



UD2

Infraestructura
y almacenamiento

COP055
Computación
en la nube

2.2 Almacenamiento
en la nube

Introducción

En la computación en la nube existen distintas formas de organizar y gestionar los datos según las necesidades de rendimiento, flexibilidad y acceso. Entre los métodos más habituales destacan el almacenamiento en bloques, el almacenamiento de objetos y el almacenamiento de archivos. Estas tres modalidades comparten un mismo objetivo: permitir que usuarios y aplicaciones accedan de forma remota a la información a través de la red. Sin embargo, cada una funciona de manera distinta y está orientada a casos de uso específicos.

Mientras que el almacenamiento de objetos se utiliza para manejar grandes volúmenes de información no estructurada, el almacenamiento en bloques ofrece un rendimiento óptimo para aplicaciones que requieren acceso rápido y consistente. El almacenamiento de archivos facilita el trabajo colaborativo mediante sistemas de archivos compartidos. Comprender las diferencias entre estos enfoques resulta esencial para elegir la solución más adecuada en entornos de virtualización y nube.

Objetivos del almacenamiento

El almacenamiento en la nube (bien sea a nivel de objetos, de bloques o de archivos) persigue una serie de objetivos más allá de los sistemas de almacenamiento clásicos. Si bien algunos aspectos como la disponibilidad pueden lograrse en sistemas de almacenamiento local, la conexión a red nos permite disfrutar de funcionalidades más avanzadas.

Disponibilidad

Los sistemas de almacenamiento en la nube garantizan que los datos permanezcan accesibles de manera continua. Para lograrlo, emplean técnicas como la replicación, que mantiene varias copias en diferentes ubicaciones y mecanismos de verificación mediante sumas de comprobación que detectan posibles daños. Estas medidas evitan interrupciones en el acceso incluso cuando se produce un fallo en parte de la infraestructura.

Accesibilidad

Estos sistemas permiten acceder a la información a través de API y redes, lo que habilita su uso remoto sin depender de un único equipo. Gracias a ello, múltiples usuarios y aplicaciones pueden consultar, modificar o compartir datos de manera eficiente. Esta capacidad facilita el trabajo colaborativo y la integración con otros servicios o plataformas.

Escalabilidad

El almacenamiento basado en objetos, bloques y archivos puede ampliarse fácilmente para satisfacer un aumento en la demanda de capacidad. Añadiendo nodos o recursos, el sistema se adapta sin comprometer el rendimiento ni generar cuellos de botella. Esta flexibilidad resulta esencial para entornos donde el volumen de datos crece de forma constante.

Seguridad de los datos

Estos sistemas incorporan cifrado tanto en reposo como en tránsito para proteger la confidencialidad y la integridad de la información. También se aplican controles de acceso que limitan quién puede ver o modificar los datos almacenados. Con estas medidas, se reduce significativamente el riesgo de uso indebido o manipulación no autorizada. Este es un objetivo especialmente relevante, pues los datos en circulación están mucho más expuestos.

Administración de metadatos

Los sistemas de almacenamiento de objetos, en bloques y de archivos en la nube suelen ofrecer formas de administrar los metadatos asociados a los datos almacenados. Los metadatos pueden incluir atributos como nombres, tamaños, marcas de tiempo y permisos de archivos u objetos. Los metadatos facilitan la organización, la búsqueda y la recuperación de los datos almacenados.

Diferentes tipos de almacenamiento

Los sistemas de almacenamiento en la nube basados en objetos, bloques o archivos funcionan de manera distinta, ya que cada uno emplea arquitecturas, estructuras y mecanismos de gestión propios. Elegir uno u otro depende de las necesidades de rendimiento, escalabilidad y tipo de aplicación que se vaya a desplegar. Veamos a continuación los fundamentos de cada sistema.

Almacenamiento de objetos

El almacenamiento de objetos organiza la información en unidades discretas llamadas objetos, que combinan tres elementos: los datos reales (por ejemplo, documentos, imágenes, vídeos o registros de datos), sus metadatos asociados y un identificador único. Los metadatos describen características del objeto, como su nombre, tamaño, fecha de creación, etiquetas personalizadas o incluso información de control de versiones, lo que facilita la búsqueda y recuperación eficiente. A diferencia de los sistemas tradicionales de archivos, el almacenamiento de objetos utiliza un espacio de nombres plano, sin carpetas ni jerarquías.

Cada objeto se localiza gracias a su identificador único, que suele generarse mediante algoritmos hash basados en el contenido. Esto garantiza integridad y evita duplicidades: dos objetos con el mismo contenido tendrán el mismo identificador. Este modelo resulta ideal para grandes volúmenes de datos no estructurados, copias de seguridad o aplicaciones que necesitan escalar mucho. Ejemplos conocidos son Amazon S3, Azure Blob Storage o Google Cloud Storage.

Almacenamiento en bloques

El almacenamiento en bloques divide los datos en fragmentos de tamaño fijo (bloques), que se tratan como unidades independientes. Estos bloques pueden medir desde unos pocos kilobytes hasta varios megabytes, y su tamaño se define normalmente durante la configuración inicial. El sistema operativo o hipervisor asigna a cada bloque una dirección única mediante un esquema llamado direccionamiento de bloques lógicos (LBA), que facilita un acceso muy rápido y directo.

Gracias a este método, se puede leer o escribir en bloques concretos sin tener que acceder al archivo completo. Esto ofrece un rendimiento excelente para cargas de trabajo que requieren bajas latencias y acceso intensivo a disco, como bases de datos, sistemas de archivos empresariales o servidores virtualizados. El almacenamiento en bloques se suele presentar como un disco virtual en bruto que el usuario debe formatear y gestionar con su propio sistema de archivos (NTFS, ext4, XFS, etc...). Ejemplos comunes de este tipo son Amazon EBS, Azure Disk Storage o Google Persistent Disk, ampliamente usados como discos de arranque o almacenamiento de alto rendimiento para máquinas virtuales.

Almacenamiento de archivos

El almacenamiento de archivos proporciona un sistema jerárquico tradicional, organizado en carpetas y subcarpetas, accesible de forma remota como si se tratara de un recurso compartido de red. Los archivos incluyen metadatos como nombre, tamaño, marcas de tiempo y permisos de acceso. Este modelo permite crear, modificar, leer o eliminar archivos de manera sencilla y ordenada, resultando muy intuitivo para usuarios y aplicaciones.

El proveedor de la nube mantiene los servidores que almacenan los datos y administra aspectos como la replicación, disponibilidad y seguridad. Además, varios usuarios o máquinas virtuales pueden acceder simultáneamente a los mismos archivos, gestionando permisos de usuario y grupo para un control granular del acceso. Es la solución idónea cuando se necesita compartir archivos entre varios sistemas o mantener aplicaciones que dependen de rutas y jerarquías de directorios. Ejemplos destacados son Amazon EFS, Azure Files o Google Filestore, que actúan como sistemas de archivos distribuidos.

Rendimiento de los diferentes sistemas

Los sistemas de almacenamiento de objetos están diseñados para priorizar la capacidad de almacenamiento por encima de la disponibilidad inmediata de los datos. Gracias a su arquitectura altamente escalable, estos sistemas permiten almacenar grandes volúmenes de datos no estructurados, como imágenes, vídeos, copias de seguridad y registros de aplicaciones. Sin embargo, este enfoque tiene un coste: el acceso a los archivos suele implicar mayor latencia en comparación con otros tipos de almacenamiento, lo que puede afectar la velocidad de recuperación de la información. Además, el rendimiento de los sistemas de almacenamiento de objetos suele ser inferior al del almacenamiento en bloques o incluso al de ciertos servicios de almacenamiento en la nube optimizados para archivos.

El almacenamiento en bloques, por su parte, destaca por ofrecer un alto rendimiento, baja latencia y rápidas velocidades de transferencia de datos. Al operar a nivel de bloques individuales, este tipo de almacenamiento permite un acceso directo y eficiente a los datos, lo que se traduce en un rendimiento de E/S (entrada/salida) muy elevado. Esto lo convierte en la opción preferida para aplicaciones que requieren un acceso rápido y constante a la información, como bases de datos transaccionales, máquinas virtuales o sistemas de análisis de datos en tiempo real.

En cuanto al almacenamiento de archivos en la nube, aunque puede ofrecer un rendimiento aceptable, su principal objetivo no es la velocidad máxima, sino la facilidad de acceso y gestión de los datos de forma intuitiva para los usuarios. Este tipo de almacenamiento se orienta más a la colaboración, el uso compartido de archivos y la creación de repositorios compartidos, facilitando el trabajo en equipo y la accesibilidad desde distintos dispositivos y ubicaciones.