

Tecnologías de Almacenamiento en la Nube

El almacenamiento en la nube ha transformado profundamente la forma en que las organizaciones gestionan y resguardan sus datos, proporcionando opciones flexibles, escalables y eficientes. Las plataformas en la nube, como Google Cloud Platform (GCP), Amazon Web Services (AWS), Microsoft Azure y otras, ofrecen diversas soluciones que permiten a las empresas elegir la opción más adecuada según sus necesidades específicas, como rendimiento, costos, escalabilidad y seguridad.

R por Kibernetum Capacitación S.A.





Preguntas de Activación

1. ¿Qué es la nube pública?
2. ¿Qué servicio permite ejecutar código sin usar servidores?
3. Nombra una ventaja de la nube.



Tipos de Almacenamiento en la Nube

Almacenamiento de Objetos

Diseñado para almacenar grandes volúmenes de datos no estructurados. Los datos se organizan como "objetos" con metadatos relacionados en contenedores llamados "buckets".

Almacenamiento de Archivos

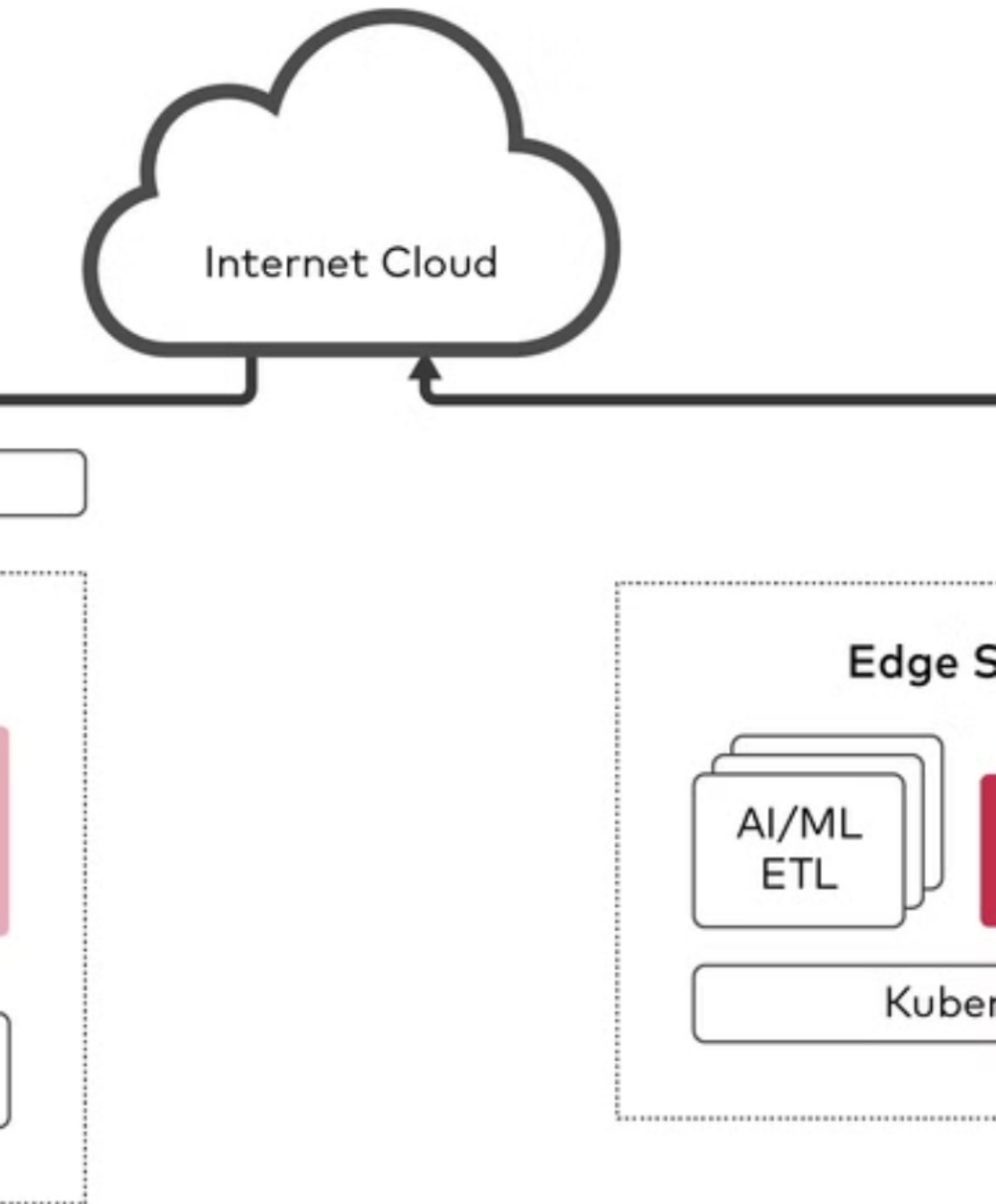
Funciona como sistemas de archivos tradicionales con jerarquía de directorios. Ideal para datos estructurados que requieren un sistema de gestión de archivos tradicional.

Almacenamiento de Bases de Datos

Opción eficiente para gestionar grandes volúmenes de datos estructurados, realizar consultas rápidas y mantener disponibilidad en tiempo real.

Almacenamiento de Bloques

Organiza la información en bloques con direcciones únicas para acceso directo y eficiente. Ideal para acceso rápido, flexible y directo a los datos.



Almacenamiento de Objetos



Características Principales

Tecnología diseñada para almacenar grandes volúmenes de datos no estructurados como archivos multimedia, documentos y backups. Los datos se organizan como "objetos" con metadatos en contenedores llamados "buckets".



Ventajas Clave

Ofrece escalabilidad ilimitada, accesibilidad global, durabilidad de 99.999999999% (11 9's), seguridad con cifrado en reposo y tránsito, y costos flexibles según frecuencia de acceso.



Ejemplos en Plataformas

Google Cloud Storage (GCS), Amazon S3 (Simple Storage Service) y Azure Blob Storage son los principales servicios de almacenamiento de objetos en sus respectivas plataformas.

Casos de Uso del Almacenamiento de Objetos



Archivos Estáticos

Almacenamiento de imágenes, videos y otros archivos que se requieren en aplicaciones web o móviles, aprovechando la alta disponibilidad y escalabilidad.



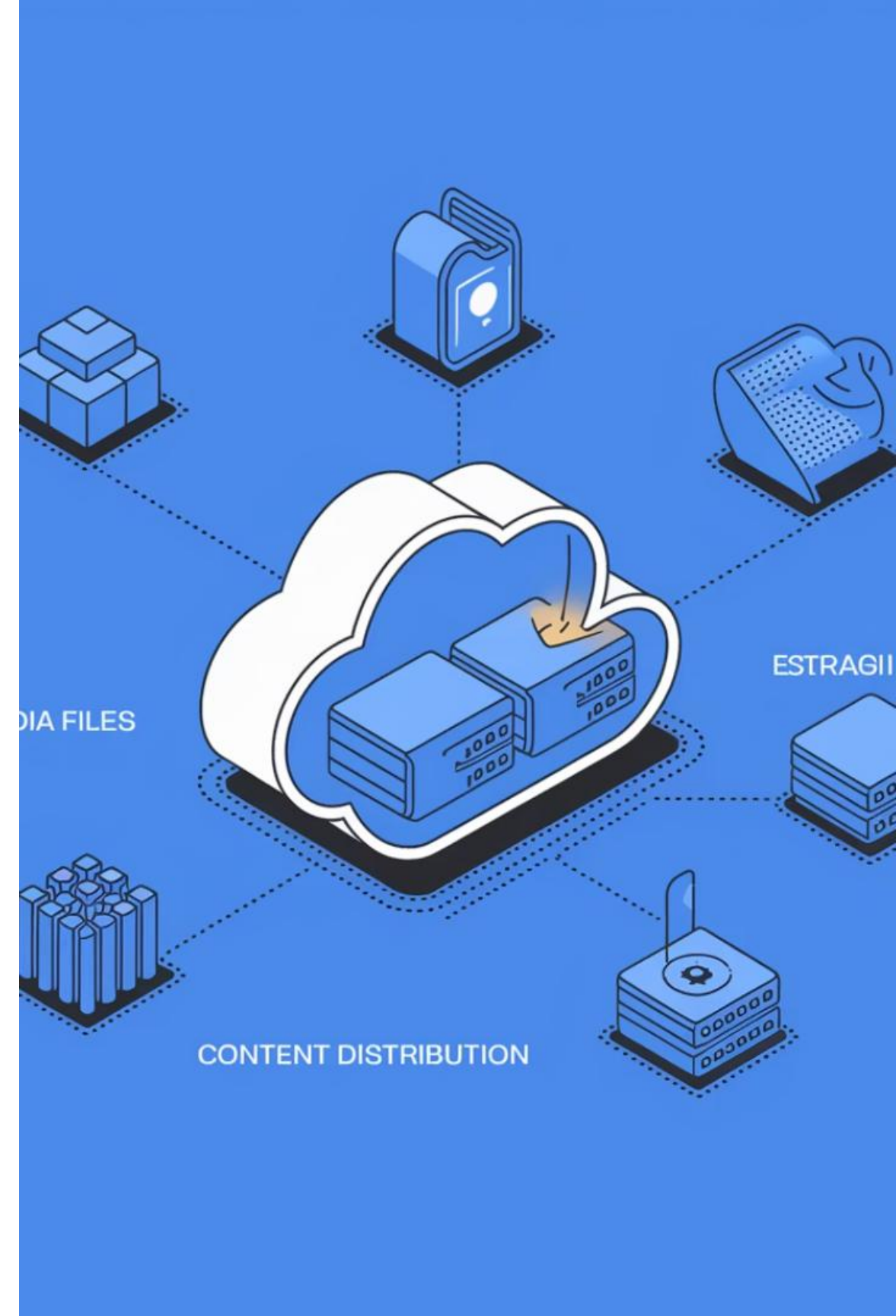
Backups y Archivos

Ideal para hacer copias de seguridad de datos o almacenar archivos a largo plazo, con costos ajustados según la frecuencia de acceso a la información.



Distribución de Contenido

Perfecto para la distribución de archivos grandes, como videos en streaming, debido a su alta disponibilidad y acceso rápido desde cualquier ubicación.



Almacenamiento de Archivos

Definición y Funcionamiento

El almacenamiento de archivos en la nube funciona de manera similar a los sistemas tradicionales, pero ubicados en la nube. Los datos se almacenan en una jerarquía de directorios y archivos, facilitando su acceso y manipulación.

Este tipo es ideal para datos estructurados que requieren un sistema de gestión de archivos tradicional, como documentos, imágenes o bases de datos.

Servicios por Plataforma

Google Filestore: Proporciona almacenamiento de archivos para aplicaciones que requieren acceso a un sistema de archivos en la nube con una latencia baja.

Amazon EFS: Ofrece almacenamiento escalable y elástico para aplicaciones que requieren estructura de directorios.

Azure Files: Proporciona almacenamiento accesible mediante el protocolo SMB, adecuado para aplicaciones que requieren una interfaz tradicional de red de archivos.

Ventajas del Almacenamiento de Archivos



Compatibilidad con aplicaciones tradicionales

Ideal para aplicaciones heredadas que requieren un sistema de archivos tradicional



Acceso en tiempo real

Permite el acceso en tiempo real para aplicaciones que necesitan modificar o leer datos continuamente



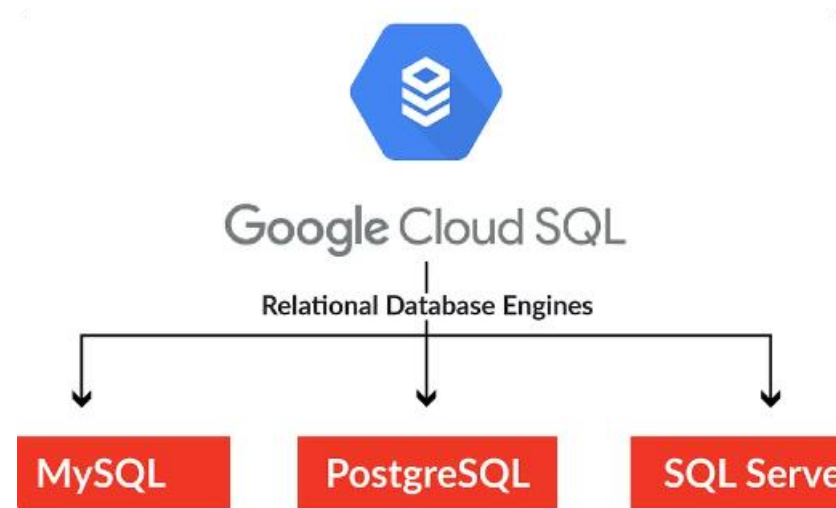
Integración con redes

Los archivos se pueden montar como discos en servidores, facilitando la integración con sistemas existentes

El almacenamiento de archivos en la nube es especialmente útil para aplicaciones empresariales que requieren compartir archivos entre múltiples usuarios o sistemas, así como para sistemas de gestión de contenidos (CMS) que necesitan estructura y fácil acceso a los documentos y archivos multimedia.

Almacenamiento de Bases de Datos en la Nube

El almacenamiento de bases de datos en la nube es una opción altamente eficiente y flexible para empresas que necesitan gestionar grandes volúmenes de datos estructurados, realizar consultas rápidas y mantener la disponibilidad de la información en tiempo real.



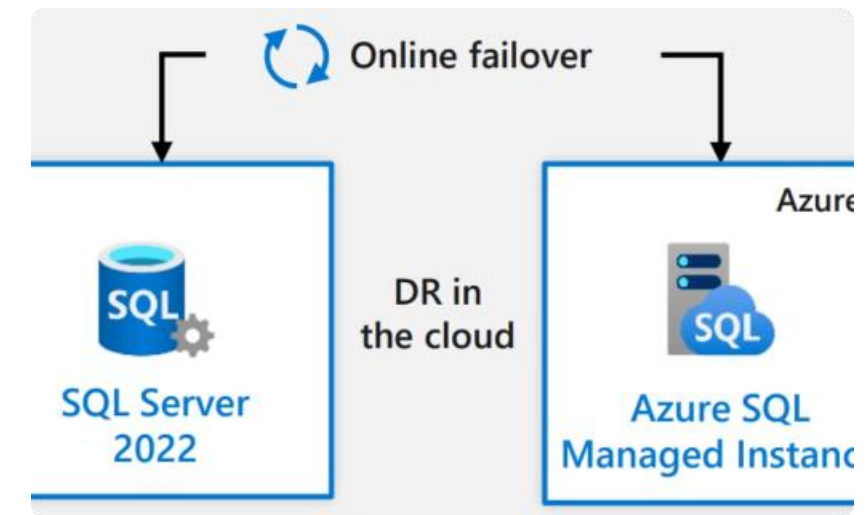
Google Cloud SQL / BigQuery

Cloud SQL es ideal para bases de datos estructuradas y admite sistemas como MySQL, PostgreSQL y SQL Server. BigQuery está diseñado específicamente para realizar consultas analíticas de grandes conjuntos de datos en tiempo real, con una capacidad de escalabilidad impresionante.



Amazon RDS

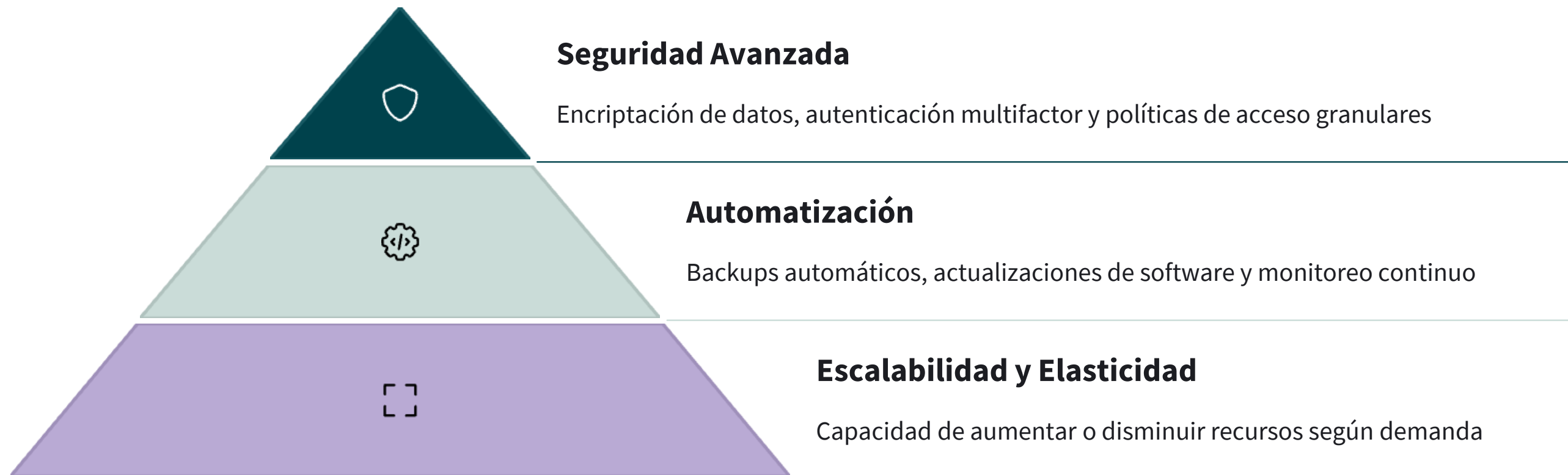
Amazon RDS es un servicio completamente gestionado que facilita la administración de bases de datos relacionales en la nube, como MySQL, PostgreSQL, Oracle, MariaDB y Amazon Aurora. Automatiza tareas como el aprovisionamiento, configuración y backups.



Azure SQL Database

Azure SQL Database es un servicio gestionado que proporciona una solución de base de datos relacional en la nube para aplicaciones empresariales. Está basado en la plataforma SQL Server de Microsoft, y ofrece alta disponibilidad, seguridad y escalabilidad.

Ventajas del Almacenamiento de Bases de Datos



Las bases de datos en la nube están optimizadas para soportar tanto aplicaciones pequeñas como grandes, permitiendo a las organizaciones ajustar la capacidad de almacenamiento y procesamiento según la demanda, de manera dinámica y eficiente. Los servicios gestionados facilitan la administración sin necesidad de gestionar la infraestructura física subyacente.

Casos de Uso de Bases de Datos en la Nube

Aplicaciones Web y Móviles

Las bases de datos en la nube son ideales para aplicaciones web y móviles que necesitan gestionar grandes cantidades de información de forma eficiente. Las empresas pueden usar bases de datos en la nube para manejar usuarios, productos, transacciones, inventarios, y más.

Este tipo de aplicaciones requiere bases de datos que sean rápidas, escalables y de fácil acceso para ofrecer una experiencia fluida a los usuarios finales. Servicios como Amazon RDS o Azure SQL Database son frecuentemente utilizados en este tipo de escenarios.

Análisis de Datos en Tiempo Real

Las bases de datos en la nube son cada vez más utilizadas para el análisis de grandes volúmenes de datos en tiempo real. Esto es crucial para sectores como el análisis financiero, la inteligencia de negocios, y el comercio electrónico, donde se necesitan informes rápidos y decisiones basadas en datos en tiempo real.

Herramientas como BigQuery (Google Cloud) permiten analizar grandes conjuntos de datos de manera casi instantánea, facilitando la toma de decisiones informadas rápidamente. La capacidad de escalar recursos según la demanda permite realizar análisis incluso con picos de tráfico.

Almacenamiento de Bloques

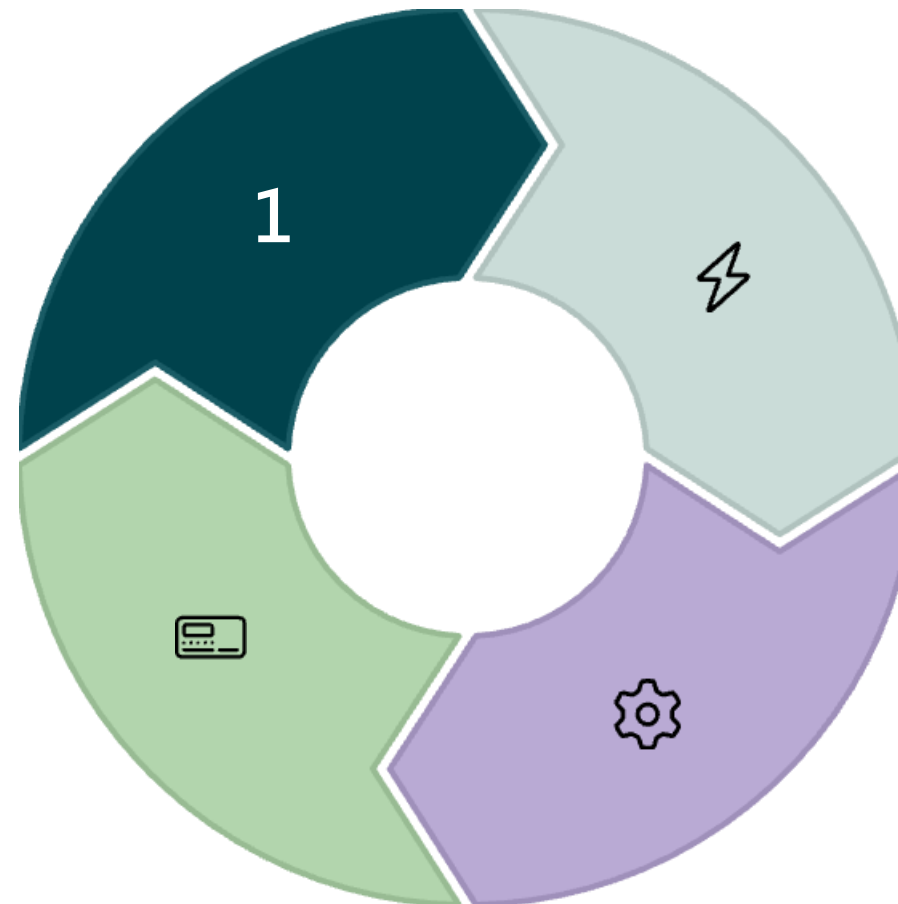
El almacenamiento de bloques es un tipo de almacenamiento de datos en el cual la información se organiza en bloques, y cada bloque tiene una dirección única que permite acceder a los datos de manera directa y eficiente.

Organización en Bloques

Los datos se dividen en unidades discretas con direcciones únicas

Alto Desempeño

Ofrece alta capacidad de IOPS y baja latencia



Acceso Rápido

Permite operaciones ágiles con control granular sobre los datos

Flexibilidad

Ideal para diferentes tipos de datos y aplicaciones

Google Cloud Filestore vs Persistent Disk vs Cloud

Google Persistent Disk



Rendimiento

Diseñado para ofrecer alto rendimiento, con baja latencia y alta capacidad de IOPS, ideal para aplicaciones como bases de datos. Los discos pueden configurarse con diferentes tipos (SSD o HDD) según las necesidades específicas de la carga de trabajo.



Escalabilidad

Los discos pueden escalarse fácilmente. Los usuarios pueden aumentar el tamaño de un disco persistente en cualquier momento sin afectar la disponibilidad de los datos, o cambiar el tipo de disco para mejorar el rendimiento según las necesidades.



Resiliencia y Disponibilidad

Los datos están replicados dentro de la infraestructura de Google Cloud, mejorando la disponibilidad y asegurando la durabilidad. Los discos persistentes se distribuyen para estar disponibles continuamente, incluso ante fallos de hardware o zona geográfica.

Casos de Uso de Google Persistent Disk

1

Bases de Datos de Alto Rendimiento

Ideal para bases de datos como MySQL, PostgreSQL o MongoDB que se benefician de la alta IOPS y baja latencia que ofrece el almacenamiento en bloques.



Sistemas Operativos de Máquinas Virtuales

Para discos de arranque y almacenamiento adicional en máquinas virtuales que se ejecutan en Google Compute Engine.



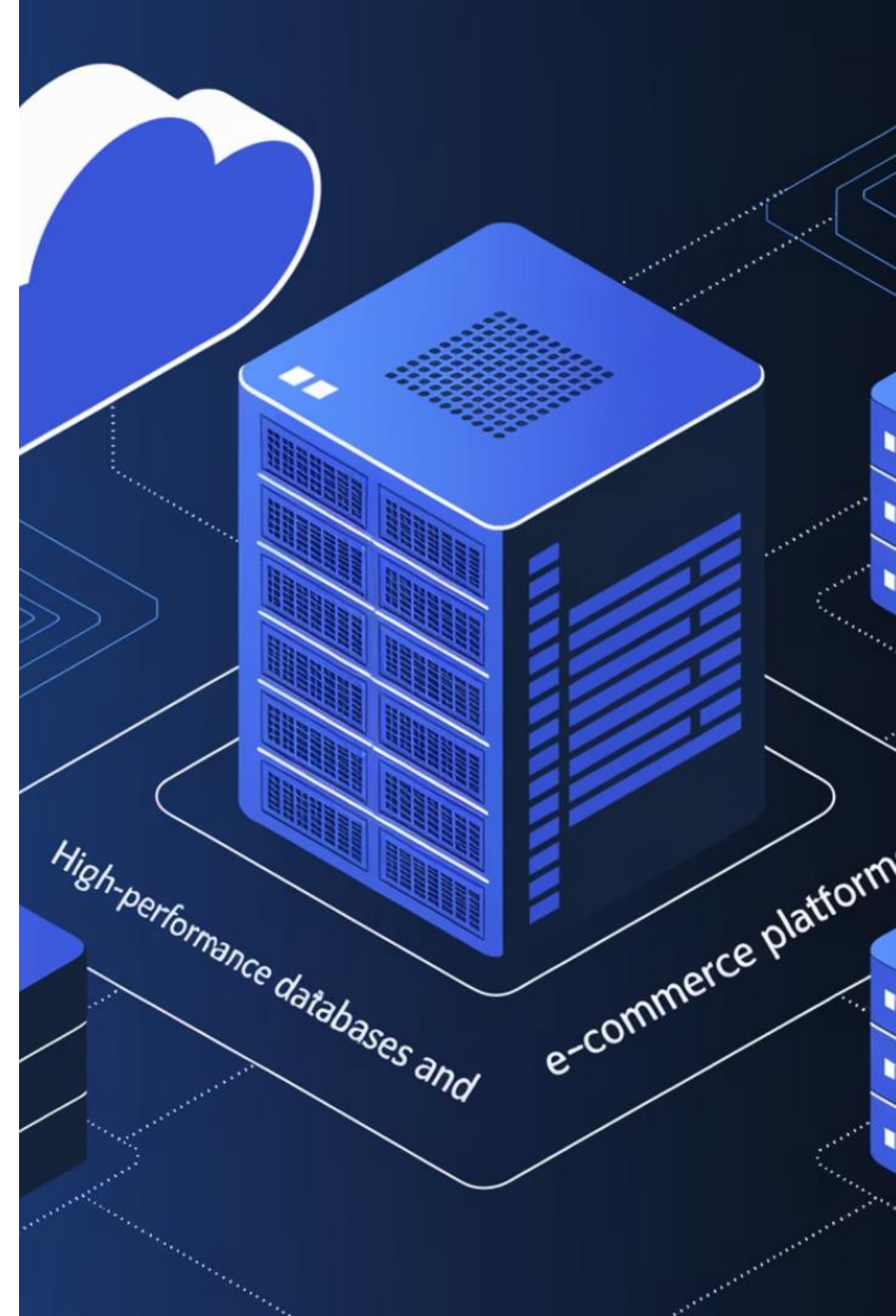
Aplicaciones Empresariales de Alta Capacidad

Para plataformas de comercio electrónico, análisis en tiempo real o sistemas ERP que necesitan alta capacidad de lectura/escritura.



Entornos de Desarrollo y Pruebas

Para desarrolladores que necesitan consistencia y rapidez en el acceso a los datos al probar diversas configuraciones.



Ventajas y Desventajas de Google Persistent Disk

Ventajas

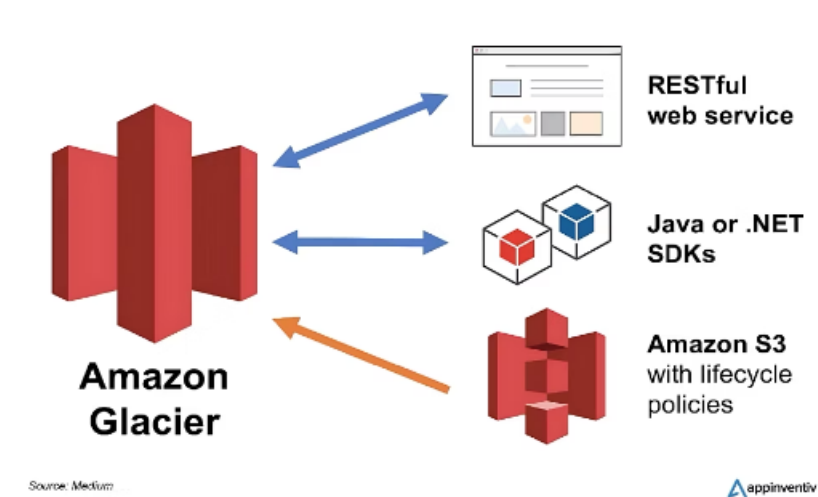
- Alto rendimiento con baja latencia y gran capacidad de IOPS
- Escalabilidad flexible permitiendo aumentar tamaño y cambiar tipo de disco
- Resiliencia y alta disponibilidad mediante replicación de datos
- Protección de datos incluso en caso de fallos de hardware o zonas geográficas

Desventajas

- Puede resultar costoso si no se aprovechan completamente las altas IOPS
- Los discos HDD tienen rendimiento inferior en latencia y capacidad de acceso
- La gestión de discos puede representar una carga administrativa
- Dependencia de Google Cloud puede ser una desventaja para soluciones multi-proveedor

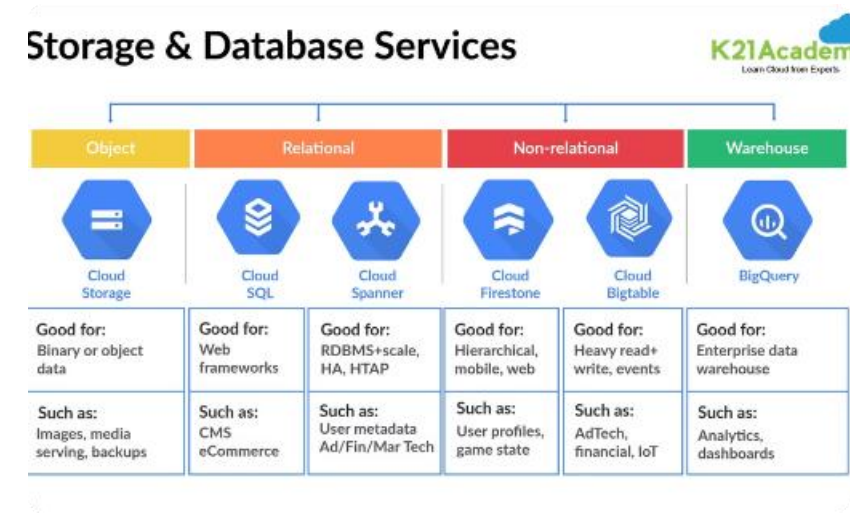
Almacenamiento a Largo Plazo

El almacenamiento a largo plazo se refiere a la solución para guardar datos que no necesitan ser accedidos con frecuencia pero que deben ser conservados por largos períodos de tiempo. Este tipo de almacenamiento está diseñado para ofrecer una forma rentable y segura de conservar datos históricos, copias de seguridad, y otros datos críticos.



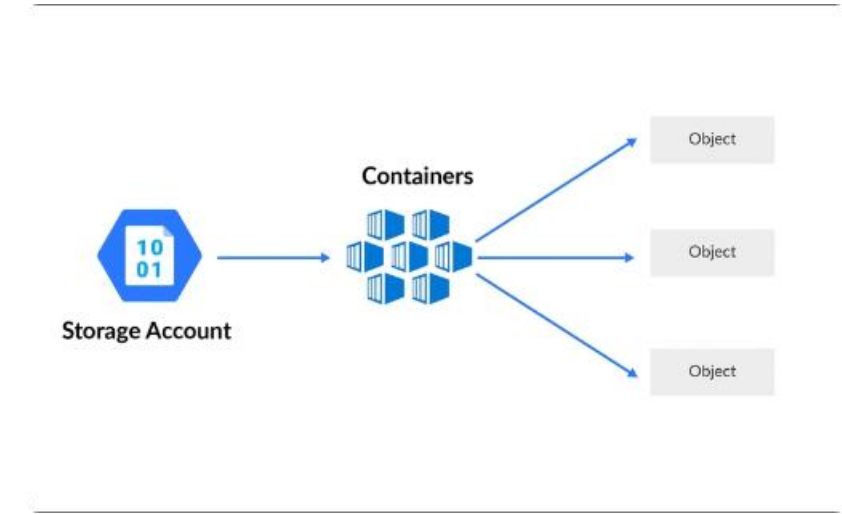
Amazon Glacier

Servicio de AWS diseñado específicamente para el archivado de datos que se acceden raramente. Ideal para almacenar datos históricos, copias de seguridad y archivos que deben conservarse por razones legales o normativas, pero que no se utilizan regularmente.



Google Cloud Coldline

Diseñado para almacenar datos que se acceden de forma poco frecuente, pero que deben ser conservados por períodos prolongados. Es más económico que las clases de almacenamiento estándar, pero tiene una latencia más alta en el acceso a los datos.



Azure Archive Storage

Opción de Microsoft para el almacenamiento a largo plazo de datos raramente accedidos. Ofrece los costos más bajos de almacenamiento con tiempos de recuperación más largos, ideal para cumplimiento normativo y archivado.



Características de Amazon Glacier



Diseñado para Archivado

Ideal para el archivado de datos que se acceden raramente. Perfecto para almacenar datos históricos, copias de seguridad y archivos que deben ser conservados por razones legales, normativas o de cumplimiento, pero que no se utilizan regularmente.



Eficiencia y Durabilidad

Glacier está diseñado para ser altamente durable. Los datos se almacenan de manera redundante en varias instalaciones físicas dentro de una región de AWS, lo que protege la información contra pérdidas.



Integración con Otros Servicios

Se integra fácilmente con otros servicios de AWS, como Amazon S3. Esto permite mover datos entre S3 y Glacier de manera fluida, utilizando políticas de ciclo de vida para automatizar el proceso de migración.



Opciones de Recuperación en Amazon Glacier



Recuperación Expedited

Permite acceder a datos en minutos, adecuada para situaciones donde se necesitan datos de forma más rápida.



Recuperación Estándar

Generalmente puede tardar varias horas, adecuada para necesidades de recuperación que no son urgentes.

3

Recuperación en Masa

Este método es la opción más económica y puede tardar entre 5 a 12 horas, ideal para la recuperación de grandes volúmenes de datos.

Amazon Glacier ofrece diferentes opciones para la recuperación de datos, permitiendo a los usuarios elegir entre velocidad y costo según sus necesidades específicas. Aunque está diseñado para el almacenamiento a largo plazo, estas opciones de recuperación proporcionan flexibilidad cuando se necesita acceder a los datos archivados.

Características de Google Cloud Coldline

Costo Efectivo

Considerablemente más barato que otras opciones de almacenamiento, lo que lo convierte en una opción ideal para datos que no se acceden con frecuencia. Su modelo de precios es ideal para almacenar grandes cantidades de datos a largo plazo.

Accesibilidad Bajo Demanda

Aunque los datos almacenados en Coldline pueden ser accedidos en cualquier momento, el tiempo de recuperación es más largo en comparación con las clases de almacenamiento estándar.

Resiliencia

Google Cloud Coldline replica los datos en múltiples ubicaciones geográficas, garantizando que los datos estén disponibles incluso en caso de fallos de infraestructura o desastres naturales (99.999999999% de durabilidad anual).

Integración

Se integra fácilmente con otros servicios de Google Cloud, como Cloud Storage y Cloud Functions, facilitando su uso dentro de arquitecturas de almacenamiento y recuperación de datos más complejas.

Casos de Uso del Almacenamiento a Largo Plazo



Archivado de Datos de Respaldo

Ideal para almacenar copias de seguridad que se utilizan solo en caso de emergencia o recuperación ante desastres. El costo bajo y la alta resiliencia permiten almacenar grandes volúmenes de datos sin incurrir en gastos significativos.



Almacenamiento de Datos Históricos

Muchas empresas deben cumplir con normativas que requieren mantener ciertos datos durante varios años. Esto incluye datos financieros, registros médicos, comunicaciones y otros documentos históricos.



Cumplimiento Normativo

En sectores como la salud, la banca y el gobierno, la retención de datos es un requisito legal. El almacenamiento a largo plazo permite cumplir con normativas de privacidad y otros marcos regulatorios.

Servicios de Almacenamiento en la Nube

GCP, AWS y Azure ofrecen una variedad de servicios de almacenamiento en la nube, cada uno diseñado para necesidades específicas de rendimiento, acceso y costo. Estos servicios permiten a las organizaciones elegir la solución más adecuada según sus requerimientos particulares.



Cada proveedor de servicios en la nube ofrece soluciones para almacenamiento de objetos, bloques, archivos y bases de datos, con características específicas que se adaptan a diferentes casos de uso y necesidades empresariales.

Amazon S3 (Simple Storage Service)



Acceso de Baja Latencia

Optimizado para ofrecer acceso rápido a los datos almacenados, esencial para aplicaciones que requieren tiempos de respuesta mínimos. Su integración con Amazon CloudFront garantiza una entrega eficiente de archivos a nivel global.



Escalabilidad y Flexibilidad

Permite almacenar desde unos pocos objetos hasta billones sin necesidad de aprovisionar infraestructura. El almacenamiento crece automáticamente con las necesidades de la aplicación y es compatible con múltiples patrones de acceso.



Durabilidad y Seguridad

Garantiza una durabilidad del 99.999999999% mediante replicación automática en múltiples zonas. Proporciona cifrado de datos tanto en reposo como en tránsito y permite aplicar políticas detalladas de acceso.

S3 Storage Classes



Infrequent access



Reduce Redundancy

Modelos de Costos en Amazon S3



S3 Standard

Diseñado para datos de acceso frecuente, ofrece alta disponibilidad y rendimiento para aplicaciones que requieren acceso rápido y constante a los datos.



S3 Intelligent-Tiering

Optimiza automáticamente los costos moviendo los datos entre dos niveles de acceso (frecuente e infrecuente) basándose en los patrones de uso.



S3 Standard-IA y One Zone-IA

Diseñados para datos de acceso infrecuente pero que requieren disponibilidad inmediata cuando se necesitan. One Zone-IA almacena datos en una sola zona de disponibilidad para reducir costos.



S3 Glacier y Glacier Deep Archive

Opciones de bajo costo para archivado a largo plazo, con tiempos de recuperación más largos. Glacier Deep Archive ofrece los costos más bajos para datos que raramente se acceden.

Casos de Uso Principales de Amazon S3



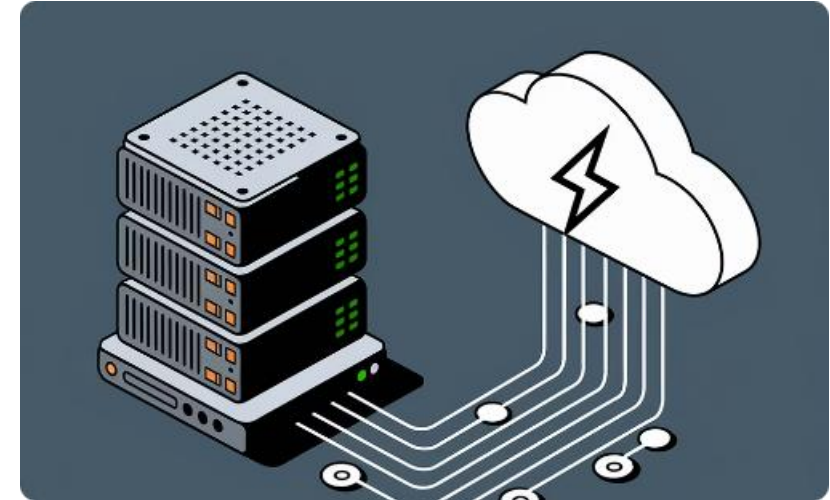
Almacenamiento de Archivos Estáticos

Ideal para alojar imágenes, videos, archivos de audio, documentos y otros contenidos estáticos que deben ser accesibles desde aplicaciones web y móviles. Su integración con CloudFront permite distribuir este contenido de forma rápida y eficiente a nivel mundial.



Manejo de Grandes Volúmenes de Datos

Ampliamente utilizado para almacenar grandes volúmenes generados por aplicaciones, dispositivos IoT, logs o sistemas. Estos pueden ser consumidos directamente por servicios como Amazon Athena, EMR o Redshift para análisis y procesamiento.



Respaldo y Recuperación ante Desastres

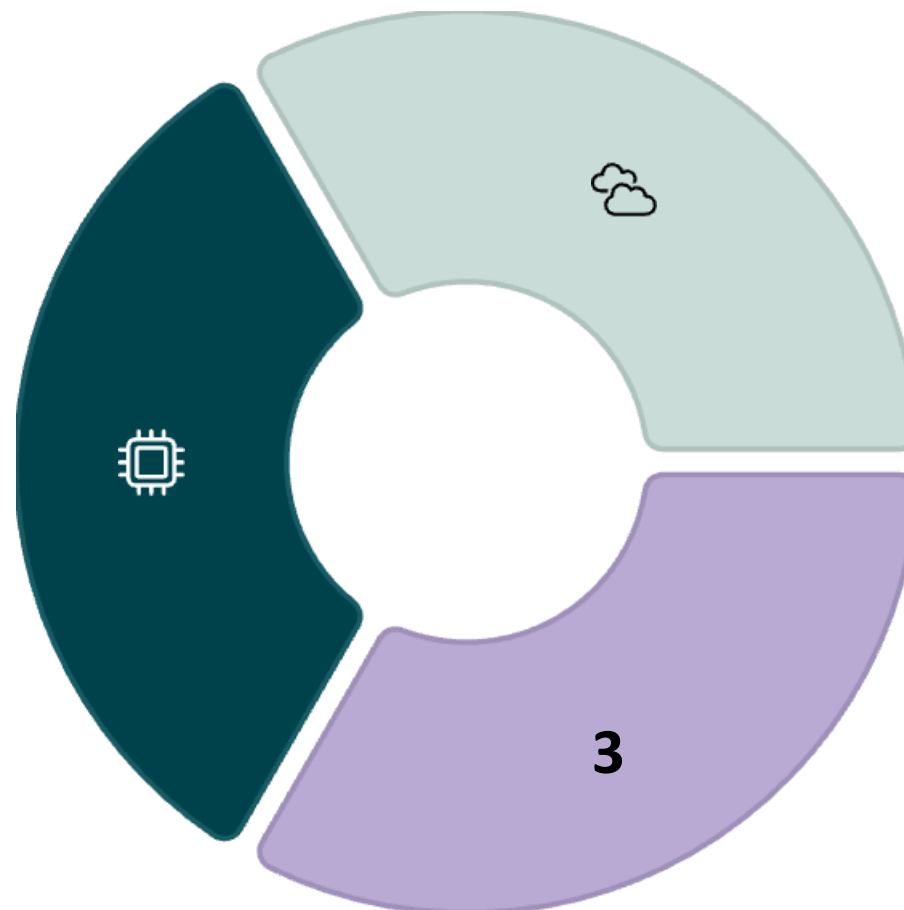
Muchas organizaciones utilizan S3 como destino para copias de seguridad y archivos críticos, gracias a su durabilidad y herramientas como AWS Backup. En caso de desastres, los datos pueden recuperarse fácilmente para garantizar la continuidad operativa.

Almacenamiento de Alto Rendimiento

El almacenamiento de alto rendimiento en la nube es fundamental para cargas de trabajo que requieren un procesamiento intensivo de datos, acceso rápido y consistente, y una alta disponibilidad.

Características Técnicas

Discos SSD y NVMe persistentes con velocidades superiores a 900 MB/s y cientos de miles de IOPS. Baja latencia y alto throughput para aplicaciones sensibles al tiempo.



Proveedores Principales

Google Cloud: Persistent Disk (Standard, SSD), Local SSD, Filestore High Scale.

AWS: EBS io2/io2 Block Express, Instance Store, FSx for Lustre.

Azure: Ultra Disk Storage, Premium SSD v2, Azure NetApp Files.

Casos de Uso

Bases de datos transaccionales, Machine Learning y Deep Learning, streaming de video en 4K/8K, sistemas financieros y bolsas electrónicas.

Amazon EBS (Elastic Block Store)

256K

IOPS Máximas

Los volúmenes EBS permiten hasta 256,000 IOPS para cargas de trabajo intensivas

4GB/s

Throughput

Rendimiento de lectura/escritura de hasta 4,000 MB/s

99.999%

Disponibilidad

Alta disponibilidad para aplicaciones críticas

Amazon Elastic Block Store (Amazon EBS) es un servicio de almacenamiento en bloques diseñado para usarse con instancias Amazon EC2. Su arquitectura permite ofrecer rendimiento elevado, baja latencia y alta disponibilidad, lo que lo convierte en una solución ideal para aplicaciones que requieren acceso rápido y constante a grandes volúmenes de datos.

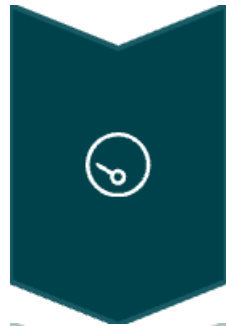
EBS Snapshots?

- Point-in-time backups
- Stored on Amazon S3
- Properties:
 - Incremental – only changes
 - Crash consistent – consistent snapshot
 - Crash consistent snapshot
1 API call for a subsecond instance
 - Can be shared and copied



amazon S3

Características de Amazon EBS



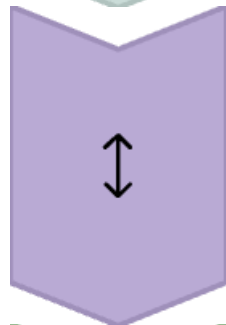
Rendimiento IOPS y throughput ajustable

Los volúmenes io2 e io2 Block Express están optimizados para cargas de trabajo intensivas en operaciones de entrada/salida por segundo (IOPS). El tipo gp3 permite ajustar IOPS y throughput de forma independiente del tamaño.



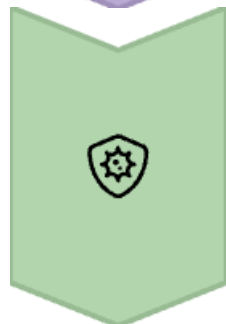
Baja latencia y alta disponibilidad

Ofrece latencias consistentes de solo unos milisegundos, ideales para sistemas sensibles al tiempo de respuesta. Los volúmenes se replican automáticamente dentro de una zona de disponibilidad.



Escalabilidad sin interrupciones

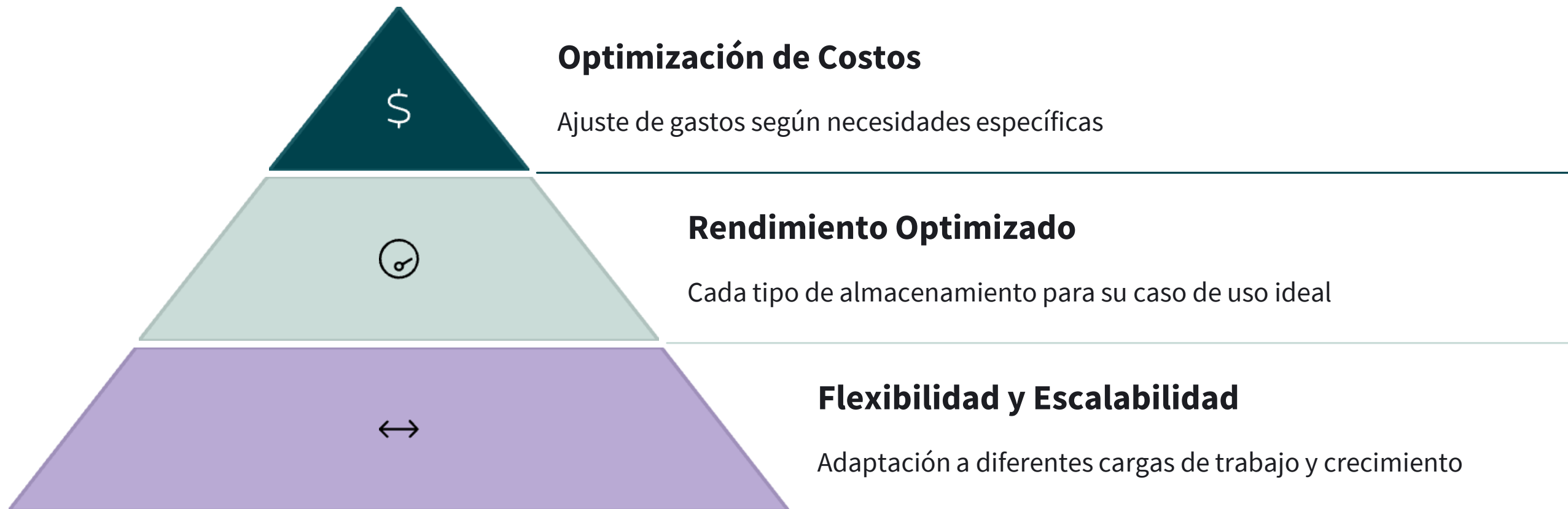
Se pueden aumentar el tamaño, el tipo o la capacidad de rendimiento de un volumen EBS sin reiniciar la instancia EC2, adaptándose a nuevas exigencias sin afectar la disponibilidad.



Seguridad y cifrado integrados

Cifrado nativo de datos en reposo usando claves administradas por AWS KMS. Protección de los datos tanto en tránsito como en almacenamiento, sin afectar el rendimiento.

Beneficios de Utilizar Almacенamientos Diferenciados



El uso de diferentes tipos de almacenamiento según las necesidades específicas de las aplicaciones es una estrategia clave para optimizar el rendimiento, la flexibilidad y la eficiencia de los recursos en la infraestructura de la nube. La elección adecuada puede marcar una diferencia significativa en términos de rendimiento y costos.

Flexibilidad y Optimización del Rendimiento

Almacenamiento de Objetos

Ideal para datos estáticos como imágenes, videos y archivos de respaldo. Se destaca por su alta durabilidad, escalabilidad y disponibilidad global. Es más económico que otros tipos de almacenamiento.

Aunque el acceso puede ser más lento, es perfecto para almacenar datos estáticos que no necesitan acceso frecuente ni latencia baja. Su rendimiento es óptimo para lecturas y escrituras simples de archivos multimedia o registros históricos.

Almacenamiento en Bloques

Ofrece alto rendimiento y es especialmente útil cuando se necesita acceso rápido y frecuente a datos. Es adecuado para almacenar bases de datos, aplicaciones críticas y sistemas operativos.

Permite lecturas y escrituras aleatorias de alta velocidad, ideal para bases de datos transaccionales y aplicaciones empresariales. Los usuarios pueden elegir entre discos SSD de alto rendimiento o HDD más económicos según sus necesidades.



Reducción de Costos y Eficiencia



Ajuste de Costos por Frecuencia de Acceso

Los diferentes tipos de almacenamiento tienen precios distintos, permitiendo ajustar costos según la frecuencia de acceso a los datos. Servicios como Coldline o Archive Storage son económicos para datos poco accedidos.



Almacenamiento Económico para Datos Inactivos

El almacenamiento en Coldline o Archive Storage es ideal para archivar datos que solo se usan ocasionalmente, optimizando el presupuesto sin perder la capacidad de recuperación.



Eficiencia con Instantáneas

El uso de instantáneas (snapshots) permite crear copias de seguridad de manera eficiente y económica, almacenándolas en tipos de almacenamiento más baratos sin duplicar los datos.

Estrategia de Uso de Diferentes Tecnologías

Tipo de Almacenamiento	AWS	Azure	GCP
Objetos	Amazon S3: Alta durabilidad, escalabilidad, acceso global	Azure Blob Storage: Escalable, con niveles de acceso	Cloud Storage: Alta disponibilidad y durabilidad
Bloques	Amazon EBS: Bajo tiempo de respuesta, ideal para bases de datos	Azure Disk Storage: Discos SSD/HDD gestionados	Persistent Disks: Rendimiento constante, expansión en caliente
Archivos	Amazon EFS: Sistema de archivos elástico, multiacceso	Azure Files: Compartido mediante SMB, integración con AD	Filestore: Sistema de archivos de alto rendimiento

La selección del almacenamiento adecuado depende del tipo de datos, la frecuencia de acceso, el rendimiento requerido y las características específicas del caso de uso. Una estrategia efectiva combina diferentes servicios para optimizar costos y rendimiento según las necesidades específicas de cada aplicación.

Enlaces de Interés

- Video: [¿Qué es el ALMACENAMIENTO en la NUBE? Ventajas y Características](#)
- Video: [¿Qué es Amazon S3?](#)
- Video: [Que son Amazon Elastic File System \(Amazon EFS\)](#)
- Enlace: [Amazon S3](#)

Preguntas clave sobre almacenamiento y análisis de datos



¿Qué tipo de almacenamiento se recomienda para archivos como imágenes y videos?



¿Qué ventaja tiene usar almacenamiento en la nube frente al local?



¿Qué servicio de Google permite analizar grandes volúmenes de datos?

