, AWS realiza el cifrado/descifrado de sus objetos mientras conserva el control total sobre las claves utilizadas para cifrar/descifrar los objetos en Amazon S3.

## Client-Side Encryption

El cifrado del lado del cliente se refiere al proceso de cifrado de los datos en el lado del cliente de la aplicación antes de que se envíen a Amazon S3. Para utilizar claves de cifrado de datos, puede seguir estas opciones:

1. Utilice una clave maestra de cliente administrada por AWS KMS.
2. Utilice una clave maestra en el lado del cliente.

Cuando se utiliza el cifrado del lado del cliente, se conserva el control total del proceso de cifrado, incluida la administración de claves de cifrado.

También puede utilizar el cifrado del lado del servidor con claves operadas por AWS (SSE-S3 o SSE-KMS) para una flexibilidad y facilidad de uso óptimas.

## Versioning

Al mantener diferentes versiones de un objeto en el bucket, marcadas por un ID de versión único, Amazon S3 Versioning ayuda a proteger los datos de la eliminación involuntaria o intencional. El control de versiones permite conservar, recuperar y restaurar cualquier versión de un objeto almacenado en el bucket de Amazon S3. Si un usuario realiza un cambio involuntariamente o incluso elimina un objeto en su bucket de S3 de forma malintencionada, puede restaurar el objeto a su origen simplemente haciendo referencia a al ID de versión junto al bucket y la clave.

Se permite el control de versiones a nivel del bucket. Hasta que esto no esté permitido, no es posible eliminar la versión del bucket, ya que no se puede suspender.

## MFA Delete

MFA Delete agrega una capa adicional de protección de datos además de la administración de versiones de bucket. MFA Delete requiere autenticación adicional para eliminar permanentemente una versión de objeto o cambiar el estado de control de versiones de un bucket. MFA Delete requiere un código de autenticación creado por un hardware o dispositivo virtual de autenticación multifactor (MFA), además de sus credenciales de seguridad habituales. Tenga en cuenta que la cuenta raíz solo puede habilitar la eliminación de MFA.

**Pre-signed URLs**

Todos los objetos de Amazon S3 son privados de forma predeterminada. Esto significa que solo el propietario tiene acceso. Sin embargo, al crear una dirección URL prefirmada, el propietario del objeto puede compartir objetos opcionalmente con otras personas, utilizando sus propias credenciales de seguridad para conceder permiso de tiempo limitado para descargar los objetos. Al crear una dirección URL predefinida para el objeto, debe proporcionar las credenciales de seguridad y proporcionar un nombre de bucket, una clave de objeto y un método HTTP (GET para descargar el objeto). Las URL prefirmadas solo son válidas durante el período de tiempo especificado. Esto es especialmente útil para formar protección contra el "raspado de contenido" de contenido web, como los archivos multimedia almacenados en Amazon S3.

**Multipart Upload**

Para admitir aún más la descarga o copia de objetos grandes, Amazon S3 proporciona la API de descarga de varias partes. Esto le permite descargar objetos grandes en partes, ya que generalmente proporciona un mejor uso de la red (a través de transferencias paralelas), la capacidad de pausar y reanudar, y la capacidad de descargar objetos con tamaños desconocidos.

La descarga multiparte es un proceso de tres pasos: iniciar, compartir, descargar y finalizar (o abortar). Las piezas se pueden descargar de forma independiente en cualquier orden, posiblemente con retransmisión. Después de todo, las piezas se cargan, Amazon S3 ensambla las piezas para crear un objeto.

En general, debe utilizar una descarga de varias partes para objetos de más de 100 MB y objetos de más de 5 GB. Si está utilizando API de bajo nivel, debe descargar el archivo en partes y realizar un seguimiento de las piezas. Cuando se utilizan las API de alto nivel y los comandos de Amazon S3 de alto nivel en la AWS CLI (AWS s3 cp, AWS s3 mv y AWS s3 sync), la carga de varias partes se requiere automáticamente para objetos grandes.

Puede configurar una política de ciclo de vida de objetos para un bucket a fin de evitar varias descargas incompletas después de un número especificado de días. Esto minimiza los costos de almacenamiento asociados con las descargas de varias partes que no se han completado.

## Range GET

Es posible descargar solo una parte de un objeto (GET) en Amazon S3 y Amazon Glacier mediante el rango GET. Mediante el uso del encabezado de intervalo HTTP en la solicitud GET o parámetros equivalentes en una de las bibliotecas contenedoras del SDK, puede especificar el intervalo de bytes para el objeto. Esto puede ser útil para tratar con objetos grandes si tiene una mala conexión o para descargar solo una parte conocida de una copia de seguridad grande de Amazon Glacier.

**Interregional Replication**

La replicación interregional es una característica de Amazon S3 que le permite replicar de forma asincrónica todos los objetos nuevos desde el bucket de origen en una región de AWS a un bucket de destino en otra región. Todos los metadatos y ACL asociados con el objeto también forman parte de la replicación. Después de configurar la replicación interregional en el bucket de origen, cualquier cambio en los datos, metadatos o ACL de un objeto desencadenará una nueva replicación en el bucket de destino. Para habilitar la replicación entre regiones, el control de versiones debe estar habilitado para los buckets de origen y destino, y debe utilizar una política de IAM para permitir que Amazon S3 replique objetos en en su nombre.

La replicación interregional se utiliza a menudo para reducir la latencia necesaria para obtener acceso a los objetos en Amazon S3 acercando los objetos a un grupo de usuarios o para cumplir los requisitos para almacenar datos de copia de seguridad fuera de los datos de origen. Si se habilita en un bucket existente, la replicación entre regiones solo replicará nuevos objetos. Los objetos existentes no se replican y deben copiarse en el nuevo compartimento a través de un comando independiente.

## Logging

Puede habilitar los registros de acceso al servidor de Amazon S3 para supervisar las solicitudes de su bucket de Amazon S3.

De forma predeterminada, el cierre de sesión está deshabilitado, pero se puede permitir fácilmente. Cuando permite el registro de bucket (el bucket de origen), debe elegir dónde se deben almacenar los logs (el bucket de destino). Los registros de acceso se pueden colocar en el mismo bucket o en otro bucket. En cualquier caso, es opcional (pero por buenas prácticas) proporcionar un prefijo como logs / o yourbucketname / logs / que facilitará la identificación de sus trazas.

Cuando está habilitado, los registros se entregan con un ligero retraso en el mejor de los casos. Los registros contienen información como:

1. Cuenta del solicitante y dirección IP
2. Nombre del bucket
3. Pide tiempo
4. Estado de respuesta o código de error
5. Acción (GET, PUT, LIST, etc.)