MODULO 09 - EJERCICIO 02-B

ALEXIS YURI M.

Diseñe una estrategia de almacenamiento escalonada para una empresa que maneja grandes volúmenes de datos (por ejemplo, videos educativos, registros de usuarios, logs del sistema y backups mensuales).

1. Tipos de Datos y Escenario.

Se analiza el caso de una "Empresa de Educación en Línea" que crea y distribuye contenido multimedia, Para este caso se definen 3 tipos de datos:

- Videos Educativos: El activo principal. Son archivos grandes y se accede a ellos con frecuencia inicialmente, pero luego el acceso disminuye con el tiempo a medida que se lanzan nuevos cursos.

- Registros de Usuarios (Logs): Datos generados por la interacción de los estudiantes con la plataforma. Son volúmenes muy grandes de datos (no estructurados) que se usan para análisis y son archivados a largo plazo.

- Backups Mensuales de la Base de Datos: Copias de seguridad de la base de datos de la plataforma, que incluyen información de usuarios y transacciones. Son críticos, se usan poco y deben conservarse por un período prolongado.

2. Tecnología de Almacenamiento.

Se alinea cada tipo de dato con la tecnología de AWS más adecuada, considerando su frecuencia de acceso, rendimiento y costo.

Videos Educativos:

Uso: Acceso frecuente a medio plazo.

Tecnología Recomendada: Amazon S3 con una estrategia de ciclo de vida.

- Los videos recién subidos se almacenan en S3 Standard para un acceso rápido e

inmediato.

- Después de 30 o 60 días sin acceso, una regla de ciclo de vida los mueve

automáticamente a S3 Standard-IA (Infrequent Access) para reducir costos.

- Después de un año, los videos muy antiguos pueden ser migrados a S3 Glacier

Flexible Retrieval para un archivo a largo plazo, ya que es muy poco probable que se

necesiten inmediatamente.

Registros de Usuarios (Logs):

Uso: Acceso poco frecuente, principalmente para análisis esporádicos. Retención a largo

plazo.

Tecnología Recomendada: Amazon S3 con una estrategia de archivo directo.

- Los logs diarios se suben a un bucket S3.

- La política de ciclo de vida mueve estos archivos a S3 Glacier Deep Archive tan

pronto como sea posible (ej. después de 7 días), dado que el acceso a estos datos es

extremadamente raro y los costos son mínimos.

Backups Mensuales de la Base de Datos:

Uso: Acceso muy poco frecuente. Se almacenan para recuperación ante desastres o

cumplimiento normativo.

Tecnología Recomendada: Amazon S3 Glacier.

- Las copias de seguridad se generan y se suben a S3 Glacier de manera programada.

El costo de este almacenamiento es muy bajo, lo que lo hace ideal para datos que

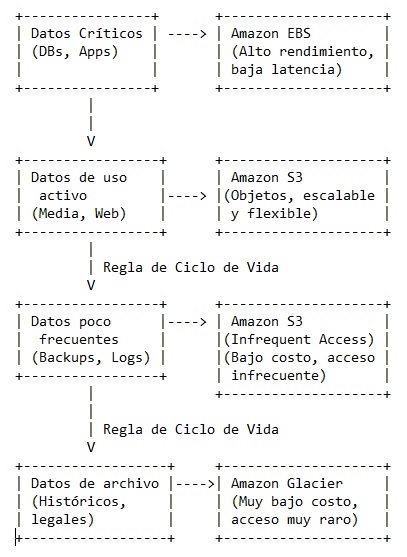
rara vez se necesitan.

- Para bases de datos de menor tamaño o que requieran un acceso más rápido en

caso de contingencia, se podría considerar S3 Standard-IA para las copias de

seguridad más recientes.

3. Representación Gráfica del Flujo de Datos.



Explicación del diagrama:

El diagrama muestra una arquitectura de almacenamiento escalonada. Cada tipo de dato se dirige a la solución de almacenamiento de AWS que mejor se adapta a sus necesidades de rendimiento, disponibilidad y costo.

- Amazon EBS: Se utiliza para datos que requieren el mayor rendimiento, como bases de datos y sistemas operativos de aplicaciones críticas. Se conecta directamente a las instancias de EC2 y funciona como un disco duro virtual.

- Amazon S3 Standard: Los datos que necesitan acceso frecuente, como archivos de sitios web o contenido multimedia recién subido, se almacenan aquí. Este servicio ofrece una alta durabilidad y disponibilidad.

- Amazon S3 Infrequent Access (S3 IA): Una vez que los datos de S3 Standard se usan con menos frecuencia, una política de ciclo de vida los mueve automáticamente a S3 IA. Esto reduce los costos mientras se mantiene la capacidad de acceso inmediato si es necesario.

- Amazon Glacier: Para los datos que ya no se usan, como copias de seguridad antiguas o información histórica, la política de ciclo de vida los traslada a Glacier. Este servicio es el más económico, pero tiene tiempos de recuperación más largos (minutos u horas).

Esta arquitectura demuestra la importancia de la estrategia de almacenamiento diferenciado para optimizar costos y garantizar el rendimiento adecuado para cada tipo de dato.