

Esquema interno

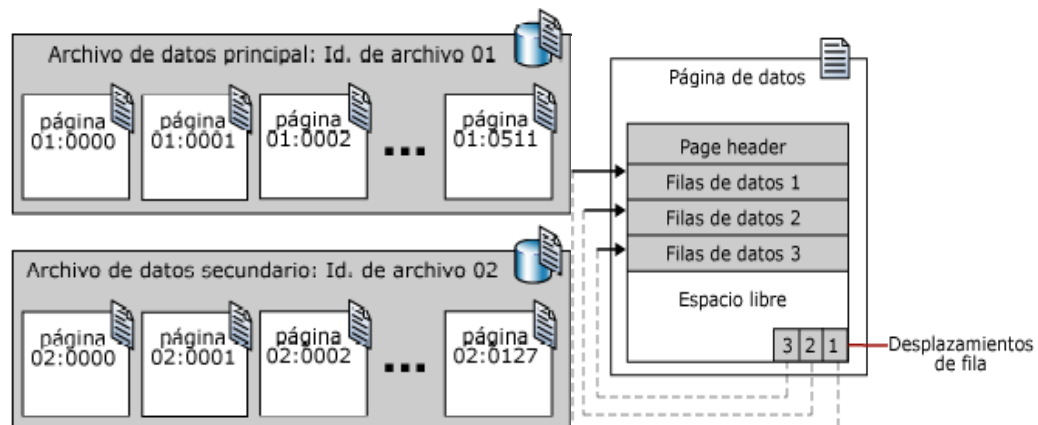
Archivos de base de datos

Las bases de datos utilizan dos tipos de archivos:

- Archivos de datos
 - Archivo de dato principal.
El archivo de datos principal es el punto de partida de la base de datos y apunta a los otros archivos de la base de datos. Cada base de datos tiene un archivo de datos principal.
 - Archivos de datos secundarios
Los archivos de datos secundarios son todos los archivos restantes. Puede que no tengan archivos de datos secundarios, o pueden tener varios archivos de datos secundarios.
 - Archivos de índices
Construidos en un o más campo(s) del archivo (campo de indexación).
- Archivos de registro
Los archivos de registro almacenan toda la información de registro que se utiliza para recuperar la base de datos. Como mínimo, tiene que haber un archivo de registro por cada base de datos, aunque puede haber varios

Páginas de archivos

La unidad fundamental del almacenamiento de datos es la página. El espacio en disco asignado a un archivo de datos de una base de datos se divide lógicamente en páginas numeradas de forma contigua de 0 a n . Las operaciones de E/S de disco se realizan en el nivel de página. Leen y graban páginas de datos enteras.



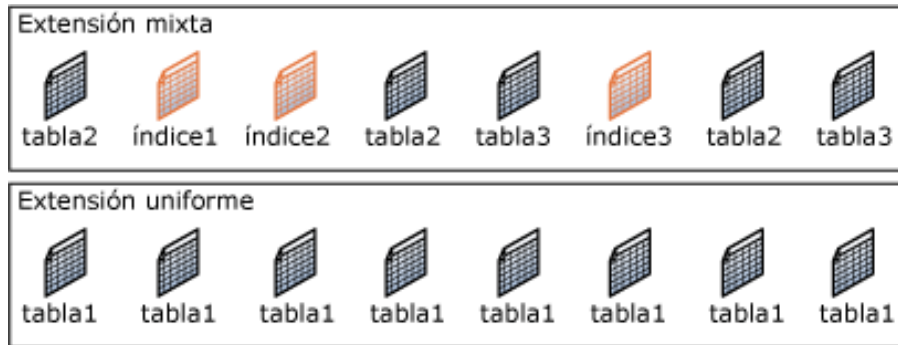
Cada página empieza con un encabezado que se utiliza para almacenar la información del sistema acerca de la página.

- el número de página.
- el tipo de página.
- el espacio disponible en la página.
- la identificación de asignación del objeto propietario de la página.

Las bases de datos transfieren los datos del almacenamiento primario al secundario en bloques nKB (8KB) denominados páginas, que corresponden a un subconjunto de registros del archivo.

Una **extensión** consta de ocho páginas contiguas físicamente, es decir 64 KB.

- Las extensiones uniformes son propiedad de un único objeto; sólo el objeto propietario puede utilizar las ocho páginas de la extensión.
- Las extensiones mixtas, que pueden estar compartidas por hasta ocho objetos. Cada una de las 8 páginas de la extensión puede ser propiedad de un objeto diferente.



Indices

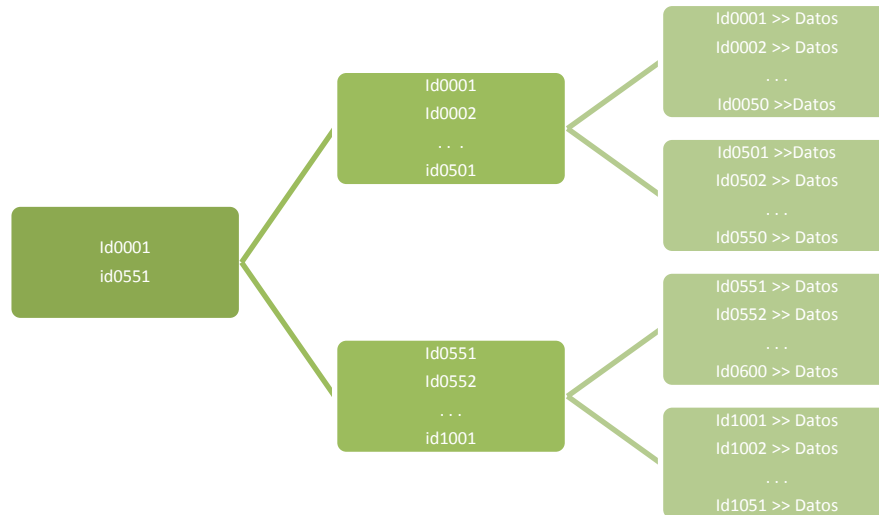
Un índice es una estructura de disco asociada con una tabla o una vista que acelera la recuperación de filas. Un índice contiene claves generadas a partir de una o varias columnas de la tabla o la vista. Dichas claves están almacenadas en una estructura (árbol b+) que permite obtener de forma rápida y eficiente la fila o filas asociadas a los valores de cada clave.

Tipos de índices:

- Agrupado
 - Los índices agrupados ordenan y almacenan las filas de los datos de la tabla o vista de acuerdo con los valores de la clave del índice. Son columnas incluidas en la definición del índice. Sólo puede haber un **índice clúster** por cada tabla, porque las filas de datos sólo pueden estar ordenadas de una forma.
 - La única ocasión en la que las filas de datos de una tabla están ordenadas es cuando la tabla contiene un índice clúster, y se denomina tabla agrupada. Si una tabla no tiene un **índice clúster**, sus filas de datos están almacenadas en una estructura sin ordenar denominada montón.
- No agrupado
 - Los índices no agrupados tienen una estructura separada de las filas de datos. Un índice no agrupado contiene los valores de clave de índice no agrupado y un puntero a la fila de datos que contiene el valor clave.
 - El puntero de una fila de **índice no agrupado** hacia una fila de datos se denomina **localizador de fila**. La estructura del localizador de filas depende de si las páginas de datos están almacenadas en un montón o en una tabla agrupada. Si están en un montón, el localizador de filas es un puntero hacia la fila. Si están en una tabla agrupada, el localizador de fila es la clave de índice agrupada.
- Único
 - Un índice único garantiza que la clave de índice no contenga valores duplicados y, por tanto, cada fila de la tabla o vista es en cierta forma única.
 - Tanto los índices agrupados como los no agrupados pueden ser únicos.

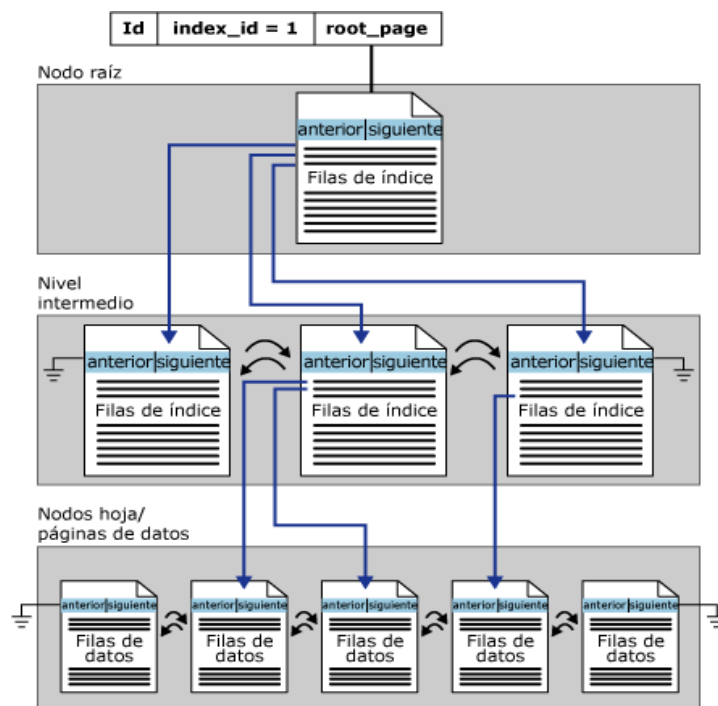
Arboles B+

Los árboles-B+ es la técnica más utilizada para la organización de archivos indizados. La principal característica de estos árboles es que todas las claves se encuentran en las hojas y por lo tanto cualquier camino desde la raíz hasta alguna de las claves, tienen la misma longitud. Los arboles-B+ ocupan un poco más de espacio que los árboles-B, y esto ocurre al existir duplicidad en algunas claves. Pero evita la operación de reorganización del árbol que es muy costosa en los arboles-B.



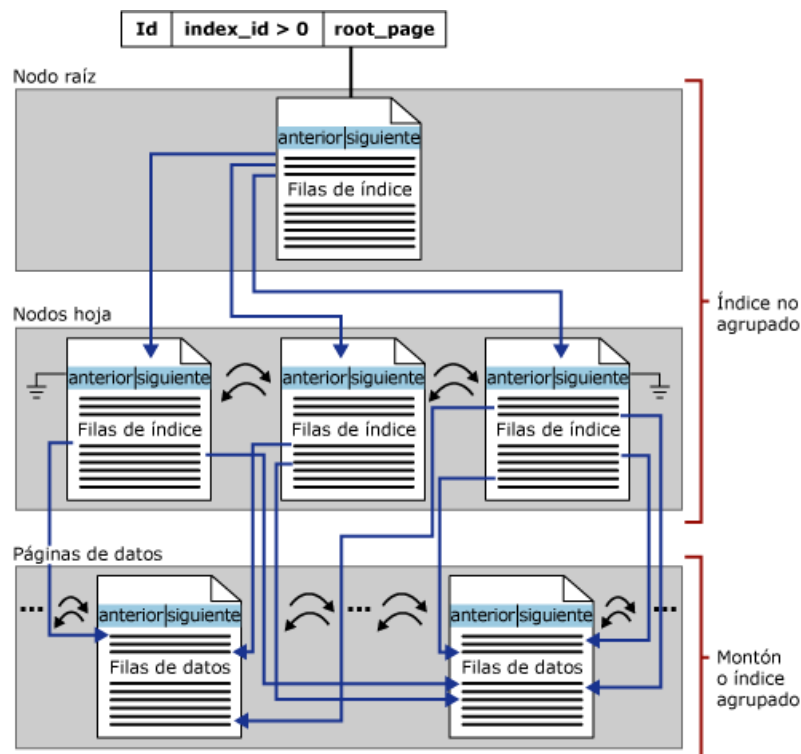
Archivos ordenados

- Los archivos ordenados son los que tienen un índice ordenado (índice cluster). Las filas de datos están almacenadas en un orden basado en la clave del índice. El índice clúster se implementa como una estructura de árbol b que admite la recuperación rápida de las filas a partir de los valores de las claves del índice.



Archivos no ordenados

- Son archivos de organización secuencial que no tienen ningún índice. Las filas de datos no se almacenan en ningún orden concreto, ni tampoco hay un orden concreto en la secuencia de las páginas de datos. Las páginas de datos no están vinculadas.
- Los índices secundarios tienen la misma estructura de árbol b+ que los índices primarios, diferenciándose por:
 - Las filas de datos de la tabla subyacente no están ordenadas ni almacenadas basándose en sus claves del índice secundario.
 - Los índices secundarios se pueden definir en una tabla con un índice primario o en un montón. Cada fila del índice secundario contiene un **valor de clave y un localizador de fila**. Este localizador apunta a la fila de datos del índice primario o el montón que contiene el valor de clave.
 - Si la tabla es un montón, lo que significa que no tiene ningún índice agrupado, el localizador de fila es un puntero a la fila. El puntero se genera a partir del identificador (Id.) de archivo, el número de página y el número de la fila dentro de la página.
 - Si la tabla tiene un índice clúster el localizador de fila es la clave del índice clúster para la fila. Si el índice clúster no es un índice único, claves duplicadas se convierten en únicas agregando un valor generado internamente denominado **valor de unicidad**.

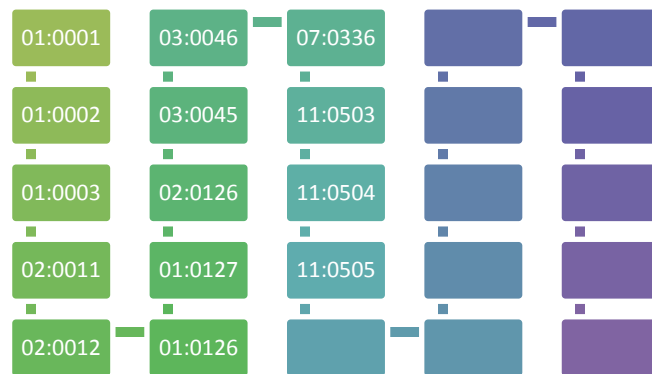


Administración del búfer

El componente de administración de búfer consta de dos partes:

1. el administrador de búfer para obtener acceso a las páginas de bases de datos y actualizarlas.
2. la caché del búfer (también conocida como *grupo de búferes*) para reducir la E/S de archivos de base de datos.

Un búfer es una página de 8 KB en memoria (el mismo tamaño que una página de índice o de datos). Por lo tanto, la caché del búfer está dividida en páginas de 8 KB.



El administrador de búfer administra las funciones para la lectura de páginas, de índice o de datos, de los archivos de disco de base de datos en la caché del búfer y para la escritura de páginas modificadas nuevamente en el disco.

Una página permanece en la caché del búfer hasta que el administrador del búfer necesite el área del búfer para leer en ella nuevos datos. Los datos sólo vuelven a escribirse en el disco si se han modificado. Los datos de la caché del búfer se pueden modificar varias veces antes de que se vuelvan a escribir en el disco.

E/S de disco

Las operaciones de E/S de disco que realiza el administrador de búfer tienen las siguientes características:

- Todas las operaciones de E/S se realizan de forma asincrónica, lo que permite que el subproceso de llamada siga con el procesamiento mientras la operación de E/S se realiza en segundo plano.
- Las operaciones de E/S de múltiples páginas se logran con E/S por dispersión y recopilación, que permite transferir datos a áreas no contiguas de memoria, o desde ellas.

Operaciones de E/S

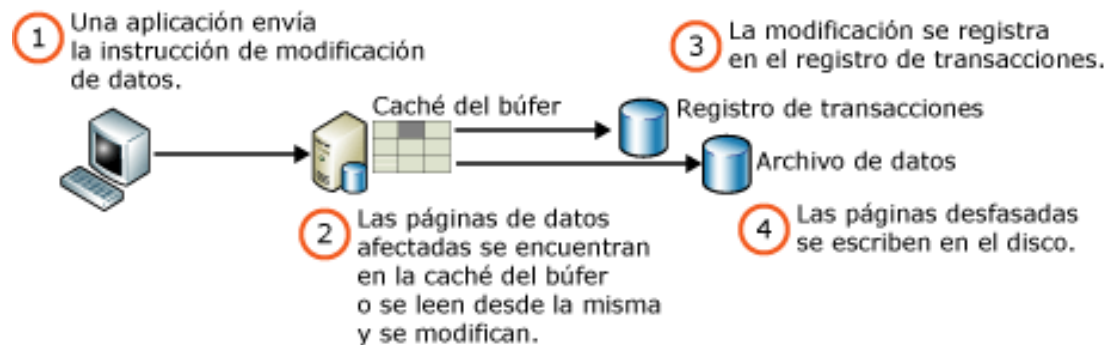
La E/S incluye lecturas lógicas y físicas.

- Se produce una **lectura lógica** cada vez que se solicita una página desde la caché del búfer. Si la página no se encuentra actualmente en la caché del búfer, una lectura física copia primero la página desde el disco a la caché.
- Se produce una **escritura lógica** cuando se modifican los datos en una página de la caché del búfer. Una escritura física se produce cuando la página se escribe de la caché del búfer al disco.

Cuando se modifica una página en la caché del búfer, no se vuelve a escribir inmediatamente en el disco; se la marca como *desfasada*. Una página puede tener más de una escritura lógica antes de que sea escrita físicamente en el disco.

Para cada escritura lógica, se inserta una entrada del registro de transacciones en la caché del registro que registra la modificación. Las entradas del registro deben escribirse en el disco antes de que la página desfasada asociada se elimine de la caché del búfer y se escriba en el disco.

Una técnica conocida como *registro de escritura anticipada* evita escribir una página desfasada antes de que la entrada de registro asociada se escriba en el disco.



Cuando el administrador del búfer escribe una página, busca páginas desfasadas adyacentes que puedan incluirse en una sola operación de escritura y recopilación. Las páginas adyacentes tienen Id. de página consecutivos y son del mismo archivo; las páginas pueden o no, ser contiguas en la memoria.

Justo antes de que se escriba una página se establece un bloqueo temporal UP (actualización) para la E/S en la página. Este bloqueo temporal evita que nadie modifique la página durante la escritura, pero permite leerla.