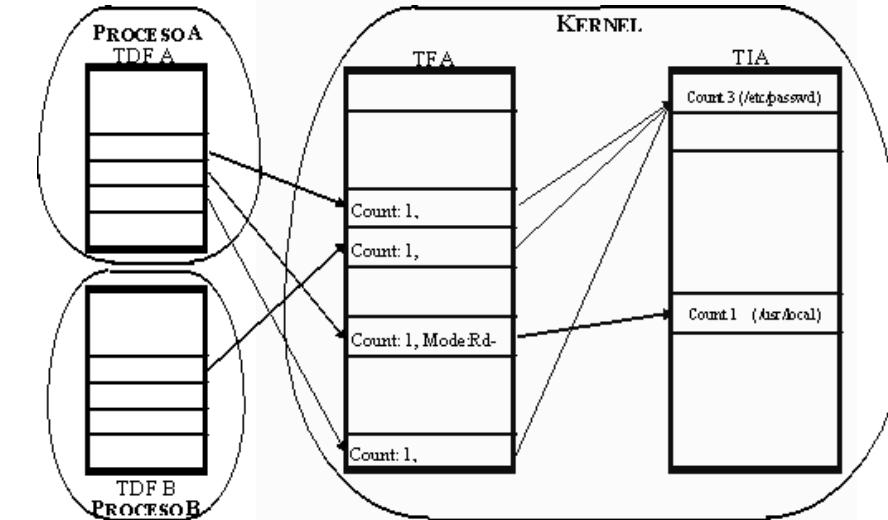


# Manejo e Implementación de Archivos



Guatemala 08 de agosto  
de 2023

Ing. David Luna



Universidad  
Rafael Landívar  
Tradición Jesuita en Guatemala

# Agenda

 Presentación

 Programa del curso

 Introducción

 Conceptos  
Fundamentales

 Cierre

# Presentación



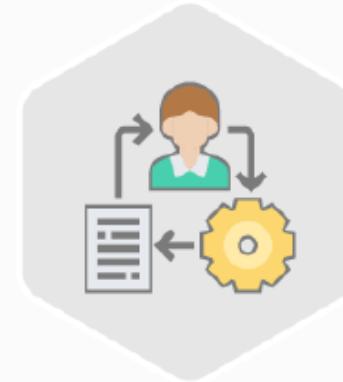
- ¿Nombre?
- ¿Porqué escogí esta carrera?
- ¿Qué me ha gustado hasta el momento?
- Actividad que hago que no tiene nada que ver con Informática y sistemas

# Programa del curso



- Descripción del curso
- Generalidades
- Temas
- Aspectos a Evaluar
- Bibliografía
- Información complementaria

# Descripción



Tener una base sólida del concepto, funciones y organizaciones de archivos en un sistema de información.

- El estudiante debe conocer las definiciones de archivos y la interrelación con los sistemas de bases de datos.
- El estudiante debe identificar las funciones y conceptos básicos de archivos.
- El estudiante debe conocer y aplicar de forma práctica, las principales organizaciones de archivos.
- El estudiante debe diseñar sistemas de archivos eficientes, dando solución a problemas específicos.

# Generalidades

- Créditos académicos: **3 Teóricos, 1 Práctico**
- Catedrático: Ing. David Fernando Luna Hernández
- Correo Electrónico: [dflunah@correo.url.edu.gt](mailto:dflunah@correo.url.edu.gt)
- Horarios:
- Martes y Jueves de 07:00 a 08:30 – T-301

# Temas





## Aspectos a evaluar

ACTIVIDAD	PUNTOS
Tareas, investigaciones y laboratorios.	10
Cortos, exposiciones y ensayos.	10
Proyectos de programación (3)	40
Exámenes Parciales (2)	20
Examen final	20
<b>TOTAL</b>	<b>100</b>

# Proyecto



Se divide en 3 fases para obtener un Sistema de Información al Final del Curso.

Las entregas de los proyectos son dependientes.

- 2do. Proyecto → 1er. Proyecto
- 3er. Proyecto → 2do. Proyecto

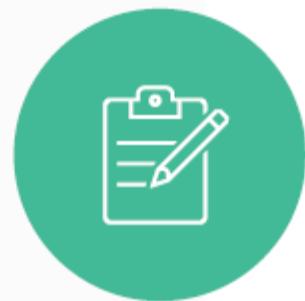
# Información Complementaria



NO SE ACEPTAN TRABAJOS  
FUERA DE TIEMPO



75% DE ASISTENCIA



NOTA 65/100



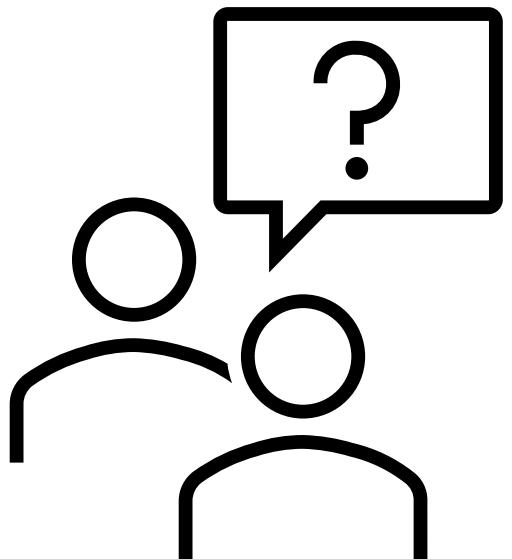
SE REQUIERE  
COMPUTADORA PERSONAL

# Introducción a los Sistemas de Información

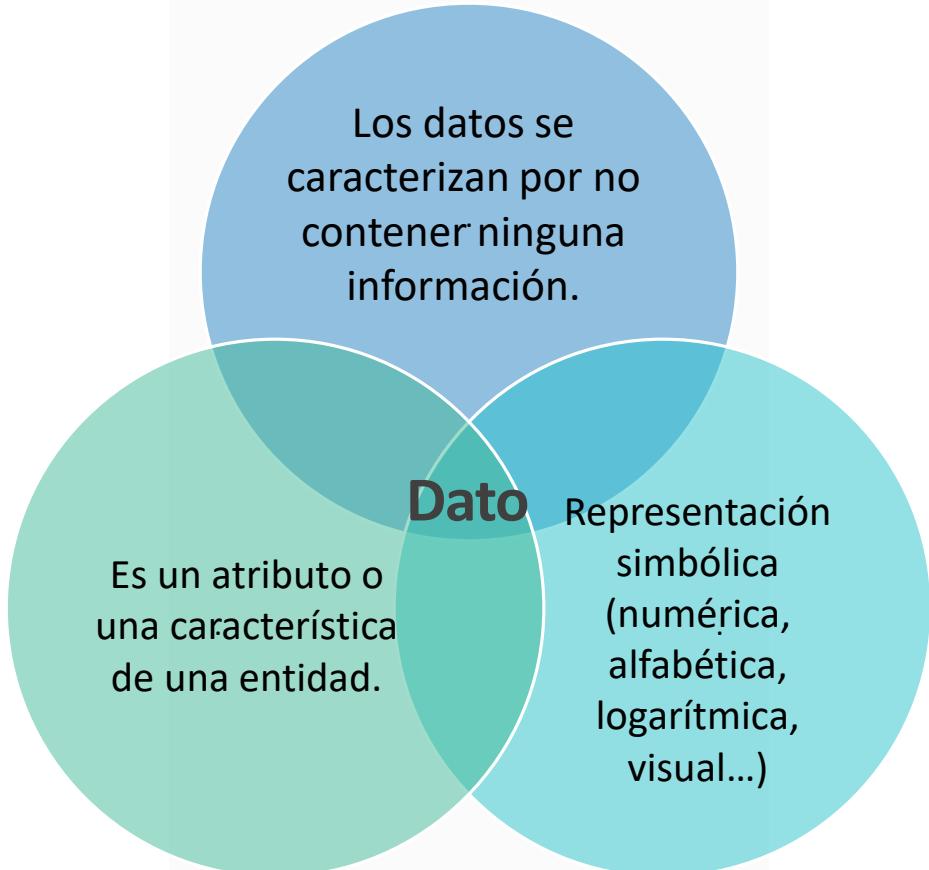


- Dato
- Información
- Cantidad vs Calidad

# Dato



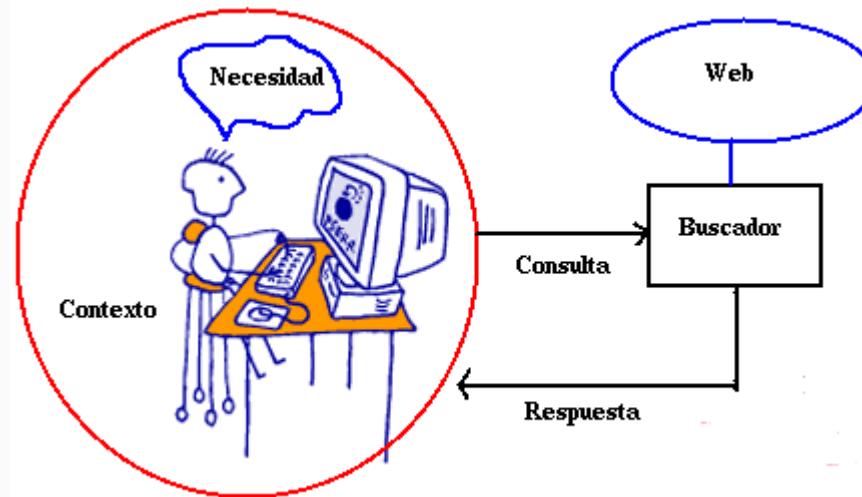
# Dato



# Dato

La importancia de los datos está en su capacidad de asociarse dentro de un contexto para convertirse en información.

Para ser útiles, los datos deben convertirse en información para ofrecer un significado, conocimiento, ideas o conclusiones.



# De datos a información

1

- Se obtienen los datos.
- Clasificados para hacer el proceso fácil y rápido

ENTRADA

# De datos a información

2

- Se convierten los datos en información significativa.
- Operación de Salida

PROCESO

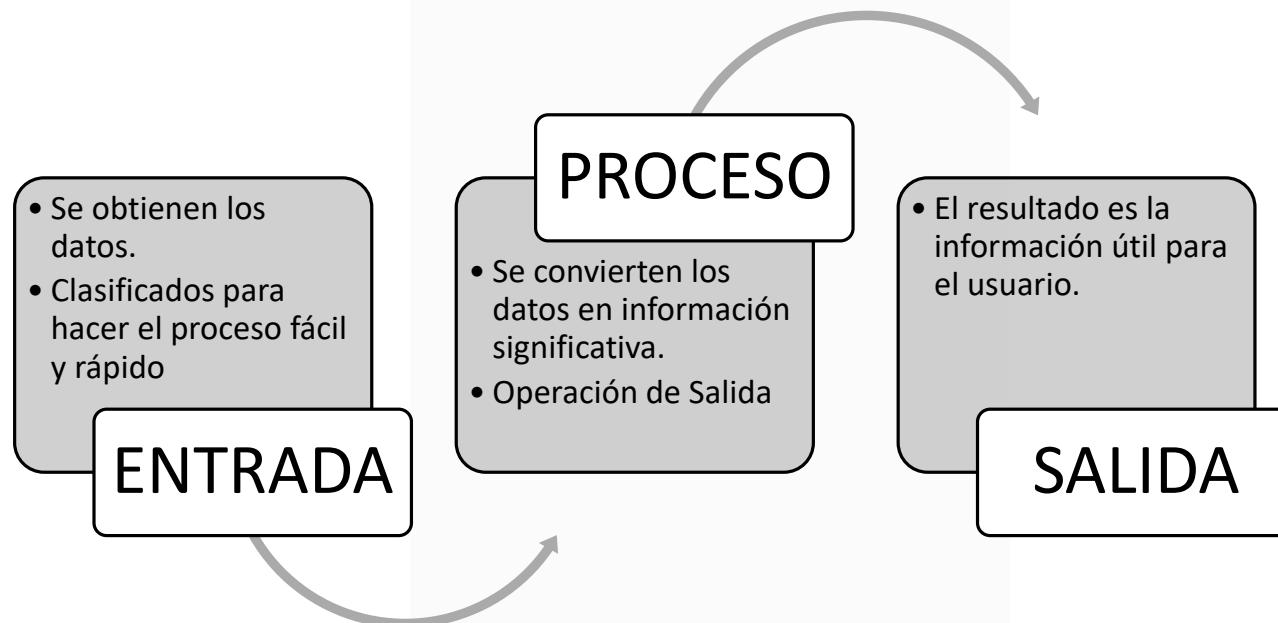
# De datos a información

3

- El resultado es la información útil para el usuario.

SALIDA

# De datos a información

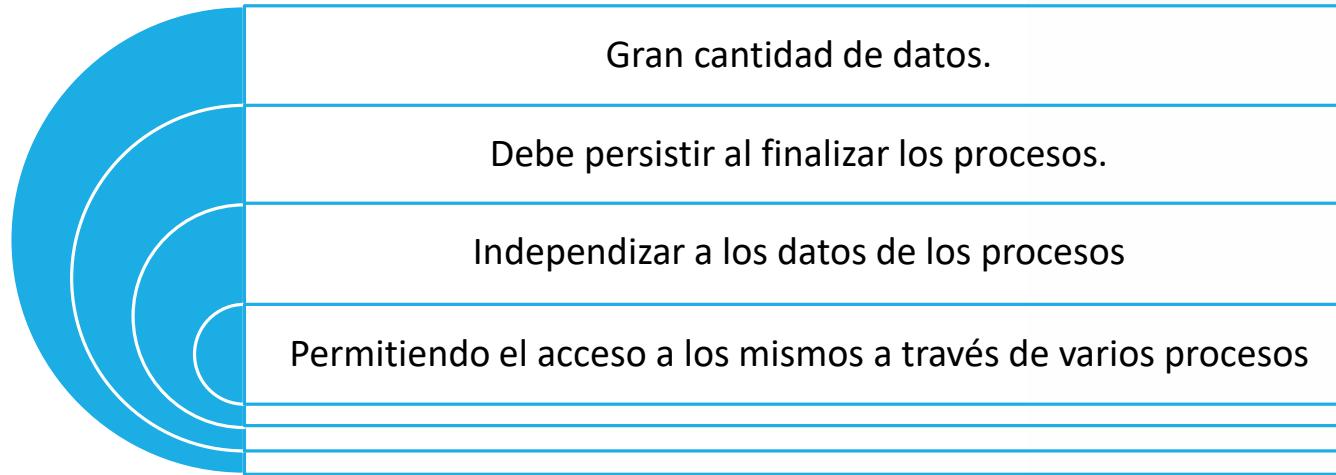


# Información



- No es un conjunto de datos cualquiera.
- Colección de **datos significativos** y pertinentes (relevantes) para la organización que los percibe.
- Para que los datos sean significativos:
  - Símbolos reconocibles.
  - Íntegros.
  - Expresar ideas no ambiguas.
- Los datos son **pertinentes** cuando pueden ser utilizados para responder a preguntas propuestas.

# ¿Por qué guardar los datos?



# Exposiciones

SAN Y NAS

SISTEMAS  
REDUNDANTES DE  
ALMACENAMIENTO

SEGURIDAD



Grupos de 5

BACKUP

CLOUD STORAGE

## Datos importantes de las Exposiciones:

- Datos claros y comprobables.
- Basados en estándares.
- Todos los integrantes del grupo exponen.
- Calificación de acuerdo a la calidad de la investigación.
- Todos los temas se tomarán en cuenta para las evaluaciones.
- Deben presentar casos de éxito.
- Simulaciones.
- Resumen para los integrantes de los otros grupos.
- Entrega de presentación un día antes de la misma.
- Bibliografía
- Apegarse a un estándar.

# Gracias

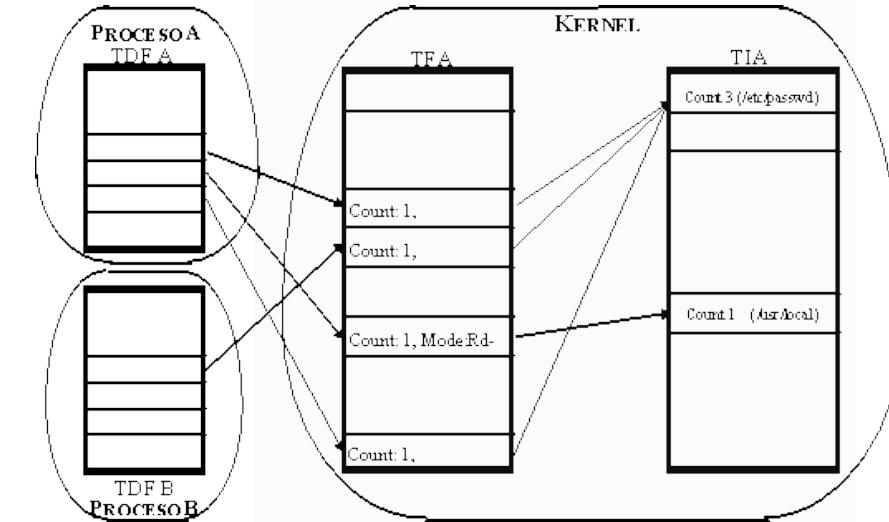
¿ALGUNA PREGUNTA?

# Manejo e Implementación de Archivos



Guatemala 10 de agosto  
de 2023

Ing. David Luna



Universidad  
Rafael Landívar  
Tradición Jesuita en Guatemala

# Conceptos Fundamentales



# Archivo:



Es una colección de bytes que representa datos y que normalmente se guarda en almacenamiento secundario, lo cual implica que tiene una ruta de acceso para recuperarlo.

# Tipos de Rutas:

## Ruta absoluta:

Señalan la ubicación de un archivo o directorio desde el directorio raíz del sistema de archivos.

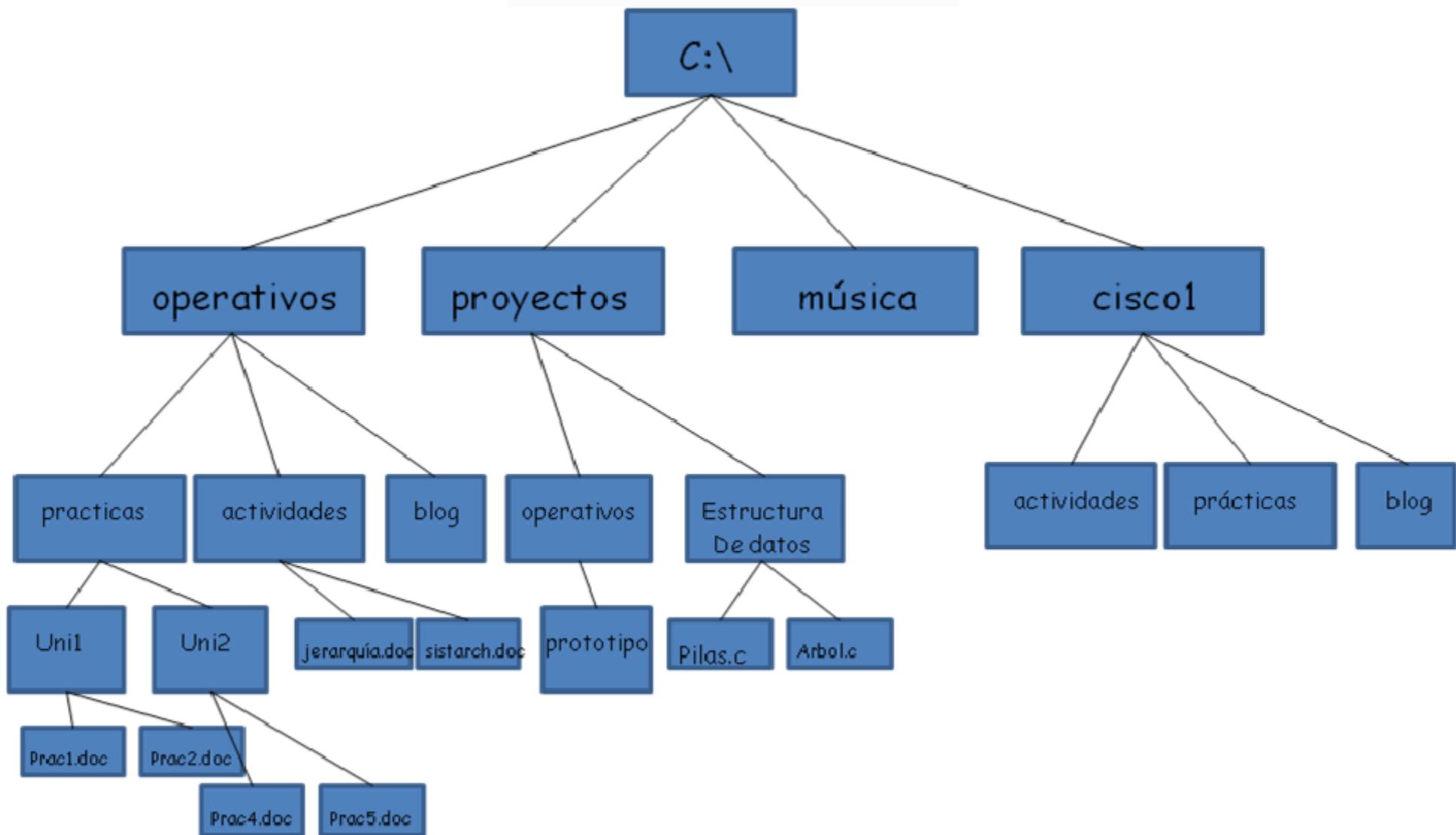
**C:\Users\dfluna\Desktop**



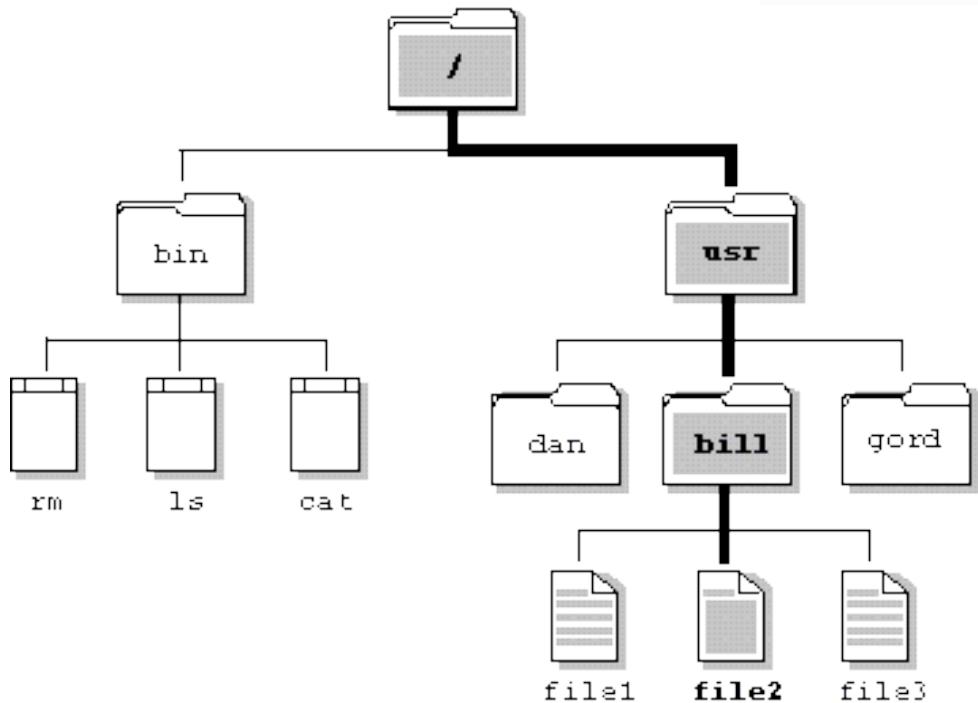
## Ruta relativa:

Señalan la ubicación de un archivo o directorio a partir de la posición actual del sistema operativo en el sistema de archivos.

**C:\Users\dfluna> cd Desktop**



# Archivo físico:



- Archivo que en realidad existe en el almacenamiento secundario.
- Es el archivo tal como lo conoce el sistema operativo y que aparece en su directorio de archivos.

# Archivo lógico:

- El archivo, visto por el programa.
- Puede decirse que un archivo lógico es como una “Vista” la cual se puede manipular de varias formas sin afectar al archivo físico.



**Archivo c:\estudiante**

**Nombre, apellido, carnet, domicilio, etc.**

Del archivo físico podemos obtener varios archivos lógicos:

- Clasificados por carnet
- Clasificados por nombre
- Clasificados por domicilio...

Esto te permite acceder rápidamente a la misma información, pero de diferente forma.

## Algunas de las propiedades de un archivo:

- Nombre
- Tamaño (volumen de datos)
- Ubicación de almacenamiento
- Tipo (programa que lo ejecuta)
- Forma de acceso
- Volatilidad
- Operaciones permitidas
- Autor
- ...y más...

# ¿Qué interesa de un archivo?

- Tipo de archivo
- Qué programa se usa para manipularlo
- Qué plataforma lo puede soportar
- Ruta de acceso
- Propiedades

# Sistema de archivos

Archivos regulares

- Contienen información del usuario.  
.txt, .pas, .c, etc.

Directorios

- Usados para llevar el control de los archivos

Archivos especiales

- Archivos especiales (De caracteres e/s, de bloques para modelar discos)

# Diseñando un sistema de archivos

## ¿Qué debemos considerar?

- Cómo almacenar los archivos (nombres, organización)
- ¿Para qué se utilizará?
- ¿Debe utilizarse luego de su procesamiento?
- ¿Qué volumen de datos contendrá?
- ¿Qué tan veloz debe ser su acceso?
- ¿Qué tan protegidos/accesibles deben estar los datos?
- ¿Con qué frecuencia cambia su contenido?
- ¿Cuántos recursos puede invertir la organización?
- Qué sucederá cuando el archivo se edite
- Qué sucede si el archivo crece bastante
- ¿Y si lo eliminan?

# Diseñando un sistema de archivos

- ✓ Que se toma en cuenta del Usuario
  - Operaciones que efectúa en los archivos
  - Aspectos de la interfaz
  
- ✓ Aspectos internos
  - Manejo de espacio en disco
  - Almacenamiento de múltiples archivos
  - Eficiencia y confiabilidad

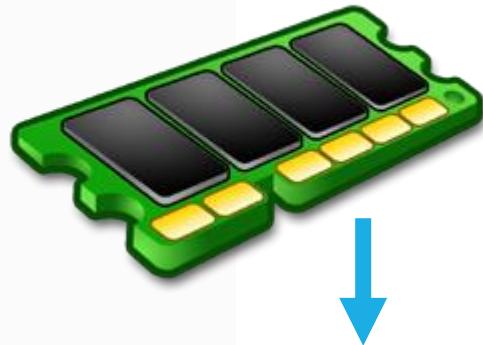
# Almacenamiento



- Se refiere a los medios y los métodos usados para guardar y mantener la información disponible para usarla más adelante. Algunas serán necesarias de inmediato, mientras que otras no serán requeridas durante largo tiempo.

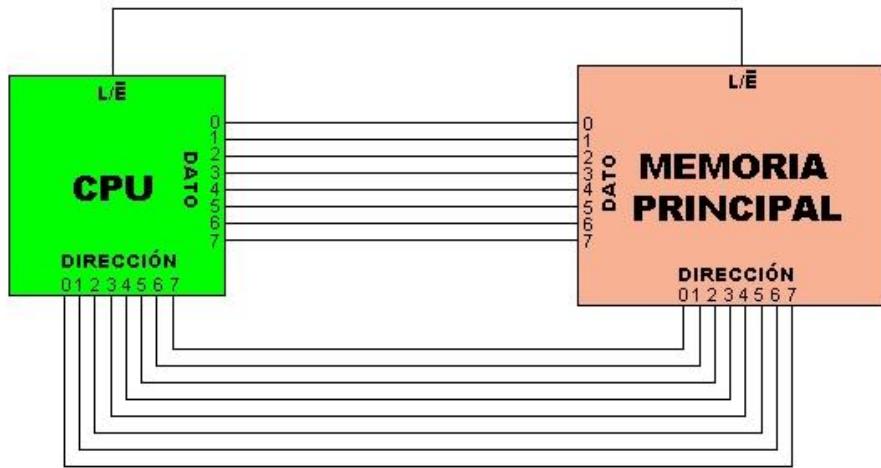
# Almacenamiento primario

- Este tipo de almacenamiento se da cuando el computador guarda información de manera temporal en memoria RAM, hasta que ésta se libera.



La memoria principal es un conjunto o colección de pequeñas celdas que almacenan información (datos e instrucciones) que se identifican únicamente por una dirección.

# Accediendo a una dirección específica

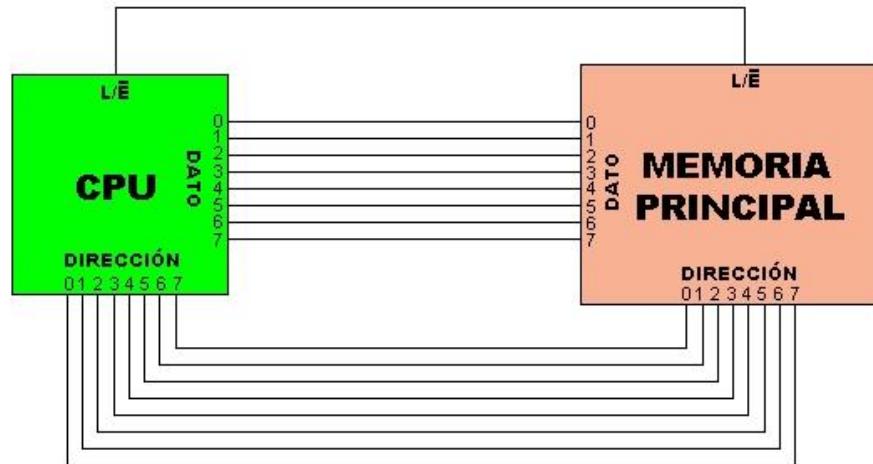


CPU manda señales en el bus de dirección



Permiten especificar a la CPU 4,296,967,296 ( $2^{32}$ ) direcciones diferentes de la memoria.

# El proceso para almacenar:



1



2



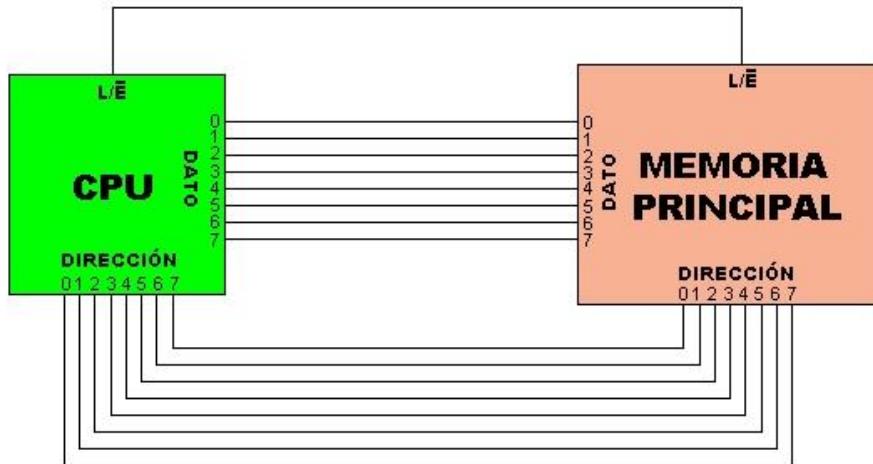
3

El procesador envía la dirección para los datos.

El controlador de la memoria encuentra la ubicación adecuada.

Por último, el procesador envía los datos a escribir.

# El proceso para leer:



1



2



3

El procesador envía la dirección de los datos solicitados.

El controlador de la memoria encuentra los bits de información contenidos en dicha dirección.

Posteriormente los envía al bus de datos del procesador

# Almacenamiento secundario



- Este tipo de almacenamiento se da cuando la información que está en almacenamiento primario (memoria RAM), pasa a un dispositivo de almacenamiento permanente.
- Es un tipo de almacenamiento masivo y permanente (no volátil), a diferencia de la memoria RAM que es volátil; pero posee mayor capacidad de memoria que la memoria principal, aunque es más lenta que ésta.

# Tecnologías de almacenamiento secundario

Para almacenar información se usan principalmente tres 'tecnologías':



En esta tecnología empleamos un laser para leer y escribir datos.

- Óptica (ej. CD, DVD, etc.)

# Tecnologías de almacenamiento secundario



Tienen muchísimas partículas magnéticas distribuidas en algún tipo de superficie. Empleando campos magnéticos, podemos polarizar estas partículas, orientándolas en una u otra dirección.

- Magnética (ej. disco duro, cintas magnéticas)

Para almacenar información se usan principalmente tres 'tecnologías':



# Tecnologías de almacenamiento secundario



- Memoria Flash (Tarjetas de Memorias Flash y Unidades de Estado sólido SSD)

Para almacenar información se usan principalmente tres 'tecnologías':

- Utiliza circuitos electrónicos para almacenar la información, los cuales no necesitan moverse para efectuar tal función.

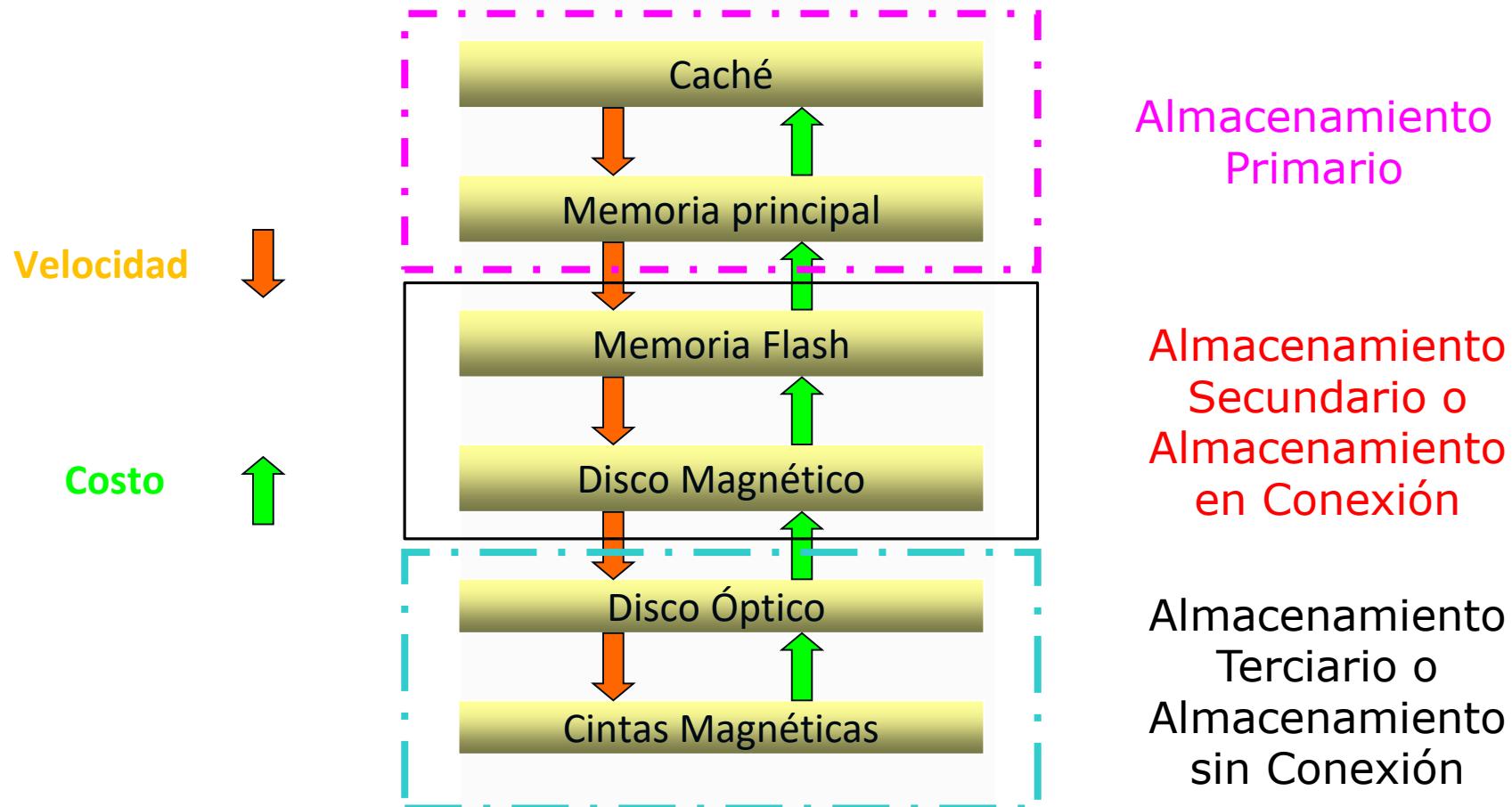
# Almacenamiento secundario

## Características



- Capacidad de almacenamiento grande.
- No se pierde información a falta de alimentación.
- Altas velocidades de transferencia de información.
- Mismo formato de almacenamiento que en memoria principal.
- Siempre es independiente del CPU y de la memoria principal.

# Jerarquía de los Medios de Almacenamiento



# Exposiciones

SAN Y NAS

SISTEMAS  
REDUNDANTES DE  
ALMACENAMIENTO

SEGURIDAD



Grupos de 5

BACKUP

CLOUD STORAGE

## Datos importantes de las Exposiciones:

- Datos claros y comprobables.
- Basados en estándares.
- Todos los integrantes del grupo exponen.
- Calificación de acuerdo a la calidad de la investigación.
- Todos los temas se tomarán en cuenta para las evaluaciones.
- Deben presentar casos de éxito.
- Simulaciones.
- Resumen para los integrantes de los otros grupos.
- Entrega de presentación un día antes de la misma.
- Bibliografía
- Apegarse a un estándar.

# Ensayo

Género literario que se caracteriza por permitir desarrollar un tema determinado de una manera libre y personal. Comúnmente, las personas escriben ensayos para manifestar alguna opinión o idea, y sin tener que preocuparse de ceñirse a una estructura rígida de redacción o documentarlo exhaustivamente.

# Tarea



- **Tema: Información**
- ¿Qué se puede incluir?
  - Temas actuales orientado a sus hobbies.
- Subirlo al portal (Martes 15 Agosto 20:00 horas).
- Mínimo 200 palabras.
- (Por cada falta de ortografía se bajarán 5 puntos)

# Gracias

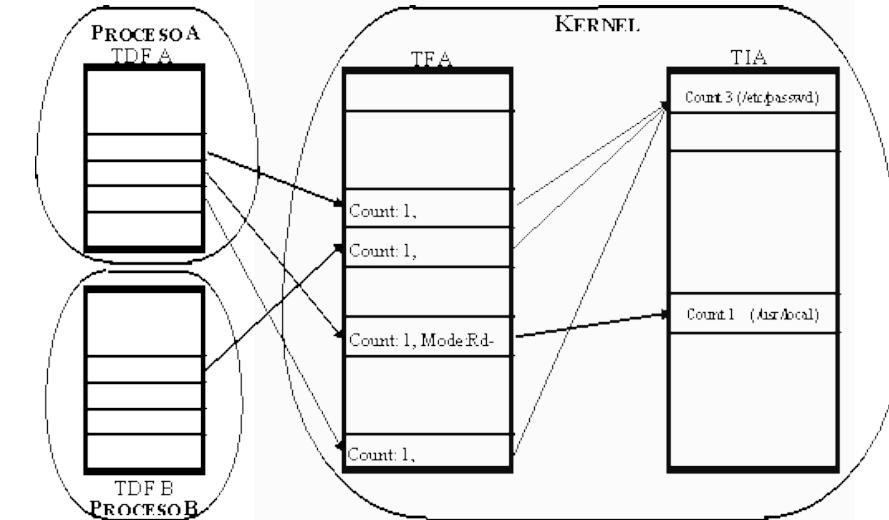
¿ALGUNA PREGUNTA?

# Manejo e Implementación de Archivos



Guatemala 17 de agosto  
de 2023

Ing. David Luna



Universidad  
Rafael Landívar  
Tradición Jesuita en Guatemala

# Agenda



**Memoria Caché**



**Almacenamiento  
según el tipo de  
Acceso**

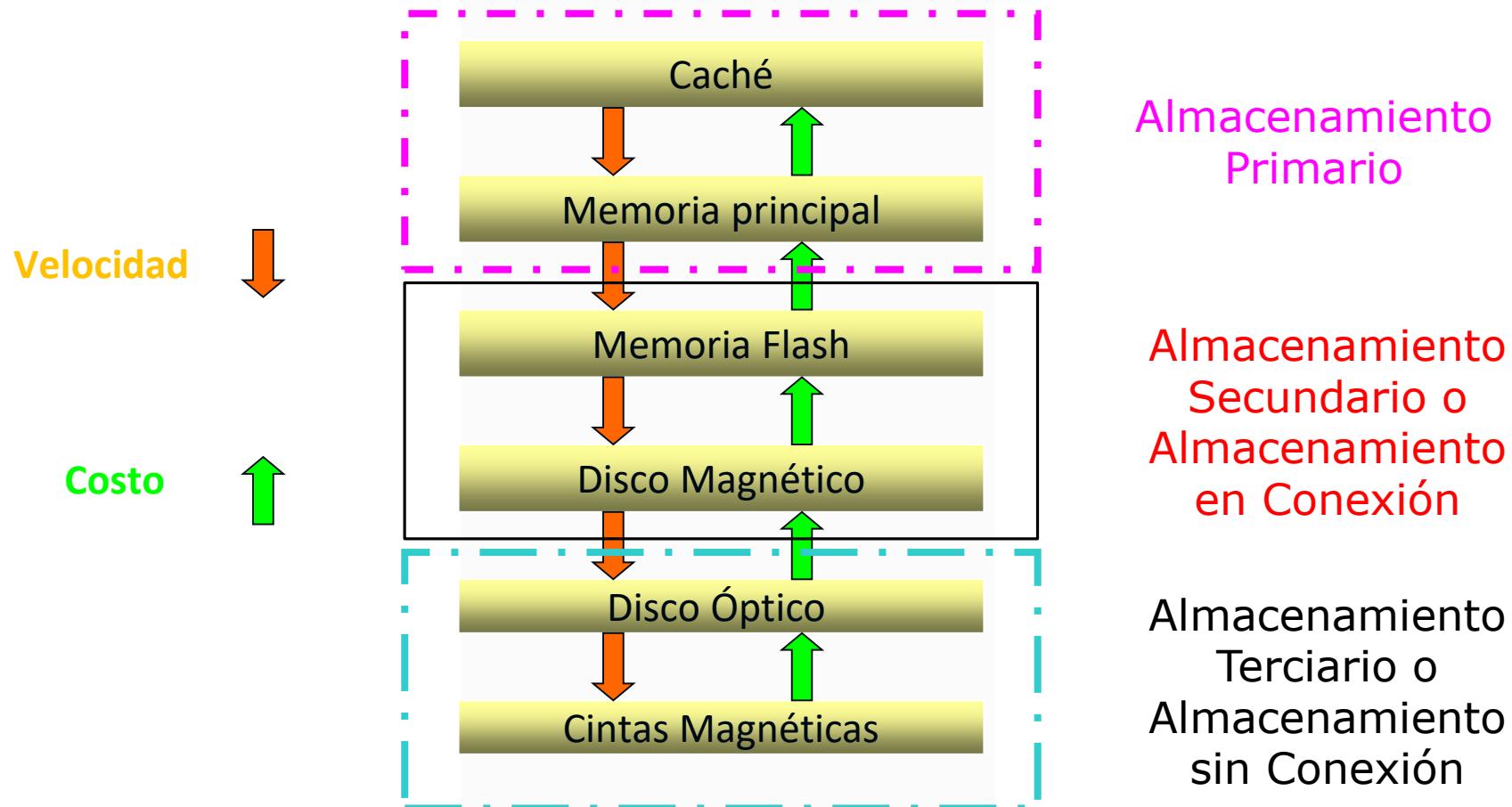


**Cierre**

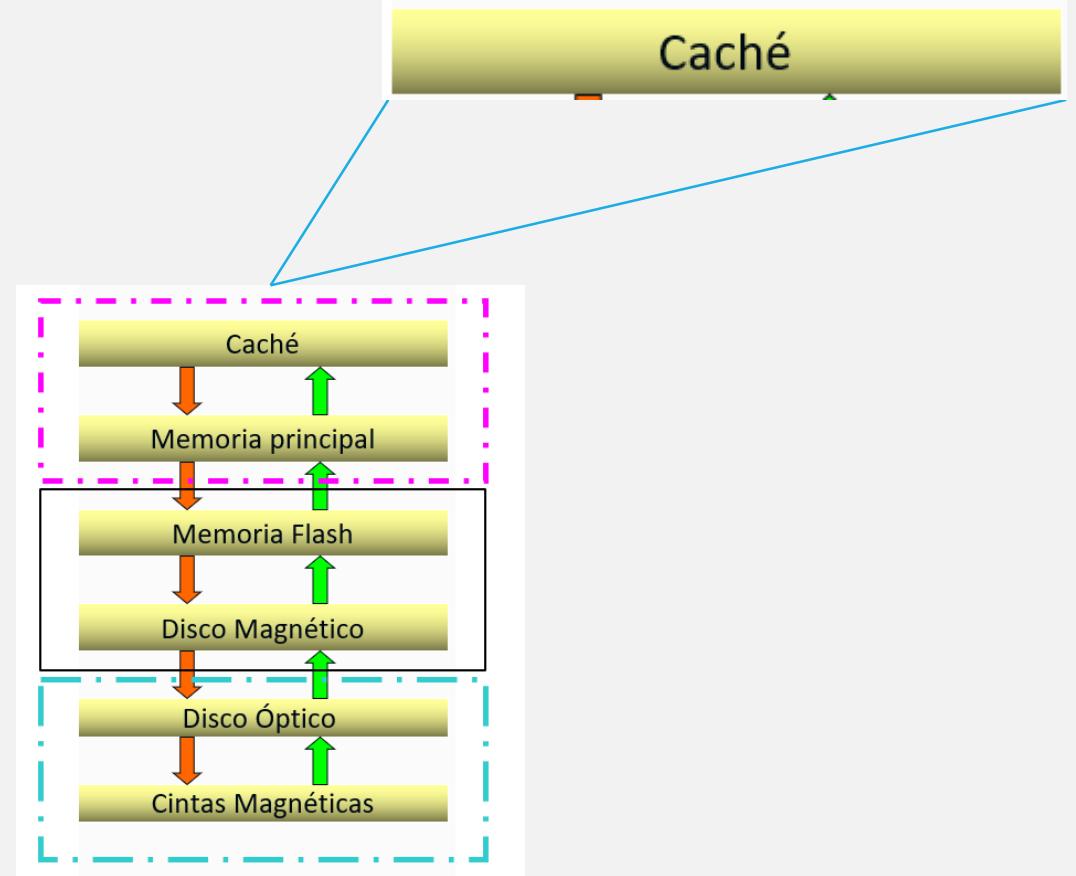
# Memoria Caché



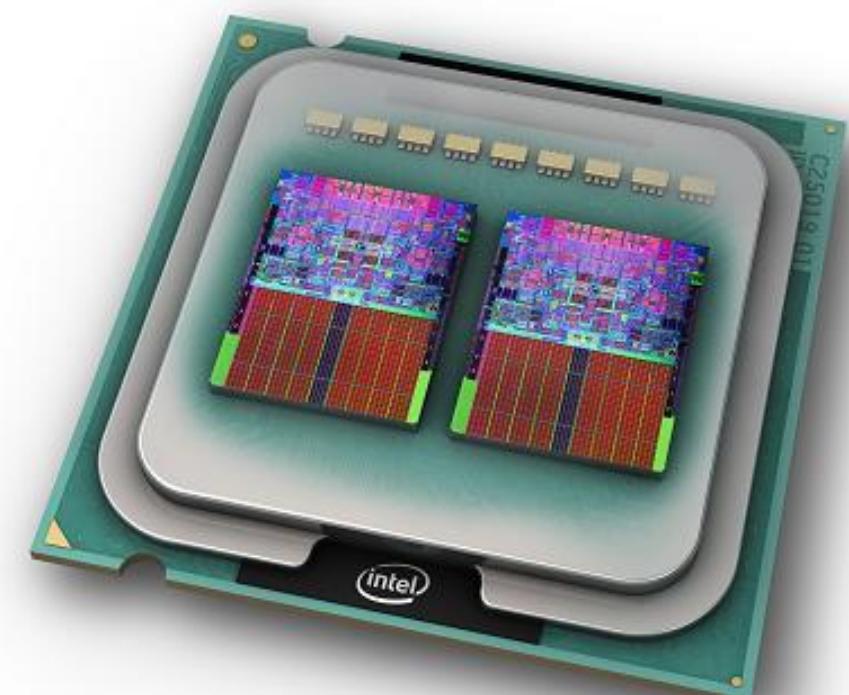
# Jerarquía de los Medios de Almacenamiento



# Memoria Caché



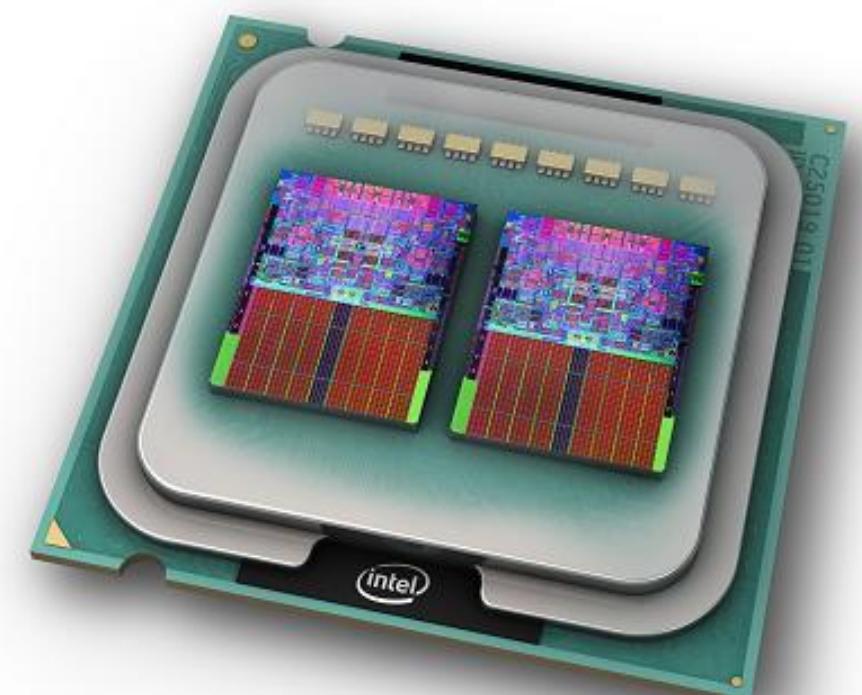
# Caché



## Objetivos:

- Almacenar una serie de instrucciones y datos a los que el procesador accede continuamente, con la finalidad de que estos accesos sean instantáneos.
- Rendimiento.

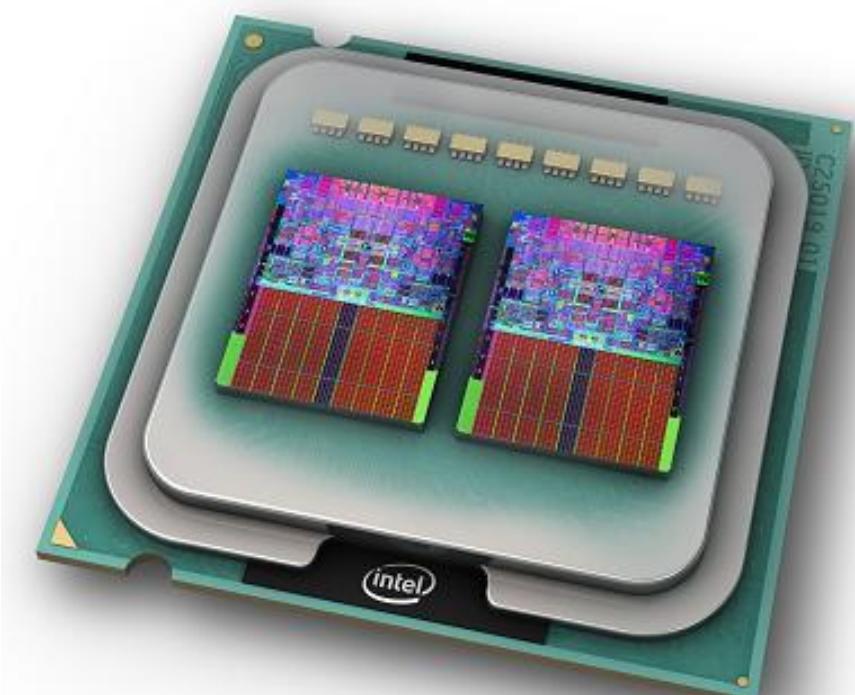
# Caché



## Características:

- Almacenamiento de datos e instrucciones para rápido acceso, que el equipo utiliza con mayor frecuencia para realizar sus tareas cotidianas.
- Tipo de memoria volátil, pero de gran velocidad.
- Tareas repetitivas.
- Integrada en el procesador.

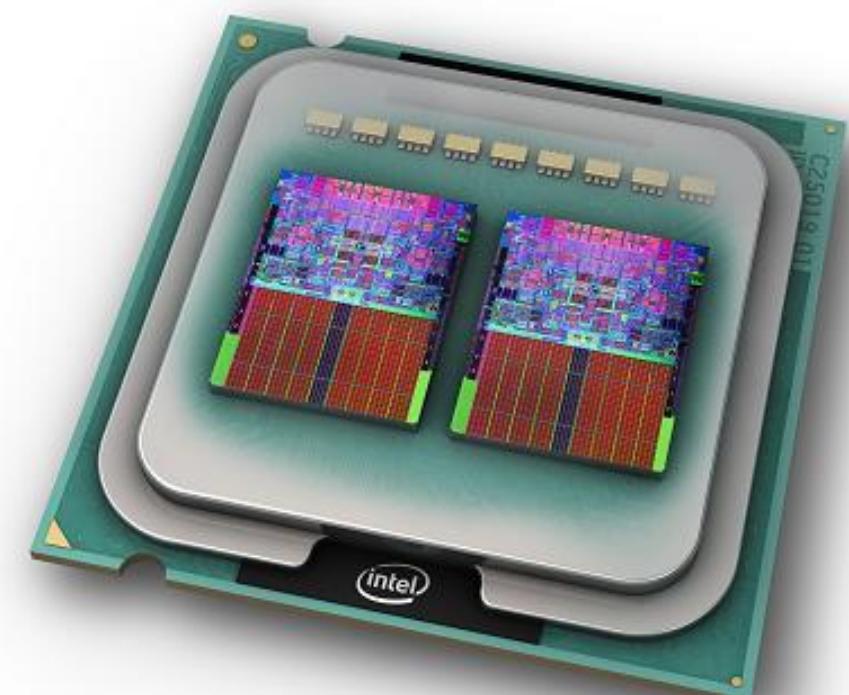
# Caché (Tipos)



## Caché de 1er nivel (L1):

- La más rápida pero la más pequeña de todas
- Integrada en el núcleo del procesador.
- La cantidad varia de un procesador a otro (Regularmente de 256KB, aunque en algunos procesadores llega a 1 MB)
- Divida en dos partes, una para direcciones (operaciones que debe realizar el procesador) y otra para datos (datos que se deben procesar).

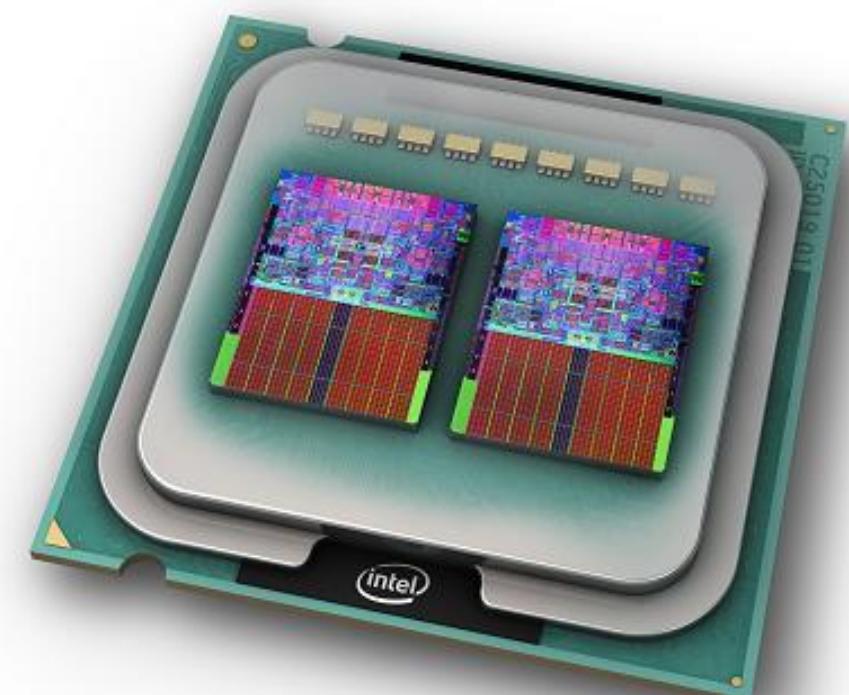
# Caché (Tipos)



## Caché de 2do. nivel (L2)

- Integrada en el procesador, aunque no directamente en el núcleo.
- Tiene las mismas ventajas que la L1, pero es mas lenta.
- La cantidad es mayor (256 KB – 8 MB)
- Utilizada para programas mas que para el sistema.
- Depende de la marca, podemos encontrarla compartida por todos los núcleos o junto en cada núcleo del procesador

# Caché (Tipos)



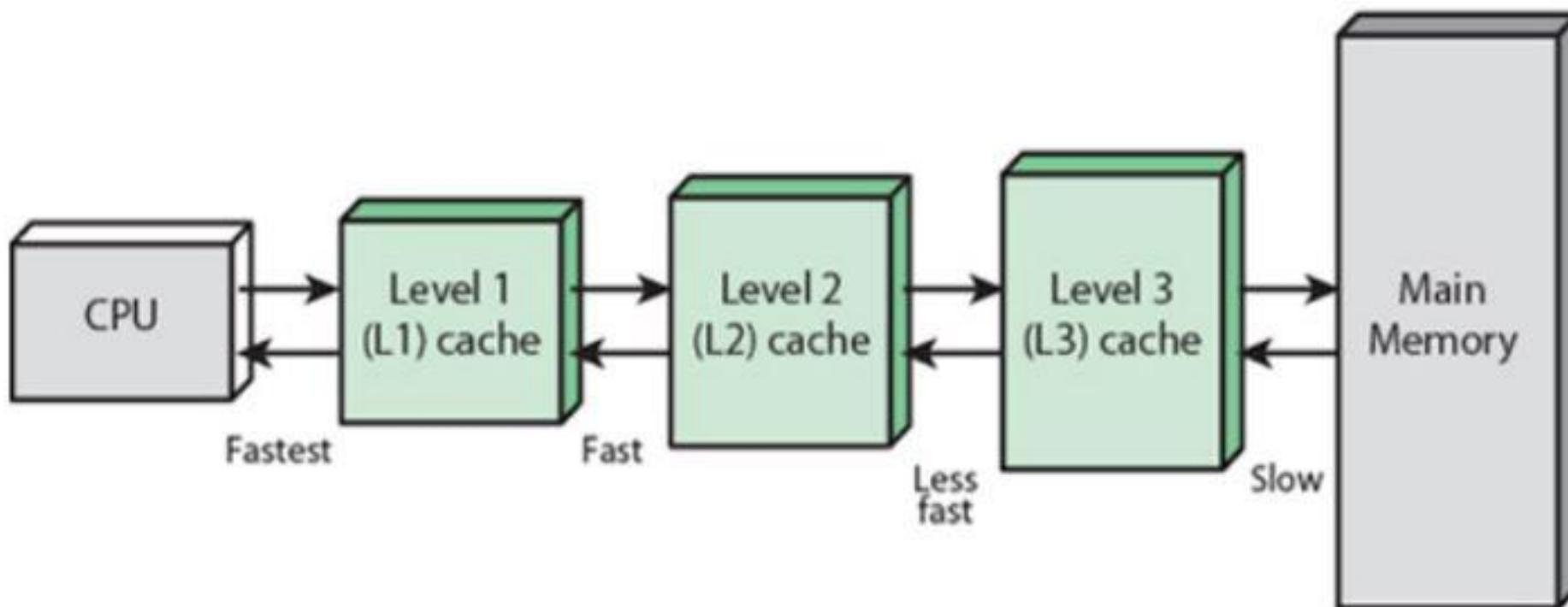
## Caché de 3er. nivel (L3)

- Más lenta, pero más grande que la anterior.
- Capacidad 4 MB – 50 MB
- Incorporada a la placa base, por lo tanto depende de la comunicación entre el procesador y la placa base.
- La comparten todos los núcleos del procesador.

# Caché

## Objetivos:

- Almacenar una serie de instrucciones y datos
- Rendimiento.



# Caché

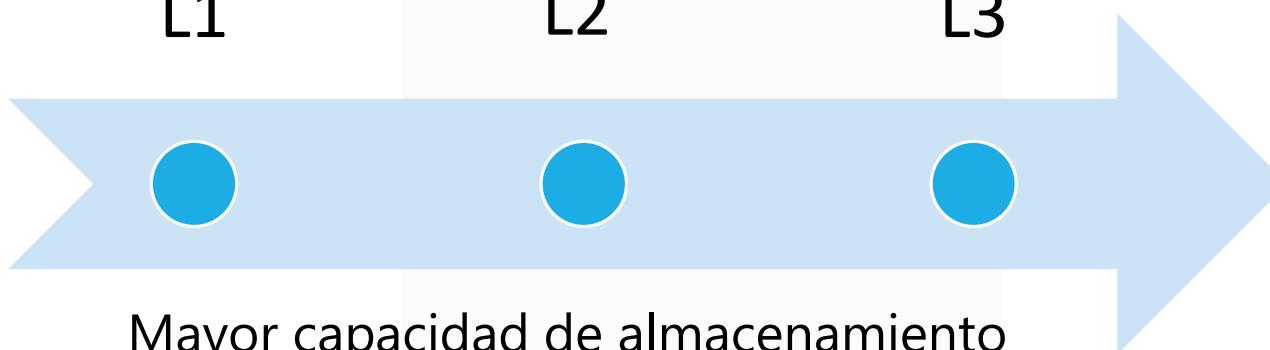
## Objetivos:

- Almacenar una serie de instrucciones y datos
- Rendimiento.

L1

L2

L3

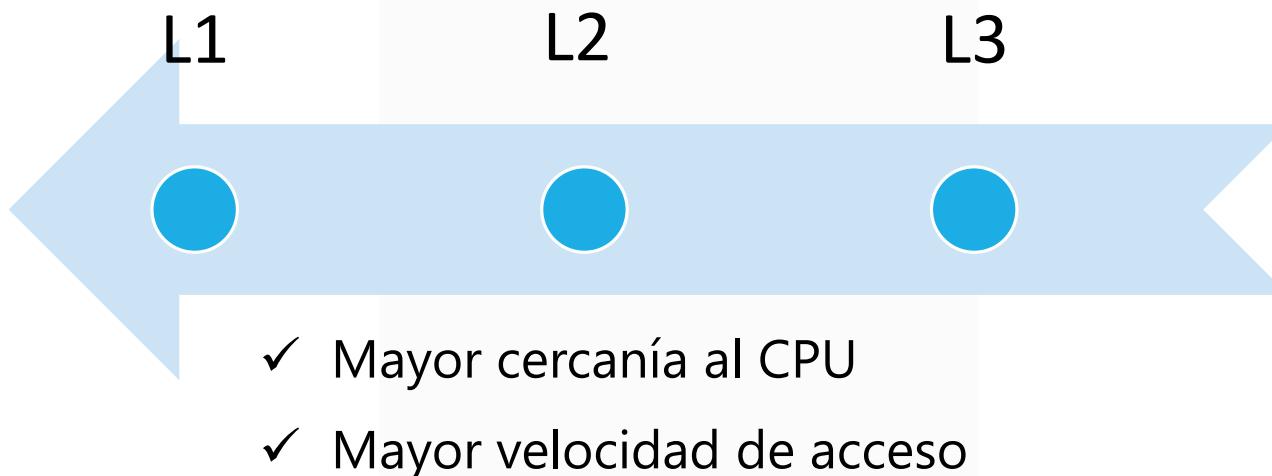


Mayor capacidad de almacenamiento

# Caché

## Objetivos:

- Almacenar una serie de instrucciones y datos
- Rendimiento.



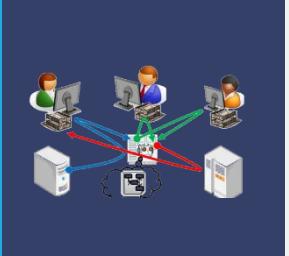
# Caché

Objetivos:

- Almacenar una serie de instrucciones y datos
- Rendimiento.

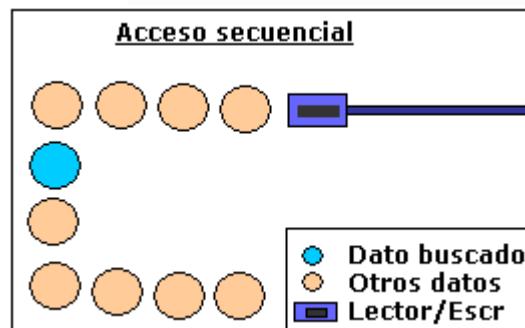
¿Es siempre conveniente  
tener mucho caché?

# Almacenamiento según el tipo de Acceso



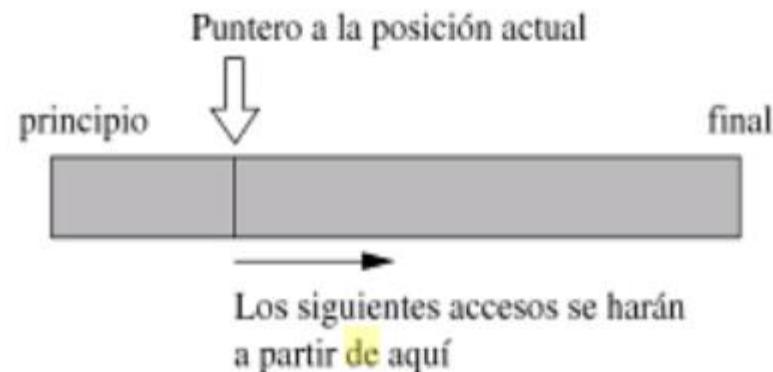
# Acceso secuencial

- No se puede leer o escribir un dato en particular hasta que todos los datos que lo preceden hayan sido leídos o escritos en orden.
- Hemos de recorrer desde el principio todas las posiciones hasta llegar a la deseada.



# Acceso secuencial

- Es la forma más simple y quizás la más intuitiva de manipular archivos.
- Cuando se abre el archivo, el sistema operativo apunta el primer bloque de información de archivo.
- Lee a partir de la posición inicial y así sucesivamente hasta que se llega al final del archivo.



# Métodos que utilizan acceso secuencial

## Cintas Magnéticas

Las cintas se presentan en una diversidad de formas, tamaños y velocidades.

Pistas sobre una banda plástica con un material magnetizado (Oxido de hierro).

Video, audio, datos.

## Acceso secuencial

1951

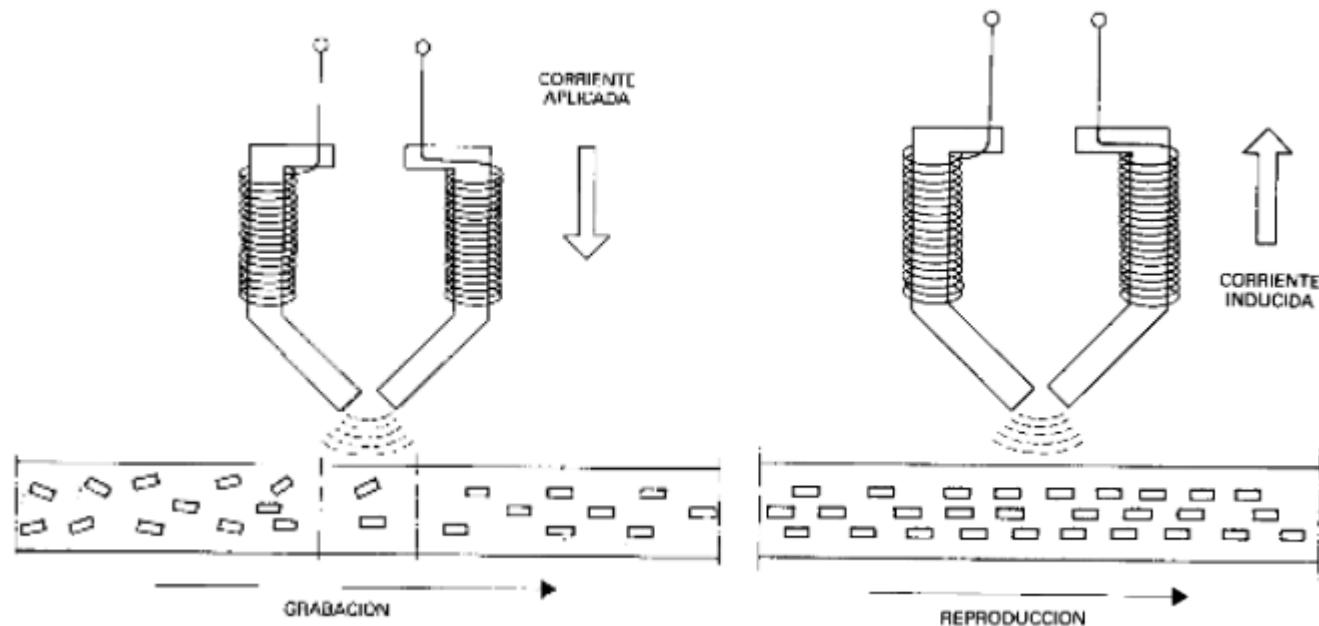
Grabación:

Corriente Aplicada > Cabezal de grabación >  
Magnetización de la cinta.

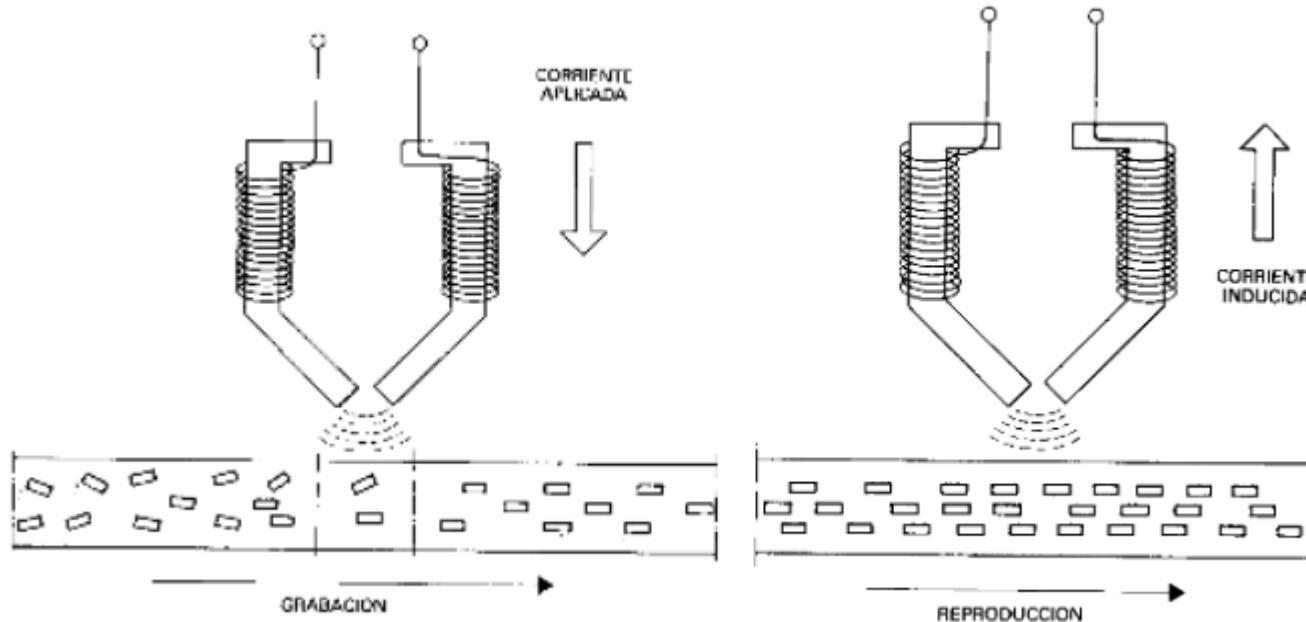
Reproducción:

Corriente Inducida> Cabezal de grabación

## Acceso secuencial



## Acceso secuencial



Parámetros Importantes:

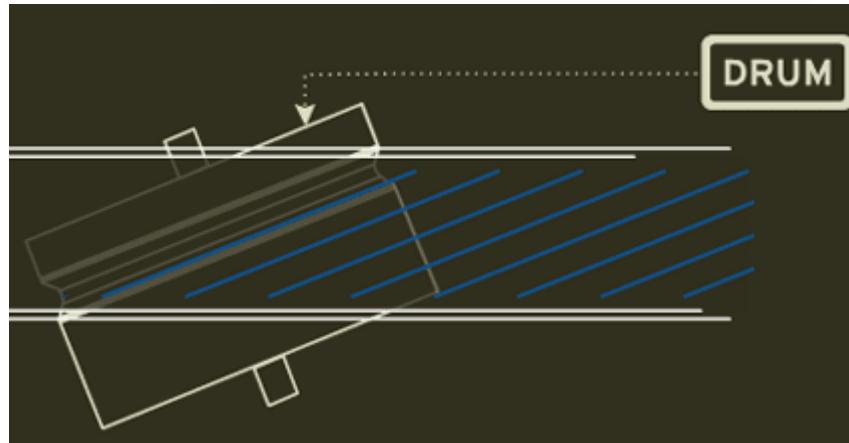
- Magnitud del entrehierro, nos limita la máxima frecuencia a grabar.
- Ancho de la cinta, cuanto mayor sea su anchura la cantidad de información a grabar se reduce.

## Acceso secuencial

Clasificación de Tecnologías de cintas magnéticas:

- Anchura de la cinta
  - Cinta de Alta Capacidad (1/2 Pul.)
- Método de grabación
  - Lineal
  - Helical

- Método de grabación
  - Helical

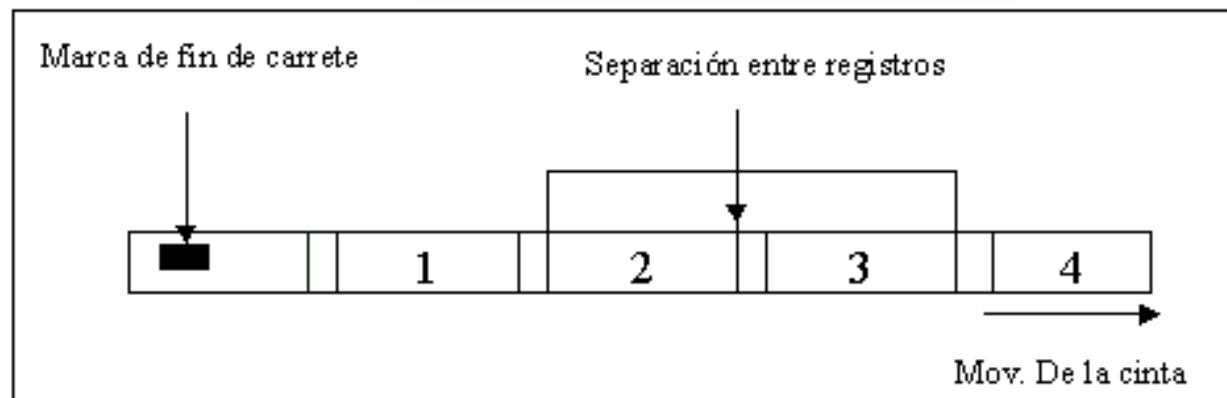


## Acceso secuencial

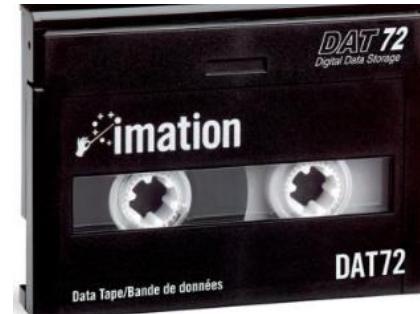
### Características Técnicas

- Distribución de los bloques.
- Separación entre registros.
  
- Tiempo de acceso.
- Latencia
  
- Compresión de los datos.
  - Algoritmos propios.
  - LZ (Lempel Ziv) la mayoría

## Acceso secuencial



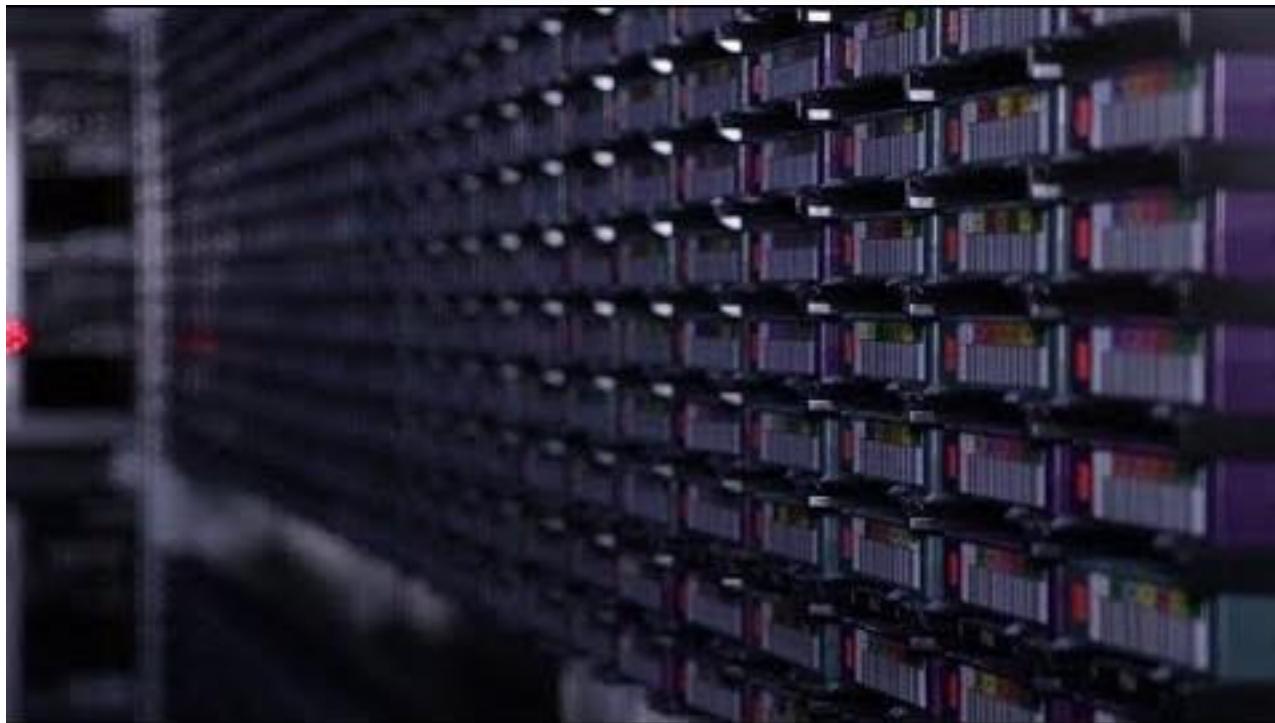
## Cintas Magnéticas





# Las cintas magnéticas en la actualidad

Relación entre precio y capacidad de **almacenamiento.**



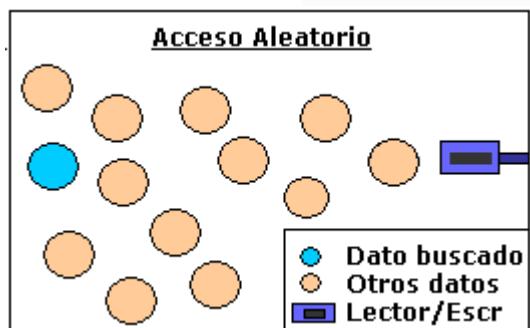
- A pesar que estos soportes ya no se utilizan de forma masiva, **las grandes empresas siguen prefiriéndolos por sobre los sistemas de discos duros** convencionales. La razón es muy sencilla: en una pulgada cuadrada se pueden almacenar 45GB de datos y esto permite crear **cartuchos de hasta 50TB**, según ha anunciado el Instituto Tecnológico de Tokio en conjunto con **Hitachi Maxell, Ltd.**
- IBM – 330 TB
  - 201 Gb por pulgada cuadrada

- Los 10 bancos más grandes del mundo, las tres compañías de telecomunicaciones más grandes del mundo y ocho de las 10 mayores farmacéuticas utilizan el almacenamiento en cinta.

<https://www.muycomputerpro.com/?p=63561?p=63561>

# Acceso directo

- Permiten el acceso justo en la posición en la que están guardados los datos.
  - ❖ Discos, memoria flash, etc.



# Almacenamiento en Disco - Magnético

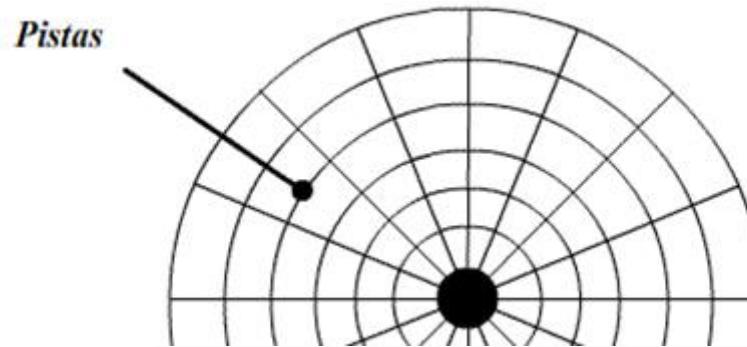
- Las unidades de disco consisten en un conjunto de cabezas de lectura y escritura, interpuestas entre uno o más platos (recubiertos de una fina película magnética).



# Almacenamiento en Disco - Magnético

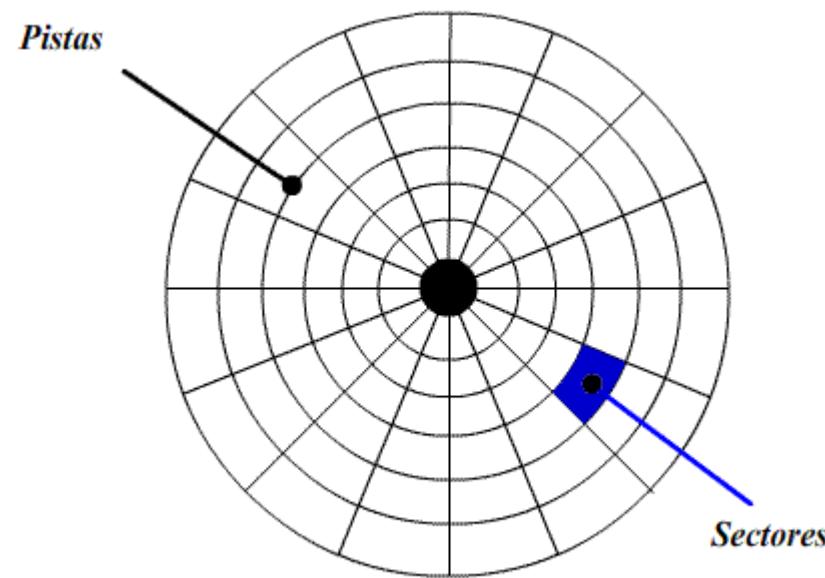
## Componentes:

**Pista:** es el conjunto de bytes en la superficie de un disco al cual puede accederse sin mover el brazo de acceso.



# Almacenamiento en Disco - Magnético

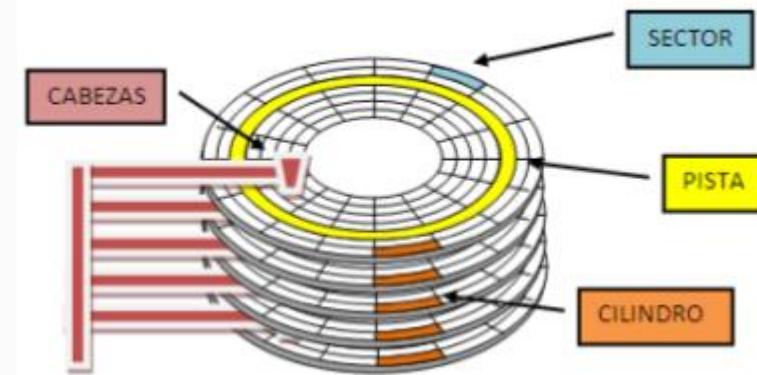
**Sector:** es la porción más pequeña, a la cual se puede hacer referencia en un disco. Cada pista está dividida en varios sectores.



# Almacenamiento en Disco - Magnético

**Cilindros:** son pistas que están directamente unas sobre otras. La importancia del cilindro es que se puede tener el acceso a toda la información almacenada en uno solo sin mover el brazo que sostiene las cabezas de lectura y escritura.

# Almacenamiento en Disco - Magnético



# Almacenamiento en Disco – Memoria Flash



Utiliza circuitos electrónicos para almacenar la información, los cuales no necesitan moverse para efectuar tal función.

Los discos SSD permiten hasta un 56% más de rapidez de respuesta del equipo en comparación a los discos duros tradicionales, son extremadamente resistente, puede soportar golpes y choques sin perder datos.

Utilizan la misma interfaz que los discos duros tradicionales

# Diversidad de Discos Duros

- Discos duros tecnología magnética
- SSD
- Sistemas de arreglos de discos, RAID  
(Redundant Arrays of Inexpensive Disks)

# Exposiciones



# Gracias

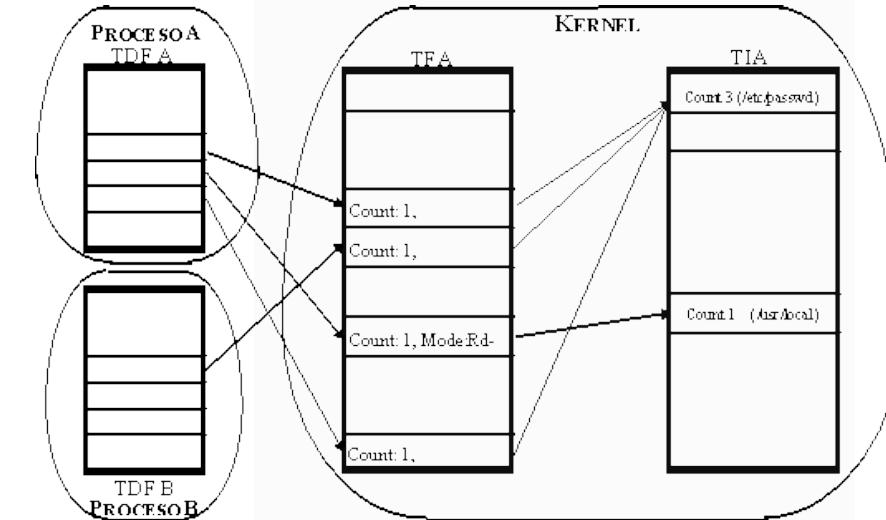
¿ALGUNA PREGUNTA?

# Manejo e Implementación de Archivos



Guatemala 22 de agosto  
de 2023

Ing. David Luna



Universidad  
Rafael Landívar  
Tradición Jesuita en Guatemala

# Agenda

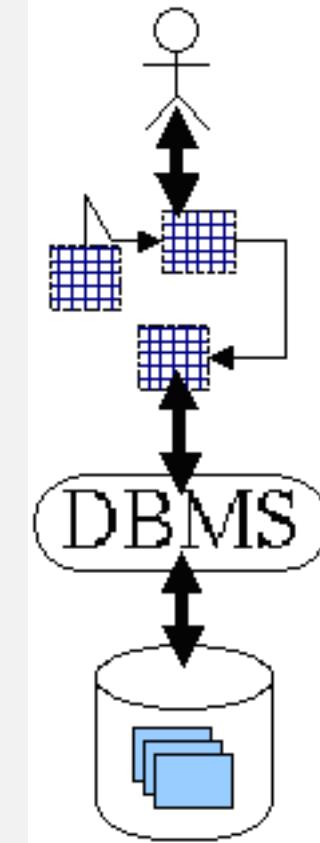


**Sistema de Archivos**



**Identificación de  
aprendizaje**

# Sistema de archivos

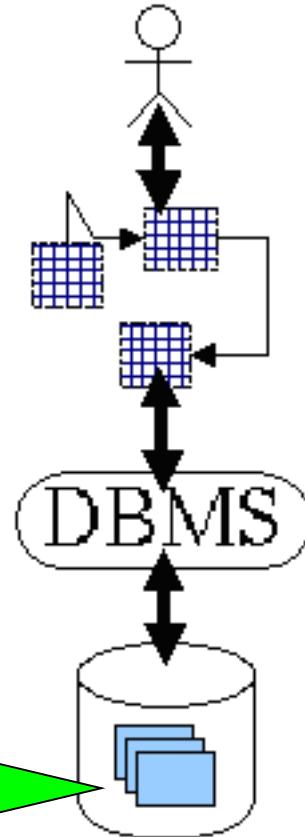


# Sistema de Archivos

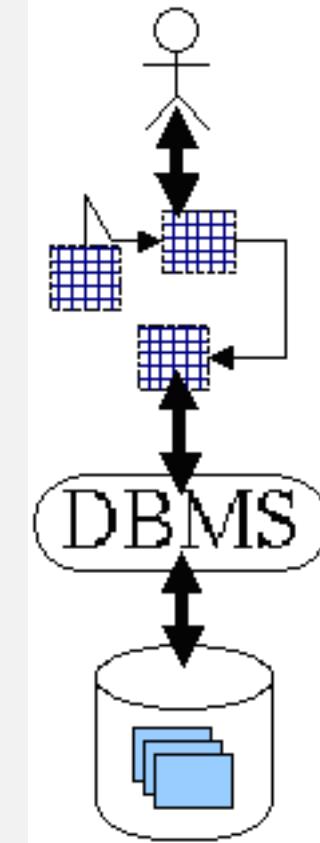
## ¿Por qué archivos?

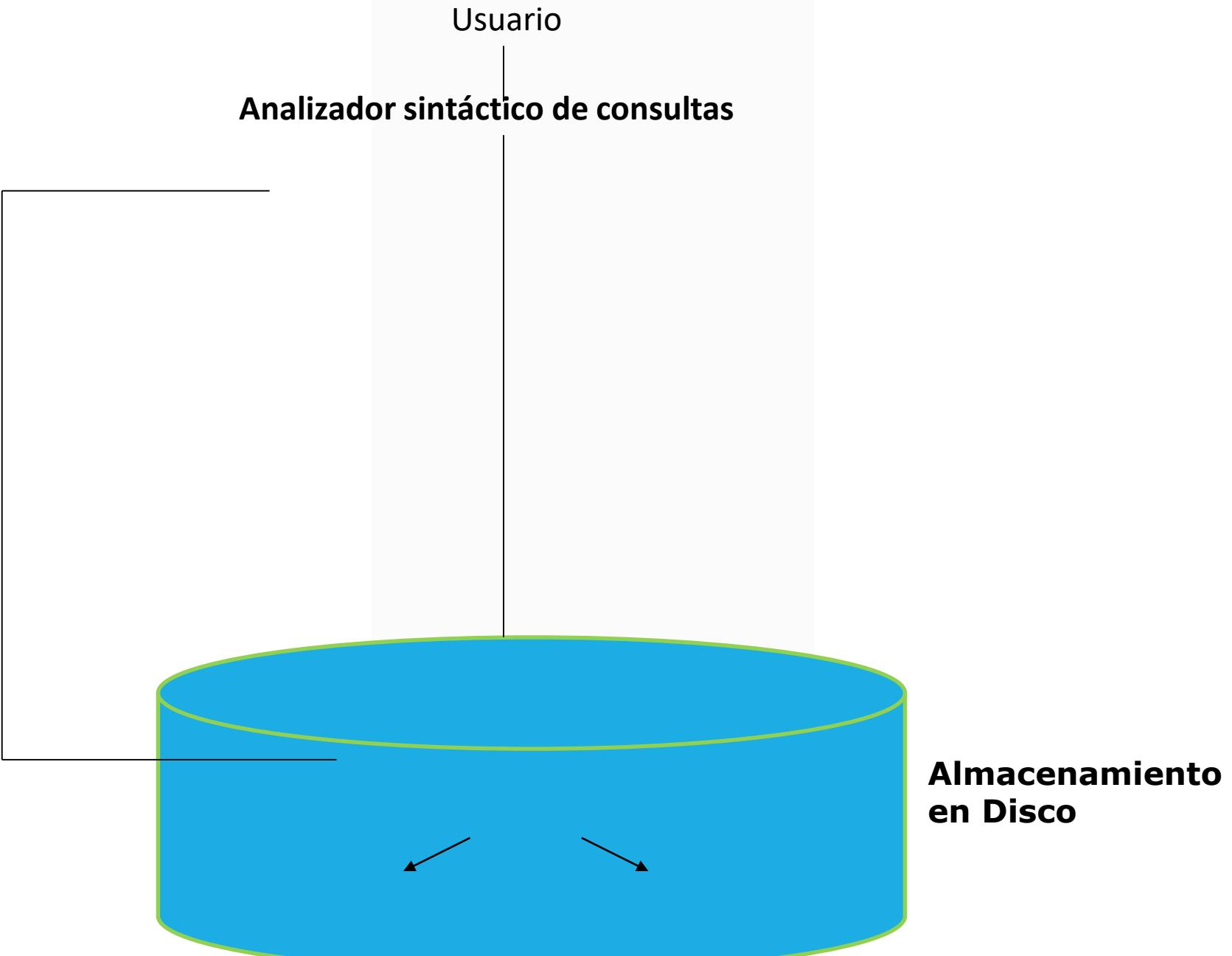
Una base de datos es un conjunto de datos relacionados

- El almacenamiento de datos se logra empleando uno o más archivos.
- Los archivos se dividen en bloques.
- Los bloques son unidades de almacenamiento y transferencia de datos.



# Diagrama de servicios de base de datos

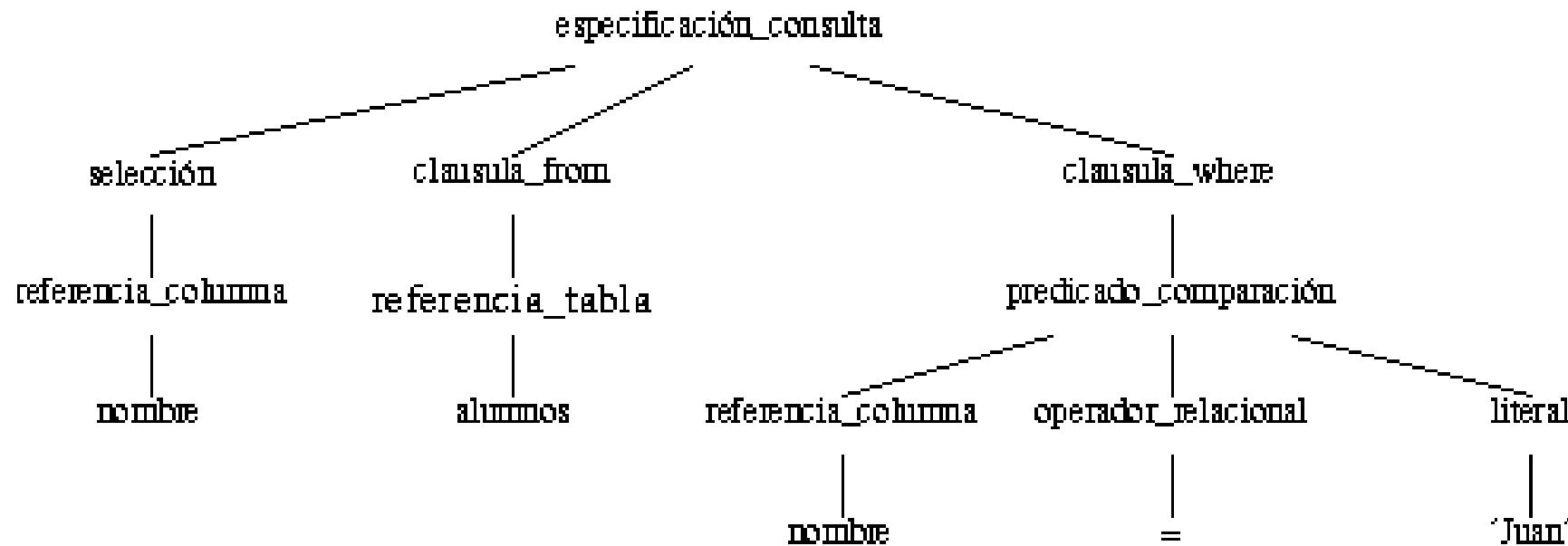




# Analizador Sintáctico

- Validación en forma de árbol de las consultas generadas por el usuario, siguiendo reglas que describen el lenguaje.
- El objetivo es validar que la consulta hecha por el usuario tenga un orden válido.
- Valida toda la cadena de “token” ingresados.

# Analizador Sintáctico Ejemplo:

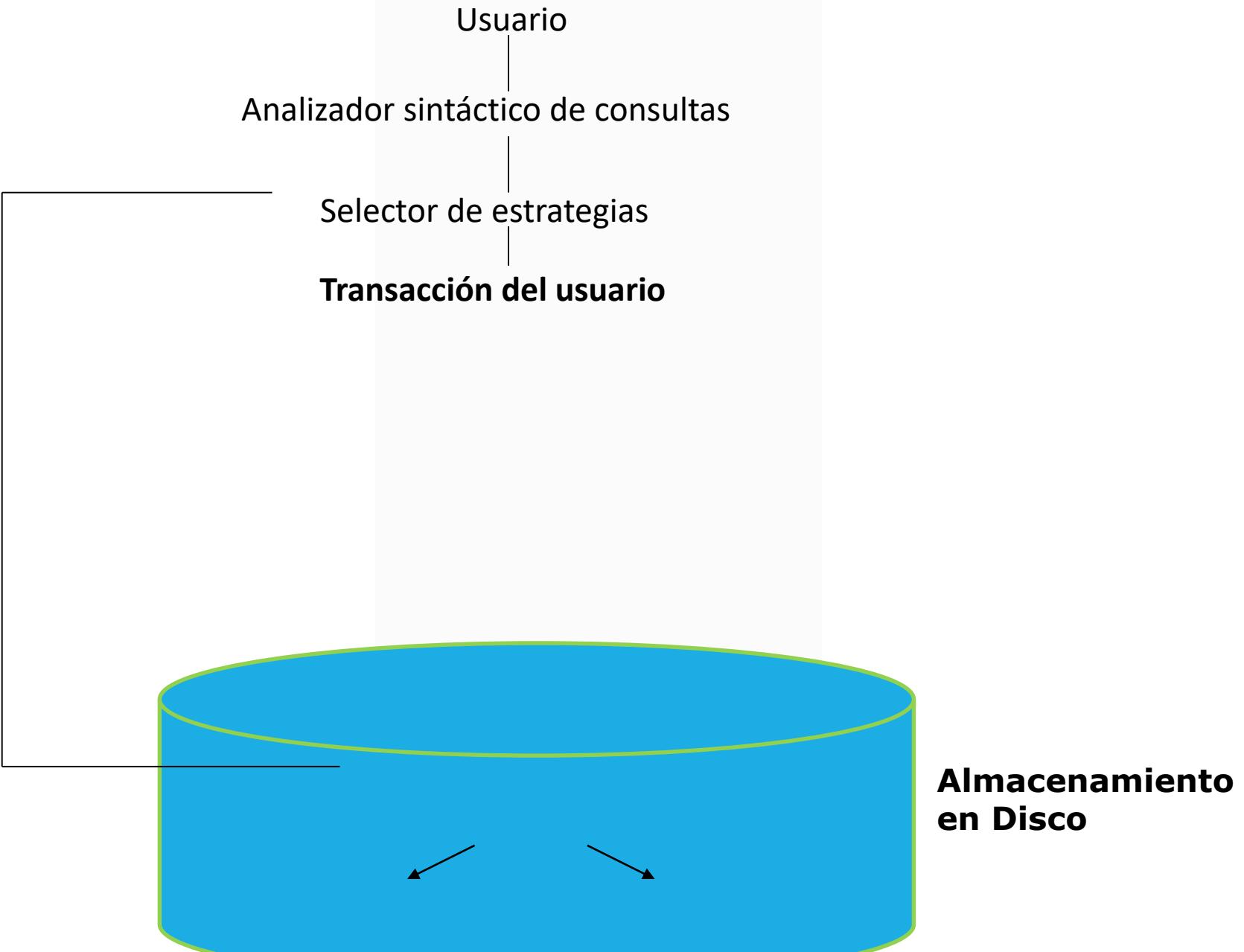


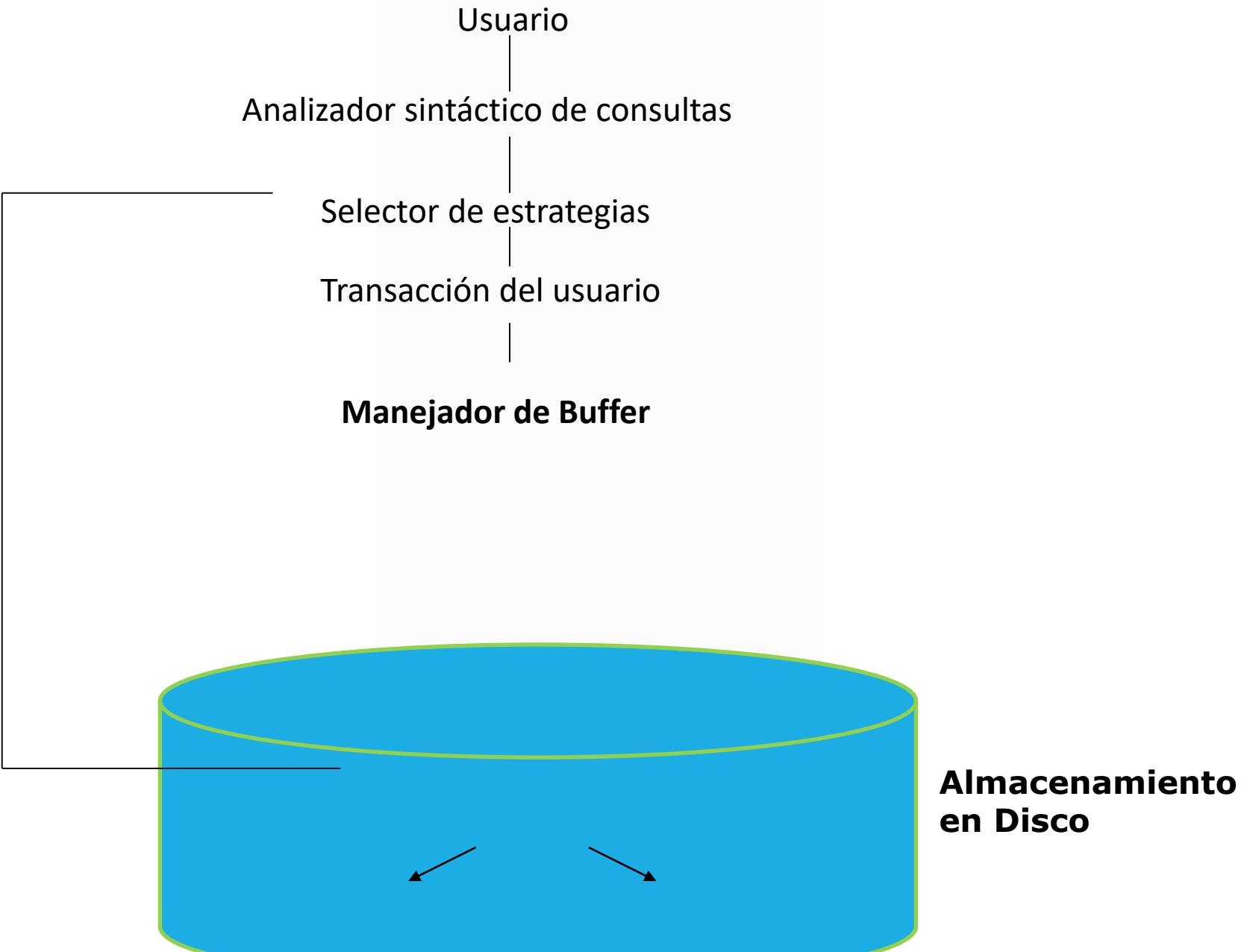


# Selector de Estrategias

Transforma una consulta del usuario en una forma efectiva para su posterior ejecución.

Traduce la orden del usuario a su forma más eficiente para su ejecución





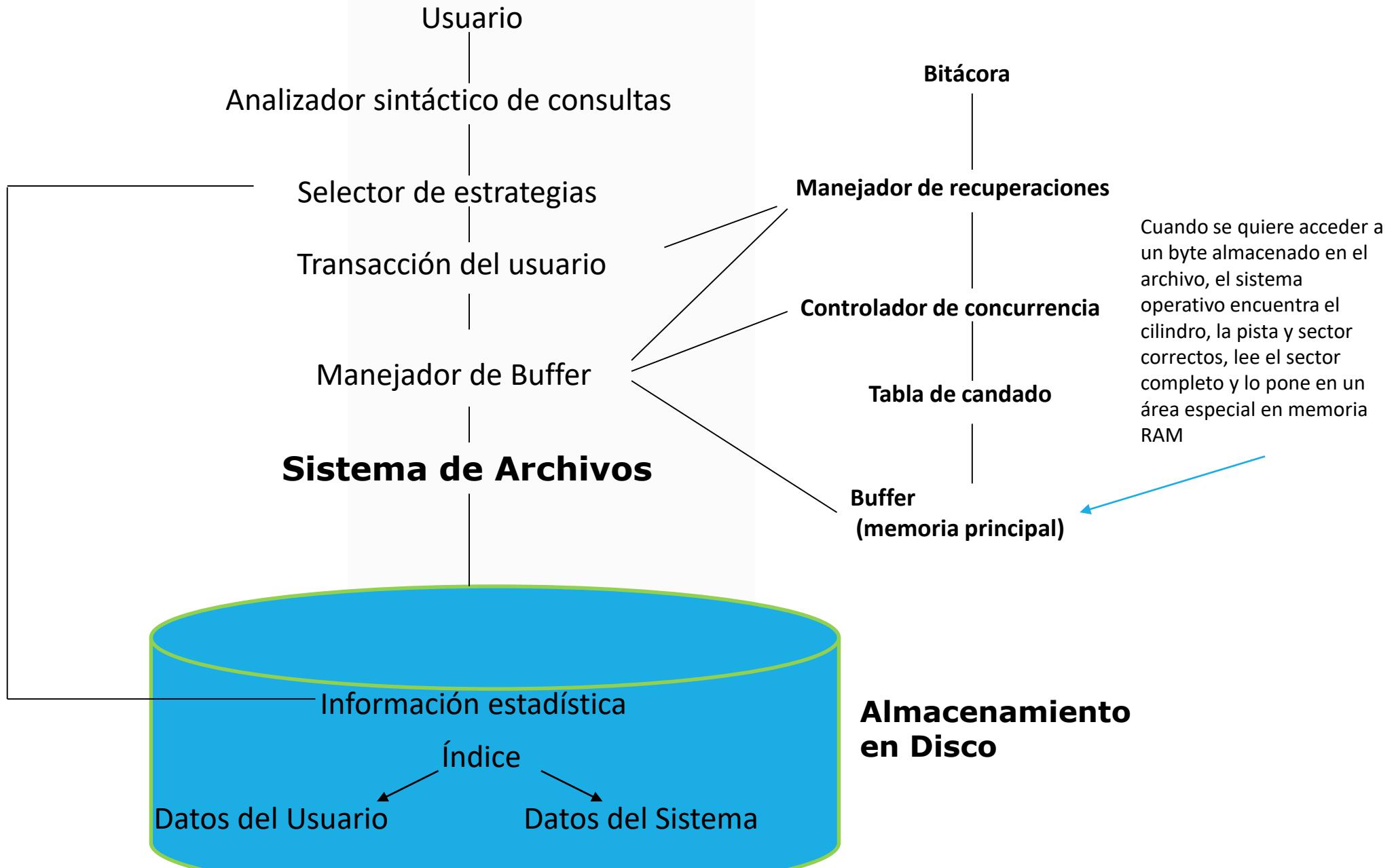
# Manejador de Buffer

- Controla el movimiento de datos entre la memoria principal y el almacenamiento en disco.
- Reservado para el almacenamiento temporal de datos, mientras que están esperando ser procesados.
- Debido a que las operaciones de E/S de disco pueden consumir muchos recursos y tardar bastante tiempo en completarse, el manejador de buffer funciona como un FIFO de todas las peticiones.



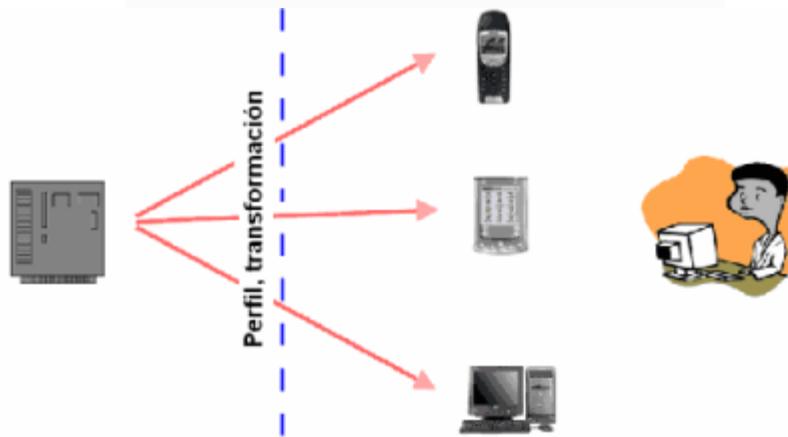
# Sistema de Archivos

- Soporte al administrador de buffer.
- Localización del almacenamiento en disco, suministrando una visión “lógica” de los datos.



# Funciones del Sistema de Archivos

- Referenciar a los archivos mediante “*Nombres Simbólicos*”, brindando “*Independencia de Dispositivos*”.



- ✓ Desde el punto de vista del usuario implica acceso universal.
- ✓ Desde el punto de vista del desarrollador, implica un único desarrollo con multitud de aplicaciones.

# Funciones del Sistema de Archivos

- Brindar una interfaz favorable al usuario
- Suministrar una “*visión lógica*” de los datos y de las funciones que serán ejecutadas, en vez de una “*visión física*”.



# Funciones del Sistema de Archivos

020081	2	0	1	3	0	13
020513	0	0	0	2	0	11
020826	2	0	0	1	0	4
021000	1	0	0	2	0	11
021013	2	0	1	1	0	4

# Funciones del Sistema de Archivos

020081	2	0	1	3	0	13
020513	0	0	0	2	0	11
020826	2	0	0	1	0	4
021000	1	0	0	2	0	11
021013	2	0	1	1	0	4

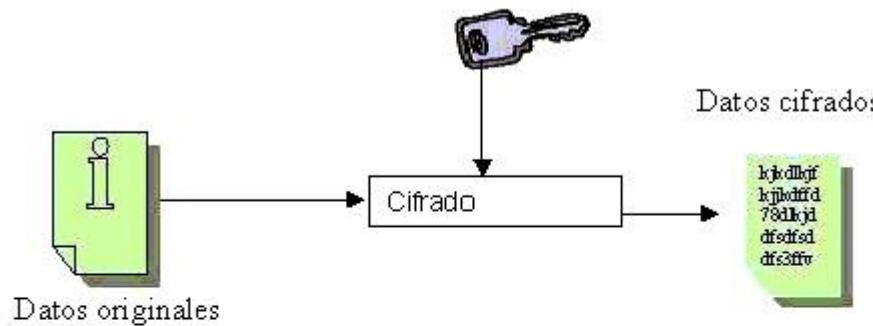
- Brindar una interfaz favorable al usuario

Curso	Créditos Teóricos	Créditos Prácticos	Prerrequisitos
▼ Primer Interciclo			
MATEMATICA DISCRETA I (11001)	4	0	EDP ESTRATEGIAS DE RAZONAMIENTO
MATEMATICA II (11001)	4	0	MATEMATICA I (11001)
▼ Segundo Ciclo			
CALCULO I (11001)	6	0	MATEMATICA II (11001)
EDP PENSAMIENTO IGNACIANO Y LANDIVARIANO	2	0	

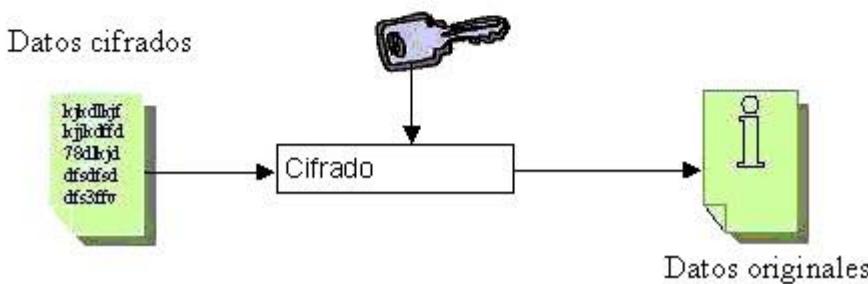
# Funciones del Sistema de Archivos

- Debe proporcionar posibilidades de “*Cifrado*” y “*Descifrado*”.

## CIFRADO



## DESCIFRADO



# Funciones del Sistema de Archivos

- El usuario no debe tener que preocuparse por:
  - ✓ Los dispositivos particulares.
  - ✓ Dónde serán almacenados los datos.
  - ✓ El formato de los datos en los dispositivos.
  - ✓ Los medios físicos de la transferencia de datos hacia y desde los dispositivos.

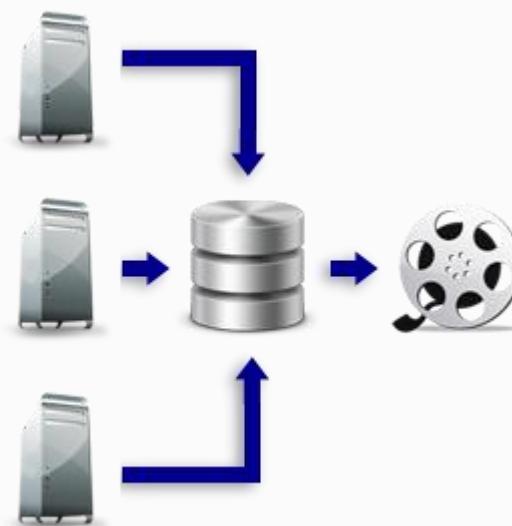
# Funciones del Sistema de Archivos

- Compartir los archivos y proporcionar varios tipos de acceso controlado:  
Ej.: “Acceso de Lectura”, “Acceso de Escritura”, “Acceso de Ejecución”, varias combinaciones de estos, etc.
- Los usuarios deben poder transferir información entre archivos.



# Funciones del Sistema de Archivos

- Se deben proporcionar posibilidades de “respaldo” y “recuperación” para prevenirse contra:
  - La pérdida accidental de información.
  - La destrucción maliciosa de información.



# Componentes del Sistema de Archivos

Métodos de acceso

Administración de Archivos

Administración del Almacenamiento Secundario/Terciario

Integridad de datos

# Componentes del Sistema de Archivos

Relacionados con la manera de acceder a los datos almacenados en archivos.

Métodos de acceso

Administración de Archivos

Administración del Almacenamiento Secundario/Terciario

Integridad de datos

# Componentes del Sistema de Archivos

Métodos de acceso

Relacionados con la manera de acceder a los datos almacenados en archivos.

Administración de Archivos

Mecanismos para que los archivos sean almacenados, referenciados, compartidos y asegurados.

Administración del Almacenamiento Secundario/Terciario

Integridad de datos

# Componentes del Sistema de Archivos

Métodos de acceso

Relacionados con la manera de acceder a los datos almacenados en archivos.

Administración de Archivos

Mecanismos para que los archivos sean almacenados, referenciados, compartidos y asegurados.

Administración del Almacenamiento Secundario/Terciario

Para la asignación de espacio a los archivos en los dispositivos de almacenamiento secundario/terciario.

Integridad de datos

# Componentes del Sistema de Archivos

Métodos de acceso

Relacionados con la manera de acceder a los datos almacenados en archivos.

Administración de Archivos

Mecanismos para que los archivos sean almacenados, referenciados, compartidos y asegurados.

Administración del Almacenamiento Secundario/Terciario

Para la asignación de espacio a los archivos en los dispositivos de almacenamiento secundario/terciario.

Integridad de datos

Para garantizar la integridad de la información almacenada.

# Componentes del Sistema de Archivos





**¿Cómo podemos identificar cuánto se aprende en una clase?**

# Aviso



- **Evaluación corta**

**24 de agosto de  
2023**

# Gracias

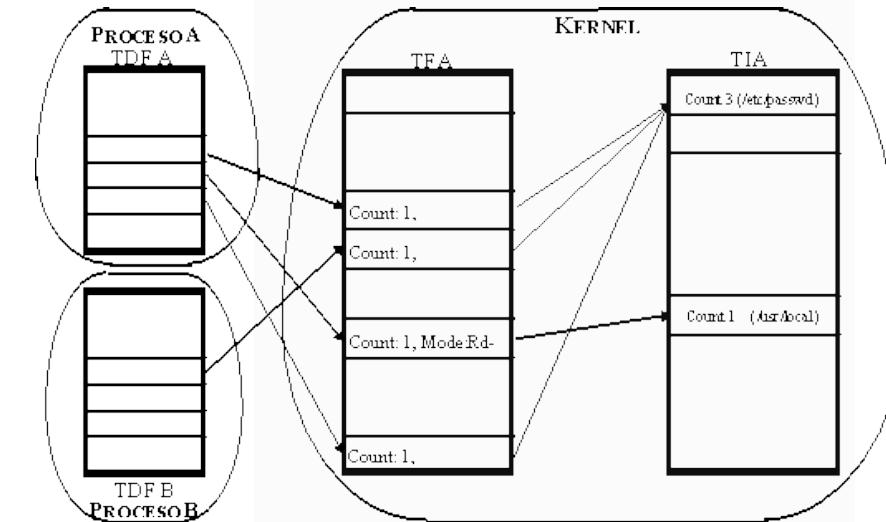
¿ALGUNA PREGUNTA?

# Manejo e Implementación de Archivos



Guatemala 29 de agosto  
de 2023

Ing. David Luna



Universidad  
Rafael Landívar  
Tradición Jesuita en Guatemala

# Agenda

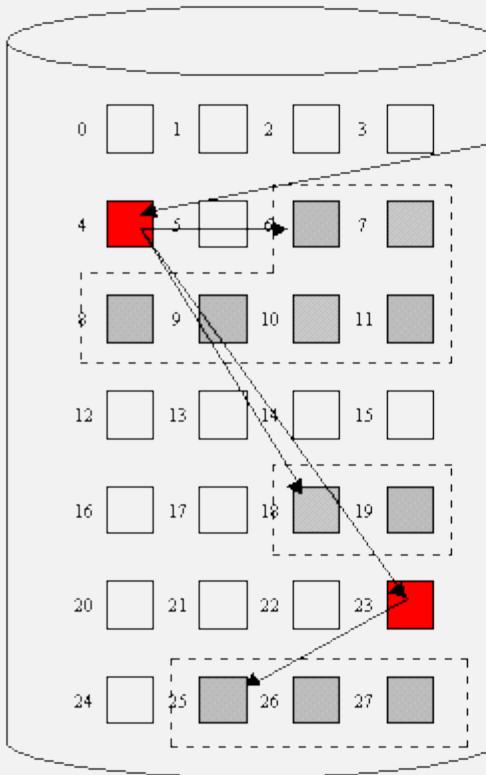


**Asignación de  
espacio**



**Cierre**

# Asignación de espacio



CABECERA 4

I

Sector 4

6	6
18	2
23	-1

Sector que contiene  
la dirección del  
siguiente bloque

# Asignación / Liberación de espacio

Aspectos a considerar:

- Forma de almacenamiento de archivos y directorios.
- Administración del espacio en disco.
- Forma de hacerlo de la mejor manera (Eficiencia y confiabilidad)

# Asignación / Liberación de espacio

Se deben tener presentes problemas tales como la “fragmentación” creciente del espacio en disco:

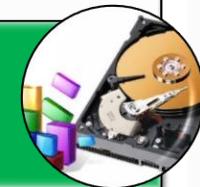
- Ocasiona problemas de rendimiento al hacer que los archivos se desperdiguen a través de bloques muy dispersos.
- Técnicas para aliviar el problema de la “fragmentación” consisten en realizar periódicamente:
  - “Condensación”: se pueden “reorganizar” los archivos expresamente o automáticamente según algún criterio predefinido.

[http://upload.wikimedia.org/wikipedia/commons/4/40/Defragmenting\\_disk.gif](http://upload.wikimedia.org/wikipedia/commons/4/40/Defragmenting_disk.gif)

- “Recolección de basura o residuos” (garbage collector): se puede hacer fuera de línea o en línea, con el sistema activo, según la implementación.

- Reorganización del espacio

Condensación



- Problemas de rendimiento

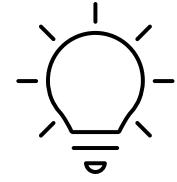
Fragmentación



- Recolección de desechos

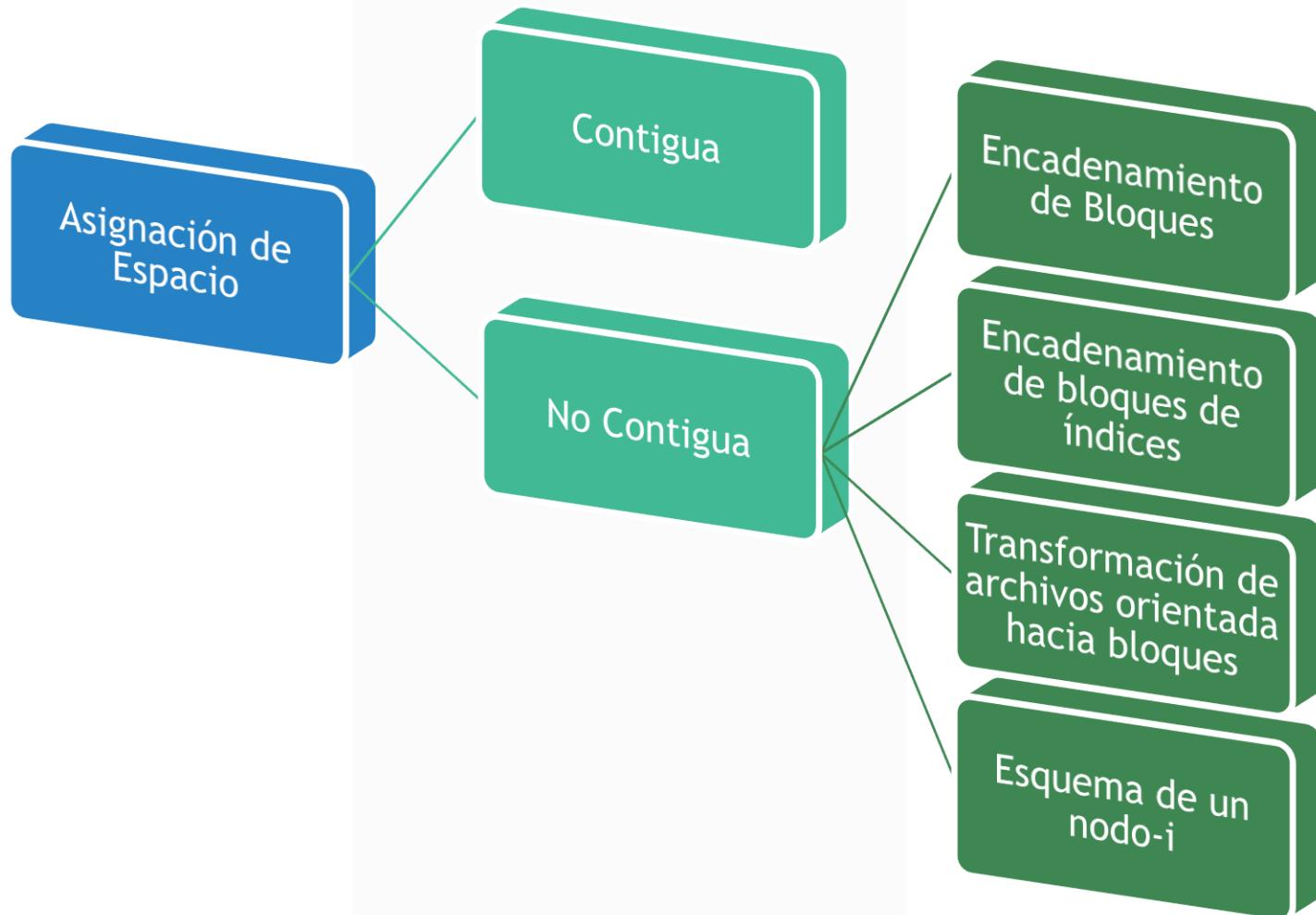
Garbage Collector





El aspecto clave de la implantación del almacenamiento de archivos es *el registro de los bloques asociados a cada archivo*.

# Implementación de Archivos



# Asignación Contigua o Adyacente.

Los archivos son asignados a áreas contiguas de almacenamiento secundario.

La principal **ventaja** es:

- Facilidad de implantación, ya que solo se precisa el número del bloque de inicio para localizar un archivo.

Los principales **defectos** son:

- Se debe conocer el tamaño máximo del archivo al crearlo.
- Produce una gran fragmentación de los discos.

# Asignación No Contigua.

Son esquemas de almacenamiento más dinámicos.

El sistema trata de asignar nuevos bloques a un archivo eligiendo bloques libres lo más próximos posible a los bloques del archivo existentes.

Se destacan los siguientes:

1. Encadenamiento de bloques
2. Encadenamiento de bloques de índices
3. Transformación de archivos orientada hacia bloques
4. Esquema de un nodo-i

## Asignación No Contigua.

### 1. *Encadenamiento de bloques o lista ligada:*

- Las entradas en el directorio apuntan al primer bloque de cada archivo.
- Cada uno de los bloques de longitud fija que forman un archivo contiene dos partes:
  - Un bloque de datos.
  - Un apuntador al bloque siguiente.
- Cada bloque contiene varios sectores.

# Asignación No Contigua.

## 1 Encadenamiento de bloques o lista ligada:

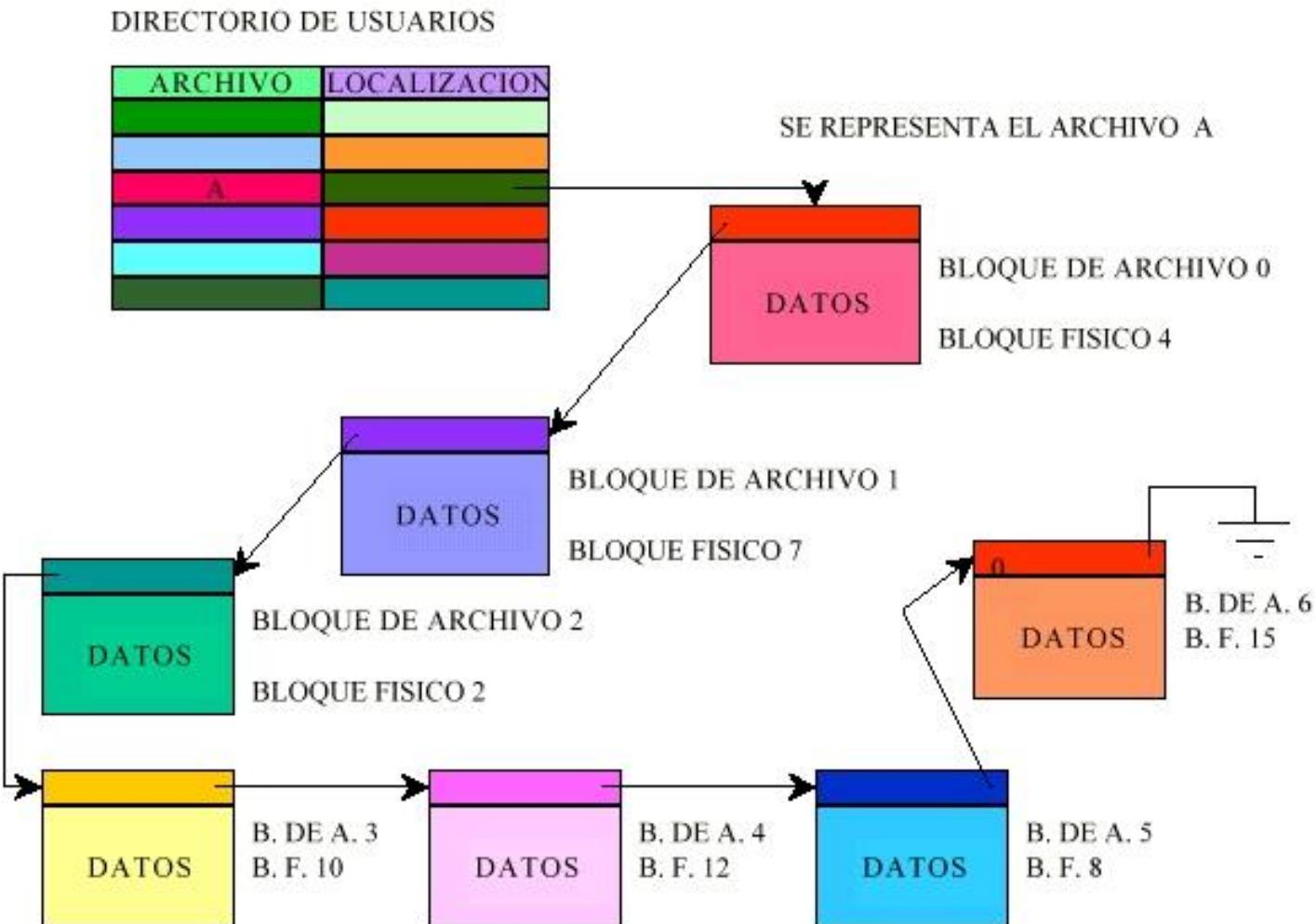


Figura ~4.4: Encadenamiento de bloques o lista ligada de bloques.

## Asignación No Contigua.

### 1 *Encadenamiento de bloques o lista ligada:*

- Localizar un registro determinado requiere:
  - Buscar en la cadena de bloques hasta encontrar el bloque apropiado.
  - Buscar en el bloque hasta encontrar el registro.
- El examen de la cadena desde el principio puede ser lento ya que debe realizarse de bloque en bloque, y pueden estar dispersos por todo el disco.
- Se pueden usar “*listas de encadenamiento doble*”, hacia adelante y hacia atrás, con lo que se facilita la búsqueda

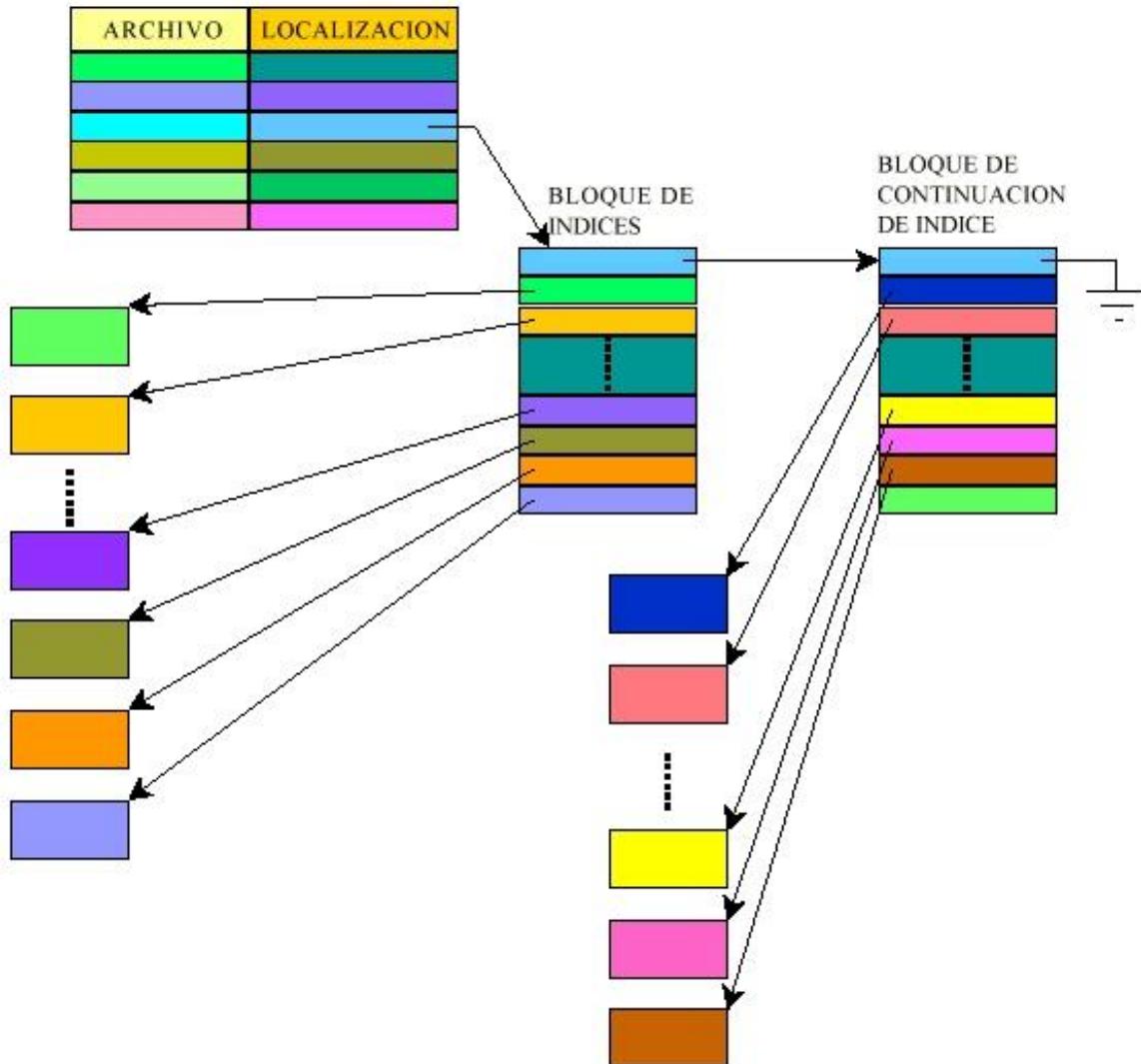
## **Asignación No Contigua.**

### **2. Encadenamiento de bloques de índice:**

- Los apuntadores son colocados en varios bloques de índices separados:
  - Cada bloque de índices contiene un número fijo de elementos.
  - Si es necesario utilizar más de un bloque de índices para describir un archivo, se encadena una serie de bloques de índices.

# Asignación No Contigua.

## 2. Encadenamiento de bloques de índice:



## Asignación No Contigua.

### 2. Encadenamiento de bloques de índice:

- La gran **ventaja** es que la búsqueda puede realizarse en los propios bloques de índices.
- Los bloques de índices pueden mantenerse juntos en el almacenamiento secundario para acortar la búsqueda.
- La principal **desventaja** es que las inserciones pueden requerir la reconstrucción completa de los bloques de índices:
  - Una posibilidad es dejar vacía una parte de los bloques de índices para facilitar inserciones futuras y retardar las reconstrucciones.
- Es suficiente que el dato del directorio contenga el número de bloque inicial para localizar todos los bloques restantes, sin importar el tamaño del archivo.

## Asignación No Contigua.

### 3. *Transformación de archivos orientada hacia bloques :*

- Se utilizan números de bloques en vez de apunadores.
- Se conserva un mapa del archivo, conteniendo una entrada para cada bloque del disco.
- Las entradas en el directorio del usuario apuntan a la primera entrada al *mapa del archivo* para cada archivo.
- Cada entrada al mapa del archivo contiene el número del bloque siguiente de ese archivo.
- La entrada al mapa del archivo correspondiente a la última entrada de un archivo determinado se ajusta a algún valor “centinela” (“null”) para indicar que se alcanzó el último bloque de un archivo.
- El sistema puede mantener una *lista de bloques libres*.
- La principal **ventaja** es que las cercanías físicas del disco se reflejan en el mapa del archivo

DIRECTORIO DE USUARIOS	
ARCHIVO	LOCALIZACION
A	8
B	6
C	2

MAPA DEL ARCHIVO	
0	22
1	NIL
2	5
3	26
4	9
5	20
6	10
7	LIBRE
8	17
9	1
10	14
11	LIBRE
12	3
13	4
14	0
15	LIBRE
16	LIBRE
17	12
18	13
19	NIL
20	23
21	LIBRE
22	18
23	19
24	LIBRE
25	LIBRE
26	NIL
27	LIBRE

*orientada hacia*

BLOQUE 0 B(4)	BLOQUE 1 B(10)	BLOQUE 2 C(1)	BLOQUE 3 A(4)	BLOQUE 4 B(8)	BLOQUE 5 C(2)	BLOQUE 6 B(1)
BLOQUE 7 LIBRE	BLOQUE 8 A(1)	BLOQUE 9 B(9)	BLOQUE 10 B(2)	BLOQUE 11 LIBRE	BLOQUE 12 A(3)	BLOQUE 13 B(7)
BLOQUE 14 B(3)	BLOQUE 15 LIBRE	BLOQUE 16 LIBRE	BLOQUE 17 A(2)	BLOQUE 18 B(6)	BLOQUE 19 C(5)	BLOQUE 20 C(3)
BLOQUE 21 LIBRE	BLOQUE 22 B(5)	BLOQUE 23 C(4)	BLOQUE 24 LIBRE	BLOQUE 25 LIBRE	BLOQUE 26 A(5)	BLOQUE 27 LIBRE

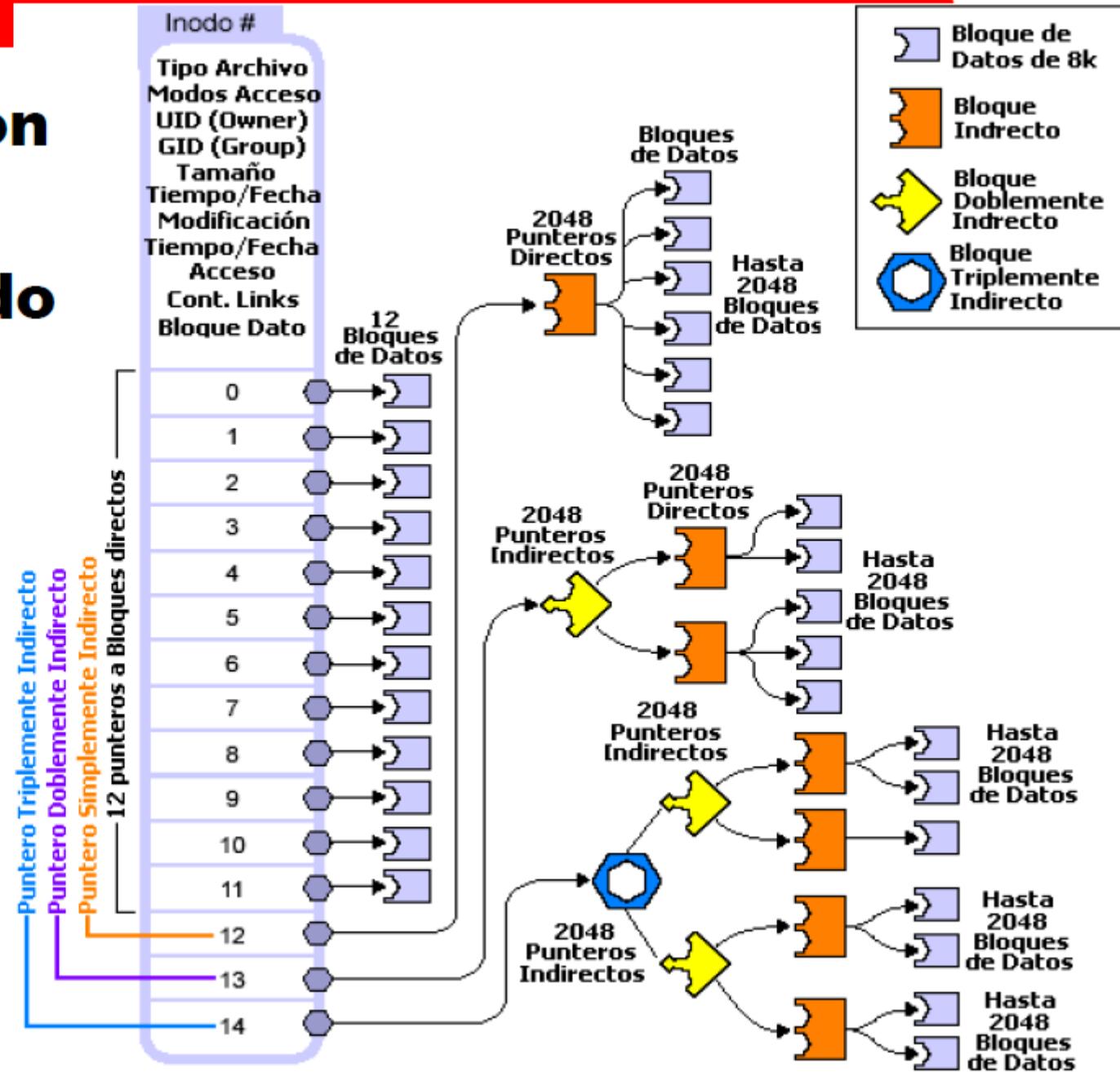
Figura~4.6: Transformación de archivos orientada hacia bloques.

## Asignación No Contigua.

### 4. Nodos-i (nodos índices) :

- Se asocia a cada archivo una pequeña **tabla**, llamada nodo-i (nodo índice):
  - Contiene los atributos y direcciones en disco de los bloques del archivo.
  - Se traslada del disco a la memoria principal al abrir el archivo.  
Si el archivo es pequeño, toda la información está en el nodo-i.  
Si el archivo es grande, una de las direcciones en el nodo-i es la dirección de un bloque en el disco llamado *bloque simplemente indirecto*:
    - Contiene las direcciones en disco adicionales.
    - Si resulta insuficiente, otra dirección en el nodo-i, el *bloque doblemente indirecto*, contiene la dirección de un bloque que presenta una lista de los bloques simplemente indirectos:
    - Cada bloque simplemente indirecto apunta a un grupo de bloques de datos.
  - De ser necesario se pueden utilizar *bloques triplemente indirectos*

# Información contenida en un i-nodo



# Seguridad y Respaldo



## Seguridad del Sistema de Archivos

Muchos esquemas de protección se basan en la hipótesis de que  
*el sistema conoce la identidad de cada usuario:*

- La identificación de los usuarios se conoce como la *autenticación de los usuarios.*

## Seguridad del Sistema de Archivos

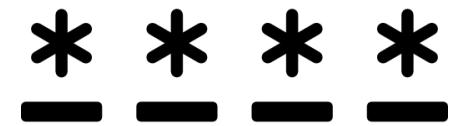
- Muchos métodos de autentificación se basan en:



•Algo que *es el usuario.*



•Algo que *posee el usuario.*



•La identificación de algo *conocido por el usuario.*

# Seguridad del Sistema de Archivos



\* \* \* \*

---

---

---

---

- **Contraseñas**
  - Son la forma de autentificación más utilizada.
  - Son de fácil comprensión e implementación.
  - Deben *almacenarse cifradas (encriptadas)*.
  - Se deben prever intentos de ataques, consistentes en *pruebas de combinaciones* de nombres y contraseñas.

## **Respaldo y recuperación**

La destrucción de la información, ya sea accidental o intencional, es una realidad y tiene distintas causas :

- Fallas de hardware y de software.
  - Fenómenos meteorológicos atmosféricos.
  - Fallas en el suministro de energía.
  - Incendios e inundaciones.
  - Robos, vandalismo (incluso terrorismo).
- 
- Esta posible destrucción de la información debe ser tenida en cuenta por:

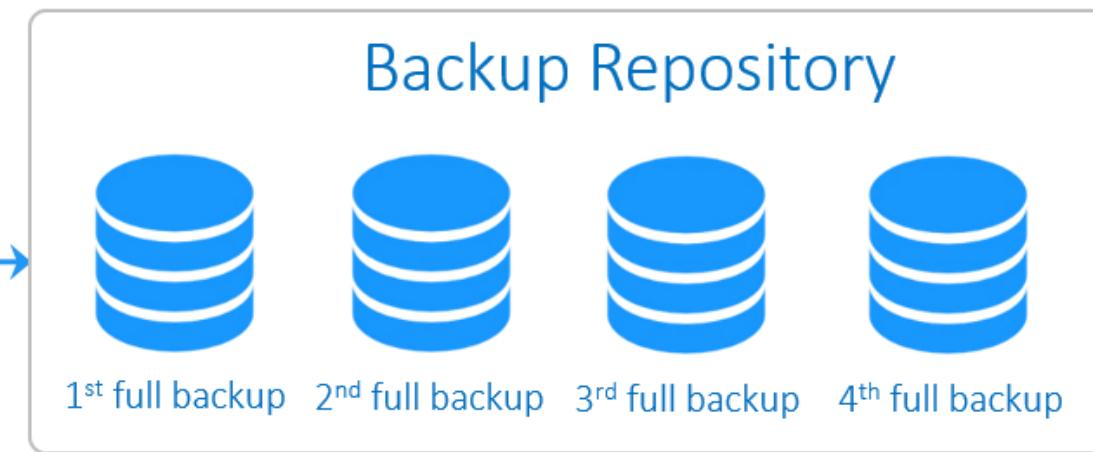
Los sistemas operativos en general.

Los sistemas de archivos en particular.

# Respaldo y recuperación

## Full Backup

Source Data



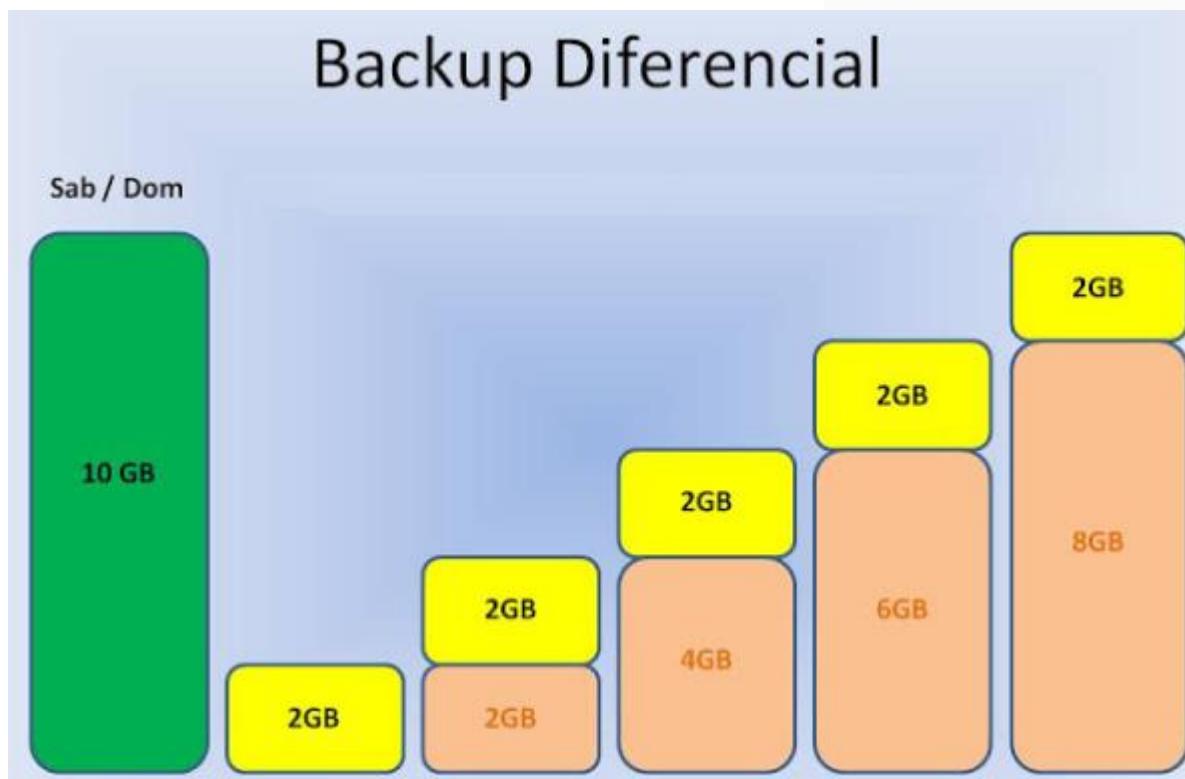
**Copia completa:** Copia todos los archivos cada vez que se lanza el proceso.

El proceso es tardado. Requiere más espacio.

La restauración es más rápida y simple.

## Respaldo y recuperación

### 2. Copia diferencial:

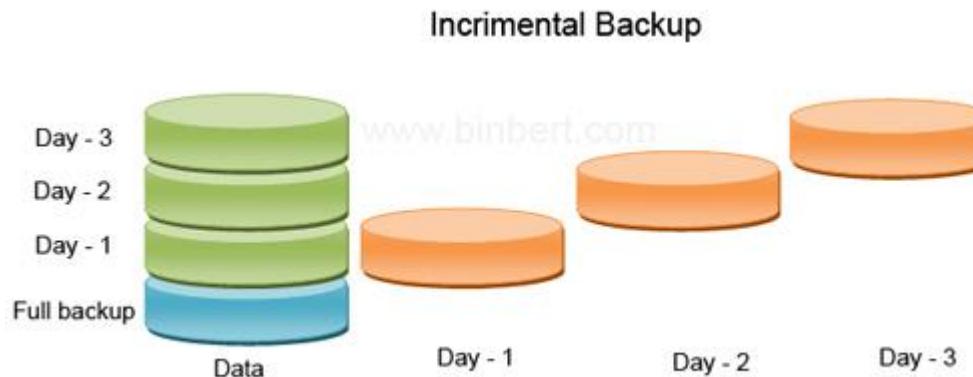


- Copia únicamente los archivos que han sido creados o modificados desde la última copia completa.
- Proceso de hacer la copia es más rápido.
- Las copias diferenciales se van haciendo más grandes ya que reflejan los cambios desde una marca fija en el tiempo.

## Respaldo y recuperación

### Copia Incremental: (o diferencial incremental).

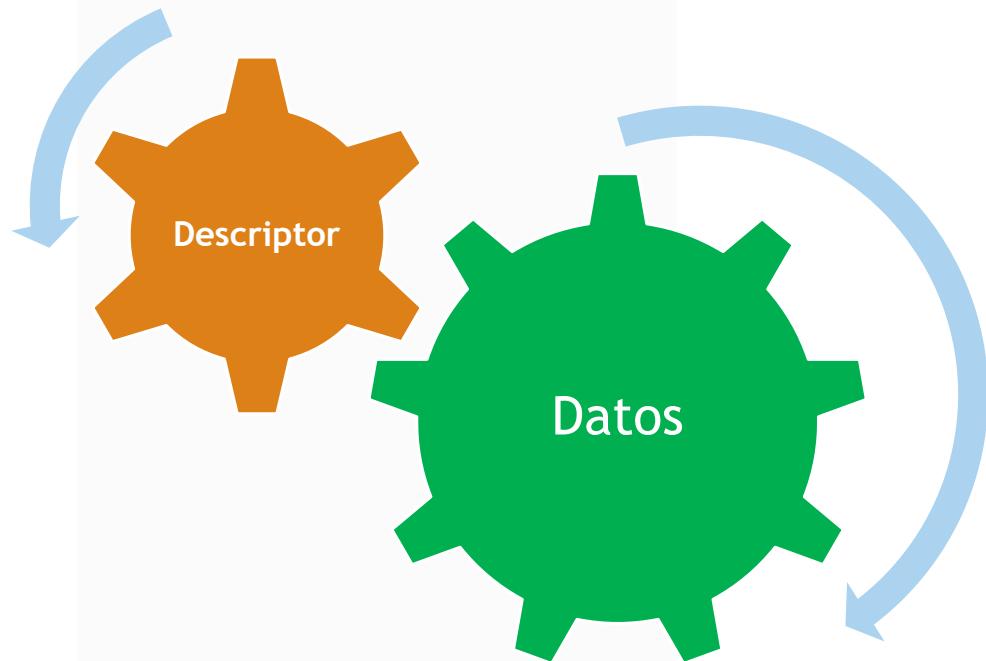
Es más avanzada (o “inteligente”) ya que sólo copia los ficheros creados o modificados desde el último backup, ya sea completo o incremental, reduciendo la cantidad de información a copiar en cada proceso.



# Componentes de un archivo



# Componentes de un Archivo



# Descriptor:

El *descriptor de archivos o bloque de control de archivos*, son los datos que el sistema necesita para administrar dicho archivo.

- Nombre simbólico del archivo
- Localización del archivo en el almacenamiento secundario
- Organización del archivo (método de organización y acceso)
- Datos de control de acceso
- Tipo (archivo de datos, programa objeto, programa fuente, etc.)
- Disposición (permanente, temporal)
- Fecha y hora de creación
- Fecha de la última modificación
- Usuario que modificó
- Estadísticas de acceso (número de lecturas, escrituras, tiempos de acceso, etc.)
- Registros activos e inactivos.
- Metadatos
- Y más...

## **Contenido del Archivo:**

Datos, que pueden estar estructurados en la forma de:

### ➤ **Registros,**

- Están compuestos por campos. Estos campos se identifican por lo que se conoce como "metadatos" (información acerca de los datos), en el caso de los campos el metadato es el nombre de dicho campo.

### ➤ **Formación abstracta,**

- El archivo tiene al inicio (a veces al final) una serie de bytes que representan los metadatos del archivo, es decir, si por ejemplo el archivo es una imagen al inicio tendremos bytes que nos digan el largo, el ancho, el número de colores, etc, estos bytes se encuentran especificados en el "formato" del archivo. Son formatos estándar de aplicaciones específicas que se emplean en todas las plataformas

# Clasificación de archivos



# Clasificación de Archivos

## Factores de Clasificación

1. Organización

- Apilo
- Secuencial
- Secuencial Indizado
- Indizado

2. Acceso

- Lectura, escritura, L/E
- Secuencial y directo

3. Longitud  
de Registro

- Longitud Fija
- Longitud Variable

# Clasificación de Archivos

## 1. Organización de Archivos:

Forma en que se disponen los datos para su almacenamiento, recuperación y procesamiento.

Normalmente es una tarea de los sistemas operativos. Para la recuperación, eficaz de los datos, estos se organizan en estructuras lógicas (pilas, listas, tablas, árboles, etc).

- 1.1 Apilo
- 1.2 Secuencial
- 1.3 Secuencial Indizado
- 1.4 Indizado

# Clasificación de Archivos

## 2. Acceso:

Forma en que puede llegarse a los datos.

2.1 Lectura, escritura, L/E

2.2 Secuencial y directo

# Clasificación de Archivos

## 3. Longitud de registro:

### 3.1 Longitud fija:

Independientemente de los datos almacenados, cada registro siempre tiene la misma cantidad de bytes.

### 3.2 Longitud variable

Los datos almacenados en un registro pueden tener mayor o menor longitud que el resto de registros.

# Selección del tipo de Archivo

¿Cómo se elije el tipo de archivo?

- Aplicación
- Tamaño del archivo
- Velocidad búsqueda
- Manejo de operaciones
- Reorganización

# Operaciones de un Archivo

¿Qué puede hacerse con un archivo?

- Creación
- Eliminación
- Inserción de datos
- Actualización de datos
- Eliminación de datos
- Reorganización
- Reestructuración

# Reorganización de un Archivo

Los registros eliminados de forma lógica, se eliminan permanentemente de forma física. Será necesario realizar varias operaciones para completarla.

Estas operaciones podrían incluir:

- Crear un archivo nuevo
- Insertar/eliminar registros
- Eliminar el archivo antiguo
- Modificar los descriptores

# Reestructuración de un Archivo

Cuando es necesario implementar cambios en la estructura de los datos almacenados en los archivos, se habla de una “reestructuración”. Puede originarse por:

- Cambios drásticos en el volumen de datos
- Cambios en las reglas del negocio
- Cambios en la normativa legal y/o políticas estatales
- Bajo desempeño del sistema de archivos

# Laboratorio

**Martes 05 de Septiembre**

**TRAER SU COMPUTADORA**

- Instalar Netbeans
  - <https://netbeans.apache.org/>
- Requisitos para instalar Netbeans:
  - [Requisitos para instalar NetBeans 2023 \(requisitomania.info\)](#)
  - Java Development Kit (JDK): NetBeans está construido en Java, por lo que es necesario tener **JDK** instalado en tu sistema

# Gracias

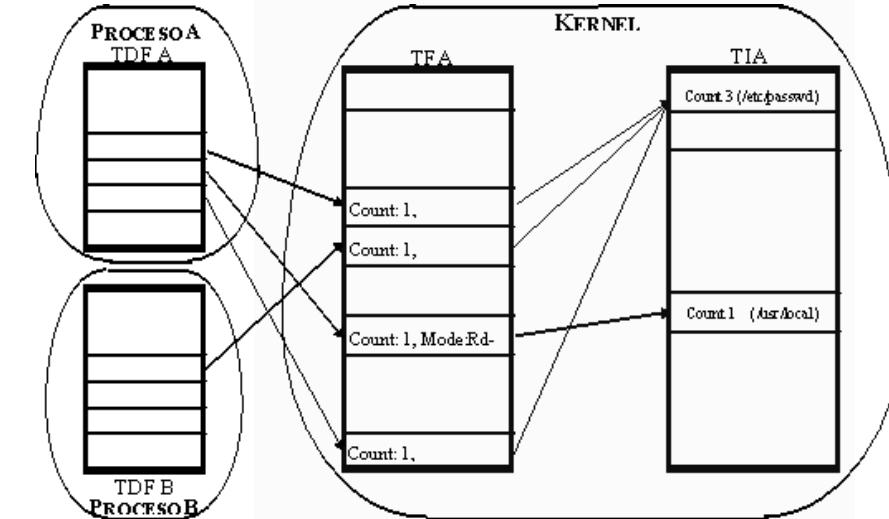
¿ALGUNA PREGUNTA?

# Manejo e Implementación de Archivos



Guatemala 19 de  
septiembre de 2023

Ing. David Luna



Universidad  
Rafael Landívar  
Tradición Jesuita en Guatemala

# Agenda

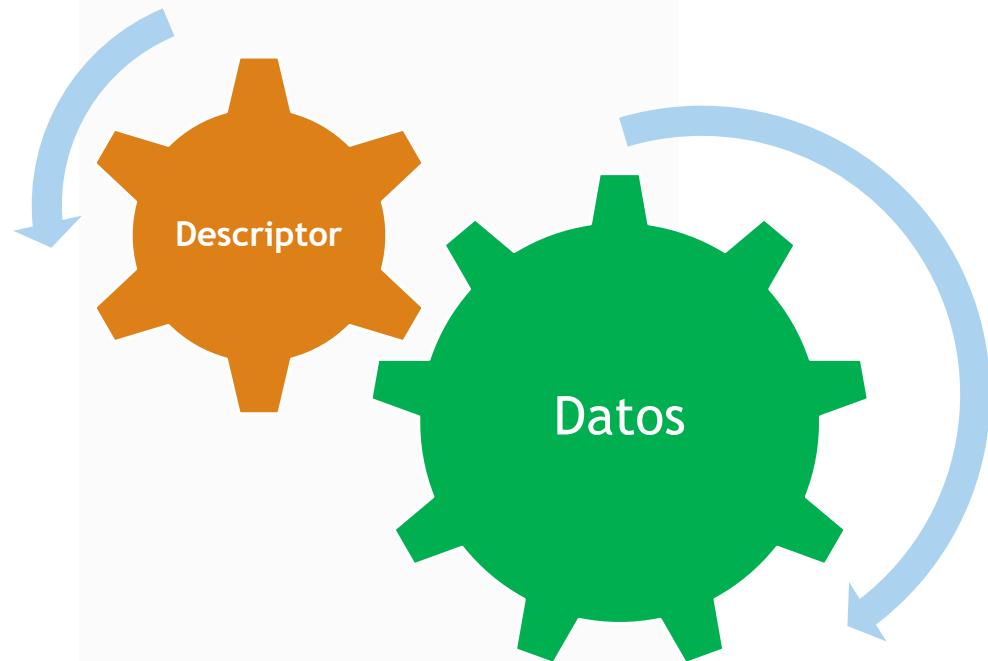


**Clasificación de  
Archivos**

# Componentes de un archivo



# Componentes de un Archivo



# Descriptor:

El *descriptor de archivos o bloque de control de archivos, son los datos que el sistema necesita para administrar dicho archivo.*

- Nombre simbólico del archivo
- Localización del archivo en el almacenamiento secundario
- Organización del archivo (método de organización y acceso)
- Datos de control de acceso
- Tipo (archivo de datos, programa objeto, programa fuente, etc.)
- Disposición (permanente, temporal)
- Fecha y hora de creación
- Fecha de la última modificación
- Usuario que modificó
- Estadísticas de acceso (número de lecturas, escrituras, tiempos de acceso, etc.)
- Registros activos e inactivos.
- Metadatos
- Y más...

## **Contenido del Archivo:**

Datos, que pueden estar estructurados en la forma de:

### ➤ **Registros,**

- Están compuestos por campos. Estos campos se identifican por lo que se conoce como "metadatos" (información acerca de los datos), en el caso de los campos el metadato es el nombre de dicho campo.

### ➤ **Formación abstracta,**

- El archivo tiene al inicio (a veces al final) una serie de bytes que representan los metadatos del archivo, es decir, si por ejemplo el archivo es una imagen al inicio tendremos bytes que nos digan el largo, el ancho, el número de colores, etc, estos bytes se encuentran especificados en el "formato" del archivo. Son formatos estándar de aplicaciones específicas que se emplean en todas las plataformas

# Clasificación de archivos



# Clasificación de Archivos

## Factores de Clasificación

1. Organización

- Apilo
- Secuencial
- Secuencial Indizado
- Indizado

2. Acceso

- Lectura, escritura, L/E
- Secuencial y directo

3. Longitud  
de Registro

- Longitud Fija
- Longitud Variable

# Clasificación de Archivos

## 1. Organización de Archivos:

Forma en que se disponen los datos para su almacenamiento, recuperación y procesamiento.

Normalmente es una tarea de los sistemas operativos. Para la recuperación, eficaz de los datos, estos se organizan en estructuras lógicas (pilas, listas, tablas, árboles, etc).

- 1.1 Apilo
- 1.2 Secuencial
- 1.3 Secuencial Indizado
- 1.4 Indizado

# Clasificación de Archivos

## 2. Acceso:

Forma en que puede llegarse a los datos.

2.1 Lectura, escritura, L/E

2.2 Secuencial y directo

# Clasificación de Archivos

## 3. Longitud de registro:

### 3.1 Longitud fija:

Independientemente de los datos almacenados, cada registro siempre tiene la misma cantidad de bytes.

### 3.2 Longitud variable

Los datos almacenados en un registro pueden tener mayor o menor longitud que el resto de registros.

# Selección del tipo de Archivo

¿Cómo se elije el tipo de archivo?

- Aplicación
- Tamaño del archivo
- Velocidad búsqueda
- Manejo de operaciones
- Reorganización

# Operaciones de un Archivo

¿Qué puede hacerse con un archivo?

- Creación
- Eliminación
- Inserción de datos
- Actualización de datos
- Eliminación de datos
- Reorganización
- Reestructuración

# Reorganización de un Archivo

Los registros eliminados de forma lógica, se eliminan permanentemente de forma física. Será necesario realizar varias operaciones para completarla.

Estas operaciones podrían incluir:

- Crear un archivo nuevo
- Insertar/eliminar registros
- Eliminar el archivo antiguo
- Modificar los descriptores

# Reestructuración de un Archivo

Cuando es necesario implementar cambios en la estructura de los datos almacenados en los archivos, se habla de una “reestructuración”. Puede originarse por:

- Cambios drásticos en el volumen de datos
- Cambios en las reglas del negocio
- Cambios en la normativa legal y/o políticas estatales
- Bajo desempeño del sistema de archivos

# Clasificación de Archivos (Organización)



## 1. Organización

- Apilo
- Secuencial
- Secuencial Indizado
- Indizado

# Archivo Apilo



# Archivo Apilo

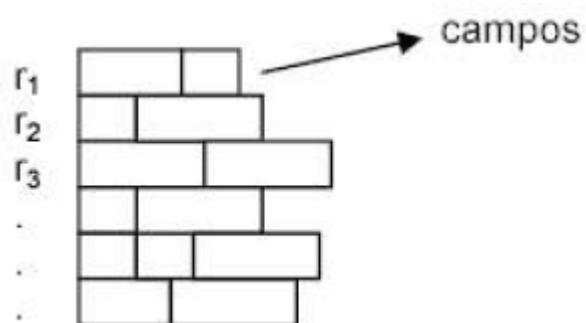
- Los datos se guardan en el orden en que llegan, sin secuencia lógica.
- **No existe una llave que identifique a los registros.**
- El tiempo para encontrar un registro es largo puesto que hay que buscar todos los registros para localizarlos.
- El tiempo que se emplea para añadir registros es mínimo ya que sólo es necesario agregarlo al final del archivo.

# Archivo Apilo

- Registros de longitud variable.
- Cada registro almacena estructura y datos.
- El orden de los campos en un registro no es necesariamente la misma.

# Archivo Apilo

Apariencia del archivo



# Operaciones en un Apilo

## a) Inserción:

- Se realiza de forma cronológica, el último registro se coloca al final del archivo.
- El archivo no tiene orden, cada registro no está identificado por una llave.
- Actualización del descriptor.

# Operaciones en un Apilo

- b) Búsqueda
  - Secuencial
  - La búsqueda arranca desde el inicio y deben compararse los campos hasta encontrar el registro o los registros que satisfacen los criterios de búsqueda.

# Operaciones en un Apilo

## c) Eliminación

- Implica la búsqueda del registro que se eliminará
- Se coloca una marca sobre el registro a eliminar, por lo cual la eliminación es lógica en lugar de física
- Actualización del descriptor.

# Operaciones en un Apilo

## d) Actualización

- Implica la búsqueda y eliminación del registro que contiene la información que se desea actualizar. Se realiza eliminación lógica del registro.
- Se inserta un nuevo registro con la información actualizada.
- Genera varios registros marcados como eliminados, debido a las actualizaciones.
- Actualización del descriptor.

# Operaciones en un Apilo

## e) Reorganización

- Se crea un archivo temporal.
- Se leen cronológicamente los registros del archivo físico y se trasladan al archivo temporal los registros NO marcados como eliminados.
- Se elimina el archivo físico y se renombra el archivo temporal, el cual pasa a ser el archivo físico.
- Deben actualizarse los datos del descriptor.

# Archivo Secuencial



# Archivo Secuencial

- Los registros de datos se guardan con una secuencia lógica, se guardan ordenados.
- Los registros de datos tienen longitud fija.
- Los registros de datos tienen los mismos atributos y todos en la misma posición; los nombres de los atributos sólo necesitan aparecer una vez en el descriptor y a cada atributo se le asocia un conjunto de datos que se denomina columna.

# Archivo Secuencial

Apariencia del archivo

Estructura de Atributos

Registros de Datos

Llave

DPI	Nombres	Apellidos
102083	Arturo	López
300310	Gabriela	Salazar
315555	Mario	López
383721	Dominic	Batres
821810	Graciela	Colindres
827274	Mariano	Gutiérrez
876659	Mario	López

# Archivo Secuencial

- Se define una llave para determinar la secuencia lógica de ordenamiento. No hay duplicados.
- Uno o más atributos se toman como la llave de cada registro y se denomina llave primaria.
- Algunas veces se agregan atributos para conformar una llave única de los registros.
- El archivo secuencial tiene dos componentes:
  - un archivo maestro y
  - la bitácora de transacciones que es un archivo tipo apilo.

# Archivo Secuencial

Bitácora de  
Transacciones(apilo)

DPI	Nombre	Apellido	Estatus
990377	Esteban	Morales	1
138309	Luisa	Keller	0
333377	Juan	Tello	1
212111	Patricia	Acevedo	1
...	...	...	

# Archivo Maestro

DPI	Nombre	Apellido	Estatus
102083	Arturo	López	1
300310	Gabriela	Salazar	1
315555	Mario	López	0
383721	Dominic	Batres	1
821810	Graciela	Colindres	0
827274	Mariano	Gutiérrez	1
876659	Mario	López	1



# Operaciones Archivo Secuencial

## a) Inserción:

- Los registros están limitados a un conjunto predeterminado de atributos.
- Debe verificarse que los datos a ingresar pertenezcan al dominio asociado a cada atributo.
- No pueden almacenarse registros de longitud mayor a la definida originalmente.
- Las inserciones se recopilan en la bitácora de transacciones (apilo), hasta que este archivo crece lo suficiente para realizar una actualización en lote del archivo maestro (una reorganización) y/o se reorganiza bajo demanda.
- Sólo se inserta un registro si se verifica que no existe la llave. Esto se hace para garantizar la unicidad de dicha llave.

## Operaciones Archivo Secuencial

### b) Búsqueda:

- Deben compararse los campos hasta encontrar el registro o los registros que satisfacen los criterios de búsqueda.
- La búsqueda puede incluir ambos archivos: el maestro y la bitácora de transacciones. La búsqueda en la bitácora puede realizarse antes de revisar el maestro y siguiendo la lógica del apilo (secuencial), o bien, pueden realizarse ambas búsquedas en paralelo.
- Algunos métodos específicos sólo buscan en el archivo maestro, lo cual implica que antes de iniciar la búsqueda debe hacerse previamente la reorganización.

# Operaciones Archivo Secuencial

b) Búsqueda (continuación):

- Puede ser secuencial o binaria.
  - Secuencial a través del archivo, es decir, registro por registro.
  - Binaria comienza con un acceso directo al centro del archivo, luego se segmenta en subsecuentes mitades el bloque donde se considere está la llave de búsqueda.

## Operaciones Archivo Secuencial

### c) Eliminación:

- Implica la búsqueda del registro que se eliminará.
- Se coloca una marca sobre el registro a eliminar en el archivo maestro o en la bitácora de transacciones, por lo cual la eliminación es lógica en lugar de física.
- Al realizarse la reorganización el registro marcado como inactivo, será físicamente eliminado.

## Operaciones Archivo Secuencial

### d) Actualización:

- Implica la búsqueda del registro que contiene la información que se desea actualizar.
- Si la actualización NO incluye la llave primaria, entonces sólo se actualizan datos, en el archivo maestro o en la bitácora de transacciones.
- Si la actualización SI incluye la llave primaria, se inserta un nuevo registro en la bitácora de transacciones y se elimina de forma lógica el registro antiguo, colocándole una marca, ya sea en el archivo maestro o la bitácora de transacciones.

# Operaciones Archivo Secuencial

## e) Reorganización:

- Consiste en intercalar ordenadamente el archivo maestro y la bitácora de transacciones, generando un nuevo archivo físico.
- Primero se ordena la bitácora de transacciones de acuerdo a las llaves de ordenamiento del archivo maestro.
- Se crea un archivo temporal.
- Se leen los registros no marcados del archivo maestro y de la bitácora de transacciones y se trasladan ordenadamente al archivo temporal.
- Se eliminan el archivo maestro y la bitácora de transacciones y se renombra el archivo temporal, el cual pasa a ser el archivo maestro. Se crea nuevamente la bitácora de transacciones (ahora vacía).
- Deben actualizarse los datos de los 2 descriptores (el de la bitácora y el del archivo maestro).

## Aplicaciones Archivo Secuencial

Procesamiento de datos orientado al manejo por lotes, cuando dicho procesamiento tiene cierta periodicidad.

Ejemplos:

- Asignaciones de un ciclo (propuestas - > Confirmadas).
- Agregar diagnósticos al historial médico de una persona.

# Gracias

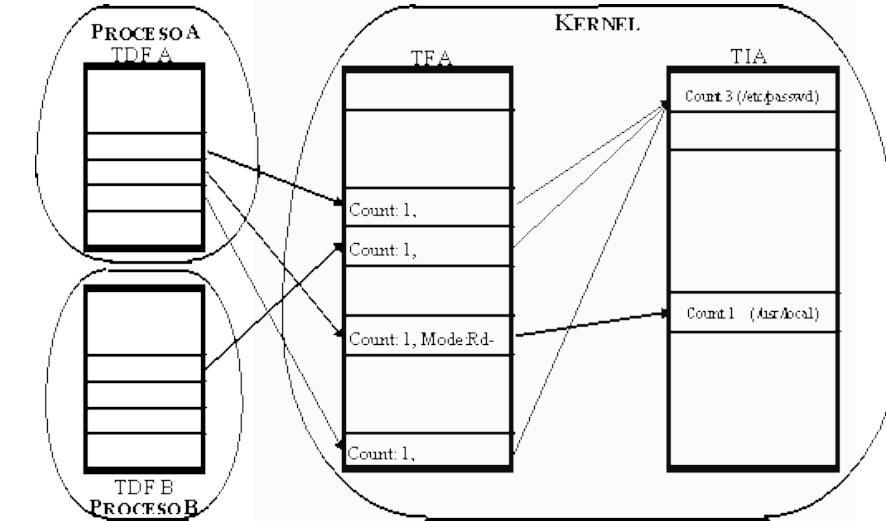
¿ALGUNA PREGUNTA?

# Manejo e Implementación de Archivos



Guatemala 17 de octubre  
de 2023

Ing. David Luna



Universidad  
Rafael Landívar  
Tradición Jesuita en Guatemala

# Agenda



**Secuencial**

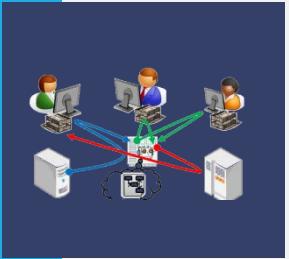


**Secuencial indizado**



**Índices**

# Índices



# Índices

## ¿Qué es?

- Estructura de datos que optimiza el acceso a los registros del archivo maestro.
- Debido a que almacena ciertos datos, suele caber en memoria principal y por ende su acceso es más rápido que acceder al archivo maestro.
- Debe almacenar, como mínimo, un campo para identificar al dato y la dirección física para localizarlo.
- Es responsabilidad del diseñador del sistema de archivos, definir de manera óptima los índices, su organización y su contenido.

# Índices

## Características generales:

- Los índices siempre deben estar clasificados (ordenados), según su atributo llave.
- Un archivo de datos puede tener asociado más de un índice.
- Un índice puede estar asociado a uno o más campos de un registro.

# Índices

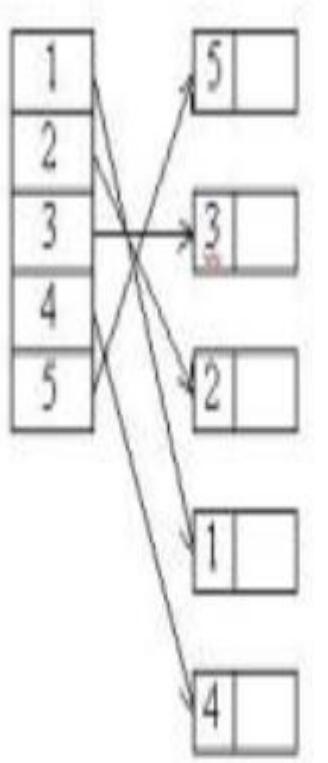
## Características generales:

- **Unicidad:** es la propiedad que exige que la llave de búsqueda sea única.
- **Duplicidad:** es la propiedad que permite tener valores repetidos para la llave de búsqueda.

# Clasificación

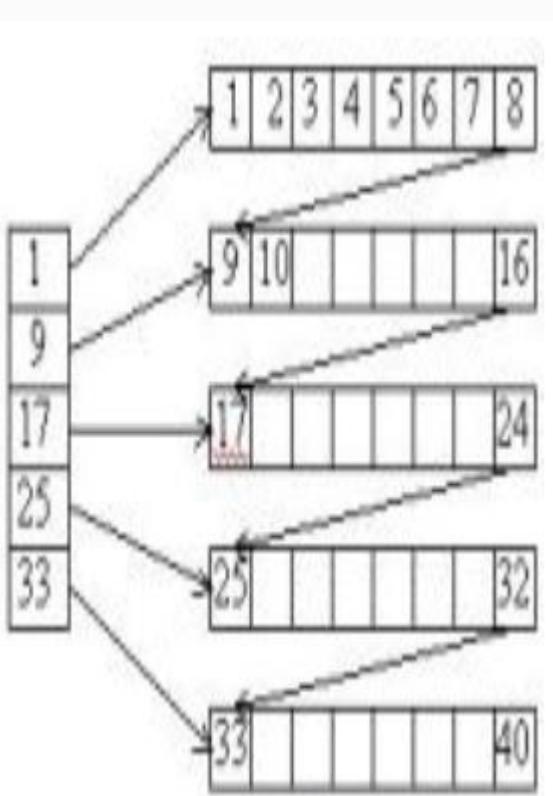
## Clasificación por densidad:

- Un índice denso o exhaustivo es el que asocia una entrada en el índice, por cada registro de datos.



## Clasificación por densidad:

- Un índice **disperso**, **escaso** o **selectivo** es el que no necesariamente asocia una entrada en el índice, por cada registro de datos. Pueden haber registros en el archivo de datos, que no tengan una correspondiente llave en el índice.



## Clasificación por frecuencia de actualización:

- Un índice en línea es el que se actualiza en sincronización con cada operación en el archivo de datos.
- Un índice en batch se usa cuando se almacena una bitácora de transacciones y en un momento determinado se lee dicha bitácora para actualizar el índice con todas las operaciones, es decir que la actualización del archivo de datos y el índice no se hace al mismo tiempo.

## Clasificación por tipo de llave:

- Un índice simple o primario

Es aquel que sólo registra la llave primaria, la cual identifica de forma única a cada registro.

El índice primario no permite la búsqueda de datos por campos que no sean la llave primaria.

## Clasificación por tipo de llave:

- **Un índice secundario**

Incluye más campos que la llave primaria y permite llaves duplicadas, lo cual es ventajoso para las búsquedas, pero puede reducir el desempeño durante las otras operaciones

-

# Secuencial Indizado

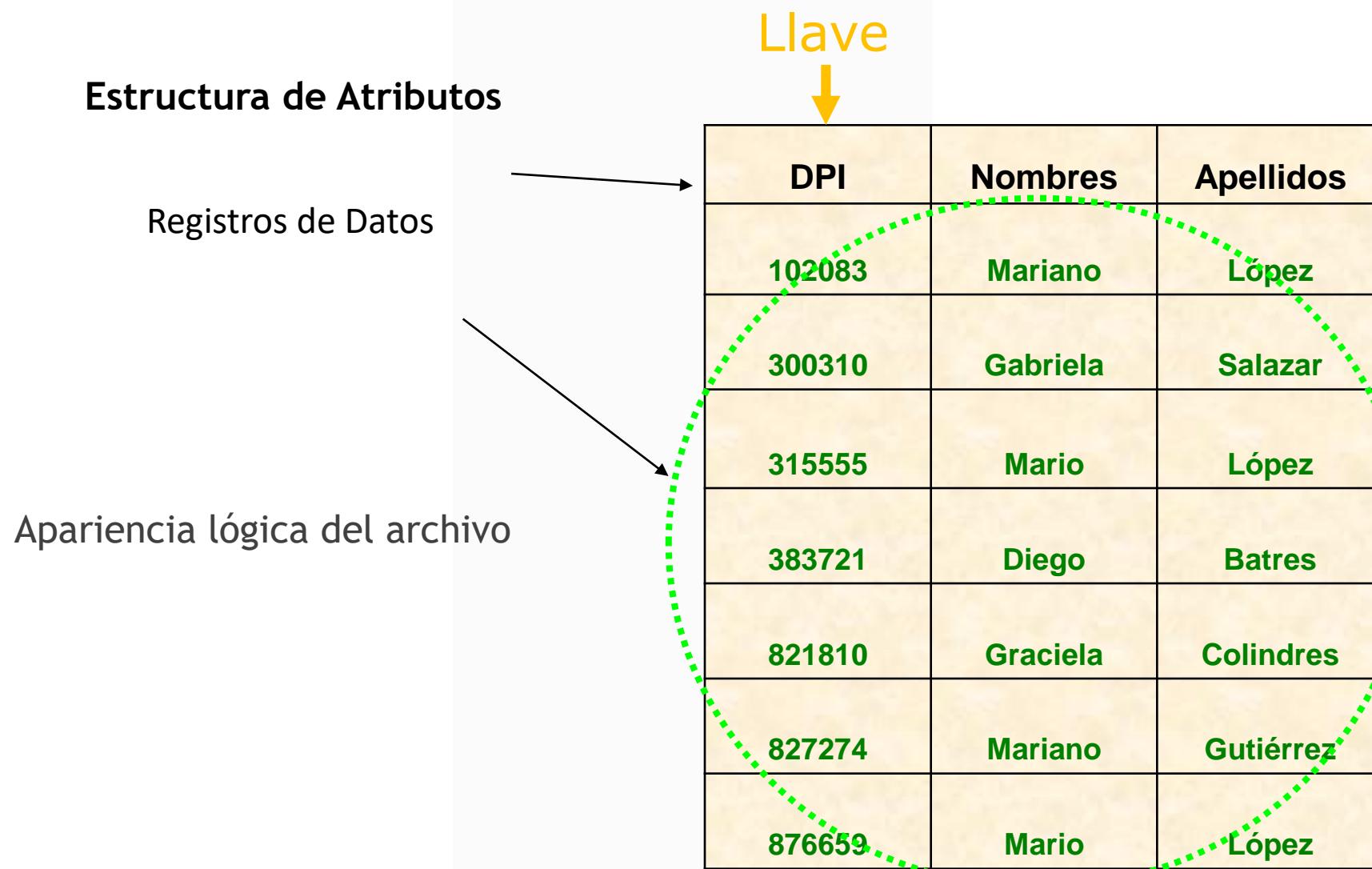


# Archivo Secuencial Indizado

Características de los Registros:

- Almacenados con una secuencia lógica.
- Longitud fija.
- Atributos en la misma posición, sólo necesitan aparecer una vez en el descriptor y a cada uno se le asocia un conjunto de datos que se denomina columna.

# Archivo Secuencial Indizado



# Archivo Secuencial Indizado

- Se define una llave para determinar la secuencia lógica de ordenamiento.
- Uno o más atributos se toman como la llave de cada registro y se denomina llave primaria.
- Si una llave primaria no es única, entonces pueden especificarse atributos llave secundarios, terciarios y así sucesivamente.

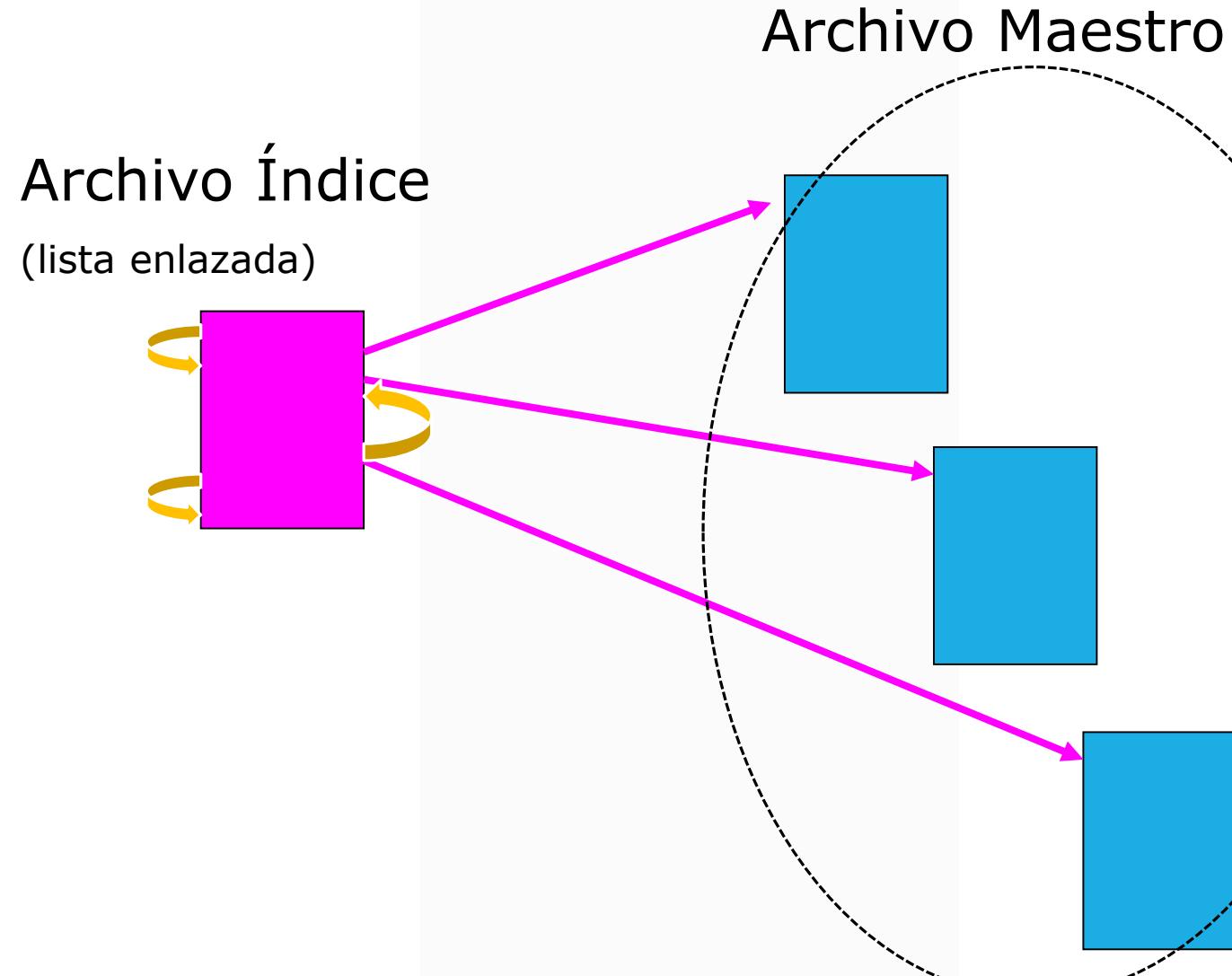
# Archivo Secuencial Indizado

Está compuesto por:



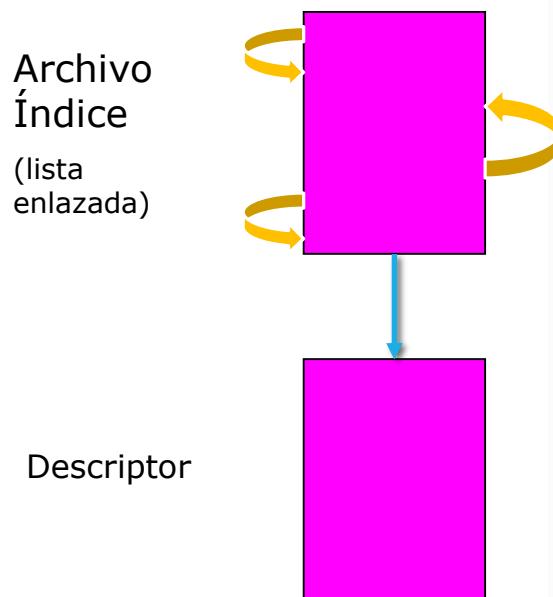
- **Objetivo:** Proveer acceso más rápido a los datos.
- Búsquedas a través de la llave.
- Puede estar compuesto por “n” bloques de longitud fija.
- Al crearse el archivo, se crea el primer bloque.
- Al llenarse un bloque, se crea uno nuevo.

# Archivo Secuencial Indizado



# Archivo Índice

- El archivo índice está formado por registros que simulan una lista enlazada; se define un descriptor para el archivo índice.



- Los registros del índice están enlazados.
- Los registros del índice contienen al menos: Registro, posición, llave, siguiente de la lista, estatus.
- La posición incluye: número de registro y número de bloque.

**Descriptor  
(índice)**  
Reg.Inicial: 2  
No.Bloques:1

**inicio**

Registro	Posición	Llave	Siguiente	Estatus
1	1.1	Alvarado Zapata	...	activo
2	1.2	Alvarado Mack	1	activo
3	...	...	...	...

## Maestro, bloque 1

**Descriptor  
(bloque 1,  
maestro)**  
No.Registros: 2  
Max.Reg: 20

	Apellidos	Nombres	Dirección	Teléfono	Estatus
1	Alvarado Zapata	Mónica María	Av. Bolívar 5-60	98737361	activo
2	Alvarado Mack	José Guillermo	1 calle 30-20	55543383	activo
	...	...	...	...	...

**Llave**

# Operaciones Archivo Secuencial Indizado

## a) Inserción:

- Los registros están limitados a un conjunto predeterminado de atributos.
- Debe verificarse que los datos a ingresar pertenezcan al dominio asociado a cada atributo.
- No pueden almacenarse registros de longitud mayor a la definida originalmente.
- **Las inserciones se hacen si no existe la llave.**
- Las inserciones se hacen en el archivo maestro, en el bloque actual (el último creado) . Se utiliza un nuevo bloque si ya se llegó al máximo de registros por bloque.
- Se agrega la llave en el índice y se actualiza el enlace de los registros (se enlazan generando el ordenamiento).
- Con cada inserción se revisa el primero de la lista.
- Se actualizan descriptores del índice y el maestro.

# Ejemplo

Se tendrá archivo secuencial indizado, que tiene las siguientes características:

Cada registro debe tener los siguientes campos:

- **No\_Carnet**
- **Nombre\_completo**
- **Correo\_Electrónico**
- **Dirección**
- **Teléfono**
- **Rol**
- **Contraseña**
- **Fecha\_Nacimiento**
- **Estatus**

La llave primaria es **No\_Carnet**



Los bloques no pueden tener mas de 4 registros

# Pasos

Identificar bloques



Identificar max para  
reorganizar



Insertar en el bloque  
que corresponde



Identificar inicio del  
índice



Insertar en índice



Actualizar punteros



Actualizar descriptores

# Operaciones Archivo Secuencial Indizado

## b) Búsqueda:

Si el criterio de búsqueda es la llave primaria, se utiliza el índice para ubicar la posición del registro y la búsqueda es directa.

Si el criterio de búsqueda NO incluye la llave primaria, se utiliza búsqueda secuencial en el archivo maestro.

# Operaciones Archivo Secuencial Indizado

## c) Eliminación:

Implica la búsqueda del registro que se eliminará.

Se cambia el estatus del registro a eliminar en ambos archivos y **se reordenan los enlaces en el archivo índice; la eliminación es lógica en lugar de física.**

Al realizarse la reorganización, los registros marcados serán físicamente eliminados.

# Operaciones Archivo Secuencial Indizado

## d) Actualización:

Implica la búsqueda del registro que contiene la información que se desea actualizar.

Si la actualización NO incluye la llave primaria, entonces sólo se actualizan datos, en el archivo maestro.

Si la actualización SI incluye la llave primaria, se inserta un nuevo registro y se elimina de forma lógica el registro antiguo, cambiándole el estatus.

# Operaciones Archivo Secuencial Indizado

## e) Reorganización:

Se crean archivos temporales: índice y maestro.

Se trasladan los registros no marcados de cada archivo y se trasladan al archivo temporal correspondiente: índice y maestro.

Se eliminan los archivos originales y se renombran los temporales.

Deben actualizarse los datos del descriptor.

# Asignación / Liberación de espacio



Se deben tener presentes problemas tales como la “fragmentación” creciente del espacio en disco:

- Ocasiona problemas de rendimiento al hacer que los archivos se desperdiguen a través de bloques muy dispersos.
- Técnicas para aliviar el problema de la “fragmentación” consisten en realizar periódicamente:
  - “Condensación”: se pueden “reorganizar” los archivos expresamente o automáticamente según algún criterio predefinido.

[http://upload.wikimedia.org/wikipedia/commons/4/40/Defragmenting\\_disk.gif](http://upload.wikimedia.org/wikipedia/commons/4/40/Defragmenting_disk.gif)

- “Recolección de basura o residuos” (garbage collector): se puede hacer fuera de línea o en línea, con el sistema activo, según la implementación.

# **Aplicaciones Archivo Secuencial Indizado**

Cuando es crucial la localización de la información.

# Ejemplo

Se tendrá archivo secuencial indizado, que tiene las siguientes características:

Cada registro debe tener los siguientes campos:

- **No\_Carnet**
- **Nombre\_completo**
- **Correo\_Electrónico**
- **Dirección**
- **Teléfono**
- **Rol**
- **Contraseña**
- **Fecha\_Nacimiento**
- **Estatus**

La llave primaria es **No\_Carnet**



Los bloques no pueden tener mas de 5 registros

# Gracias

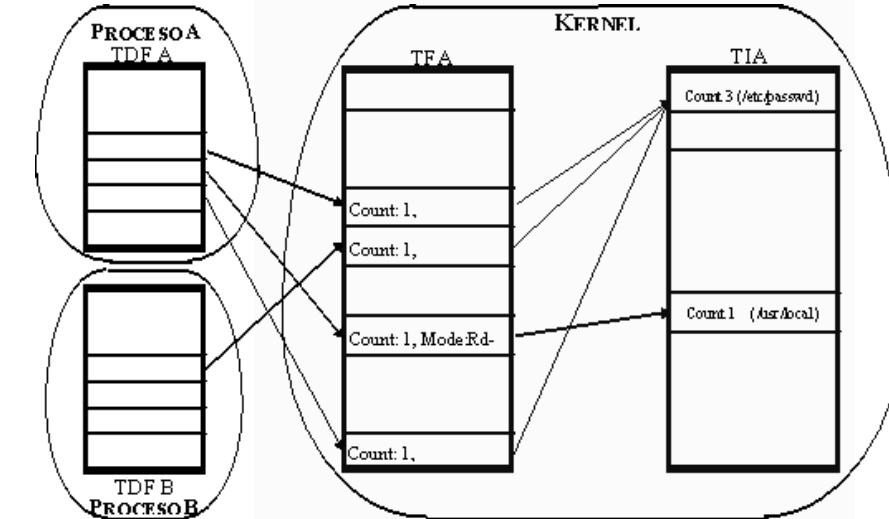
¿ALGUNA PREGUNTA?

# Manejo e Implementación de Archivos



Guatemala 31 de octubre  
de 2023

Ing. David Luna



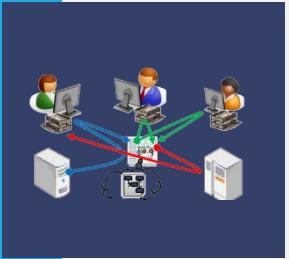
Universidad  
Rafael Landívar  
Tradición Jesuita en Guatemala

# Agenda



Índices

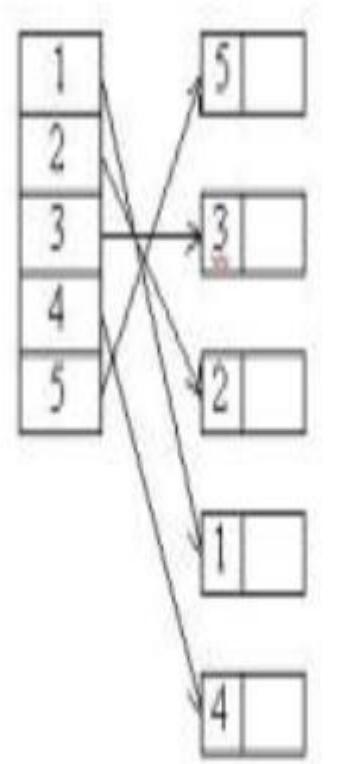
# Índices



# Clasificación

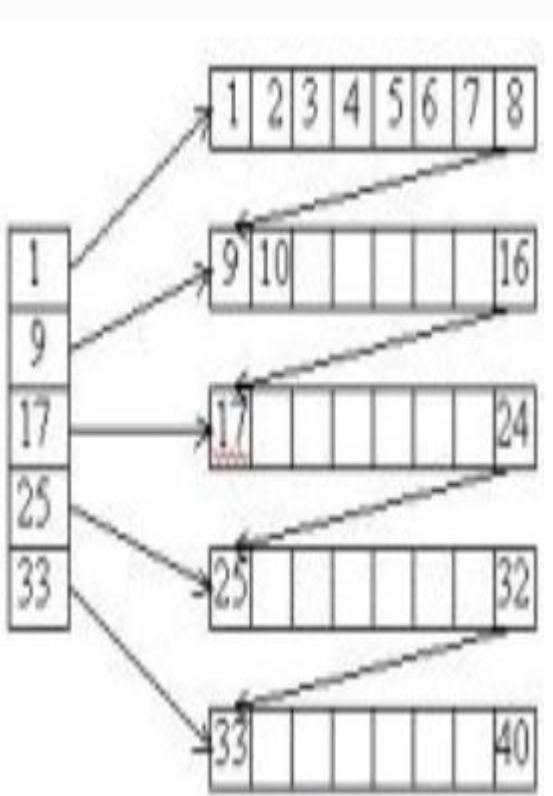
## Clasificación por densidad:

- Un índice denso o exhaustivo es el que asocia una entrada en el índice, por cada registro de datos.



## Clasificación por densidad:

- Un índice **disperso**, **escaso** o **selectivo** es el que no necesariamente asocia una entrada en el índice, por cada registro de datos. Pueden haber registros en el archivo de datos, que no tengan una correspondiente llave en el índice.



# Archivos Indizados con estructura de árbol



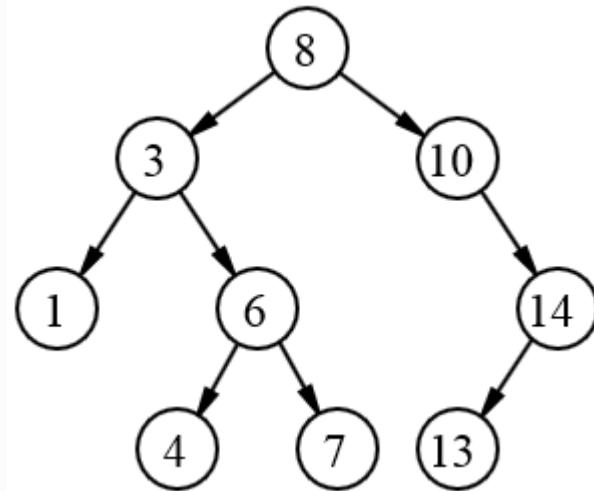
# **Archivo Indizado con estructura de árbol**

Árbol binario de búsqueda.

# Archivo Indizado con estructura de árbol

## Árbol binario de búsqueda

- La mayoría de los árboles binarios son de búsqueda
- El subárbol izquierdo de cualquier nodo (si no está vacío) contiene valores menores que el que contiene dicho nodo.
- el subárbol derecho (si no está vacío) contiene valores mayores.



# Archivo Indizado con estructura de árbol

## Árbol binario de búsqueda

Tamaño:

9

Profundidad:

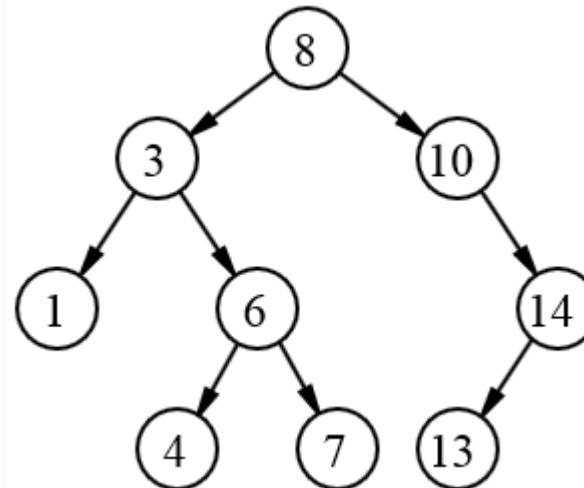
3

Raíz:

8

Hojas:

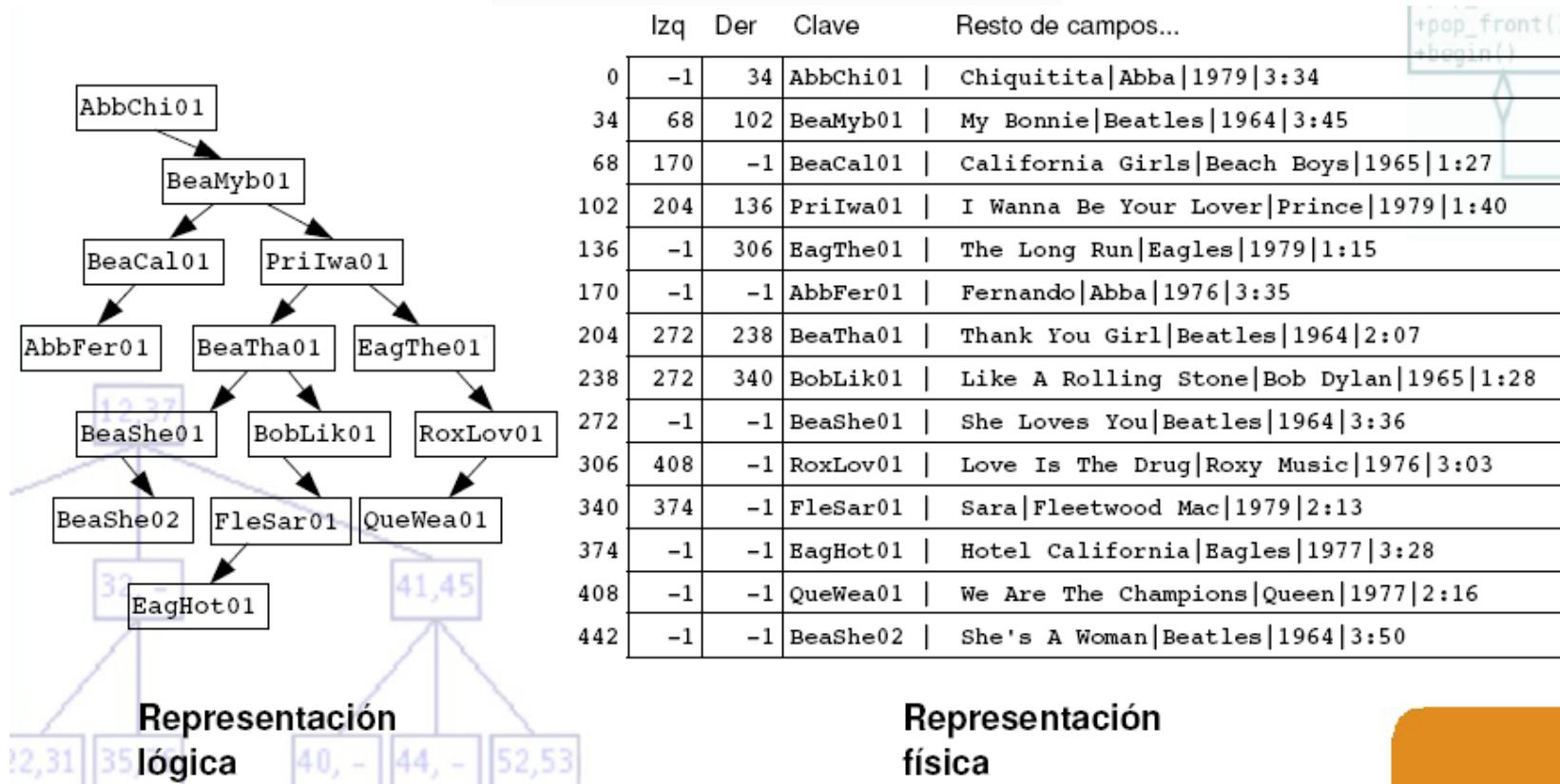
1, 4, 7 y 13



Un árbol binario de búsqueda de tamaño 9 y profundidad 3, con raíz 8 y hojas 1, 4, 7 y 13

# Archivo Indizado con estructura de árbol

Ordenamiento lógico de los datos, siguiendo una estructura de árbol binario de búsqueda.



# Archivos de árboles binarios

## Ventajas

- Una pequeña cabecera en el archivo indica la posición del nodo raíz del árbol.
- Los nodos hoja son aquellos cuyos índices izquierda y derecha apuntan a -1.
- Representación muy compacta, pues tanto los datos como los índices están almacenados en un único archivo.
- No es necesario ordenar físicamente los datos. Los datos se insertan al final y se enganchan al nodo hoja correspondiente.

# Archivos de árboles binarios

## Desventajas

- Desequilibrado del árbol tras varias inserciones y borrados.
- La búsqueda binaria no es suficientemente eficiente, se requieren muchos accesos para encontrar un dato.

# Ejemplo:

Se tiene el siguiente Archivo con estructura de árbol binario en donde se almacenarán los datos de los estudiantes de MEIA:

Izq.	Der.	Carnet	Nombre

Se insertan de acuerdo al listado del portal.

# Ejemplo:

Insertamos carnet 1299720 :

Rep. Lógica



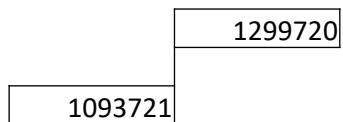
Rep. Física

Iz	Der	Carnet	Nombre
	-1	-1	1299720 AGUSTIN GABRIEL JULIO ANDRÉS

# Ejemplo:

Insertamos carnet 1093721:

Rep. Lógica



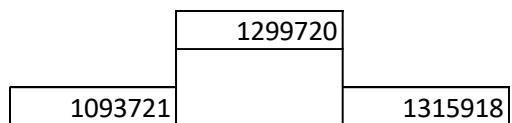
Rep. Física

Iz	Der	Carnet	Nombre
	-1	1299720	AGUSTIN GABRIEL JULIO ANDRÉS
2	-1	1093721	ARMAS RODRIGUEZ MARIO ROBERTO

# Ejemplo:

Insertamos carnet 1315918:

Rep. Lógica

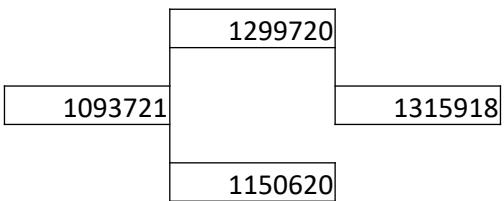


Rep. Física

Iz	Der	Carnet	Nombre
	2	3	2120316 ALVARADO ERICK DAVID
-1		-1	1018419 ALVAREZ MAZARIEGOS RAFAEL ANDRÉS
-1		-1	2530019 ARAGÓN LÓPEZ DIANA ALEJANDRA

# Ejemplo:

Insertamos carnet 1150620:  
Rep. Lógica

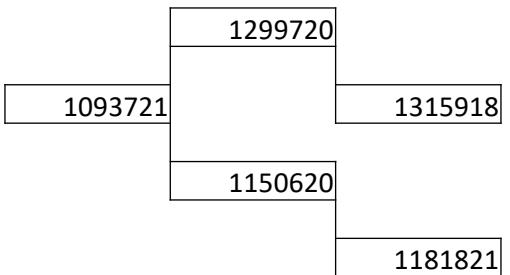


Rep. Física

Iz	Der	Carnet	Nombre
2	3	1299720	AGUSTIN GABRIEL JULIO ANDRÉS
-1	4	1093721	ARMAS RODRIGUEZ MARIO ROBERTO
-1	-1	1315918	AVALOS PÉREZ JOSÉ CARLOS
-1	-1	1150620	BARILLAS CONTRERAS EMILIO ANTONIO

# Ejemplo:

Insertamos carnet 1015920 :  
Rep. Lógica



Rep. Física

Iz	Der	Carnet	Nombre
2	3	1299720	AGUSTIN GABRIEL JULIO ANDRÉS
-1	4	1093721	ARMAS RODRIGUEZ MARIO ROBERTO
-1	-1	1315918	AVALOS PÉREZ JOSÉ CARLOS
-1	5	1150620	BARILLAS CONTRERAS EMILIO ANTONIO
-1	-1	1181821	BAUTISTA CRUZ DIEGO ANDRES

# Proyecto

¿?

# Gracias

¿ALGUNA PREGUNTA?

## **Resumen sobre Backups**

Es una copia de información importante, un duplicado de datos e información que dispone de una organización, esto con el objetivo de guardar y tener a mano esta información en casos de emergencia, donde se pierda la información o datos.

### **Tipos de copia de seguridad**

Existe una gran cantidad de tipos de copias de seguridad, que se diferencian por su manera de copiar los datos, velocidad y requerimientos de espacio.

- **Completa.**

Crea una copia completa de todos los datos de un sistema. Este puede ser una copia de los datos de un ordenador portátil o un ejemplo sería una copia de todas las nóminas digitales del año.

**VENTAJAS:** La restauración de una copia de seguridad completa es rápida

**DESVENTAJAS:** cada ejecución es lenta y ocupa más espacio con respecto a las otras tipologías.

- **Incremental.**

Primero se realiza una copia de seguridad completa y las siguientes copias incluyen únicamente los cambios realizados desde la última copia de seguridad.

**VENTAJAS:** Es mucho más rápida que una copia de seguridad completa y requiere menos espacio

**DESVENTAJAS:** la restauración es más lenta que con una copia de seguridad completa o diferencial.

- **Diferencial.**

Se realiza una copia de seguridad de todos los cambios realizados desde la última copia de seguridad completa.

**VENTAJAS:** Es mucho más rápida y requiere menos espacio de almacenamiento que una copia de

seguridad completa,

**DESVENTAJA:** Las restauraciones son más lentas que con una copia de seguridad completa, pero más rápidas que con copias de seguridad incrementales.

Además de la velocidad de realización y de restauración y la seguridad, el espacio es un requisito fundamental a la hora de planificar una estrategia de copias de seguridad

- **Espejo.**

Es un reflejo fiel de la fuente que se está respaldando, lo que implica que un archivo eliminado en el origen, también se eliminará en la copia de seguridad.

- **Sintética completa.**

Reconstruye la imagen de copia de seguridad completa usando todas las copias incrementales o diferenciales. Puede almacenarse en cintas en localizaciones externas, con la ventaja de que se reduce el tiempo de restauración.

- **Backup incremental inverso.**

Es una copia de seguridad incremental de los cambios realizados entre dos instancias de una copia espejo. Después de la copia completa inicial, cada copia sucesiva aplica los cambios a la anterior completa, creando una nueva copia de seguridad sintética completa cada vez, mientras se mantiene la capacidad de volver a las versiones anteriores

- **Protección de datos continua (CDP).**

es un proceso en el que cada cambio realizado en un documento, archivo o carpeta almacenado activa automáticamente una copia de seguridad de esa fuente de datos.

## Tipos de copias de seguridad según su destino

- **Locales**

cuando el medio de almacenamiento se mantiene a mano o en el mismo edificio que la fuente.

Puede tratarse de discos duros o unidades de almacenamiento conectado en red (NAS).

- **Externas**

cuando el medio de almacenamiento se mantiene en una ubicación geográfica diferente de la fuente (otra oficina, otro edificio o ubicaciones externas). De esta manera se consigue protección adicional contra robos, incendios, inundaciones y otros desastres naturales.

- **Remotas**

cuando, además de ser externas, es posible acceder, restaurar o administrar las copias de seguridad sin estar físicamente presente en la instalación de almacenamiento de respaldo.

- **En línea**

cuando se realizan en un medio de almacenamiento que siempre está conectado de forma segura a una red o conexión a Internet. Es un servicio ofrecido hoy en día por muchos centros de datos. Es también llamado copia de seguridad en la nube, si es proporcionado como un servicio en la nube.

## Dispositivos utilizados para copias de seguridad

### **Discos Duros Externos:**

Los discos duros externos son unidades de almacenamiento que se conectan a través de puertos USB o Thunderbolt a una computadora o servidor. Vienen en diferentes capacidades, desde gigabytes hasta varios terabytes.

Usuarios comunes: Tanto empresas como usuarios individuales utilizan discos duros externos. Son populares para copias de seguridad personales y para pequeñas empresas debido a su accesibilidad y costo asequible.

**Uso para backup:** Los discos duros externos son ideales para realizar copias de seguridad de datos importantes, ya que ofrecen una forma rápida y sencilla de hacerlo. Sin embargo, no son la mejor opción para la retención a largo plazo debido a la posibilidad de daños físicos o pérdida.

**Tipo de Backup Favorecido:** Copias de seguridad incrementales y diferenciales.

**Ventajas:**

Capacidad de almacenamiento significativa.

Acceso rápido a los datos.

Relativamente económicos.

**Desventajas:**

Vulnerables a daños físicos y fallas.

No ideales para la retención a largo plazo.

Pueden ser robados o extraviados.

### **Unidades de Cinta:**

Las unidades de cinta utilizan cintas magnéticas para almacenar datos. Están diseñadas para una retención a largo plazo y se utilizan principalmente en entornos empresariales.

**Usuarios comunes:** Las empresas grandes y organizaciones gubernamentales son los principales usuarios de unidades de cinta debido a su capacidad de almacenamiento masivo y confiabilidad.

**Uso para backup:** Las unidades de cinta son ideales para copias de seguridad empresariales y archivado de datos a largo plazo debido a su capacidad y resistencia a la degradación con el tiempo. Sin embargo, son menos adecuadas para copias de seguridad frecuentes debido a la velocidad de acceso más lenta.

**Tipo de Backup Favorecido:** Copias de seguridad completas

**Ventajas:**

Excelente para almacenamiento a largo plazo.

Gran capacidad de almacenamiento.

Menor costo por gigabyte en comparación con otros medios.

**Desventajas:**

Velocidad de acceso lenta.

Requiere hardware especializado.

No es adecuado para copias de seguridad frecuentes.

### **Dispositivos de Almacenamiento en Red (NAS):**

Descripción: Un NAS es un dispositivo de almacenamiento conectado a una red que permite el acceso compartido a datos y la configuración de copias de seguridad automáticas. Puede contener múltiples discos duros.

Usuarios comunes: Empresas de todos los tamaños y usuarios avanzados en el hogar que necesitan un almacenamiento centralizado y compartido.

Uso para backup: Los NAS son ideales para copias de seguridad empresariales y personales debido a su capacidad de acceso remoto, redundancia y escalabilidad.

Tipo de Backup Favorecido: diferenciales

Ventajas:

Acceso compartido a través de la red.

Redundancia de datos para mayor seguridad.

Escalabilidad fácil al agregar discos adicionales.

Desventajas:

Requiere configuración y mantenimiento.

Vulnerable a fallas de red y ataques.

Costoso en comparación con soluciones individuales.

### **Servicios de Almacenamiento en la Nube:**

Descripción: Los servicios de almacenamiento en la nube son plataformas en línea que permiten a los usuarios almacenar y acceder a sus datos a través de Internet. Los datos se almacenan en servidores remotos gestionados por proveedores de servicios en la nube.

Usuarios comunes: Tanto individuos como empresas utilizan servicios de almacenamiento en la nube. Son muy populares para respaldos personales y empresariales debido a su facilidad de uso y flexibilidad.

Uso para backup: Los servicios en la nube ofrecen copias de seguridad automáticas y programables, lo que los hace ideales para usuarios que desean acceso desde cualquier lugar y una mayor seguridad de los datos. También son útiles para colaboración en línea y recuperación ante desastres.

Tipo de Backup Favorecido: Copias de seguridad Incrementales y Diferenciales, pero también puede manejar Copias de seguridad Completas si es necesario.

Ventajas:

Acceso desde cualquier lugar con conexión a Internet.

Alta escalabilidad y flexibilidad.

Copias de seguridad automáticas y programables.

Desventajas:

Dependencia de proveedores externos.

Posibles problemas de seguridad y privacidad.

Costos recurrentes a largo plazo.

### **Unidades de Estado Sólido (SSD):**

Descripción: Las unidades de estado sólido utilizan memoria flash para almacenar datos y no tienen partes móviles. Son rápidas y resistentes a golpes.

Usuarios comunes: Principalmente, usuarios avanzados y empresas que necesitan alta velocidad de acceso a datos.

Uso para backup: Los SSD se utilizan para copias de seguridad que requieren velocidad y resistencia, como copias de seguridad de servidores o aplicaciones críticas. Sin embargo, debido a su costo por gigabyte más alto, a menudo se combinan con otros dispositivos de respaldo.

Tipo de backup favorecido: completa e incremental

Ventajas:

Velocidades de lectura y escritura extremadamente rápidas.

Resistentes a golpes y vibraciones.

Baja latencia.

Desventajas:

Costo por gigabyte más alto que otros dispositivos.

Vida útil limitada en comparación con algunos medios.

### **Dispositivos de Almacenamiento Óptico (por ejemplo, Discos Blu-ray):**

Descripción: Los discos ópticos utilizan láseres para grabar y leer datos en discos. Son resistentes a daños físicos y ofrecen una forma de archivar datos a largo plazo.

Usuarios comunes: Usuarios individuales y algunas empresas que desean archivar datos importantes de forma segura.

Uso para backup: Los discos ópticos son ideales para el almacenamiento a largo plazo y la protección contra daños físicos. Sin embargo, son menos adecuados para copias de seguridad frecuentes debido a su lentitud en la grabación y recuperación de datos.

Tipo de Backup Favorecido: Completa

Ventajas:

Resistencia a daños físicos y al agua.

Ideal para archivar datos importantes a largo plazo.

Portabilidad.

Desventajas:

Capacidad de almacenamiento limitada en comparación con otros medios.

Lentitud en la grabación y recuperación de datos.

Requiere hardware compatible.

### **Políticas de respaldo de información**

Las políticas de respaldo son un conjunto de normas con el objetivo de proteger la información contra una amplia gama de amenazas para asegurar la continuidad del servicio y minimizar los daños, procurando la preservación de la confidencialidad, disponibilidad e integridad de la información. Estas políticas son esenciales para prevenir la pérdida de datos debido a fallas tecnológicas, errores humanos, desastres naturales o ataques cibernéticos. Las políticas de respaldo de información pueden variar significativamente de una empresa a otra. Esto se debe a que las necesidades, recursos, objetivos y entornos tecnológicos de cada organización son diferentes. Algunos de los factores que influyen en las diferencias entre las políticas de respaldo de una empresa a otra incluyen:

- Tamaño y tipo de empresa: Las pequeñas empresas pueden tener políticas de respaldo más simples en comparación con las grandes corporaciones.
- Recursos disponibles: Las empresas con mayores recursos pueden invertir en tecnologías y soluciones de respaldo más avanzadas, mientras que las empresas más pequeñas pueden depender de soluciones más asequibles.
- Tecnología y sistemas utilizados: El entorno tecnológico de una empresa, incluyendo el tipo de sistemas operativos, bases de datos y aplicaciones utilizados, puede influir en la elección de las soluciones de respaldo.
- Riesgos y amenazas: Las empresas pueden enfrentar diferentes riesgos y amenazas, como ciberataques, desastres naturales o errores humanos, lo que puede influir en la estrategia de respaldo.
- Cambios en la infraestructura y la tecnología: Las empresas que adoptan nuevas tecnologías o cambian su infraestructura pueden necesitar ajustar sus políticas de respaldo en consecuencia.
- Criticidad de los datos: La importancia de los datos varía según la empresa y puede influir en la frecuencia y los métodos de respaldo. Por ejemplo, una empresa de servicios financieros puede considerar sus datos financieros como críticos y respaldarlos con mayor frecuencia.

Dado que las necesidades y circunstancias varían, es fundamental que cada empresa desarrolle una política de respaldo de información que se adapte a sus características específicas. Sin embargo, todas las empresas deben asegurarse de que sus políticas de respaldo cumplan con los estándares de seguridad y buenas prácticas de gestión de datos para proteger adecuadamente sus activos digitales. Algunas de las políticas que se pueden encontrar en varias empresas son:

**1.COBIT:** Es un conjunto de reglas y guías para que las empresas administren su tecnología de manera efectiva. Ayuda a asegurarse de que las personas obtengan lo que necesitan de la tecnología y de que se use de manera segura.

- **Definir una política de respaldo:** Debería haber una política clara que establezca los requisitos de respaldo para la organización, incluyendo la frecuencia de respaldos, la retención de datos, la ubicación de los respaldos, etc.
- **Documentar procedimientos de respaldo:** Los procedimientos para realizar respaldos, así como la recuperación de datos, deben documentarse y seguirse de manera consistente.
- **Automatizar cuando sea posible:** Se debe utilizar la automatización para programar respaldos de manera regular y garantizar su cumplimiento.
- **Proteger los respaldos:** Los datos de respaldo son valiosos y deben protegerse adecuadamente contra el acceso no autorizado y los desastres, como incendios o inundaciones.
- **Realizar pruebas de recuperación:** Deben realizarse pruebas periódicas para garantizar que los respaldos sean recuperables y que los tiempos de recuperación sean aceptables.
- **Gestionar la retención de datos:** COBIT sugiere que las organizaciones establezcan políticas claras de retención de datos para determinar cuánto tiempo deben mantenerse los respaldos y cuándo deben ser eliminados de manera segura.
- **Revisión y mejora continua:** Los procesos de respaldo deben revisarse regularmente y mejorarse según sea necesario para mantenerse alineados con los objetivos de la organización y los cambios en el entorno de TI.

**2.ISO/IEC 27001:** La norma ISO 27001 es una norma internacional que establece los requisitos para establecer, implementar, mantener y mejorar un sistema de gestión de seguridad de la información (SGSI) en una organización. Esta norma se utiliza para ayudar a las organizaciones a proteger sus activos de información y garantizar la confidencialidad, integridad y disponibilidad de la información.

- **Alcance del SGSI para Copias de Seguridad:** La organización debe definir claramente el alcance del SGSI para incluir la gestión de copias de seguridad. Esto implica identificar los activos de información relevantes que requieren respaldo y las partes interesadas relacionadas.
- **Política de Seguridad de la Información para Copias de Seguridad:** La organización debe establecer una política de seguridad de la información que incluya objetivos y compromisos específicos relacionados con la gestión de copias de seguridad. Esto asegura que se aborden los

aspectos críticos, como la frecuencia de las copias, los procedimientos de recuperación y la retención de datos.

- **Análisis de Riesgos para Copias de Seguridad:** Se debe realizar una evaluación de riesgos enfocada en las amenazas y vulnerabilidades que pueden afectar la integridad y disponibilidad de las copias de seguridad. Esto incluye considerar posibles fallos en el proceso de respaldo y ataques ciberneticos dirigidos a los datos de respaldo.
- **Tratamiento de Riesgos para Copias de Seguridad:** Un plan de tratamiento de riesgos debe ser desarrollado para las copias de seguridad. Esto implica la implementación de controles adecuados, como cifrado de datos de respaldo y sistemas de autenticación, para mitigar o aceptar los riesgos identificados.
- **Cumplimiento Legal y Regulatorio para Copias de Seguridad:** La organización debe identificar y cumplir con todos los requisitos legales y regulaciones relacionados con la gestión de copias de seguridad, como las leyes de privacidad de datos y retención de registros.
- **Objetivos y Controles de Seguridad para Copias de Seguridad:** Establecer objetivos específicos para la gestión de copias de seguridad, tales como tiempos de recuperación, y aplicar controles técnicos y procedimentales para alcanzar esos objetivos.
- **Procesos de Gestión de la Seguridad para Copias de Seguridad:** La norma ISO 27001 incluye procesos de gestión como la planificación de respaldos, monitoreo de copias y recuperación ante desastres, que son fundamentales para garantizar la seguridad de las copias de seguridad.
- **Auditoría y Revisión para Copias de Seguridad:** Realizar auditorías internas periódicas para evaluar el cumplimiento de los requisitos y revisar la eficacia de los procedimientos de copias de seguridad y recuperación.
- **Mejora Continua para Copias de Seguridad:** La organización debe buscar constantemente la mejora en la gestión de copias de seguridad, corrigiendo deficiencias, implementando acciones correctivas y preventivas, y revisando regularmente los resultados.
- **Documentación para Copias de Seguridad:** Mantener documentación adecuada que respalte el SGSI en relación con las copias de seguridad, incluyendo políticas, procedimientos y registros de respaldo.

**3.NIST:** Es un instituto que establece estándares para la seguridad de la información y la tecnología.

Ofrece pautas para proteger sistemas informáticos, contraseñas y enseña sobre seguridad en línea.

Algunas de estas son:

- **Copia de seguridad (Backup):** La copia de seguridad se refiere al proceso de hacer copias de los datos y la información crítica de una organización para su preservación y recuperación en caso de pérdida, daño o eliminación accidental. Las copias de seguridad se utilizan para garantizar la disponibilidad y la integridad de los datos.
- **Restauración (Restore):** La restauración es el proceso de recuperar datos y sistemas desde las copias de seguridad almacenadas. Cuando ocurre una pérdida de datos o un incidente, la restauración se utiliza para devolver los sistemas y datos a un estado operativo normal.

- **Política de Copias de Seguridad:** Una política de copias de seguridad es un conjunto de directrices y procedimientos que una organización establece para gestionar las copias de seguridad de sus datos. Esto incluye la frecuencia de las copias de seguridad, la retención de datos, la ubicación de almacenamiento y otros aspectos relacionados con la gestión de la información respaldada.
- **Copia de Seguridad Completa:** Una copia de seguridad completa implica la copia de todos los datos y archivos en un sistema o dispositivo en un momento dado. Estas copias de seguridad son útiles para restaurar completamente un sistema, pero pueden ser intensivas en tiempo y recursos.
- **Copia de Seguridad Incremental:** Una copia de seguridad incremental solo copia los datos que han cambiado desde la última copia de seguridad, lo que ahorra espacio de almacenamiento y tiempo de respaldo. Para restaurar completamente un sistema, es necesario combinar múltiples copias de seguridad incrementales.
- **Copia de Seguridad Diferencial:** Una copia de seguridad diferencial copia todos los datos que han cambiado desde la última copia de seguridad completa. Esto significa que solo se necesitan dos conjuntos de copias de seguridad (la completa y la diferencial más reciente) para restaurar completamente un sistema.
- **Plan de Continuidad del Negocio:** Un plan de continuidad del negocio incluye estrategias y procedimientos para garantizar la continuidad de las operaciones de una organización en caso de interrupciones, desastres o incidentes que afecten la disponibilidad de datos y sistemas. Las copias de seguridad y la restauración son componentes clave de estos planes.
- **Pruebas de Restauración:** Las pruebas de restauración son pruebas regulares que se realizan para verificar que las copias de seguridad sean efectivas y que los datos se puedan restaurar correctamente en caso de necesidad. Estas pruebas son esenciales para garantizar la capacidad de recuperación de una organización.

**4.EITIL:** Es una biblioteca de mejores prácticas para administrar servicios de tecnología. Ayuda a las empresas a gestionar y entregar servicios tecnológicos de manera organizada y eficiente, enfocándose en la satisfacción del cliente y la mejora continua.

- **Enfocarse en el valor:** Asegúrate de que los backups se realicen de manera que agreguen valor real a la organización al garantizar la disponibilidad y la recuperación de datos críticos en caso de pérdida o desastre.
- **Empezar desde donde ya estamos:** Utiliza las infraestructuras y sistemas de backup existentes para mejorar y optimizar la gestión de respaldo en lugar de comenzar desde cero.
- **Avanzar iterativamente y retroalimentarse:** Implementa mejoras continuas en tus procesos de respaldo. Realiza revisiones periódicas de la efectividad de los backups y ajusta tus estrategias en función de los resultados y los comentarios recibidos.
- **Colaborar y promover la visibilidad:** Trabaja en colaboración con los equipos de TI y los usuarios finales para comprender sus necesidades de respaldo. Asegúrate de que los procedimientos y políticas de respaldo sean transparentes y comprensibles para todos los involucrados.
- **Visión Holística:** Considera todos los aspectos de la gestión de backups, desde la planificación y la implementación hasta la monitorización y la recuperación. Asegúrate de que los backups se ajusten a la estrategia general de gestión de servicios de TI de la organización.

- **Mantenerlo simple y práctico:** Diseña y administra tus políticas y procedimientos de respaldo de manera que sean simples y fáciles de seguir. Evita la complejidad innecesaria que pueda dificultar la recuperación de datos.
- **Optimizar y automatizar:** Utiliza herramientas y soluciones de backup automatizadas para optimizar la eficiencia de tus procesos de respaldo. Esto puede ayudar a reducir errores humanos y garantizar la consistencia en las copias de seguridad.

## **CASOS DE ÉXITO**

### **T-Mobile 2020**

T Mobile sufrió una pérdida de datos debido al error de un empleado el cual por accidente eliminó el directorio raíz causando así la completa eliminación de los datos de producción, esto causó interrupción del servicio y fallos en la red de los usuarios. Al tener una buena implementación de backup se lograron recuperar los datos.

### **Play Station Network 2011**

Play Station durante 2011 sufrió un ataque de Ransomware debido a una brecha de seguridad, pero al ser demasiada información no era enviada a los atacantes por lo cual se pudo restaurar la información eliminando la información actual, las consecuencias de este inconveniente fueron la interrupción de servicios, para enmendar su problema Play Station Network regaló 2 juegos de su plataforma.

## **CASOS DE FRACASO**

### **British Airways 2017**

British Airways durante 2017 sufrió un accidente debido a una falla en la línea eléctrica lo cual causó una eliminación en la información principal y el backup se encontraba realizándose en ese momento quedando comprometido también, como resultado no se pudo recuperar la información.

### **Gitlab 2017**

Gitlab en 2017 debido a un error por parte de uno de los empleados el cual eliminó el directorio raíz borrando así toda la información, como ultimo recurso les quedaba recurrir a la copia de seguridad, pero para su mala suerte el equipo encargado del backup no había verificado que el backup se haya realizado correctamente, por lo cual se perdieron todos los datos.

## **CASO DE ÉXITO Y FRACASO**

### **Pixar 1999**

Durante la producción de Toy Story 2 una de las encargadas de modelado 3D de personajes se encontraba en proceso de embarazo el equipo le preparó una computadora con la copia entera del proyecto para que pudiera trabajar desde su casa, un día de producción uno de los empleados ejecutó un comando en la terminal lo cual hizo que borrara todos los archivos del proyecto, en ese momento no se tenía el backup actualizado y la única copia con el trabajo más próximo a la fecha era el de la compañera que estaba embarazada, en ese momento la computadora que le prepararon paso a valer mil millones de dólares y luego de recuperar la información se continuó con normalidad, decimos que es un caso de éxito porque se logró recuperar la información pero de fracaso porque no estaban bien implementado el respaldo de información.

## **Implementación de respaldo de la información**

se refiere al proceso de poner en práctica una estrategia de copias de seguridad (backups) para proteger y respaldar los datos críticos de una organización. Implica la configuración y ejecución de los procedimientos y sistemas necesarios para asegurar que los datos se copien, almacenen y puedan recuperarse de manera eficiente y segura en caso de pérdida, corrupción, fallos del sistema, ataques ciberneticos u otros incidentes.

A continuación, se detallan algunos aspectos clave de la implementación de respaldo de la información:

- 1. Selección de tecnologías y herramientas:** Esto implica elegir el hardware y el software adecuados para realizar copias de seguridad. Esto puede incluir la elección de dispositivos de almacenamiento como discos duros externos, cintas magnéticas, servidores de almacenamiento en red (NAS) o soluciones de copia de seguridad en la nube. También se seleccionan aplicaciones y software de copia de seguridad que se ajusten a las necesidades de la organización.
- 2. Políticas de respaldo:** Se definen políticas específicas que establecen qué datos se deben respaldar, con qué frecuencia se deben realizar las copias de seguridad y cuánto tiempo se deben retener. Estas políticas deben estar alineadas con las necesidades de la organización y las regulaciones aplicables.
- 3. Programación de copias de seguridad:** Se establece un horario para la realización de las copias de seguridad. Esto puede incluir copias de seguridad diarias, semanales o mensuales, según los requisitos de la organización y la importancia de los datos.
- 4. Seguridad de datos:** Se implementan medidas de seguridad para proteger los datos de copia de seguridad contra el acceso no autorizado. Esto puede incluir la encriptación de datos de copia de seguridad y el control de acceso a los sistemas de copias de seguridad.
- 5. Pruebas y verificación:** Se realizan pruebas periódicas para asegurarse de que las copias de seguridad se estén realizando correctamente y que los datos se puedan recuperar según lo planeado en caso de necesidad.

**6. Monitoreo y gestión:** Se establecen procedimientos para supervisar el estado de las copias de seguridad, detectar problemas y tomar medidas correctivas si es necesario. Esto puede incluir la supervisión de registros de copias de seguridad y la implementación de alertas.

**7. Capacitación del personal:** Se capacita al personal responsable de la gestión de las copias de seguridad para que comprendan los procedimientos y las políticas, y puedan responder adecuadamente en caso de incidentes.

**Enlaces de los simuladores:**

Simulador 1

[http://lasdpc.icmc.usp.br/~ssc640/grad/ec2015/backup\\_simulator/](http://lasdpc.icmc.usp.br/~ssc640/grad/ec2015/backup_simulator/)

Simulador 2

<https://pbs.proxmox.com/docs/prune-simulator/>

Simulador 3

<https://calculator.veeam.com/vbr/>

**Manejo e implementación de archivos**

# **Grupo 3: Cloud Storage**

Diego Cosillo - 1136222

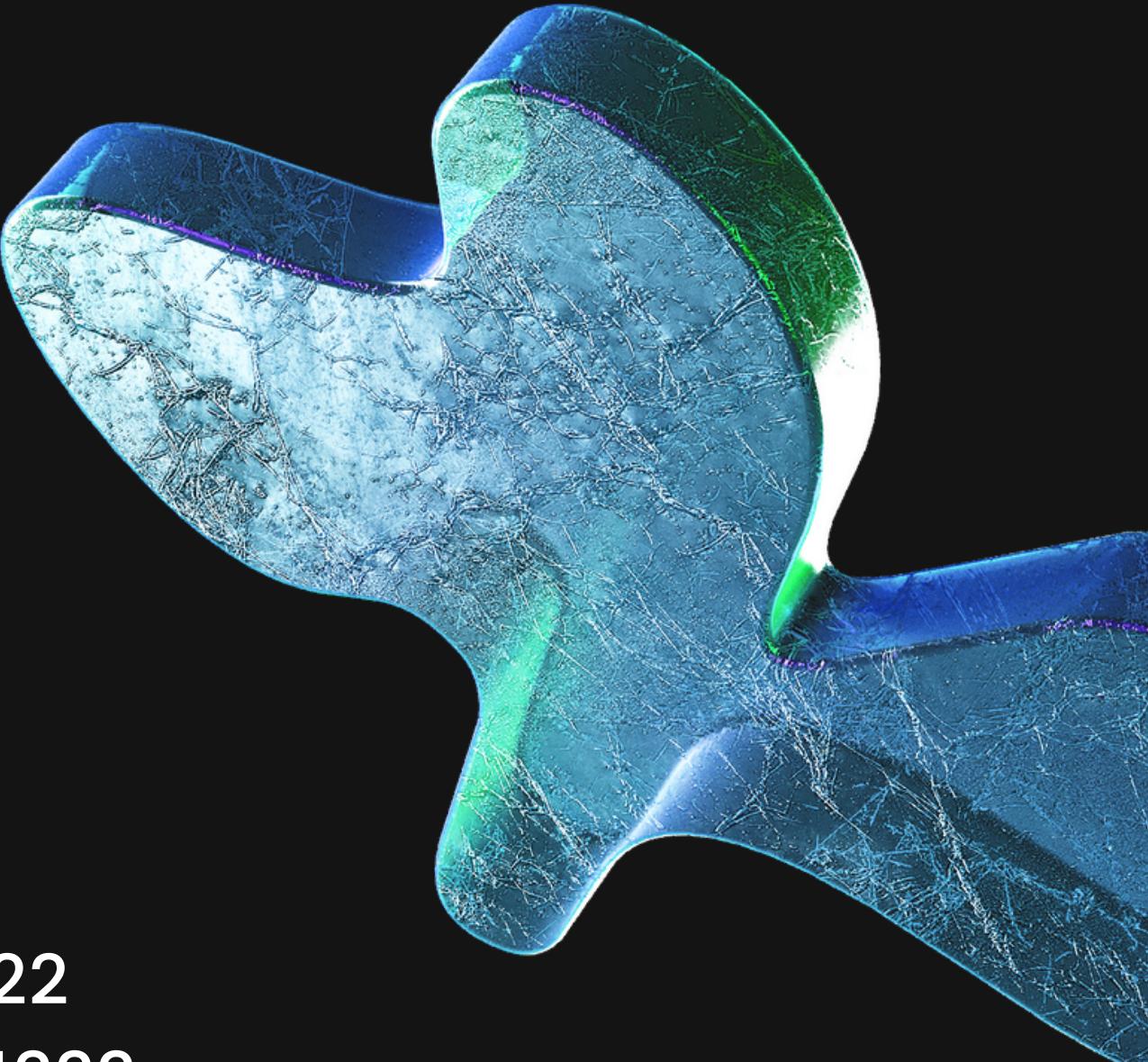
Diego Azurdia - 1010821

José Alvarado - 1047222

Javier Godínez - 1179222

Katherine Mayen - 1129222

Sebastian Cuevas - 1034222



# ¿Qué es Cloud Storage?



El almacenamiento en la nube es un modelo de depósito de datos en el que la información digital, como documentos, fotografías, videos y otras formas de medios, se almacena en servidores virtuales o en la nube alojados por terceros. Le permite transferir datos a un sistema de almacenamiento externo y acceder a ellos cuando sea necesario.

# ¿Cómo funciona?

El almacenamiento en la nube funciona como un centro de datos virtual. Ofrece a los usuarios finales y a las aplicaciones una infraestructura de almacenamiento virtual que se puede ampliar según los requisitos de la aplicación.

Generalmente opera a través de una API basada en web implementada de forma remota a través de su interacción con la infraestructura interna de almacenamiento en la nube.

El almacenamiento en la nube incluye al menos un servidor de datos al que un usuario puede conectarse a través de Internet.



**1**

El usuario envía archivos al servidor de datos, que reenvía el mensaje a múltiples servidores, de forma manual o automatizada, a través de Internet.

**2**

Luego se puede acceder a los datos almacenados a través de una interfaz basada en web.

**3**

Para garantizar la disponibilidad constante de los datos, los sistemas de almacenamiento en la nube implican una gran cantidad de servidores de datos.

Por lo tanto, si un servidor requiere mantenimiento o falla, el usuario puede estar seguro de que los datos se han movido a otra parte para garantizar la disponibilidad.



# Tipos de Cloud Storage

Privada

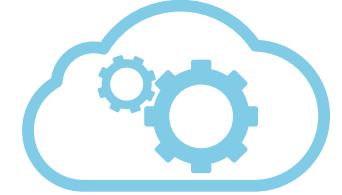
---

Pública

---

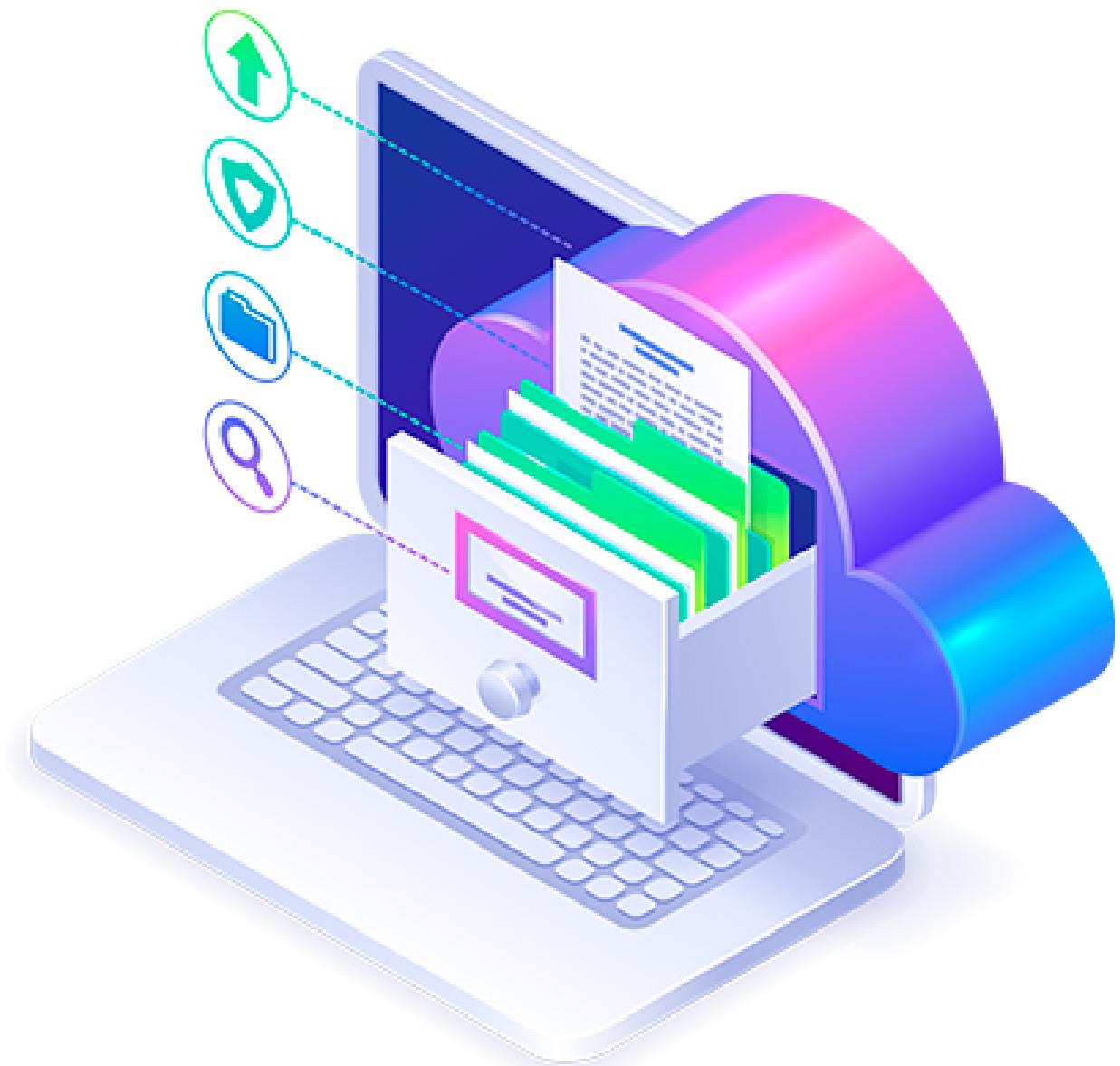
Híbrida

---



# Privada

- El almacenamiento en la nube privada también se conoce como almacenamiento en la nube interna o empresarial. En este caso, los datos se almacenan en la intranet de la empresa u organización. Estos datos están protegidos por el propio firewall de la empresa.
- El almacenamiento en la nube privada es una excelente opción para empresas con centros de datos costosos y que pueden administrar la privacidad de los datos internamente.
- Una de las principales ventajas de guardar datos en una nube privada es que ofrece un control total al usuario.
- Uno de los mayores inconvenientes del almacenamiento en la nube privada es el costo y el esfuerzo de mantenimiento y actualizaciones. La responsabilidad de gestionar el almacenamiento en la nube privada recae en la empresa anfitriona.





# Pública

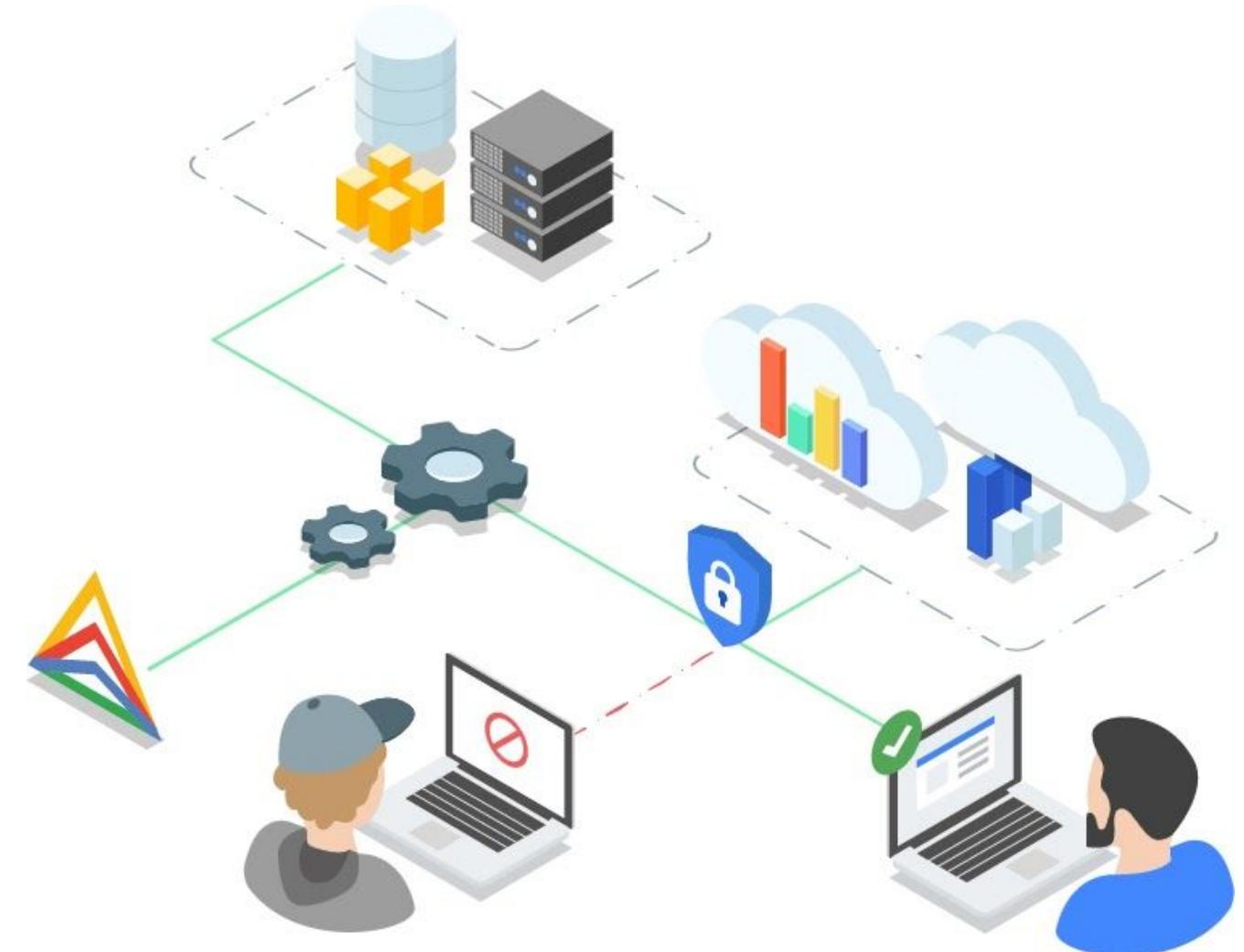
- El almacenamiento en la nube pública requiere pocos controles administrativos y el usuario y cualquier otra persona que el usuario autorice puede acceder en línea.
- Con el almacenamiento en la nube pública, el usuario/empresa no necesita mantener el sistema.
- El almacenamiento en la nube pública está alojado en diferentes proveedores de soluciones, por lo que hay muy pocas posibilidades de personalizar los campos de seguridad, ya que son comunes para todos los usuarios. Amazon Web Services (Aws), IBM Cloud, Google Cloud y Microsoft Azure son algunos proveedores populares de soluciones de almacenamiento en la nube pública.
- El almacenamiento en la nube pública es fácilmente escalable, asequible, confiable y ofrece un monitoreo perfecto y cero mantenimiento.





# Híbrida

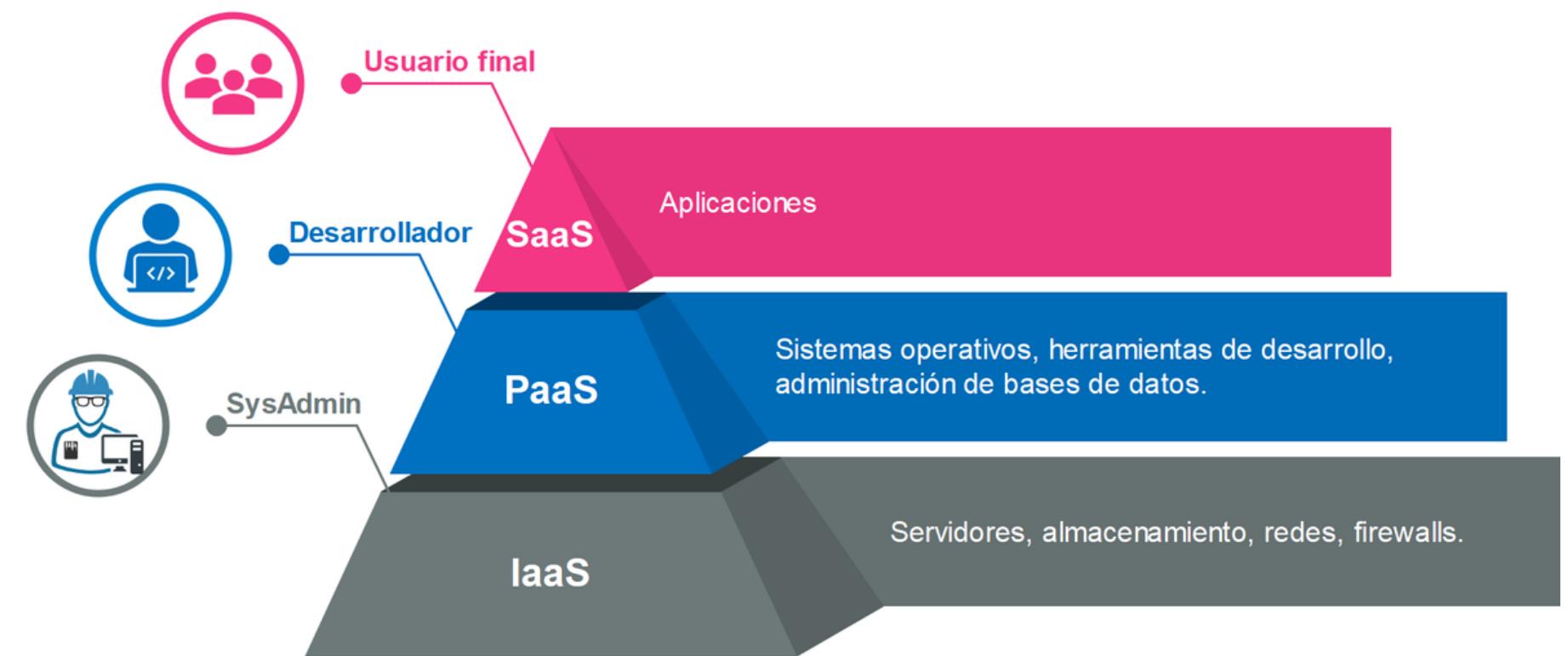
- El almacenamiento en la nube híbrida es una combinación de almacenamiento en la nube pública y privada. Este tipo ofrece al usuario lo mejor de ambos mundos: la seguridad de una nube privada y la personalización de una nube pública.
- En una nube híbrida, los datos se pueden almacenar en la nube privada y las tareas de procesamiento de información también se pueden asignar a la nube pública, con la ayuda de servicios de computación en la nube.
- El almacenamiento en la nube híbrida es accesible y ofrece una fácil personalización y un mayor control del usuario.

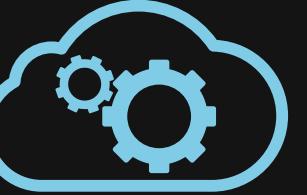




# Servicios

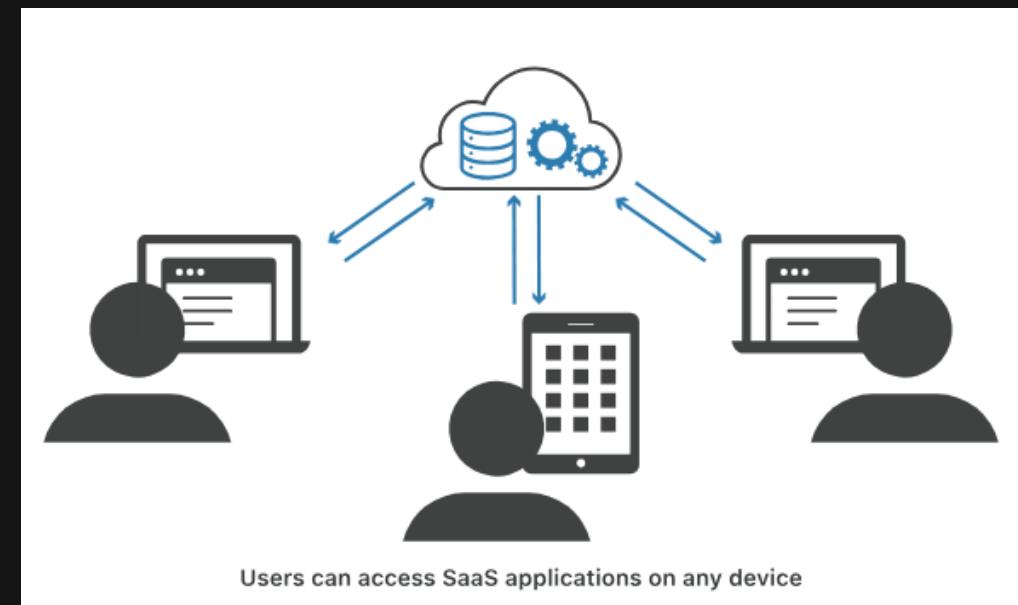
La elección del modelo de servicio en la nube es una decisión estratégica que debe basarse en las necesidades específicas de cada empresa. No hay una talla única para todos en la nube, y las organizaciones deben evaluar cuidadosamente sus objetivos, recursos y requisitos para tomar decisiones informadas sobre cómo adoptar la computación en la nube. Ya sea externalizando solo ciertas partes de la infraestructura o combinando varios modelos de servicio, la nube ofrece flexibilidad para adaptarse a las circunstancias únicas de cada empresa.





# 1 SaaS (Software as a Service)

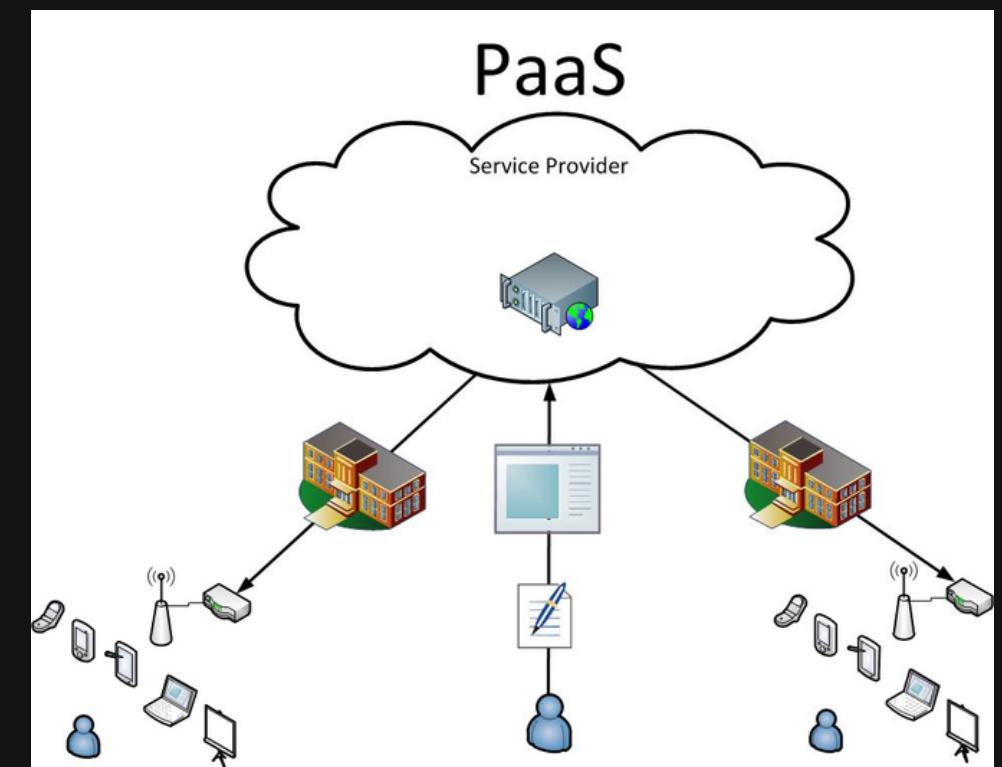
- **SaaS facilita la colaboración entre equipos y el acceso a las aplicaciones de software.**
- **Definición:** SaaS, que significa "Software como Servicio", es un modelo de servicio en la nube que proporciona acceso a aplicaciones y software a través de Internet sin necesidad de instalar ni mantener el software localmente.
- **Características Principales:**
  - Accesibilidad desde Cualquier Lugar: Los usuarios pueden acceder a las aplicaciones SaaS desde cualquier dispositivo con conexión a Internet, lo que facilita la colaboración y el trabajo remoto.
  - Actualizaciones Automáticas: El proveedor de SaaS se encarga de las actualizaciones de software y el mantenimiento, lo que libera a los usuarios de la carga de gestionar y mantener el software.
  - Pago por Uso: Los usuarios suelen pagar por suscripción o según el uso, lo que permite un modelo de costos más predecible.





## 2 PaaS (Platform as a Service)

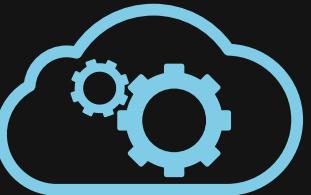
- PaaS hace que el proceso de desarrollo y despliegue de aplicaciones sea mucho más simple, eficiente y competitivo.
- Definición: PaaS, que significa "Plataforma como Servicio", es un modelo de servicio en la nube que ofrece un entorno de desarrollo y alojamiento para aplicaciones, eliminando la necesidad de preocuparse por la infraestructura subyacente.
- Características Principales:
  - Herramientas de Desarrollo: PaaS proporciona herramientas y servicios que simplifican el desarrollo de aplicaciones, lo que acelera el ciclo de desarrollo y permite a los desarrolladores centrarse en la creación de software en lugar de la administración de infraestructura.
  - Gestión de Bases de Datos: Los servicios PaaS a menudo incluyen bases de datos gestionadas, lo que facilita el almacenamiento y la recuperación de datos.
  - Despliegue Fácil: Los desarrolladores pueden implementar sus aplicaciones en entornos PaaS sin preocuparse por la configuración de servidores y la administración de la infraestructura subyacente.





### 3 IaaS (Infrastructure as a Service)

- **IaaS es el modelo de servicio cloud que ofrece un mayor nivel de control respecto a una solución on-premises.** IaaS es uno de los modelos fundamentales de servicios en la nube que ofrece a las organizaciones acceso a recursos de infraestructura virtualizada.
- **Definición:** IaaS, que significa "Infraestructura como Servicio", es un modelo de servicio en la nube que proporciona a las empresas acceso a componentes de infraestructura virtualizada a través de Internet.
- **Características Principales:**
  - **Escalabilidad Bajo Demanda:** Los usuarios pueden escalar recursos de manera rápida y sencilla según las necesidades de su negocio. Esto significa que pueden aumentar o reducir la capacidad de procesamiento, el almacenamiento y la red según sea necesario.
  - **Flexibilidad:** Los usuarios tienen control sobre la configuración de la infraestructura virtual, lo que les permite instalar sistemas operativos y aplicaciones personalizadas.
  - **Eliminación de Inversión en Hardware:** Al utilizar IaaS, las organizaciones no necesitan comprar ni mantener su propia infraestructura física, lo que puede reducir costos significativamente.



# Resumen de Servicios

En el sitio	IaaS	PaaS	SaaS
Aplicaciones	Aplicaciones	Aplicaciones	Aplicaciones
Datos	Datos	Datos	Datos
Tiempo de ejecución	Tiempo de ejecución	Tiempo de ejecución	Tiempo de ejecución
Software intermedio	Software intermedio	Software intermedio	Software intermedio
Sistema operativo	Sistema operativo	Sistema operativo	Sistema operativo
Virtualización	Virtualización	Virtualización	Virtualización
Servidores	Servidores	Servidores	Servidores
Almacenamiento	Almacenamiento	Almacenamiento	Almacenamiento
Redes	Redes	Redes	Redes

Tú administras       El proveedor de servicios administra

# Proveedores de servicios de Cloud



# Proveedores

- **Infraestructura como Servicio (IaaS):**

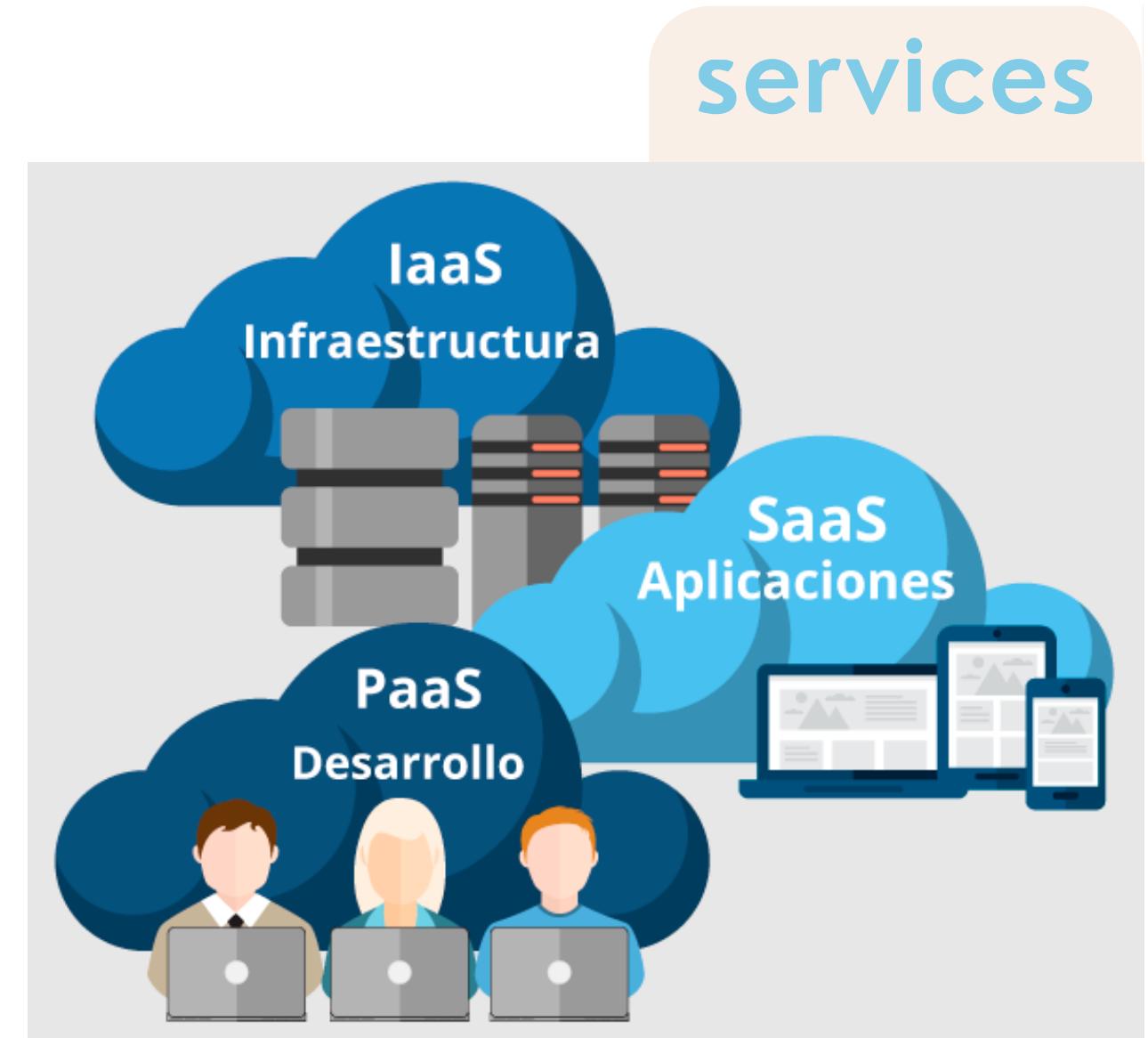
Ejemplos de Proveedores: Algunos ejemplos de proveedores de IaaS populares incluyen Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform (GCP) y IBM Cloud.

- **Plataforma como Servicio (PaaS):**

Ejemplos de Proveedores: Algunos ejemplos de proveedores de PaaS incluyen Heroku, Google App Engine, Azure App Service y Red Hat OpenShift.

- **Software como Servicio (SaaS):**

Ejemplos de Aplicaciones SaaS: Algunos ejemplos de aplicaciones SaaS populares incluyen Microsoft Office 365, Salesforce, Dropbox y Google Workspace.



# Principales proveedores

→ Amazon Web  
Services



→ Google Cloud  
Platform

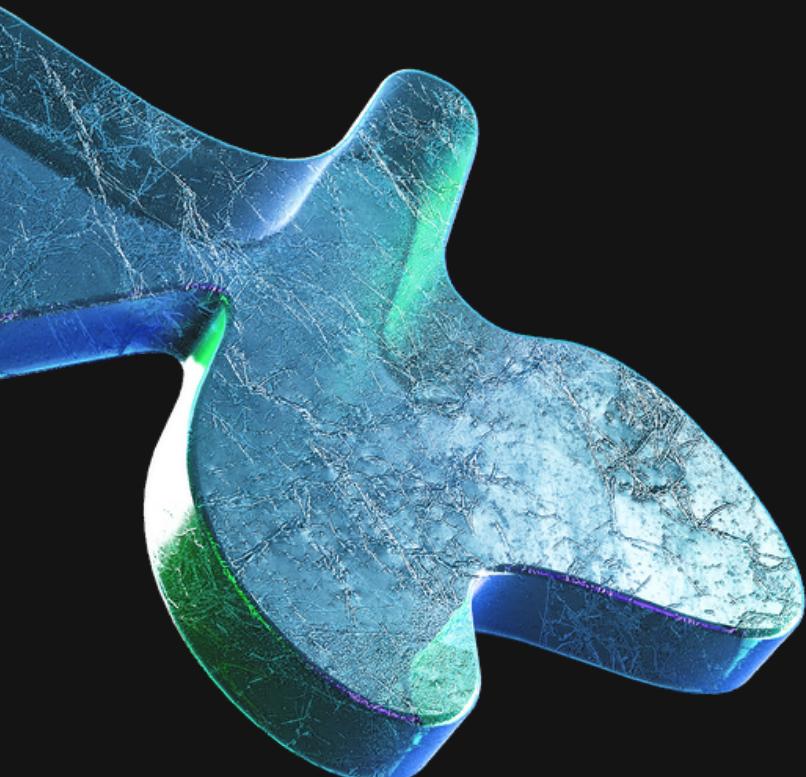


Google Cloud

→ Microsoft  
Azure



# Políticas de cloud



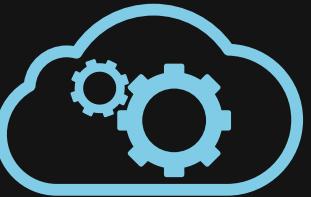


# Estándares a utilizar

## ISO/IEC 27001

---

Es un estándar internacional que se centra en la gestión de la seguridad de la información en general y que nos puede servir de guía para desarrollar políticas de seguridad de la información que sean aplicables al almacenamiento en la nube. La norma ISO 27001 requiere que las organizaciones realicen una evaluación de riesgos para identificar las amenazas y vulnerabilidades relacionadas con la seguridad de la información, incluido el almacenamiento en la nube. La norma también enumera una serie de controles de seguridad que pueden aplicarse al almacenamiento en la nube. Estos controles incluyen la gestión de accesos, la encriptación de datos, la supervisión y el registro de eventos, la gestión de proveedores de servicios en la nube y la continuidad del negocio.



## ISO 27017

---

Es un estándar internacional que proporciona pautas específicas para la seguridad de la información en la nube y es parte de la serie ISO 27000, que incluye estándares relacionados con la gestión de la seguridad de la información.

ISO 27017 se centra en los aspectos de seguridad de la información específicos de los servicios en la nube, y su objetivo es ayudar a las organizaciones, proveedores de servicios en la nube y usuarios a comprender y abordar los riesgos de seguridad asociados con la adopción y el uso de servicios en la nube.

El estándar clarifica las responsabilidades de seguridad compartidas entre los proveedores de servicios en la nube y los clientes. Esto es crucial para garantizar que ambas partes entiendan quién es responsable de qué aspecto de la seguridad.



## ISO 27018

---

ISO 27018 es un estándar internacional que se centra en la protección de la información personal en la nube. Código de práctica para la protección de información de identificación personal (PII) en la nube que actúa como procesador de PII. Este estándar es una extensión de ISO 27001, que es un estándar ampliamente reconocido para la gestión de la seguridad de la información.

La ISO 27018 establece directrices y principios específicos para los proveedores de servicios en la nube que actúan como procesadores de datos personales (PII) para proteger la privacidad de los datos personales almacenados en la nube.

Aunque no es una norma de certificación, es una guía que ayuda a las organizaciones a cumplir con las regulaciones de privacidad y protección de datos, como el Reglamento General de Protección de Datos (RGPD) de la Unión Europea.



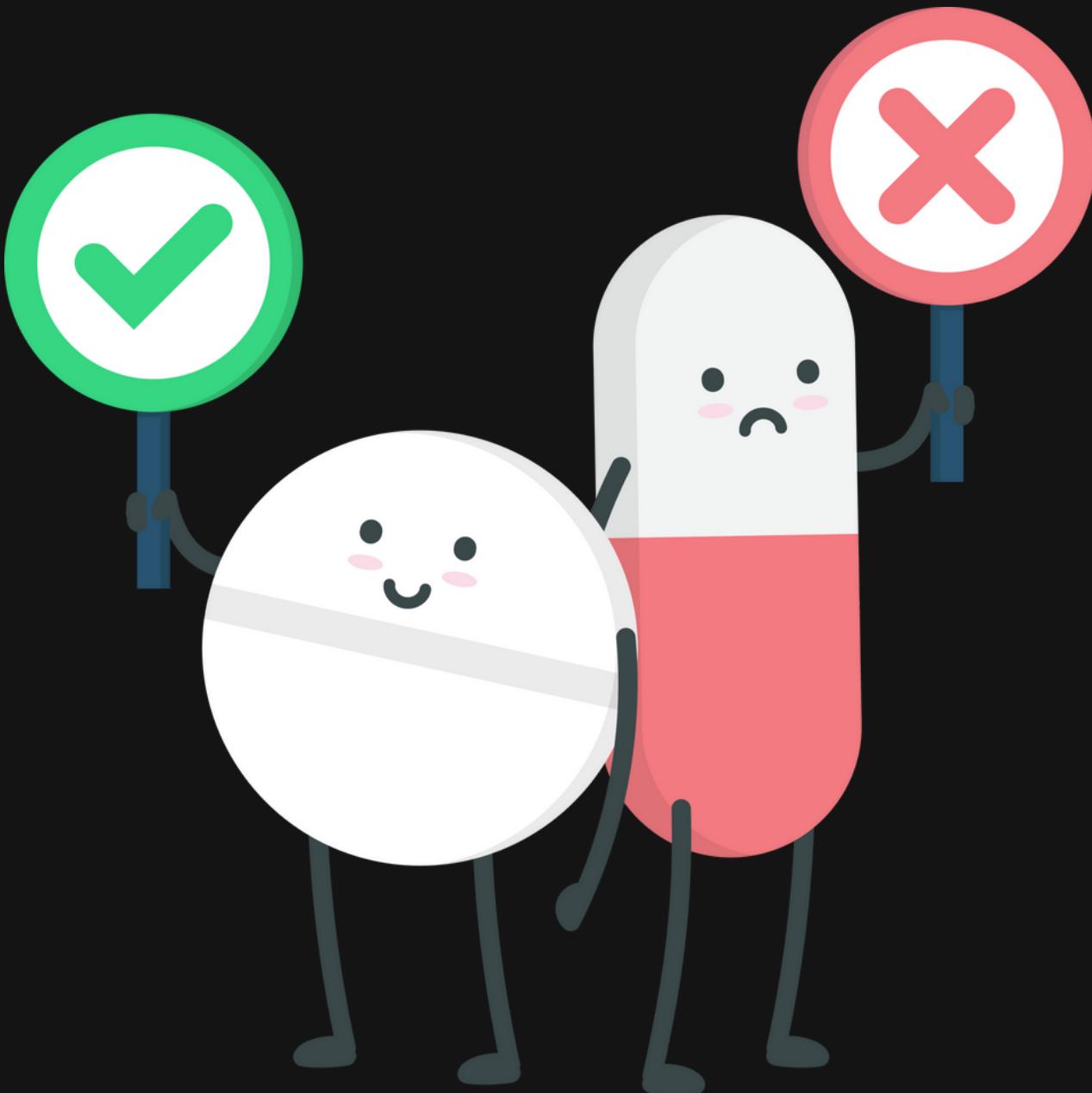
## ISO 22301

---

Es un estándar internacional que se enfoca en la gestión de la continuidad del negocio. La ISO 22301 establece un marco de referencia para que las organizaciones desarrollen, implementen, gestionen y mantengan un sistema de gestión de la continuidad del negocio efectivo. Este sistema de gestión tiene como objetivo ayudar a las organizaciones a prepararse y responder de manera adecuada a situaciones de interrupción, como desastres naturales, fallos tecnológicos, ciberataques, crisis económicas u otros eventos que puedan afectar su capacidad para operar de manera normal. Al implementar los principios y requisitos de ISO 22301, las organizaciones pueden fortalecer su capacidad para proteger los datos en la nube y mantener la continuidad de sus operaciones en caso de eventos adversos que puedan afectar la seguridad de esos datos.

# Casos de éxito y fracaso al utilizar cloud

- **Implementaciones de servicios en Cloud**
- Noticias de optimización de recursos al utilizar Cloud



# Casos de éxito

---



## 1- GRUPO BIMBO

Grupo Bimbo ha apostado por utilizar el servicio de cloud de Oracle, para gestión de datos y procesos de datos, se reporta que desde que utilizan este servicio, se logró reducir en un 75% los abandonos de llamada, y mejorando este aspecto, han logrado definir un pronóstico de operaciones que ha beneficiado a la empresa





## 2- UNAM

La UNAM en 2020 al enfrentar la pandemia, necesitaba de un servicio el cual permitiera a sus estudiantes conectarse a un entorno 100% digital con la mínima conexión y velocidad, para lograr esto, se utilizo el servicio de AWS, que es el servicio de nube de Amazón.





### 3- NETFLIX

Netflix anuncio en el año 2016 que había finalizado totalmente su migración a la nube AWS, menciono que al utilizar la nube, las horas de streaming mensuales de los clientes aumento en factor de 8, también menciono que debido la elasticidad de la nube, su expansión en mas de 130 países fue posible. y gracias a ello son una red de televisión por internet global.



# Casos de fracaso

---

1

En el año 2014 un hacker fue capaz de acceder al servicio de nube de apple, y robo fotos comprometedoras de algunas estrellas de cine como Jennifer Lawrence, Kirsten Dunst, Kate Upton, o Rihanna.

2

En el año 2019 usuarios de la nube de Google fueron victimas de phishing, y así varios usuarios les fue robada información personal la cual almacenaban y dicha nube.



iCloud



Google Cloud

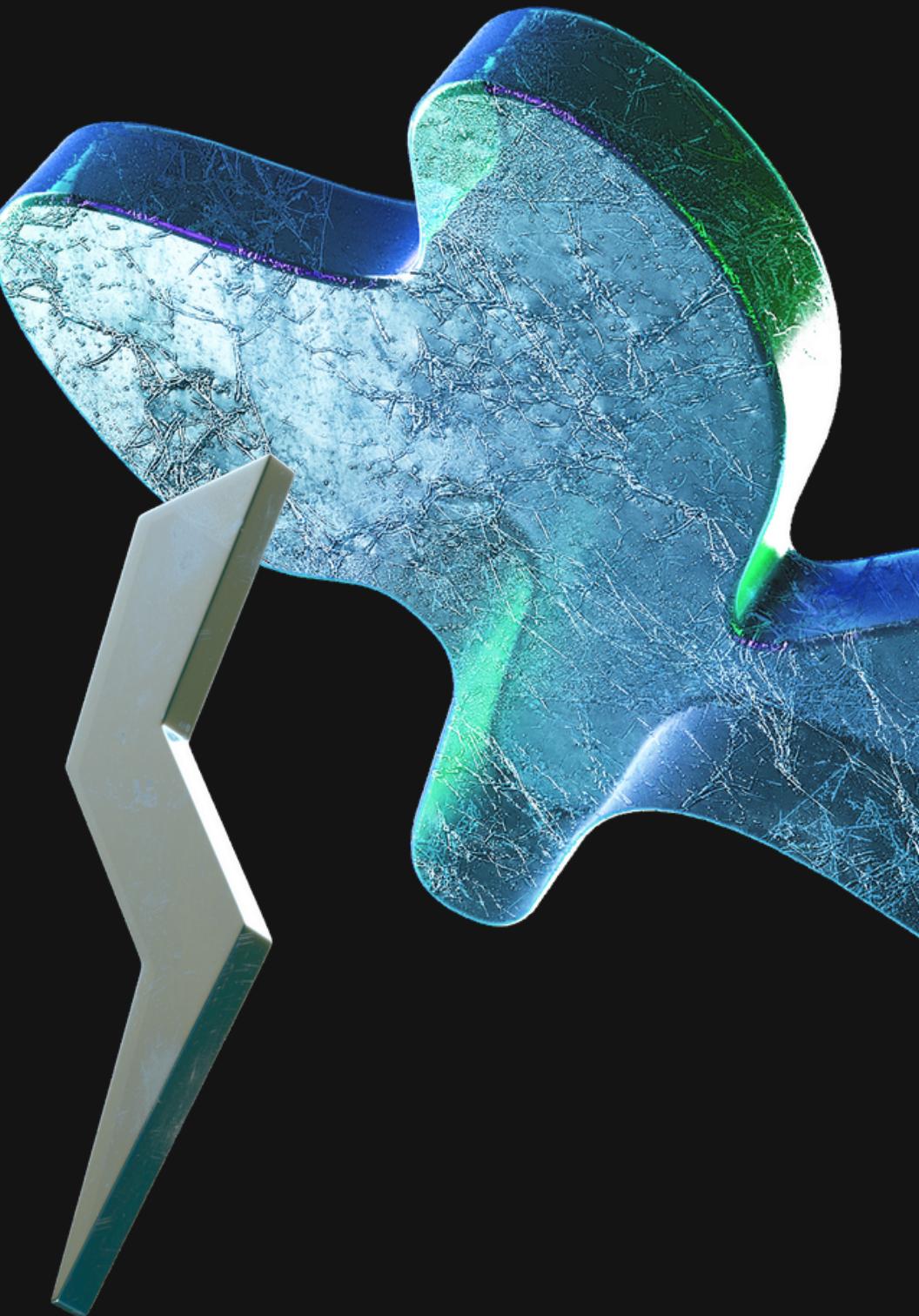
# Simuladores

# Calculadora de costos

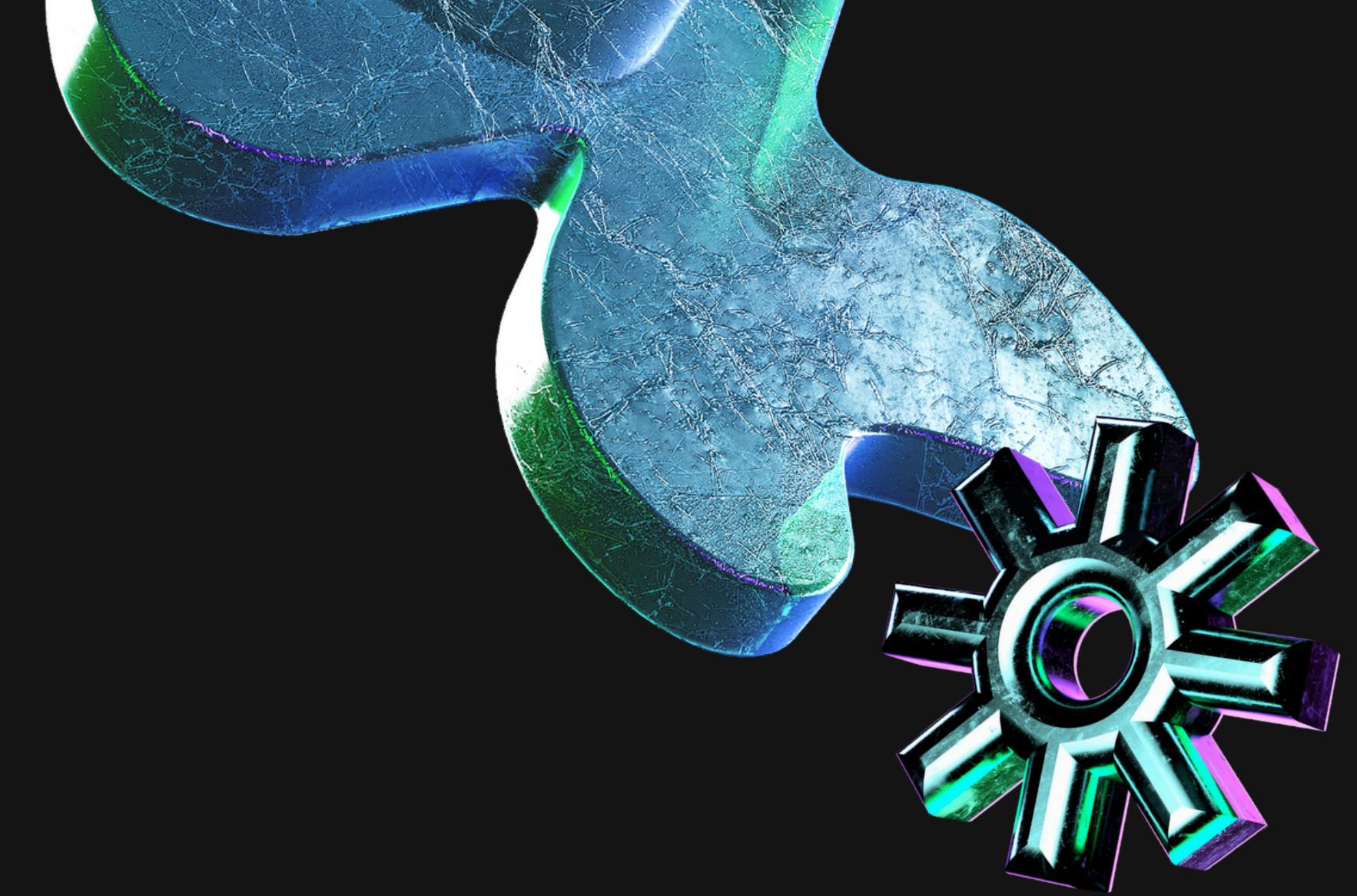
<https://cloud.google.com/products/calculator#id=c6f8b13d-ba80-4eca-8df4-3df14b17b867>

# Simulador de la nube

[https://console.cloud.google.com/storage/browser/meia\\_bucket;tab=objects?hl=es-419&project=potent-shade-401518&prefix=&forceOnObjectsSortingFiltering=false](https://console.cloud.google.com/storage/browser/meia_bucket;tab=objects?hl=es-419&project=potent-shade-401518&prefix=&forceOnObjectsSortingFiltering=false)



**¡Gracias!**



PRESENTACIÓN

SISTEMAS  
REDUNDANTES DE  
ALMACENAMIENTO  
(RAID)

DIEGO AZURDIA 2528119

CÉSAR SILVA 1184519

CÉSAR BOCEL 1094921

BENJAMIN IZQUIERDO 1321220

JULIO RUIZ 1284719

EMILIO ALFARO 1181320

UNIVERSIDAD RAFAEL LANDIVAR

- 
- 03 DEFINICION Y OBJETIVOS**
  - 04 NIVELES DE RAID**
  - 07 BIT DE PARIDAD**
  - 08 ¿CÓMO SE IMPLEMENTA UNA RAID?**
  - 09 POLÍTICAS DE RAID**
  - 10 CASOS DE EXITO Y FRACASO**
  - 12 SIMULADORES**
  - 13 CONCLUSIONES**

## **DEFINICIÓN**

Un RAID (Redundant Array of Independent Disks, conjunto redundante de discos independientes en inglés), es un sistema de almacenamiento de datos que permite configurar varios discos duros de un ordenador en una única unidad lógica.

Según su configuración, se pueden conseguir diversas ventajas, como mayor integridad, mayor tolerancia a fallos, mayor rendimiento o mayor capacidad.

## **OBJETIVO**

El principio base de un sistema RAID es almacenar datos de manera redundante, lo que garantiza que la integridad y funcionalidad de la matriz completa no queden comprometidas por un fallo en un disco duro.

# NIVELES DE RAID

## RAID 0 (conjunto dividido)

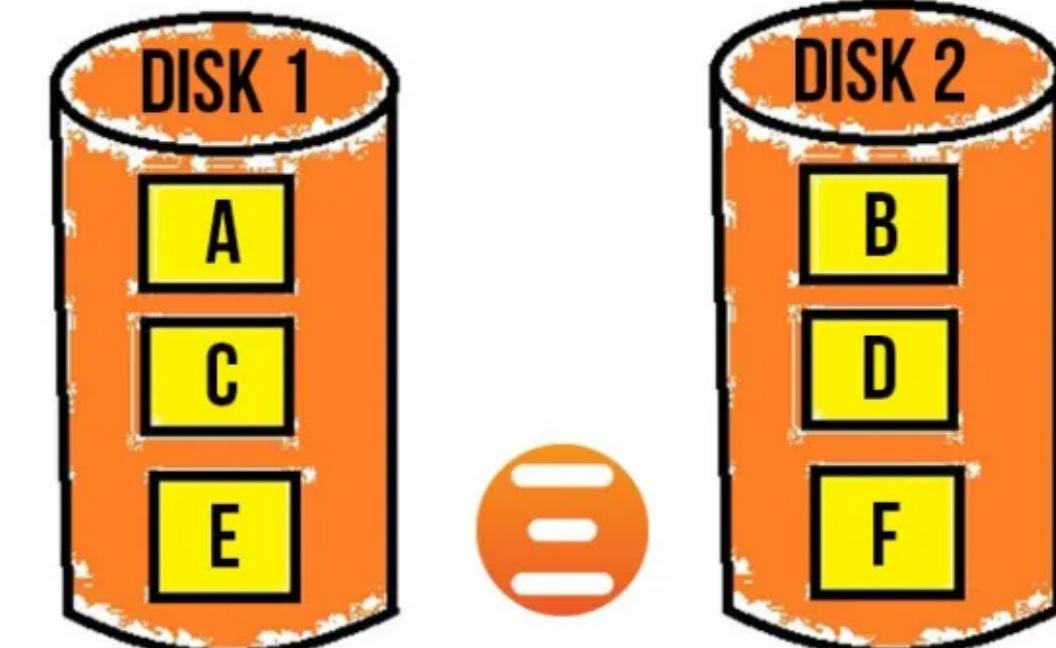
### Ventajas

- El rendimiento es muy alto
- Agrega rendimiento para cada disco que sea añadido sin muchos conflictos (dependiendo de la controladora RAID)

### Contras

- Sin redundancia significa que si se pierde algún dato, no se puede recuperar.
- Mayor riesgo de pérdida de datos para cada disco de la matriz (más unidades = más riesgo)
- No se recomienda para sistemas operativos ni datos valiosos

- Requiere un mínimo de 2 discos para funcionar adecuadamente
- No ofrece redundancia alguna (tampoco duplicación ni paridad)



RAID 0 - BLOCKS STRIPED (NO PARITY)

# NIVELES DE RAID

## RAID 1 (conjunto en espejo o mirror)

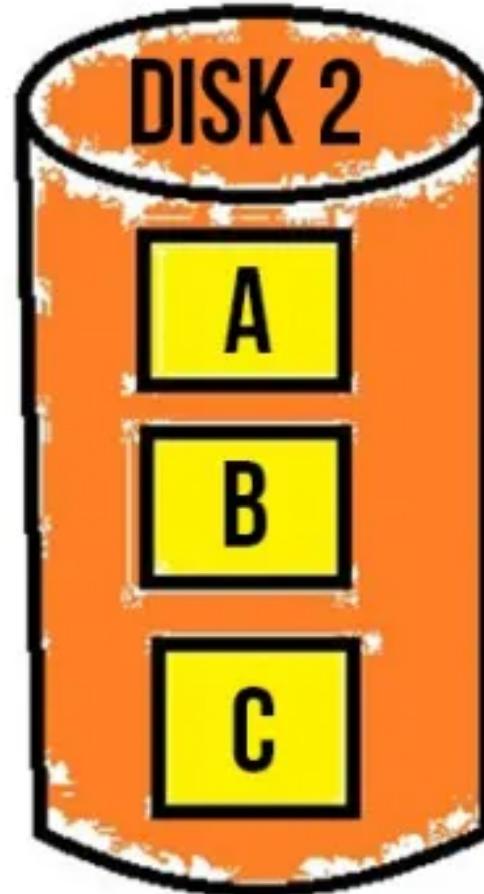
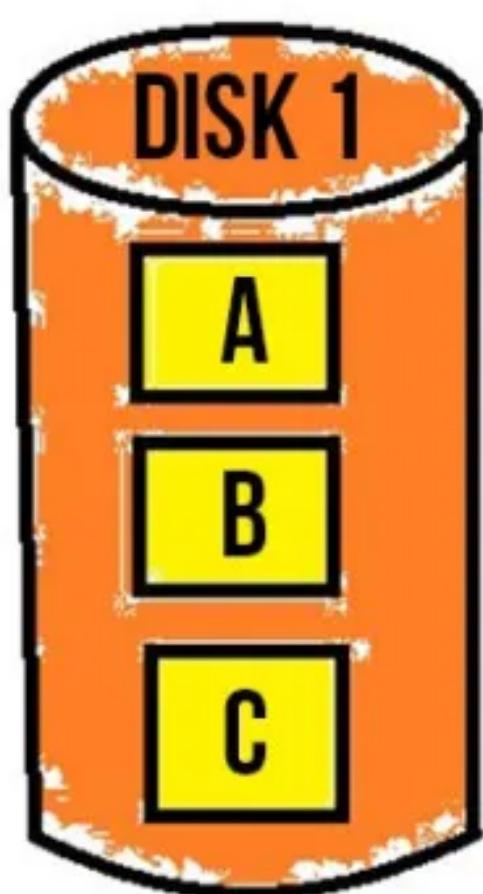
- Al igual que RAID 0, requiere un mínimo de 2 discos para funcionar correctamente.
- Ofrece una buena redundancia gracias a RAID 1 que utiliza una unidad duplicada

### Ventajas

- Altas velocidades de lectura (a veces mejores que RAID 0)
- Proporciona a los usuarios redundancia a través de la duplicación (esencialmente, la 2<sup>a</sup> unidad es un clon del disco maestro)

### Contras

- La velocidad de escritura es muy deficiente
- La capacidad de almacenamiento se “refleja”, lo que significa que la capacidad de almacenamiento no aumenta a expensas de la redundancia



RAID 1 - BLOCKS MIRRORED (NO PARITY)

# NIVELES DE RAID

## RAID 5 (conjunto dividido con paridad distribuida)

- Requiere un mínimo de 3 discos para la configuración
- Proporciona un nivel agregado de redundancia a través de la paridad

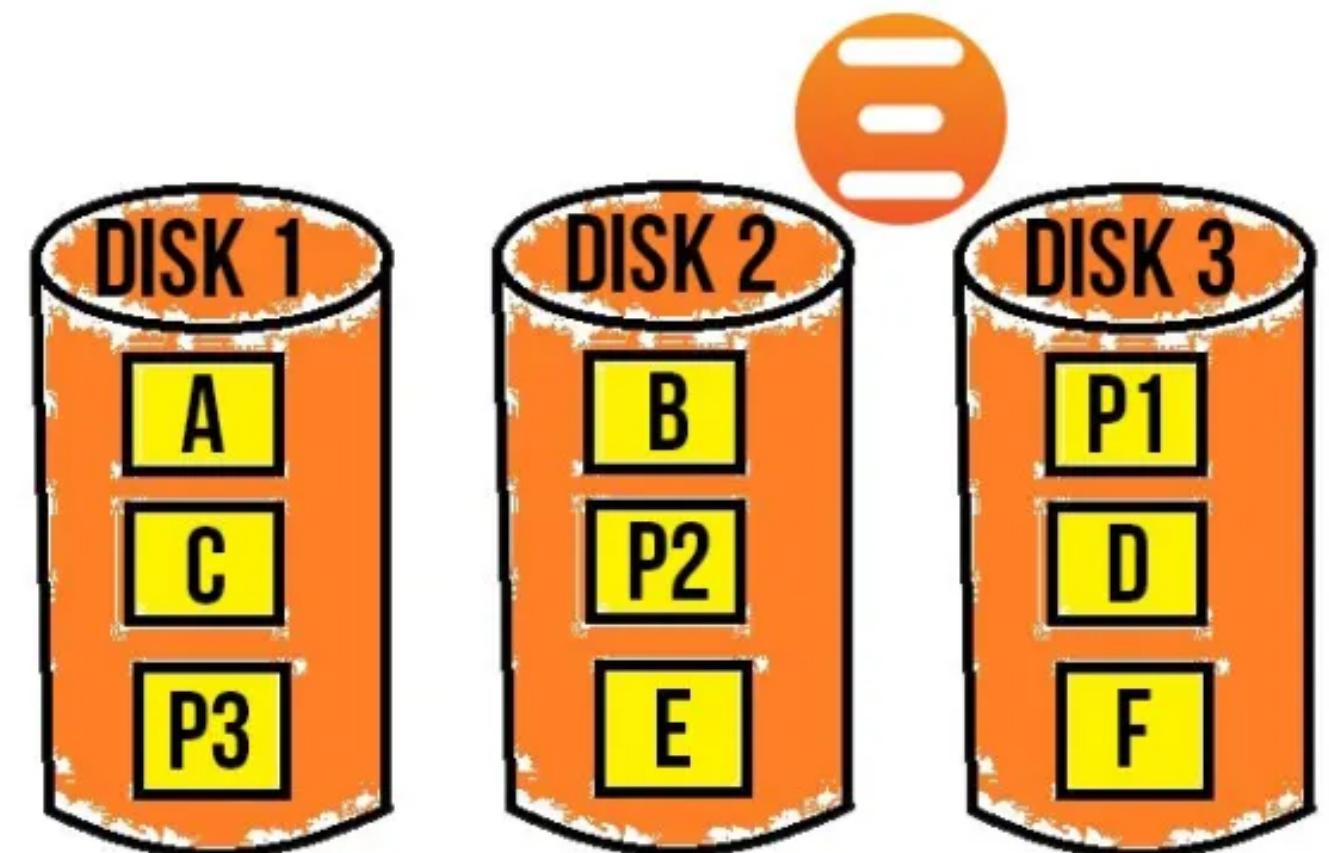
### Ventajas

Ofrece un rendimiento de lectura decente gracias a las bandas (como RAID 0)

La paridad ofrece un nivel de redundancia en cada una de las unidades (la paridad ofrece la capacidad de reparar bloques de datos rotos)

### Contras

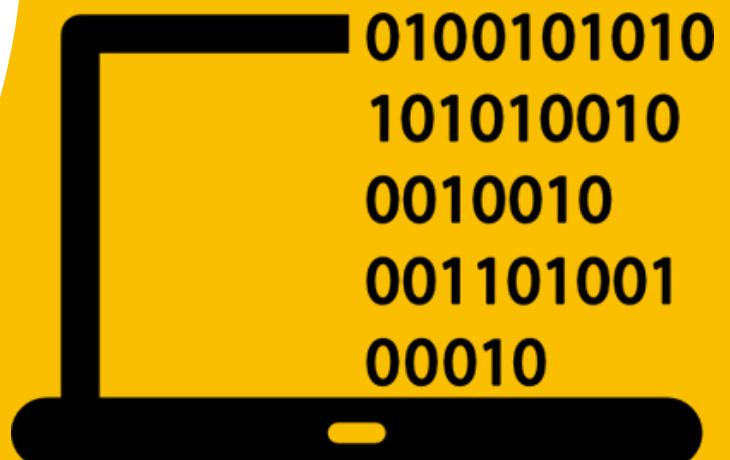
Rendimiento de escritura deficiente



RAID 5 - BLOCKS STRIPED - SPLIT PARITY

# BIT DE PARIDAD

EL BIT DE PARIDAD ES UNA TÉCNICA QUE ACTÚA COMO UN GUARDIÁN DE DATOS. SU FUNCIÓN PRINCIPAL ES DETECTAR Y CORREGIR ERRORES EN LOS DATOS ALMACENADOS EN UNA MATRIZ DE DISCOS EN UN SISTEMA RAID.



LA PARIDAD SE CALCULA MEDIANTE LA SUMA DE LOS VALORES DE LOS DISPOSITIVOS EN LA MATRIZ. ES COMO LA FORMA EN QUE COMPROBAMOS SI LA SUMA DE UN GRUPO DE NÚMEROS ES PAR O IMPAR, PERO A UNA ESCALA MUCHO MAYOR.



# POLITICAS DE RAID

## ISO/IEC 27001

EL PROCESO DE IMPLEMENTACIÓN DE LA NORMA ISO 27001 SE DIVIDE EN CUATRO FASES:

- PLANIFICACIÓN
- IMPLEMENTACIÓN
- EVALUACIÓN
- MEJORA CONTINUA

## ISO/IEC 27000

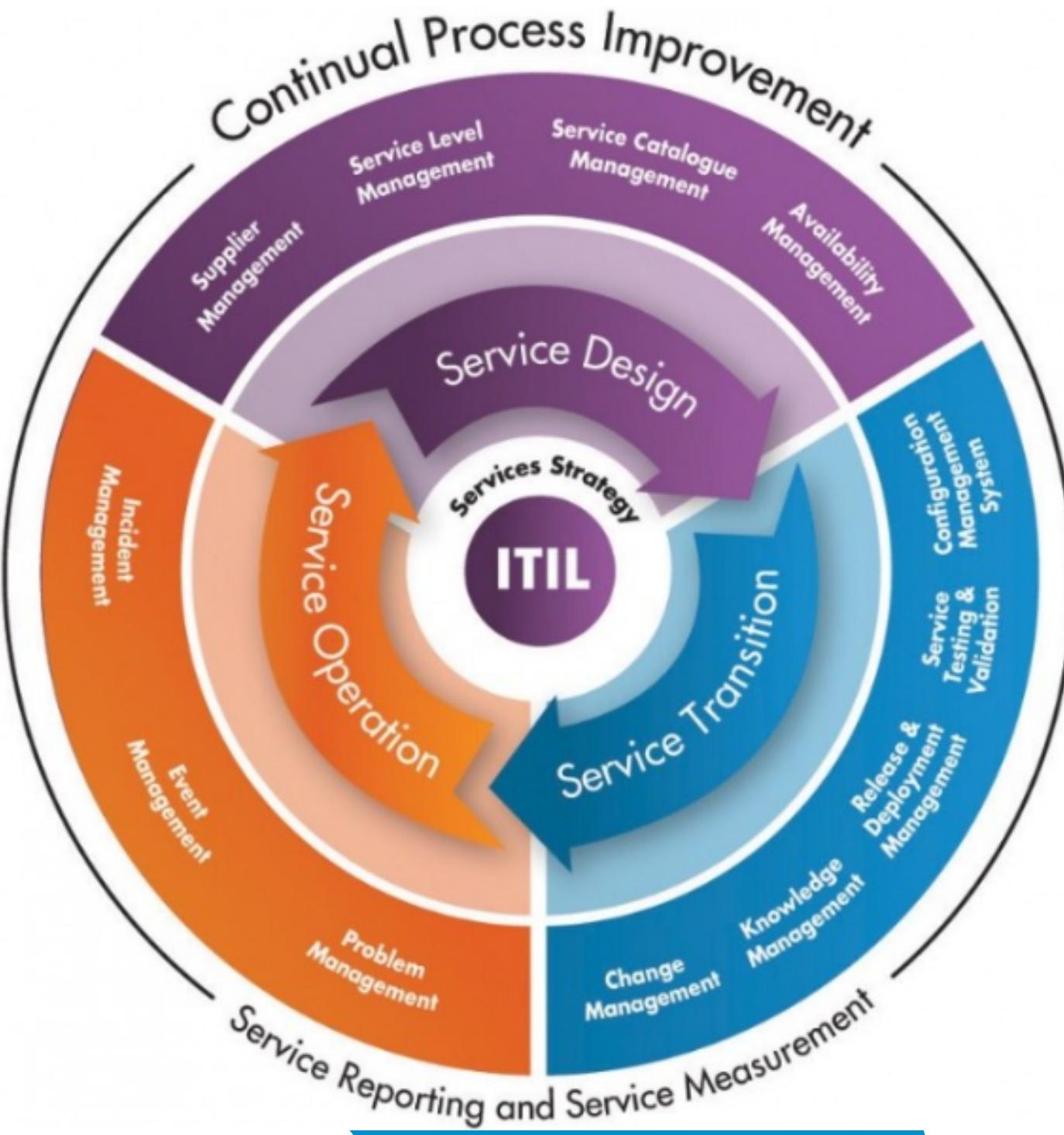
ORIENTADAS AL ESTABLECIMIENTO DE BUENAS PRÁCTICAS EN RELACIÓN CON LA IMPLANTACIÓN, MANTENIMIENTO Y GESTIÓN DEL SISTEMA DE GESTIÓN DE SEGURIDAD DE LA INFORMACIÓN (SGSI)

## ITIL

INFORMATION TECHNOLOGY INFRASTRUCTURE LIBRARY.

- UNOS DE LOS MARCOS DE REFERENCIA MÁS UTILIZADOS DEL MUNDO.
- SE ENFOCA EN LA MEJORA CONTINUA DE LOS PRODUCTOS DE SOFTWARE.

# ITIL



## CADENA DE VALOR DE SERVICIOS DE ITIL

ES UNA COMBINACIÓN DE 6 ACTIVIDADES QUE TRABAJAN EN CONJUNTO CREANDO VALOR PARA LA ORGANIZACIÓN O LOS USUARIOS, A TRAVÉS DE LA ENTREGA DE UN PRODUCTO O SERVICIO.

## GOBIERNO DE ITIL

LA ALTA CALIDAD Y MEJORA CONTINUA SON LA ESENCIA PARA AGREGAR VALOR.

## PRÁCTICAS DE ITIL

PROMUEVEN CALIDAD, SATISFACCIÓN AL CLIENTE Y EL LOGRAR LOS OBJETIVOS.

# ¿CÓMO SE IMPLEMENTA UN RAID?

## IMPLEMENTACIÓN RAID POR HARDWARE

### SELECCIÓN DE HARDWARE:

- Tarjeta controladora RAID: Con procesador y memoria propios.
- Discos duros: Seleccionar según capacidad, velocidad y tipo. Preferiblemente del mismo modelo.



### CONFIGURACIÓN E INICIALIZACIÓN (HARDWARE)

- Configuración:
  - Durante POST, opción para configurar RAID (ej. "Press Ctrl+R...").
  - Crear, gestionar y eliminar arrays.
- Inicialización:
  - Prepara los discos para uso. Tiempo varía según cantidad y tamaño.

# ¿CÓMO SE IMPLEMENTA UN RAID?

## IMPLEMENTACIÓN RAID POR SOFTWARE

- Preparativos:
- Asegurarse de que los discos estén conectados y reconocidos.
- Necesidad de herramientas específicas según el sistema operativo.
- Creación del Array:
- Iniciar herramienta RAID.
- Seguir instrucciones y seleccionar nivel de RAID.



## FORMATO, CONFIGURACIÓN Y MONITOREO (SOFTWARE)

- Formateo y Configuración:
- El array aparece como nueva unidad.
- Formatear y usar para almacenamiento o aplicaciones.
- Gestión y Monitoreo:
- Herramientas ofrecen vista detallada del estado de discos.
- Importancia de seguimiento para integridad de datos.

# EL FUTURO ECONÓMICO

## Crisis financiera mundial.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut a enim nec nisl ullamcorper eleifend. Praesent risus leo, fringilla et nulla at, egestas euismod orci. Suspendisse porttitor diam eu condimentum aliquam.

**2022**

**2023**

**2024**

**2025**

**2026**

## Inicio de la recuperación.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut a enim nec nisl ullamcorper eleifend. Praesent risus leo, fringilla et nulla at, egestas euismod orci. Suspendisse porttitor diam eu condimentum aliquam. Fusce interdum cursus nisl ut rutrum.

## Las grandes potencias mundiales.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut a enim nec nisl ullamcorper eleifend. Praesent risus leo, fringilla et nulla at, egestas euismod orci. Suspendisse porttitor diam eu condimentum aliquam. Fusce interdum cursus nisl ut rutrum.

## Estudio mundial económico.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut a enim nec nisl ullamcorper eleifend. Praesent risus leo, fringilla et nulla at, egestas euismod orci. Suspendisse porttitor diam eu condimentum aliquam.

## Previsión de aumento económico.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut a enim nec nisl ullamcorper eleifend. Praesent risus leo, fringilla et nulla at, egestas euismod orci. Suspendisse porttitor diam eu condimentum aliquam. Fusce interdum cursus nisl ut rutrum.



## CASOS DE EXITO

### Cloud Storage

Cloud Storage es un modo de almacenamiento de datos de computadora en el que los datos digitales se almacenan en servidores ubicados fuera de las instalaciones. Un proveedor externo responsable de alojar, administrar y proteger los datos almacenados en su infraestructura mantiene los servidores.



### QNAP TS-433

El QNAP TS-433 es un NAS de cuatro bahías pensando tanto para uso doméstico como para pequeñas empresas que busquen un sistema con multitud de opciones de almacenamiento y todo el catálogo de software y funciones que integra el sistema operativo QTS 5 de QNAP.



### Intel® Virtual RAID on CPU (Intel® VROC)

Combina las características de protección de datos de RAID con el alto desempeño de las SSD NVMe, todo sin necesidad de un adaptador de bus host RAID (HBA) tradicional.



# CASOS DE FRACASO

## 70 Millones, el ejercito de EEUU. 2009.

Una perdida de datos que afecta a decenas de millones de registros sobre veteranos militares de los Estados Unidos. El problema involucra un disco duro defectuoso



## Contrato de empresa de pérdida de datos eliminado

Una compañía que perdió los detalles de miles de delincuentes retenidos en un disco, ha rescindido su contrato de £1.5m



## Ameritrade advierte a 200.000 clientes de datos perdidos

Ameritrade Inc. ha informado a 200,000 clientes actuales y anteriores que se ha perdido la información cinta de respaldo de computadora que contiene su información personal. La cinta contenía información que abarcaba los años 2000-2003 e incluía tanto a los consumidores actuales como a los anteriores del corredor en línea



# **SIMULADORES CALCULADORAS RAID**

Esta proporciona una estimación del almacenamiento con unidades mixtas y configuraciones de RAID.

Los resultados son estimaciones destinadas a representar nuevas capacidades de grupo de almacenamiento. El tamaño máximo de volumen individual varía según la arquitectura de la CPU y puede ser: 16 TB, 108 TB, 200 TB o 1 PB



# ¿QUÉ ES EL SHR?

## SYNOLOGY HYBRID RAID

Es un sistema de gestión RAID automatizado, es decir podrá utilizar discos duros de diferentes tamaños en su servidor NAS. En donde no es posible en otros tipos de RAID.

## ¿CÓMO FUNCIONA?

Una matriz RAID tradicional utiliza sólo el disco más pequeño de un pool de almacenamiento, lo que significa que los discos más grandes siempre tendrán por defecto el espacio de almacenamiento del disco duro más pequeño cuando se crea el RAID. La única forma de aumentar el tamaño del pool de almacenamiento sería hacer que todos los discos tuvieran el mismo tamaño.

## LA GRAN VENTAJA

El SHR ofrece una gran ventaja, ya que la mezcla de discos duros de diferente capacidad le da la posibilidad de utilizar discos duros más pequeños al principio y con el tiempo actualizarlos a discos duros más grandes si lo desea.

El SHR es diferente porque divide el espacio de almacenamiento del disco duro en trozos más pequeños que permiten utilizar el espacio de almacenamiento adicional.

# LAS FASES

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut a enim nec nisl ullamcorper eleifend. Praesent risus leo, fringilla et nulla at, egestas euismod orci.

Suspendisse porttitor diam eu condimentum aliquam. Fusce interdum cursus nisl ut rutrum. Donec et sapien sit amet nisl pretium efficitur.

## FASE 1 - RECOPILACIÓN

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut a enim nec nisl ullamcorper eleifend. Praesent risus leo, fringilla et nulla at, egestas euismod orci.

## FASE 2 - EXPORTACIÓN DE DATOS

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut a enim nec nisl ullamcorper eleifend. Praesent risus leo, fringilla et nulla at, egestas euismod orci.

## FASE 3 - ESTUDIO

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut a enim nec nisl ullamcorper eleifend. Praesent risus leo, fringilla et nulla at, egestas euismod orci.

# LAS PRINCIPALES EMPRESAS MUNDIALES

Empresa 1

## **Soto y Ochoa**

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut a enim nec nisl ullamcorper eleifend. Praesent risus leo, fringilla et nulla at, egestas euismod orci.

Empresa 4

## **Fauget**

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut a enim nec nisl ullamcorper eleifend. Praesent risus leo, fringilla et nulla at, egestas euismod orci.

Empresa 2

## **Industrias Ariova**

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut a enim nec nisl ullamcorper eleifend. Praesent risus leo, fringilla et nulla at, egestas euismod orci.

Empresa 5

## **Rodríguez y familia**

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut a enim nec nisl ullamcorper eleifend. Praesent risus leo, fringilla et nulla at.

Empresa 3

## **Rimbeiro y Asociados**

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut a enim nec nisl ullamcorper eleifend. Praesent risus leo, fringilla et.

Empresa 6

## **Multinacional González**

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut a enim nec nisl ullamcorper eleifend. Praesent risus leo, fringilla et nulla at, egestas euismod orci.

# **GRACIAS**

SANDRA HARO  
HOLA@UNSITIOGENIAL.ES

UNIVERSIDAD ENSIGNA

**UNIVERSIDAD RAFAEL LANDÍVAR**  
**FACULTAD DE INGENIERÍA**  
**MANEJO E IMPLEMENTACIÓN DE ARCHIVOS**  
**SECCIÓN 01 VESPERTINA**  
**MGTR. DAVID FERNANDO LUNA HERNÁNDEZ**

# **SISTEMAS REDUNDANTES DE ALMACENAMIENTO (RAID)**

**César Adrian Silva Pérez 1184519**  
**Julio Anthony Engels Ruiz Coto 1284719**  
**Benjamin Izquierdo Echeverria 1321220**  
**Diego Andrés Azurdia Ortiz 2528119**  
**Herbert Emilio Alfaro Montes 1181320**

**GUATEMALA DE LA ASUNCIÓN, OCTUBRE 26 DE 2023**  
**CAMPUS CENTRAL**

## **DEFINICIÓN Y OBJETIVO**

Un RAID (Redundant Array of Independent Disks, conjunto redundante de discos independientes en inglés), es un sistema de almacenamiento de datos que permite configurar varios discos duros de un ordenador en una única unidad lógica. Según su configuración, se pueden conseguir diversas ventajas, como mayor integridad, mayor tolerancia a fallos, mayor rendimiento o mayor capacidad.

El principio base de un sistema RAID es almacenar datos de manera redundante, lo que garantiza que la integridad y funcionalidad de la matriz completa no queden comprometidas por un fallo en un disco duro.

## **NIVELES DE RAID**

### **RAID 0 (conjunto dividido)**

- Requiere un mínimo de 2 discos para funcionar adecuadamente
- No ofrece redundancia alguna (tampoco duplicación ni paridad)

#### Ventajas

- El rendimiento es muy alto
- Agrega rendimiento para cada disco que sea añadido sin muchos conflictos (dependiendo de la controladora RAID)

#### Contras

- Sin redundancia significa que si se pierde algún dato, no se puede recuperar.
- Mayor riesgo de pérdida de datos para cada disco de la matriz (más unidades = más riesgo)
- No se recomienda para sistemas operativos ni datos valiosos

### **RAID 1 (conjunto en espejo o mirror)**

- Al igual que RAID 0, requiere un mínimo de 2 discos para funcionar correctamente.
- Ofrece una buena redundancia gracias a RAID 1 que utiliza una unidad duplicada

#### Ventajas

- Altas velocidades de lectura (a veces mejores que RAID 0)
- Proporciona a los usuarios redundancia a través de la duplicación (esencialmente, la 2<sup>a</sup> unidad es un clon del disco maestro)

#### Contras

- La velocidad de escritura es muy deficiente
- La capacidad de almacenamiento se “refleja”, lo que significa que la capacidad de almacenamiento no aumenta a expensas de la redundancia

### **RAID 5 (conjunto dividido con paridad distribuida)**

- Requiere un mínimo de 3 discos para la configuración
- Proporciona un nivel agregado de redundancia a través de la paridad

#### Ventajas

- Ofrece un rendimiento de lectura decente gracias a las bandas (como RAID 0)

- La paridad ofrece un nivel de redundancia en cada una de las unidades (la paridad ofrece la capacidad de reparar bloques de datos rotos)

Contras

- Rendimiento de escritura deficiente

## ¿CÓMO SE IMPLEMENTA UN RAID?

### RAID POR HARDWARE

#### SELECCIÓN DE HARDWARE:

- Tarjeta controladora RAID: Es el componente clave para la implementación de RAID por hardware. Estas tarjetas vienen con su propio procesador y memoria para manejar las operaciones de RAID, liberando al CPU principal de esta tarea.
- Discos duros: Deben ser seleccionados teniendo en cuenta la capacidad, velocidad y tipo (HDD, SSD). A menudo, se recomienda usar discos del mismo modelo y capacidad para evitar incompatibilidades y problemas de rendimiento.

#### CONFIGURACIÓN:

- Al arrancar el sistema, durante el proceso de POST (Power-On Self-Test), se mostrará una opción para entrar en la configuración de la tarjeta RAID (por ejemplo, "Press Ctrl+R to enter RAID setup").
- Una vez en la utilidad de configuración, se podrá crear, gestionar y eliminar arrays RAID. Selecciona la opción para crear un nuevo array y elige el nivel de RAID deseado (RAID 0, RAID 1, RAID 5, etc.).
- Añade los discos que formarán parte del array y confirma la creación.

#### INICIALIZACIÓN Y FORMATEO:

- Despues de crear el array, este debe ser inicializado. Este proceso prepara los discos para su uso y puede tomar un tiempo dependiendo del tamaño y número de discos.
- Una vez inicializado, el array RAID será detectado por el sistema operativo como una única unidad de disco. Puedes formatear y particionar este disco desde el sistema operativo como lo harías con cualquier otro disco.

### RAID POR SOFTWARE

#### Preparativos:

- Discos duros: Asegurarse de que los discos que se desean usar para el RAID estén conectados y sean reconocidos por el sistema.
- Software: Dependiendo del sistema operativo, necesitarás herramientas o utilidades específicas para gestionar el RAID. Estas herramientas son las que permitirán crear, configurar y monitorear los arrays RAID.

#### Creación del Array:

- Se inicia la herramienta o utilidad RAID proporcionada por el sistema operativo.

- Se siguen las instrucciones o el asistente para crear un nuevo array. Durante este proceso, se le pedirá que selecciones el nivel de RAID (como RAID 0, RAID 1, RAID 5, etc.).
- Añade los discos que formarán parte del array. Estos discos serán agrupados y tratados como una única unidad lógica por el sistema operativo.

#### **Formateo y Configuración:**

- Después de crear el array, este aparecerá como una nueva unidad de almacenamiento en el sistema operativo.
- Formatear esta nueva unidad con el sistema de archivos que se desee, al igual que con cualquier otra unidad de almacenamiento.
- Una vez formateado, se puede comenzar a usar el array RAID para almacenar datos, instalar aplicaciones o cualquier otra función que deseas.

#### **Gestión y Monitoreo:**

- Utilizar las herramienta o utilidad RAID para monitorear regularmente el estado y salud del array. La mayoría de las herramientas ofrecen una vista detallada del estado de cada disco dentro del array, así como alertas o notificaciones en caso de fallos o problemas.
- Es fundamental mantener un seguimiento regular para garantizar que el sistema RAID funcione correctamente y que los datos estén protegidos.

## **Bit de paridad**

#### **Función del Bit de Paridad:**

- El bit de paridad es una técnica que permite detectar y corregir errores en los datos almacenados en una matriz de discos en un sistema RAID. La paridad se calcula como la suma de los valores de los dispositivos en la matriz. Cuando se produce un error en uno de los discos, es posible recuperar los datos correctos al comparar la paridad con los datos restantes en la matriz. Este proceso se realiza a nivel de bloques de datos, no solo a nivel de bits. Aunque en teoría los bloques pueden ser de 1 bit en tamaño, en la práctica, el tamaño mínimo común suele ser de 512 bytes, que coincide con el tamaño de un sector en la mayoría de los dispositivos de almacenamiento de bloques.

#### **Aplicación en Niveles de RAID:**

- El bit de paridad se utiliza en varios niveles de RAID, como RAID 2, 3, 4 y 5. En estos niveles, la paridad se utiliza para garantizar la redundancia y la integridad de los datos. Cuando se produce un fallo en un dispositivo en una matriz RAID con bit de paridad, el dato de paridad puede utilizarse para reconstruir los datos correctos. Por ejemplo, en el caso de RAID 5, donde se distribuye la paridad en todos los discos, un disco dañado puede recuperarse mediante la paridad y los datos en los otros discos.

## **Políticas de RAID**

### **- ISO/IEC 27000**

Son un conjunto de estándares creados y gestionados por la *Organización Internacional para la Estandarización* (ISO) y la *Comisión Electrónica Internacional* (IEC).

Estas normas están orientadas al establecimiento de buenas prácticas en relación con la implantación, mantenimiento y gestión del *Sistema de Gestión de Seguridad de la Información* (SGSI). (Solutions, 2023)

Dichas reglas tienen como objetivo establecer las mejores prácticas en relación con diferentes aspectos vinculados a la gestión de la seguridad de la información, con una fuerte orientación a la mejora continua y la mitigación de riesgos.

De entre los 47 miembros de la familia 27000 la “ISO/IEC-27001” es considerada la principal de la familia, ya que en ella es donde se especifican los requerimientos necesarios para implantar, mantener y gestionar un SGSI, dentro del proceso de mejora continua conocido como Ciclo Deming o PDCA, acrónimo de Plan-Do-Check-Act, en relación con las fases de Planificar, Hacer, Verificar y Actuar. (Solutions, 2023)

### **- ISO/IEC 27001**

Es un estándar internacional que establece los requisitos para la implementación, mantenimiento y mejora continua de un Sistema de Gestión de la Seguridad de la Información (SGSI).

La norma proporciona un marco para la seguridad de la información que ayuda a las organizaciones a identificar y gestionar sus riesgos de seguridad de la información de manera efectiva. (Solutions, 2023)

El proceso de implementación de la norma ISO 27001 se divide en cuatro fases: planificación, implementación, evaluación y mejora continua.

#### **- Fase de planificación**

Durante la fase de planificación, la organización identifica sus requisitos de seguridad de la información y establece un plan para implementar el SGSI.

#### **- Fase de implementación**

La fase de implementación incluye la creación de políticas, procedimientos y controles para proteger la información.

#### **- Fase de evaluación**

Durante la fase de evaluación, la organización evalúa la eficacia de su SGSI e identifica áreas de mejora.

#### **- Fase de mejora continua**

La fase de mejora continua implica la identificación y aplicación de mejoras a los procesos y controles del SGSI.

La implantación de la norma ISO 27001 a través de un software puede mejorar significativamente la eficiencia, la eficacia y la transparencia de la gestión de la seguridad de la información, lo que a su vez puede ayudar a mitigar los riesgos y proteger la información crítica. (Solutions, 2023)

- **ITIL (Information Technology Infrastructure Library)**

Es una librería de buenas prácticas para la gestión de servicios de tecnología. Se enfoca en la mejora continua de los productos de software. Se basa en el principio de cómo TI agrega valor a los usuarios finales o a los objetivos de un negocio.

ITIL se ha elaborado para abarcar campos como la infraestructura, desarrollo y operaciones(Monitoreo, control o mantenimiento) del software, consta de 4 pilares:

- Procesos: Para la gestión de TI y su alineación a los objetivos de la organización.
- Calidad: Entrega al cliente del producto o servicio óptimo, que incluya los requerimientos acordados.
- Cliente: Satisfacción es el objetivo de la mejora de los productos de TI.
- Independencia: Siempre deben mantenerse buenas prácticas a pesar de los métodos para cada proceso y de los proveedores.

### ***Cadena de valor de servicios de ITIL***

Es una combinación de 6 actividades que trabajan en conjunto creando valor para la organización o los usuarios, a través de la entrega de un producto o servicio.

1. Planificación (crear políticas, obtener requerimientos funcionales y no funcionales, etc.)
2. Mejora (servicios de tecnología)
3. Compromiso (con el proyecto para la mejora continua)
4. Diseño y Transición
5. Obtener
6. Entrega (agrupado con soporte, se enfocan en la garantía de producto o servicio. Ponerlo en producción)
7. Soporte (Monitorear y estar atentos ante cualquier problema)

### ***Gobierno de ITIL***

- Evaluación: Verificar que se cumplan los objetivos.
- Dirección: Ofrecer la guía con tips con los 7 principios de ITIL.
- Supervisión: Monitorear que las buenas prácticas se cumplan.

- Objetivo: Garantizar que la cadena de valor del servicio y las prácticas de la organización funcionen alineadas con los objetivos de la empresa.

La alta calidad y mejora continua son la esencia para agregar valor.

### ***Prácticas de ITIL***

Promueven calidad, satisfacción al cliente y el lograr los objetivos.

Diseño → Qué servicios la organización de TI debe ofrecer para los clientes indicados

Operaciones → Que siga funcionando

Transiciones → Monitorear cambios y dar soporte

## **CASOS DE ÉXITO Y FRACASO**

### **Casos de Éxito**

#### **· Cloud Storage, 2008**

Cloud Storage usa servidores remotos para guardar datos, como archivos, datos de la empresa, videos o imágenes. Los usuarios suben los datos a los servidores a través de una conexión a Internet, donde se guardan en una máquina virtual en un servidor físico. Para mantener la disponibilidad y proporcionar redundancia

#### **· El QNAP TS-433, 2022**

El QNAP TS-433 es un modelo con un procesador ARM Cortex-A55 de cuatro núcleos, 4 GB de RAM y 4GB de almacenamiento interno. Ofrece cuatro bahías para HDD y SSD de 2,5 o 3,5" con sistema hot-swap de extracción en caliente. Destaca por la posibilidad de configurar distintos tipos de RAID, como RAID 0 para rendimiento máximo, RAID 1 para redundancia de datos, y modos intermedios como RAID 5 o RAID 10.

#### **· Intel® Virtual RAID on CPU (Intel® VROC)**

Intel® Virtual RAID on CPU (Intel® VROC) es una solución RAID híbrida empresarial diseñada para unidades de estado sólido (SSD) NVMe conectadas directamente a la CPU. Utiliza la función Intel® Volume Management Device (Intel® VMD) en procesadores Intel® Xeon® escalables para mejorar la confiabilidad de conexiones NVMe a través de 48 canales PCIe\*. La principal ventaja de Intel VROC es la capacidad de conectar SSD NVMe directamente a la CPU, permitiendo la creación de matrices RAID sin necesidad de un adaptador de bus host RAID (HBA), lo que simplifica el rendimiento y la potencia de las SSD NVMe. Intel VROC está diseñado específicamente para estas unidades y apoya la transición del mercado hacia tecnologías más rápidas.

## **Casos de Fracaso**

- 70 millones, el ejército de EEUU. 2009.**

El inspector general de la Administración Nacional de Archivos y Registros (NARA) investigó una posible violación de datos que afecta a registros de 70 millones de veteranos militares de EE. UU. Esto ocurrió debido a un disco duro defectuoso enviado para reparación y reciclaje sin borrar los datos. El disco duro formaba parte de una base de datos con información detallada de veteranos, incluyendo números de Seguro Social. Al final no se logró recuperar la información que se encontraba.

- Contrato de empresa de pérdida de datos eliminado. 10 septiembre 2008**

PA Consulting, perdió datos de delincuentes retenidos en sus discos y rescindió su contrato de £1.5 millones después de una investigación. Se descubrió que los datos se perdieron después de que se transfirieron de forma “segura” a la empresa. PA Consulting se disculpó y asumió la responsabilidad.

- Ameritrade advierte a 200.000 clientes de datos perdidos. 2005**

Ameritrade Inc. ha informado a 200,000 clientes actuales y anteriores que se había perdido la información de una cinta de respaldo que contenía su información personal, abarcando los años 2000-2003. La información incluía detalles de cuentas y datos personales, como números de cuenta y números de Seguro Social.

## **SIMULADORES**

## **CALCULADORAS RAID**

Esta proporciona una estimación del almacenamiento con unidades mixtas y configuraciones de RAID.

Los resultados son estimaciones destinadas a representar nuevas capacidades de grupo de almacenamiento. El tamaño máximo de volumen individual admitido por Synology NAS varía según la arquitectura de la CPU y puede ser: 16 TB, 108 TB, 200 TB o 1 PB.

## **¿QUÉ ES EL SHR?**

Es un sistema de gestión RAID automatizado, es decir podrá utilizar discos duros de diferentes tamaños en su servidor NAS. En donde no es posible en otros tipos de RAID.

## **¿CÓMO FUNCIONA?**

Una matriz RAID tradicional utiliza sólo el disco más pequeño de un pool de almacenamiento, lo que significa que los discos más grandes siempre tendrán por defecto el espacio de almacenamiento del disco duro más pequeño cuando se crea el RAID. La única forma de aumentar el tamaño del pool de almacenamiento sería hacer que todos los discos tuvieran el mismo tamaño.

## **LA GRAN VENTAJA**

El SHR ofrece una gran ventaja, ya que la mezcla de discos duros de diferente capacidad le da la posibilidad de utilizar discos duros más pequeños al principio y con el tiempo actualizarlos a discos duros más grandes si lo desea.

El SHR es diferente porque divide el espacio de almacenamiento del disco duro en trozos más pequeños que permiten utilizar el espacio de almacenamiento adicional.

## **BIBLIOGRAFÍA**

Bonshor, G. (2016, April 28). What is RAID? – RAID 0, 1, 5 & 10 Explained With Images | Play3r.

<https://play3r.net/reviews/storage-reviews/what-is-raid-raid-0-1-5-10-explained-with-images/>  
(Bonshor, 2016)

Solutions, GlobalSuite (2023), *ISO 27000 y el conjunto de estándares de Seguridad de la información.*

(Solutions, 2023)

<https://www.globalsuitesolutions.com/es/la-familia-de-normas-iso-27000/>

Solutions, GlobalSuite (2023), *¿ISO 27001 y para qué sirve?*

(Solutions, 2023)

<https://www.globalsuitesolutions.com/es/que-es-la-norma-iso-27001-y-para-que-sirve/>

Fuente Propia, *ITIL-Ingeniería de Software I 2023*

# **SAN Y NAS**

## **MANEJO E IMPLEMENTACIÓN DE ARCHIVOS**



# **INDÍCE:**

<b>INTRODUCCIÓN</b>	<b>1</b>
<b>SAN</b>	<b>2</b>
<b>NAS</b>	<b>3</b>
<b>DIFERENCIAS ENTRE SAN Y NAS</b>	<b>4</b>
<b>IMPLEMENTACIÓN DE UNA SAN O NAS</b>	<b>5</b>
<b>POLÍTICAS ISO</b>	<b>6-7</b>
<b>PROVEEDORES IAAS</b>	<b>8</b>
<b>CASOS DE ÉXITO Y FRACASO</b>	<b>9</b>
<b>BIBLIOGRAFÍA</b>	<b>10</b>

# INTRODUCCIÓN

El almacenamiento en red es esencial para cualquier empresa que maneje grandes cantidades de datos. Las tecnologías de almacenamiento en red más comunes son SAN y NAS. Una SAN es una red independiente de servidores que proporciona almacenamiento de alto rendimiento y alta disponibilidad para aplicaciones críticas. Por otro lado, una NAS funciona como un servidor de archivos que permite compartir archivos y recursos de manera más sencilla. Los proveedores de servicios en la nube también pueden ofrecer servicios de almacenamiento en la nube que utilizan tecnologías similares a SAN y NAS para proporcionar almacenamiento a sus clientes, permitiéndoles acceder y administrar datos de manera eficiente a través de la infraestructura de la nube. La elección entre SAN y NAS dependerá de las necesidades específicas de cada empresa, como el rendimiento, la escalabilidad, la facilidad de uso y el costo.

# SAN:

SAN significa Storage Area Network, que es una red independiente de servidores que proporciona almacenamiento de alto rendimiento y alta disponibilidad para aplicaciones críticas. Los componentes de una SAN incluyen switches, adaptadores de host y dispositivos de almacenamiento, como discos duros y unidades de estado sólido. Una controladora RAID se encarga de combinar varios discos en una sola unidad de disco lógica. Una SAN es más rápida y escalable que un NAS, y ofrece una mayor flexibilidad en términos de configuración y administración. Sin embargo, una SAN es más compleja y costosa que un NAS. Las ventajas de una SAN incluyen alto rendimiento, escalabilidad y flexibilidad, mientras que las desventajas incluyen alto costo y complejidad. La elección entre SAN y NAS dependerá de las necesidades específicas de cada empresa, como el rendimiento, la escalabilidad, la facilidad de uso y el costo.

# NAS:

NAS significa Network Attached Storage, que es un dispositivo de almacenamiento dedicado conectado a una red local. Funciona como un servidor de archivos que permite compartir archivos y recursos de manera más sencilla. Los componentes de un sistema NAS incluyen dispositivos de almacenamiento, procesador, memoria, sistema operativo y una controladora RAID que se encarga de combinar varios discos en una sola unidad de disco lógica. En comparación con una SAN, un NAS es más fácil de usar y más económico, pero tiene un menor rendimiento y menos flexibilidad. Sin embargo, un NAS es escalable y puede ser una buena opción para pequeñas y medianas empresas que necesitan almacenamiento en red.

# DIFERENCIAS ENTRE SAN Y NAS:

La principal diferencia entre SAN y NAS es que una SAN es independiente de servidores, mientras que un NAS funciona como un servidor de archivos. Una SAN proporciona almacenamiento de alto rendimiento y alta disponibilidad para aplicaciones críticas, mientras que un NAS es más fácil de usar y más económico. Una SAN es más rápida y escalable que un NAS, pero también es más compleja y costosa. La elección entre SAN y NAS dependerá de las necesidades específicas de cada empresa, como el rendimiento, la escalabilidad, la facilidad de uso y el costo.

# IMPLEMENTACIÓN DE UNA SAN O NAS:

La implementación de una SAN o NAS implica varios pasos, que incluyen evaluar las necesidades de almacenamiento, seleccionar los componentes adecuados, configurar la red y los dispositivos de almacenamiento, configurar los servidores y los clientes, y probar y ajustar el sistema. Es importante tener en cuenta las medidas de seguridad necesarias para proteger los datos almacenados en la red, como la autenticación, el control de acceso, la confidencialidad de datos, la integridad de datos y el no repudio.

# POLÍTICAS ISO:

Existen tres políticas ISO relevantes para la gestión de la seguridad de la información y el almacenamiento en red:

1. Estándar ISO/IEC 27001: Este estándar ayuda a las organizaciones a establecer la política y los objetivos de gestión de la seguridad de la información y a comprender cómo se pueden gestionar los aspectos importantes, aplicar los controles necesarios y establecer objetivos claros para mejorar la seguridad de la información.
2. Estándar ISO/IEC 27040: Este estándar ofrece una visión general de los conceptos de seguridad de almacenamiento y las definiciones relacionadas. Incluye orientación sobre las amenazas, el diseño y los aspectos de control asociados con escenarios de almacenamiento típicos y áreas de tecnología de almacenamiento. Además, proporciona referencias a otras Normas Internacionales e informes técnicos que abordan prácticas y técnicas existentes que se pueden aplicar a la seguridad de almacenamiento.

# POLÍTICAS ISO:

3. Estándar ISO/IEC 7498-2: Este estándar proporciona un modelo de referencia que establece servicios y mecanismos relacionados con la seguridad en las comunicaciones. Se compone de cinco clases de servicios de seguridad: Integridad de Datos, Autenticación, Confidencialidad de Datos, Control de Acceso y No Repudio. Estos servicios de seguridad se implementan mediante mecanismos de seguridad específicos y generalizados, que incluyen cifrado, firma digital, control de acceso y otros. La elección de los mecanismos depende de los niveles del modelo OSI y de los requisitos de seguridad específicos.

# PROVEEDORES IAAS:

En un entorno de IaaS (Infraestructura como Servicio), es común utilizar SAN y NAS para satisfacer las necesidades de almacenamiento de los servidores virtuales. Los servidores pueden conectarse a dispositivos SAN para acceder a almacenamiento de alto rendimiento y alta disponibilidad para aplicaciones críticas. Al mismo tiempo, se pueden utilizar dispositivos NAS para compartir archivos y recursos de manera más sencilla.

# CASOS DE ÉXITO Y FRACASO:

Un caso de éxito es el de Walgreens, que implementó tecnología SAN para gestionar grandes volúmenes de datos de manera más eficiente, mejorando la velocidad de acceso y la disponibilidad de la información. Además, se garantizó que los datos críticos estuvieran protegidos y disponibles en todo momento, lo que permitió a Walgreens optimizar sus operaciones y ofrecer servicios más efectivos a sus clientes.

Un caso de fracaso es el de Target, que sufrió una violación de seguridad en 2013 que comprometió los datos personales de millones de clientes. La violación se debió a una vulnerabilidad en el sistema de almacenamiento de la empresa, que permitió a los atacantes acceder a los datos de las tarjetas de crédito de los clientes. Target no había implementado medidas adecuadas de seguridad de la información y no había detectado la violación de seguridad a tiempo, lo que resultó en una pérdida significativa de confianza de los clientes y una disminución en las ventas.

# BIBLIOGRAFÍA

IONOS Digitalguide. “Qué es el Network Attached Storage (NAS).” IONOS Digitalguide. Accedido el 30 de octubre de 2023. <https://www.ionos.es/digitalguide/servidores/know-how/que-es-el-network-attached-storage-nas/>.

IONOS Digitalguide. “SAN: Almacenamiento para grandes cantidades de datos.” IONOS Digitalguide. Accedido el 30 de octubre de 2023.

<https://www.ionos.es/digitalguide/servidores/know-how/san-almacenamiento-para-grandes-cantidades-de-datos/>.

Infodisa. “Comparativa SAN vs NAS.” Infodisa. Accedido el 30 de octubre de 2023. <https://www.infodisa.com/es/comparativa-san-vs-nas/>.

[Nombre del creador del video, si está disponible]. “Título del video” [Vídeo]. YouTube. Fecha de publicación en YouTube, si está disponible. URL del video <https://www.youtube.com/watch?v=3yZDDr0JKVc>.

VMware. “Storage Area Network (SAN).” VMware. Accedido el 30 de octubre de 2023. <https://www.vmware.com/es/topics/glossary/content/storage-area-network-san.html>.