



## 6. INTRODUCCIÓN AL ALMACENAMIENTO Y RECUPERACIÓN DE DATOS MULTIMEDIA

### 6. INTRODUCCIÓN AL ALMACENAMIENTO Y RECUPERACIÓN DE DATOS MULTIMEDIA..1

6.1. Bases de datos multimedia.....	1
6.2. Datos multimedia.....	2
6.2.1. Tipos de datos multimedia.....	2
6.2.2. Características de los datos multimedia.....	3
6.2.3. Compresión de datos.....	4
6.2.4. Consultas en bases de datos multimedia.....	4
6.3. Situación actual, trabajo futuro.....	4
6.4. Object store.....	5
6.4.1. Arquitectura (ORACLE).....	5
6.4.2. Beneficios.....	6
6.4.3. Limitaciones.....	6
6.4.4. Casos de uso.....	6

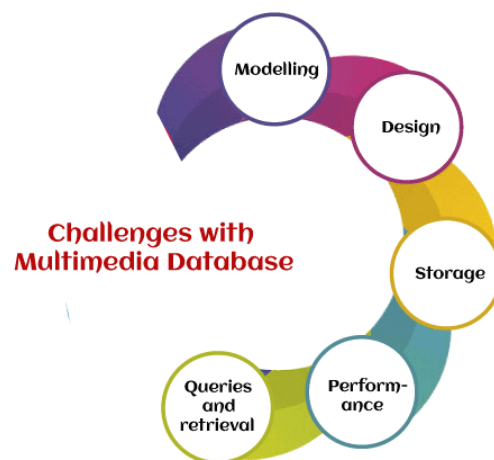
### 6.1. Bases de datos multimedia

Una base de datos multimedia está constituida por datos multimedia:

- texto
- imágenes
- gráficos
- video
- audio

Un DBMS multimedia proporciona soporte para la creación, almacenamiento, consulta y control de acceso para datos multimedia. Los principales requerimientos que este tipo de DBMS deben cumplir son:

- Modelado de datos multimedia
- Almacenamiento eficiente de datos multimedia
- Indexado de objetos multimedia
- Soporte de consultas para datos multimedia.



En general los datos multimedia se representan como señales digitales que se usan en diversas aplicaciones como son:

- Videoconferencias
- Bibliotecas digitales
- Realidad virtual
- Telemedicina, etc.

Los primeros medios empleados para distribuir y almacenar objetos multimedia son discos ópticos y magnéticos. Posteriormente, a través del Internet, uso de LANs como principales fuentes de distribución. El desarrollo y evolución de los navegadores web permiten la posibilidad de visualizar y reproducir datos multimedia. Surgen tecnologías para lograrlo:

- Tecnologías de compresión
- Servidores de video (video file systems)
- Técnicas de transmisión de datos
- Bases de datos Multimedia.

Respecto a las bases de datos multimedia, se requieren nuevos tipos de datos: text, image, graphic objects, secuencias de animación, video, audio.

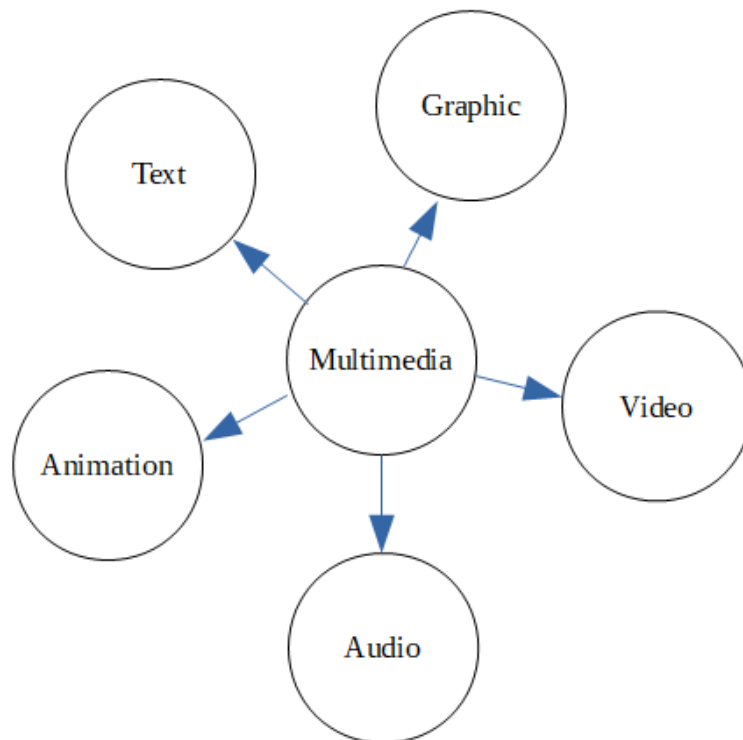
Estos tipos de datos requieren la definición de métodos o funciones especiales para realizar un almacenamiento, acceso, indexado y consultas eficientes.

## 6.2. Datos multimedia

Los datos multimedia pueden ser divididos en 2 tipos:

- Datos continuos, por ejemplo audio y video. Estos tipos de datos cambian respecto al tiempo
- Datos discretos, cuyos cambios son independientes al tiempo: texto con o sin formato, imágenes y gráficos.

### 6.2.1. Tipos de datos multimedia



- Texto
- Graphics: dibujos e ilustraciones codificadas con descripciones de alto nivel como son pic, archivos postscripts
- Imágenes: fotografías o imágenes codificadas con algún formato estándar: JPEG, GIF, etc.
- Animaciones
- Videos
- Audio

### 6.2.2. Características de los datos multimedia

- *Falta de estructura:* Los datos multimedia tienden a ser **no estructurados**. Esto implica que métodos tradicionales para realizar su administración como es el caso de datos relacionales, no aplican para este tipo de datos.
- *Temporalidad:* Videos, audios o secuencias de animación tienen requerimientos de temporalidad los cuales afectan su almacenamiento, manipulación, presentación.
- *Altos volúmenes de almacenamiento:* En especial, audio y video pudieran requerir dispositivos con una gran capacidad de almacenamiento.
- *Logística:* Datos multimedia con formato no estándar puede complicar su procesamiento.

### 6.2.3. Compresión de datos

Se refiere a la reducción del número de bloques requeridos para su almacenamiento, típicamente a través de la eliminación de partes sin uso o redundantes. Algunos ejemplos de métodos de compresión:

- Redundancia en escala
- Redundancia en frecuencia
- Redundancia espacial

Algunas técnicas de codificación que permite realizar compresión:

- *Entropía*: Define el número mínimo de bits necesarios para representar la datos multimedia sin perder información (lossless process). Ejemplos de esta técnica: codificación Huffman, codificación aritmética.
- *Fuente*: Permite distinguir entre datos irrelevantes y datos relevantes. A diferencia de entropía, esta técnica si puede perder ciertos datos (lossy process)
- *Híbrida*: Por ejemplo. JPEG, videos MPEG, emplean una combinación de los 2 métodos anteriores.

La mayoría de los videos se almacenan en una base de datos empleando algún método de compresión. JPEG, MPEG, etc. El principal reto es realizar la manipulación y acceso de estos datos sin tener que descomprimir los datos Este requerimiento sigue en constante mejora y estudio.

### 6.2.4. Consultas en bases de datos multimedia

Representa una de las funcionalidades más importantes en este tipo de bases de datos. Para realizar estas consultas se emplean lenguajes específicos: *multimedia query language*.

Estos lenguajes deben ser capaces de manejar las relaciones de datos multimedia como son, relaciones temporales y espaciales manteniendo la independencia entre la base de datos y la aplicación. Las consultas dirigidas a datos multimedia pueden clasificarse de la siguiente manera:

- Keyword querying
- Semantic querying
- visual query
- video query

Por ejemplo, query video se emplea para encontrar videos basados en su contenido (semántica del video).

## 6.3. Situación actual, trabajo futuro

Las características y requerimientos para dar soporte a aplicaciones que hacen uso de datos multimedia, en especial aplicaciones web, aún están en constante estudio y

desarrollo. Dentro de las principales requerimientos que requieren ser implementados están:

- Capacidad de realizar consultas basadas en el contenido de los datos multimedia
- Administración eficiente de buffers de datos multimedia
- Métodos efectivos para analizar datos espaciales y temporales
- Métodos eficientes para desarrollo de lenguajes de consulta
- Desarrollo de métodos eficientes de indexado de datos multimedia. Hasta ahora no es posible indexar datos multimedia basados en su contenido, y más aún, sin tener que descomprimir el formato de los datos multimedia.

La siguiente sección muestra uno de los principales mecanismos que se emplean para almacenar datos multimedia, basado principalmente en servicios en la nube

## 6.4. Object store

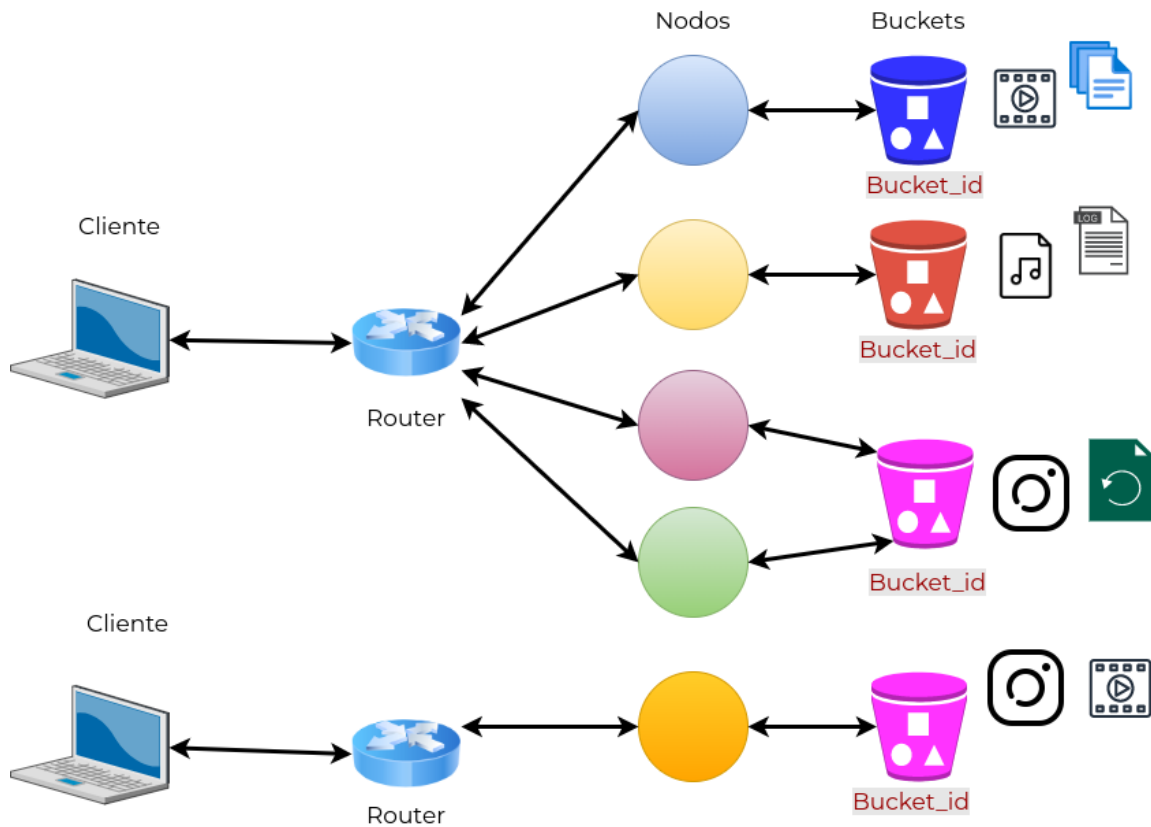
- Método de almacenamiento en cloud utilizado frecuentemente para almacenar datos no estructurados.
- También se le conoce como almacenamiento basado en objetos (object-based storage).
- Ofrece ventajas para backups o archivar datos que no se actualizan con mucha frecuencia.

### 6.4.1. Arquitectura (ORACLE).

La arquitectura está formada por los siguientes elementos:

- **Buckets:** Contenedores en los que se guardan los datos, metadatos y un identificador único para el bucket.
- **Storage nodes:** Administran buckets, los nodos están ligados a un router. El router contiene una lista de lo que contiene cada nodo.
- **Router:** El router contiene una serie de reglas para asignar identificadores a los objetos dentro de los nodos.

No existe una organización jerarquía en los buckets, en todos los buckets se puede leer o escribir peticiones que se envían a través del router. El router identifica al nodo que le corresponde dicha petición



#### 6.4.2. Beneficios.

- Consultas rápidas hacia los datos debido a la estructura router/node/bucket.
- Niveles altos de redundancia en los nodos, si un nodo se cae, los datos siguen estando disponibles en otro nodo.
- Escalabilidad sencilla agregando más nodos.
- Ideal para almacenar datos multimedia o backups.



#### 6.4.3. Limitaciones.

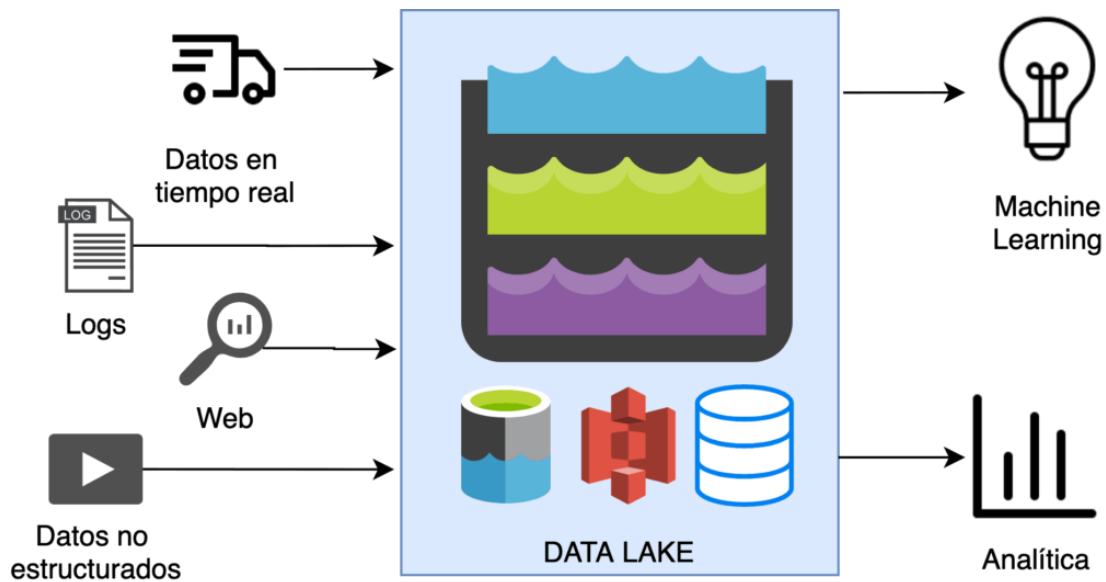
- No es recomendable actualizar los datos una vez que ya están almacenados. Debido a que se requiere un procesamiento de recuperación-actualización-reescritura.
- Recomendado para archivos grandes, pero no para un gran volumen de registros, en ese caso se recomienda utilizar una base de datos.



#### 6.4.4. Casos de uso.

- Archivado: Guardar archivos que no requieran actualizarse.
- Backup: Object storage puede generar varios niveles de redundancia debido a la facilidad de escalamiento. Por lo que es recomendable utilizar object store para almacenar backups.
- Media files: Debido a que son archivos pesados y son generados por IoT, la facilidad de escalamiento hace que se puedan guardar varios de estos archivos multimedia.

- Data lake: También es posible guardar datos estructurados o semiestructurados a parte de los datos no estructurados.



Bibliografía: <https://www.oracle.com/cloud/storage/object-storage/what-is-object-storage/>