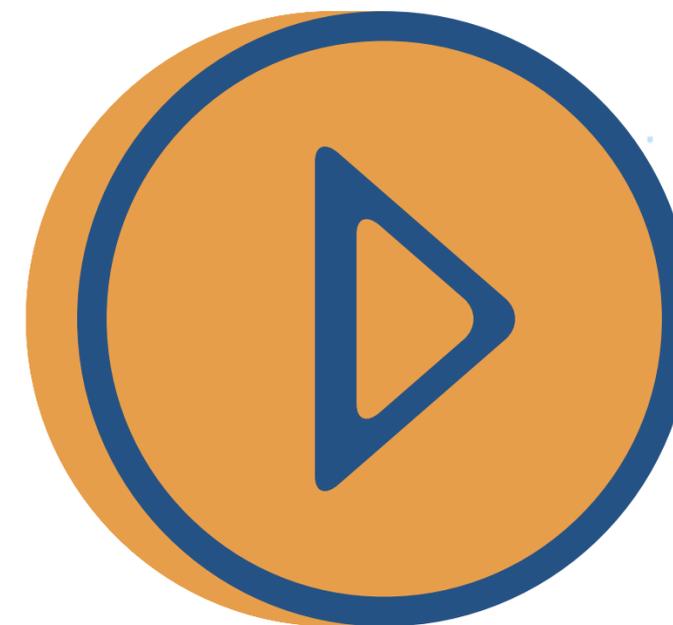


LECCIÓN 4: AMAZON S3 GLACIER





Amazon S3 Glacier es un servicio de almacenamiento en la nube seguro, duradero y de muy bajo costo para archivar datos y hacer respaldos a largo plazo.

ALMACENAMIENTO



**Amazon
S3
Glacier**





Amazon S3 Glacier es un servicio de archivado de datos diseñado para ofrecer seguridad, durabilidad y un costo extremadamente bajo.

Amazon S3 Glacier está diseñado para brindar once nueves de durabilidad para objetos.

Admite el cifrado de datos en tránsito y en reposo con Secure Sockets Layer (SSL, capa de conexión segura) o Transport Layer Security (TLS, seguridad de capa de transporte).

La función de bloqueo de almacenes impone el cumplimiento mediante una política.

El diseño muy económico es ideal para archivar a largo plazo.

Proporciona tres opciones de acceso a los archivos (rápido, estándar y masivo).
Los tiempos de recuperación oscilan entre unos minutos y varias horas.

Cuando usa Amazon S3 Glacier para archivar datos, puede almacenar los datos a un costo muy bajo (incluso en comparación con Amazon S3), pero no puede recuperarlos de inmediato cuando lo desea.

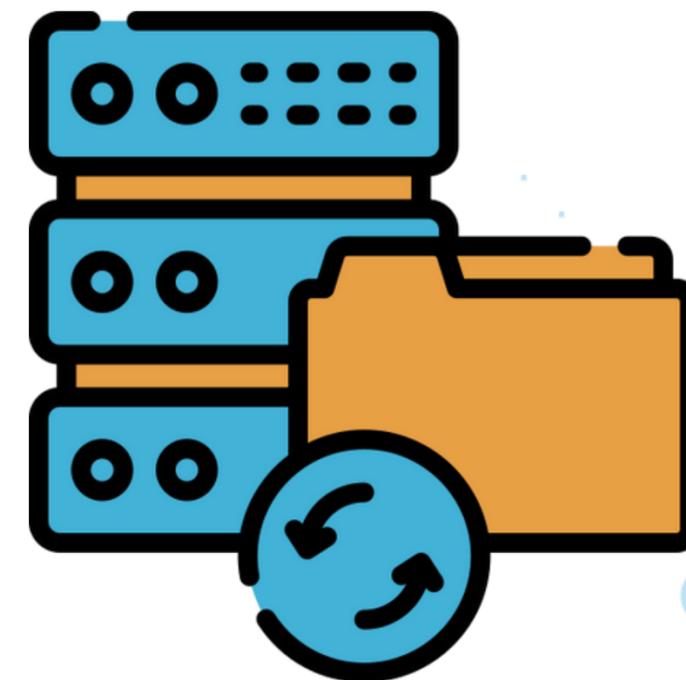
Los datos almacenados en Amazon S3 Glacier pueden tardar varias horas en recuperarse, por lo que este servicio es útil para el archivado.

Existen tres términos clave de Amazon S3 Glacier que debería conocer:



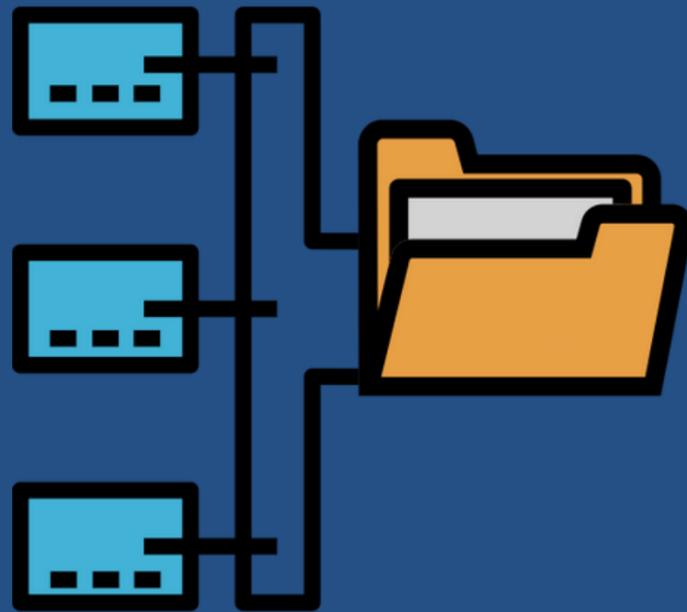
ARCHIVO:

Cualquier objeto (como una foto, un video, un archivo o un documento) que almacene en Amazon S3 Glacier. Es la unidad base de almacenamiento de Amazon S3 Glacier. Cada archivo tiene su propio identificador único y también puede tener una descripción.



ALMACÉN:

Un contenedor para almacenar archivos. Cuando se crea un almacén, se especifica el nombre del almacén y la región en la que desea que este se encuentre.





POLÍTICA DE ACCESO AL ALMACÉN:

Determina quién puede y quién no puede acceder a los datos almacenados en el almacén y también qué operaciones pueden o no pueden realizar los usuarios. Se puede crear una política de permisos de acceso a almacenes para cada uno a fin de administrar los permisos de acceso de dicho almacén. También puede utilizar una política de bloqueo de almacenes para asegurarse de que un almacén no se pueda modificar. Cada almacén puede tener asociada una política de acceso a almacenes y una política de bloqueo de almacenes.



DISPONE DE TRES OPCIONES PARA RECUPERAR LOS DATOS, CADA UNA CON DIVERSOS PLAZOS DE ACCESO Y COSTOS:

 **Las recuperaciones aceleradas suelen estar disponibles en un plazo de 1 a 5 minutos (el costo más alto).**

Las recuperaciones estándar suelen completarse en un plazo de 3 a 5 horas (menos que las aceleradas, más que las masivas).

Las recuperaciones masivas suelen completarse en un plazo de 5 a 12 horas (el costo más bajo). 

Puede comparar estas opciones con la elección del costo de envío de un paquete mediante el método más económico para sus necesidades.

AMAZON S3 GLACIER

El archivado de datos de Amazon S3 Glacier implica que, aunque pueda almacenar los datos a un costo muy bajo (incluso en comparación con Amazon S3), no puede recuperarlos de inmediato cuando lo desea.

Los datos almacenados en Amazon S3 Glacier pueden tardar varias horas en recuperarse.

- Servicio de almacenamiento para el archivado de datos a bajo costo con respaldo a largo plazo
- Posibilidad de configurar el archivado del ciclo de vida del contenido de Amazon S3 a Amazon S3 Glacier
- Opciones de recuperación:
 - Estándar: de 3 a 5 horas
 - Masivo: de 5 a 12 horas
 - Acelerado: de 1 a 5 minutos



CASOS PRÁCTICOS DE AMAZON S3 GLACIER

ARCHIVADO DE ACTIVOS DE MEDIOS DIGITALES

Los activos de medios digitales, como videos y grabaciones de noticias, requieren almacenamiento duradero y, con el tiempo, pueden aumentar hasta alcanzar muchos petabytes. Amazon S3 Glacier permite archivar contenido de medios digitales más antiguos de manera asequible y, luego, migrarlo a Amazon S3 para su distribución cuando sea necesario.



ARCHIVADO DE INFORMACIÓN DE SANIDAD



Para cumplir con los requisitos normativos, los sistemas hospitalarios deben retener petabytes de registros de pacientes, como información del Low-Income Subsidy (LIS, subsidio por ingresos bajos), datos del picture archiving and communication system (PACS, sistema de archivado y comunicación de imágenes) o Electronic Health Records (EHR, registros médicos electrónicos), durante décadas. Amazon S3 Glacier puede ayudarlo a archivar de forma fiable y segura datos de registros del paciente a un costo muy bajo.

ARCHIVADO RELACIONADO CON NORMATIVAS Y CUMPLIMIENTO

Muchas empresas, como aquellas de servicios financieros y sanidad, deben retener archivos normativos y de cumplimiento por tiempos prolongados. El bloqueo de almacenes de Amazon S3 Glacier puede ayudarlo a establecer controles de cumplimiento, de modo que pueda contribuir a la consecución de sus objetivos de cumplimiento, como la norma 17a-4(f) que establece la Securities and Exchange Commission (SEC, Comisión de Bolsa y Valores) de EE. UU.





ARCHIVADO DE DATOS CIENTÍFICOS



Las organizaciones de investigación generan, analizan y archivan grandes cantidades de datos. Gracias al uso de Amazon S3 Glacier, se pueden reducir las complejidades de la administración del hardware e instalaciones y de la planificación de capacidades.





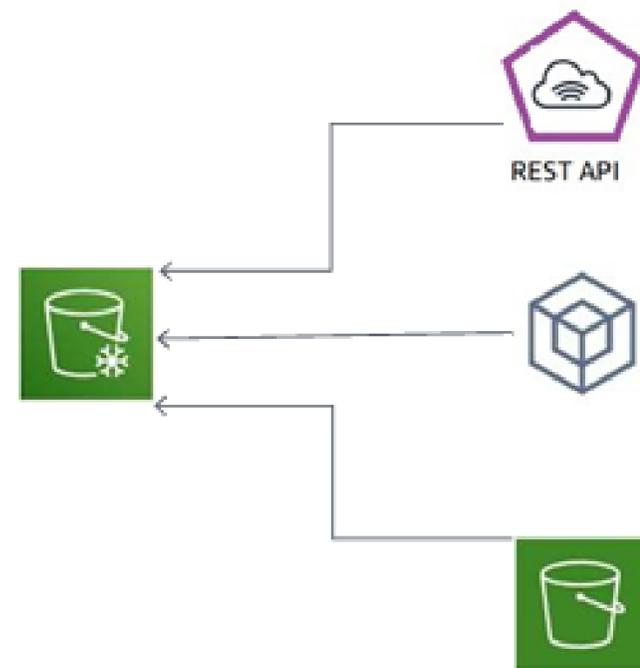
CONSERVACIÓN DIGITAL

Las bibliotecas y agencias gubernamentales deben encarar desafíos relacionados con la sustitución de cintas magnéticas.

Las bibliotecas de citas en las instalaciones o fuera del sitio pueden disminuir los costos de almacenamiento, pero pueden requerir grandes inversiones iniciales y un mantenimiento especializado. Amazon S3 Glacier no tiene un costo inicial y reduce el costo y la carga del mantenimiento.



UTILIZAR AMAZON S3 GLACIER



Servicios web
RESTful

Java o SDK de
.NET

Amazon S3 con
políticas de ciclo de
vida

© 2022, Amazon Web Services, Inc. o sus filiales. Todos los derechos reservados.



 Para almacenar datos y acceder a ellos en Amazon S3 Glacier, puede utilizar la consola de administración de AWS. Sin embargo, en la consola solo están disponibles unas pocas operaciones, como crear o eliminar almacenes y crear o administrar políticas de archivado.

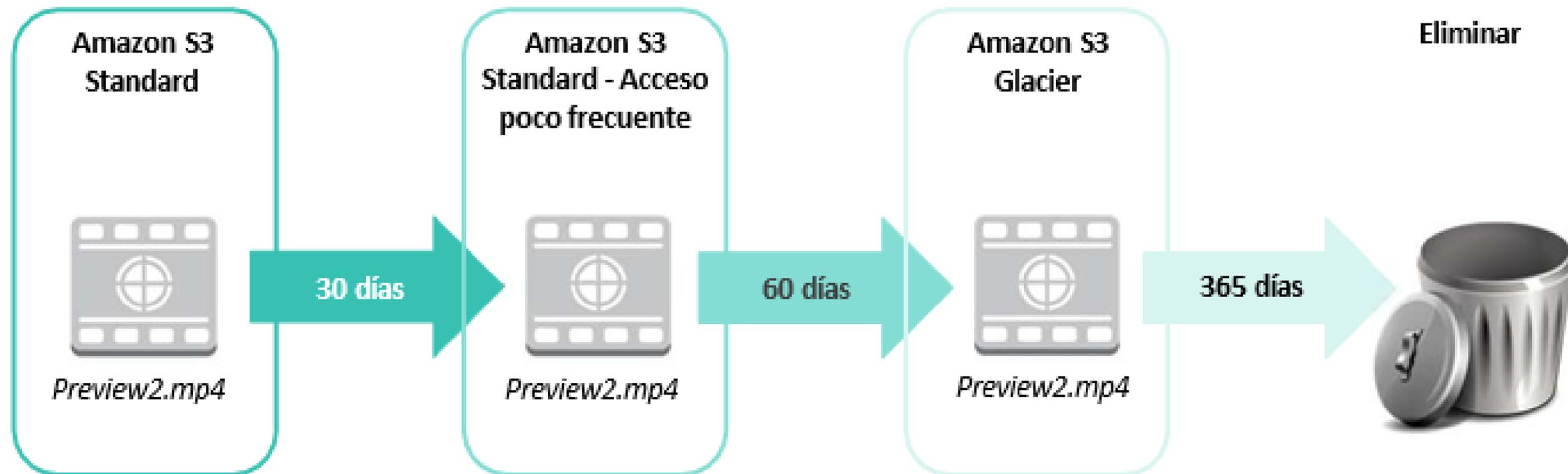
Para casi todas las demás operaciones e interacciones con Amazon S3 Glacier, debe usar las API de REST de Amazon S3 Glacier, los SDK de Java o .NET de AWS o la AWS CLI.

También puede usar políticas de ciclo de vida para archivar datos en Amazon S3 Glacier. A continuación, aprenderá sobre políticas de ciclo de vida.



POLÍTICAS DE CICLO DE VIDA

Las políticas de ciclo de vida de Amazon S3 permiten eliminar o migrar objetos en función de su antigüedad.



Debe automatizar el ciclo de vida de los datos que se almacenan en Amazon S3. Mediante el uso de políticas de ciclo de vida, puede hacer que los datos se ciclen a intervalos regulares entre los distintos tipos de almacenamiento de Amazon S3. Esta automatización reduce los costos generales porque paga menos por los datos que se vuelven menos importantes con el tiempo.

Además de establecer reglas de ciclo de vida por objeto, también puede establecer reglas de ciclo de vida por bucket.

Considere un ejemplo de política de ciclo de vida que mueve los datos a medida que envejecen de Amazon S3 Standard a Amazon S3 Standard - Acceso poco frecuente y, por último, a Amazon S3 Glacier antes de que se eliminen. Suponga que un usuario carga un video en la aplicación y esta genera una vista previa en miniatura del video. Esta vista previa de video se almacena en Amazon S3 Standard, porque es probable que el usuario quiera acceder a ella de inmediato.

Los datos de uso indican que no se accede a la mayoría de las vistas previas de miniaturas transcurridos 30 días. La política de ciclo de vida tomará estas vistas previas y las trasladará a Amazon S3 - Acceso poco frecuente después de 30 días. Después de que hayan transcurrido otros 30 días, es poco probable que se vuelva a acceder a la vista previa. Entonces la vista previa se traslada a Amazon S3 Glacier, donde permanece durante 1 año. Después de 1 año, se elimina la vista previa. Lo importante es que la política de ciclo de vida gestiona todo este movimiento de forma automática.

Para obtener más información sobre la administración del ciclo de vida de los objetos, consulte

<http://docs.aws.amazon.com/AmazonS3/latest/dev/object-lifecycle-mgmt.html>

COMPARACIÓN DE ALMACENAMIENTO

	Amazon S3	Amazon S3 Glacier
Volumen de datos	Sin límite	Sin límite
Latencia media	ms	minutos/horas
Tamaño de elemento	5 TB máximo	40 TB máximo
Costo por GB al mes	Costo más alto	Costo más bajo
Solicitudes facturadas	PUT, COPY, POST, LIST y GET	SUBIR y recuperar
Precios de recuperación	¢ Por solicitud	¢¢ Por solicitud y por GB

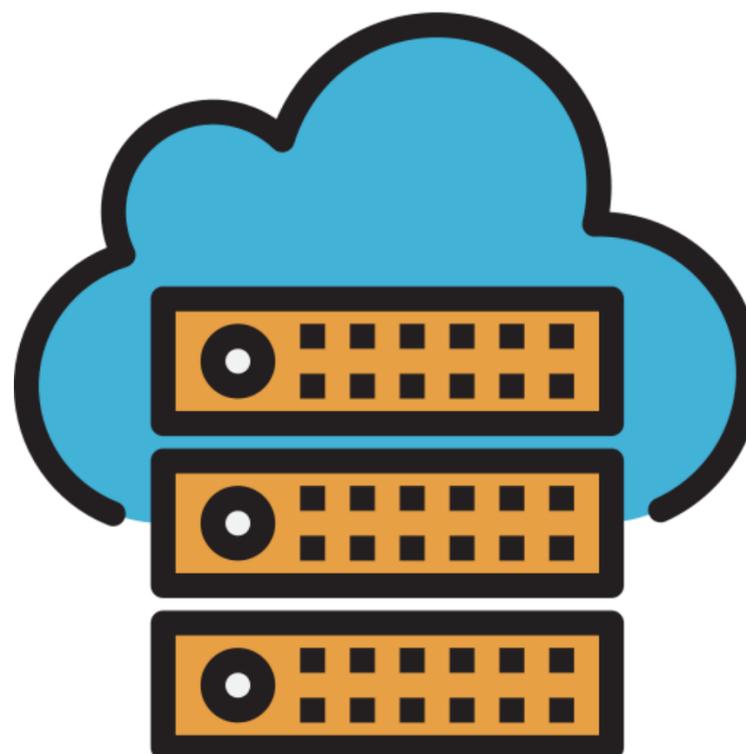
aws

© 2022, Amazon Web Services, Inc. o sus filiales. Todos los derechos reservados.

51



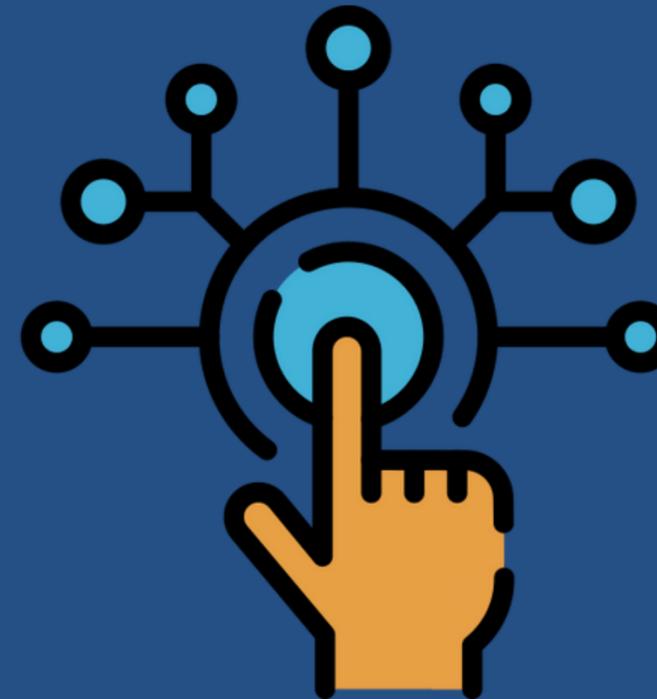
Aunque Amazon S3 y Amazon S3 Glacier son soluciones de almacenamiento de objetos que permiten almacenar una cantidad de datos prácticamente ilimitada, tienen algunas diferencias fundamentales. En el cuadro, se describen algunas de estas diferencias.





1. Tenga cuidado al decidir qué solución de almacenamiento es la adecuada para sus necesidades. Estos dos servicios satisfacen necesidades de almacenamiento muy diferentes.

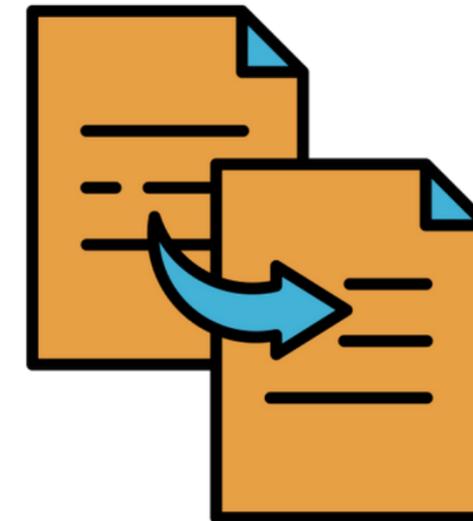
Amazon S3 está diseñado para proporcionar acceso frecuente y de baja latencia a los datos, mientras que Amazon S3 Glacier está diseñado para un almacenamiento a largo plazo y de bajo costo de los datos a los que se accede con poca frecuencia.





2. El tamaño máximo del elemento en Amazon S3 es de 5 TB, pero Amazon S3 Glacier puede almacenar elementos de hasta 40 TB.

3. Debido a que Amazon S3 le brinda un acceso más rápido a los datos, el costo de almacenamiento por gigabyte es mayor que con Amazon S3 Glacier.



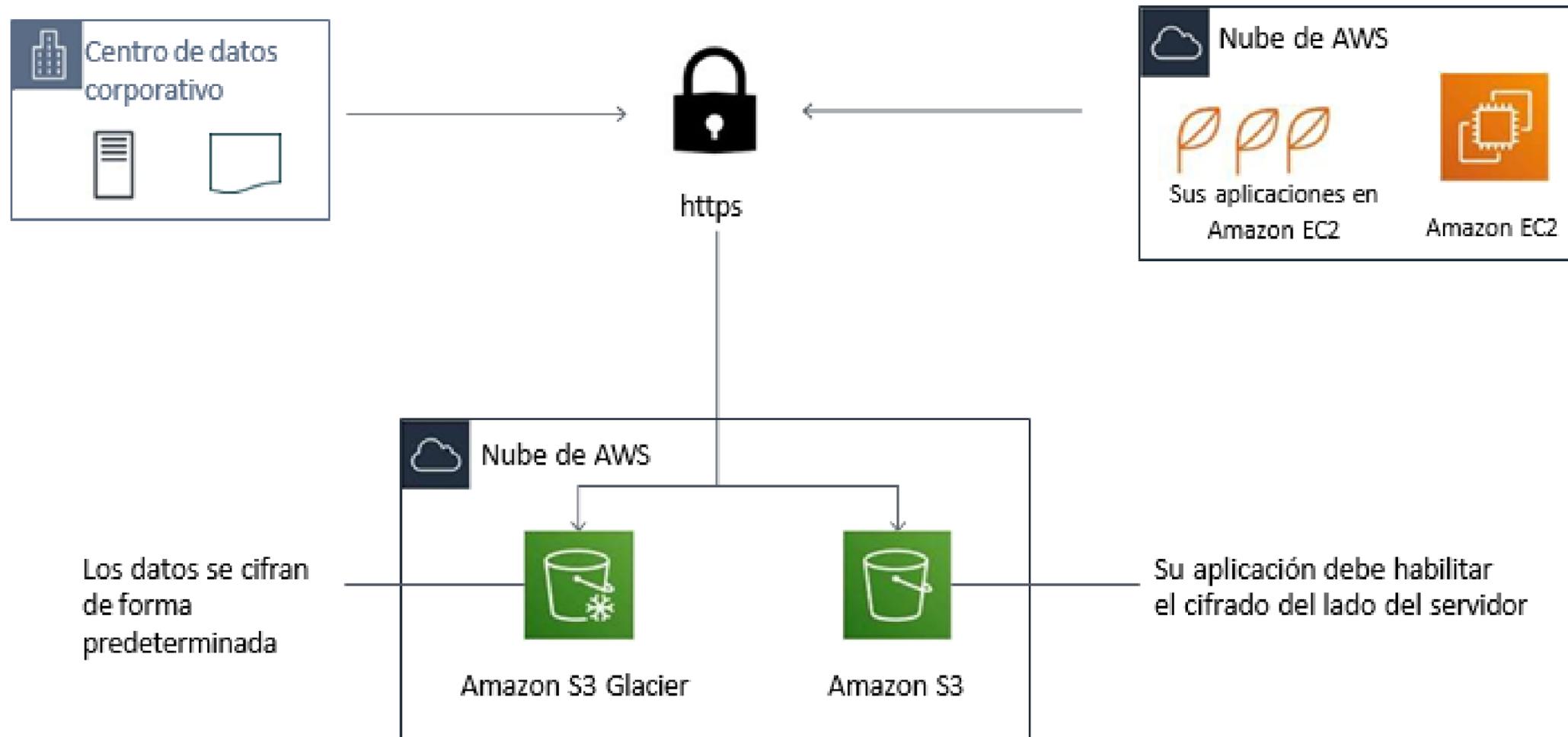
4. Aunque ambos servicios tienen cargos por solicitud, Amazon S3 cobra por las operaciones PUT, COPY, POST, LIST y GET. Por el contrario, Amazon S3 Glacier cobra por las operaciones UPLOAD y de recuperación.



5. Debido a que Amazon S3 Glacier se ha diseñado para un acceso menos frecuente a los datos, cada solicitud de recuperación cuesta más que en Amazon S3.



CIFRADO DEL LADO DEL SERVIDOR



© 2022, Amazon Web Services, Inc. o sus filiales. Todos los derechos reservados.

52

Otra diferencia importante entre Amazon S3 y Amazon S3 Glacier es cómo se cifran los datos. El cifrado del lado del servidor se enfoca en proteger los datos en reposo. Con ambas soluciones, puede transferir los datos de forma segura mediante HTTPS. Los datos que se archivan en Amazon S3 Glacier se cifran de forma predeterminada. Con Amazon S3, la aplicación debe iniciar el cifrado en el servidor. Puede lograr el cifrado del lado del servidor en Amazon S3 de varias formas.





El cifrado del lado del servidor con claves de cifrado administradas por Amazon S3 (SSE-S3) emplea un cifrado multifactor seguro. Amazon S3 cifra cada objeto a partir de una clave única. Como medida de seguridad adicional, cifra la propia clave con una clave principal que rota periódicamente. La encriptación del lado del servidor de Amazon S3 utiliza uno de los métodos de cifrado de bloques más potentes en el sector, el estándar de cifrado avanzado de 256 bits (AES-256), para cifrar los datos.

1.



Usar el cifrado del lado del servidor con claves de cifrado proporcionadas por el cliente (SSE-C) permite establecer sus propias claves de cifrado. Se proporciona la clave de cifrado como parte de la solicitud, y Amazon S3 administra tanto el cifrado (al escribir en los discos) como el descifrado (cuando accede a los objetos).

1.



Usar el cifrado del lado del servidor con AWS Key Management Service (AWS KMS) es un servicio que combina hardware y software seguros y de alta disponibilidad para proporcionar un sistema de administración de claves que se escala para la nube. AWS KMS utiliza Customer Master Keys (CMK, claves principales del cliente) para cifrar los objetos de Amazon S3. Se utiliza AWS KMS en la lección de claves de cifrado en la consola de IAM. También puede acceder a AWS KMS con la API para crear claves de cifrado de forma centralizada, definir las políticas que controlan cómo se pueden usar las claves y auditar el uso de las claves para demostrar que se estén utilizando de manera correcta.

Puede utilizar estas claves para proteger los datos de los depósitos de Amazon S3.



SEGURIDAD CON AMAZON S3 GLACIER



Amazon S3
Glacier



Controle el acceso con
IAM



Amazon S3 Glacier cifra sus
datos con AES-256



Amazon S3 Glacier
administra sus claves
en su nombre



© 2022, Amazon Web Services, Inc. o sus filiales. Todos los derechos reservados.

53

De forma predeterminada, solo usted puede acceder a sus datos. Puede habilitar y controlar el acceso a los datos en Amazon S3 Glacier usando IAM. Ha configurado una política de IAM que especifica el acceso de los usuarios.

Ha completado una introducción a Amazon S3 Glacier, que incluyó las diferencias principales entre Amazon S3 y Amazon S3 Glacier.

Para obtener más información sobre Amazon S3 Glacier, consulte Glacier.

 **Ahora tómese un momento para ver la demostración de Amazon Glacier. La grabación dura un poco más de 2 minutos y refuerza muchos de los conceptos que mencionamos en esta lección de la unidad.** 

En la demostración, se muestra cómo configurar los siguientes recursos con la Consola de administración de AWS. En la demostración, se muestra cómo hacer lo siguiente:

- **Crear un almacén de Amazon Glacier.**
- **Subir elementos archivados al almacén usando una herramienta de interfaz gráfica de terceros.**

ACTIVIDAD: CASOS PRÁCTICOS DE ALMACENAMIENTO

En esta actividad orientada por el instructor, se le pedirá que inicie sesión en la Consola de administración de AWS.

Las instrucciones de la actividad se encuentran en la siguiente diapositiva.

Se le solicitará que responda cinco preguntas. El instructor orientará a la clase en un análisis de cada pregunta y revelará las respuestas correctas.





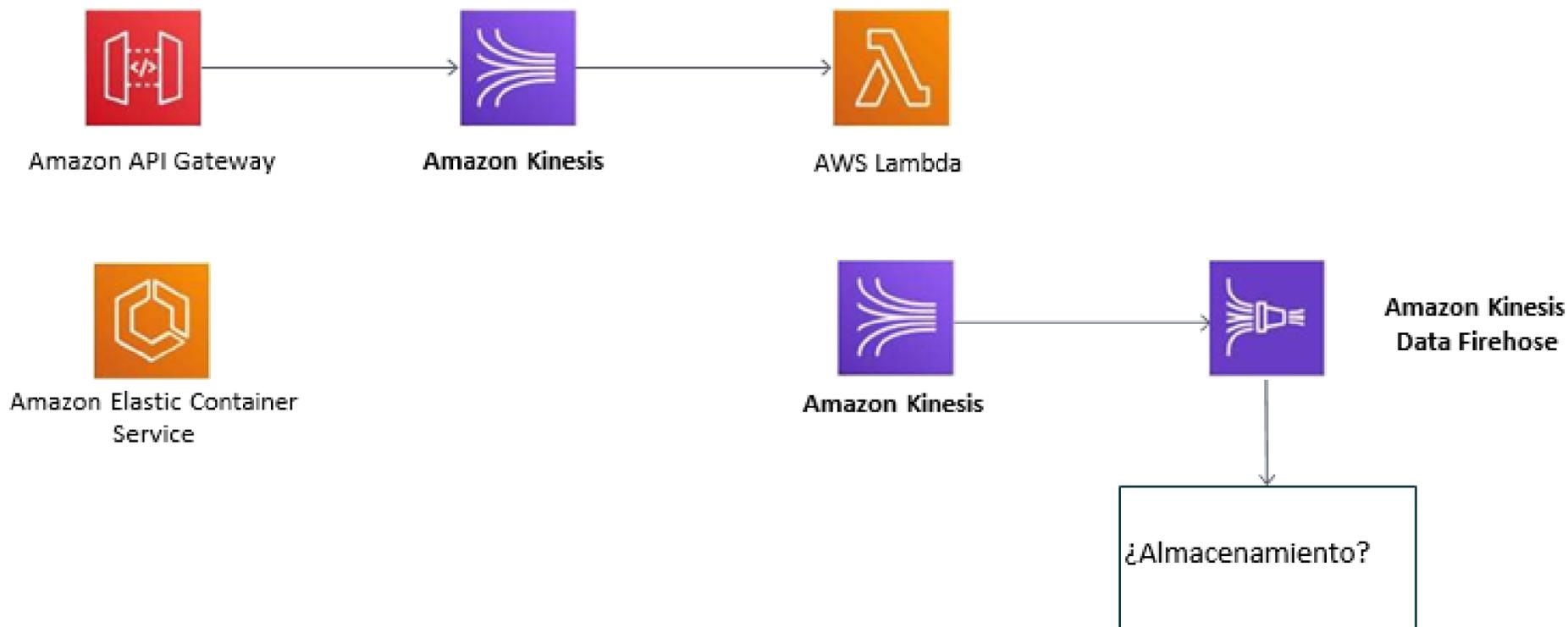
ACTIVIDAD DEL CASO PRÁCTICO DE ALMACENAMIENTO (1 DE 3)

Caso 1

Una empresa de análisis de datos para sitios de viajes debe almacenar miles de millones de eventos de clientes por día.

Utilizan los servicios de análisis de datos que se muestran en el diagrama. En el siguiente diagrama, se ilustra su arquitectura.





© 2021, Amazon Web Services, Inc. o sus filiales. Todos los derechos reservados.

Formen grupos de cuatro o cinco integrantes.

Repasen el caso práctico asignado.

Creen una presentación que describa la mejor solución de almacenamiento para la organización que se menciona en el caso de su grupo.

Su presentación debe contener los factores clave que se consideraron cuando se seleccionó la tecnología de almacenamiento y cualquier otro factor que pudiera cambiar la recomendación.

1.



ACTIVIDAD DEL CASO PRÁCTICO DE ALMACENAMIENTO (2 DE 3)

Caso 2

Una empresa de software de colaboración procesa el correo para clientes empresariales. Tienen más de 250 clientes empresariales y más de medio millón de usuarios.

Deben almacenar petabytes de datos para sus clientes. En el siguiente diagrama, se ilustra su arquitectura.



1. **Formen grupos de cuatro o cinco integrantes. Repasen el caso práctico asignado. Creen una presentación que describa la mejor solución de almacenamiento para la organización que se menciona en el caso de su grupo. Su presentación debe contener los factores clave que se consideraron cuando se seleccionó la tecnología de almacenamiento y cualquier otro factor que pudiera cambiar la recomendación.**
- 2.
- 3.

ACTIVIDAD DEL CASO PRÁCTICO DE ALMACENAMIENTO (3 DE 3)

Caso 3

Una empresa de protección de datos debe ingerir y almacenar grandes cantidades de datos de clientes y ayudarlos a satisfacer los requisitos de cumplimiento.

Utilizan Amazon EC2 para hacer cómputos escalables y Amazon DynamoDB para datos duplicados y búsquedas de metadatos. En el siguiente diagrama, se ilustra su arquitectura.



Formen grupos de cuatro o cinco integrantes. Repasen el caso práctico asignado. Creen una presentación que describa la mejor solución de almacenamiento para la organización que se menciona en el caso de su grupo. Su presentación debe contener los factores clave que se consideraron cuando se seleccionó la tecnología de almacenamiento y cualquier otro factor que pudiera cambiar la recomendación.