

INDICE

1.	INTRODUCCION	3
2.	CONCEPTOS BÁSICOS	3
3.	ÍNDICES ORDENADOS.....	3
3.1	Índice primario	4
3.1.1	Índices densos y dispersos	4
3.1.2	Índices multinivel	5
3.1.3	Actualización del índice	6
3.2	Índices secundarios	6

1. INTRODUCCION

La utilización de las técnicas de indexación y asociación sirven para realizar consultas en una base de datos por medio de código SQL.

Muchas consultas hacen referencia sólo a una pequeña parte de los registros de un archivo. Por ejemplo, la pregunta «Buscar todas las cuentas de la sucursal Pamplona» o «Buscar el saldo del número de cuenta C-101» hace referencia solamente a una fracción de los registros de la relación cuenta.

No es eficiente para el sistema tener que leer cada registro y comprobar que el campo nombre- sucursal contiene el nombre «Pamplona» o el valor C-101 del campo número-cuenta. Lo más adecuado sería que el sistema fuese capaz de localizar directamente estos registros. Para facilitar estas formas de acceso se diseñan estructuras adicionales que se asocian con archivos.

2. CONCEPTOS BÁSICOS

- Tiempo de acceso. El tiempo que se tarda en buscar un determinado elemento de datos, o conjunto de elementos, usando la técnica en cuestión.
- Tiempo de inserción. El tiempo empleado en insertar un nuevo elemento de datos. Este valor incluye el tiempo utilizado en buscar el lugar apropiado donde insertar el nuevo elemento de datos, así como el tiempo empleado en actualizar la estructura del índice.
- Tiempo de borrado. El tiempo empleado en borrar un elemento de datos. Este valor incluye el tiempo utilizado en buscar el elemento a borrar, así como el tiempo empleado en actualizar la estructura del índice.
- Espacio adicional requerido. El espacio adicional ocupado por la estructura del índice. Como normalmente la cantidad necesaria de espacio adicional suele ser moderada, es razonable sacrificar el espacio para alcanzar un rendimiento mejor.

3. ÍNDICES ORDENADOS

Para permitir un acceso directo rápido a los registros de un archivo se puede usar una estructura de índice. Cada estructura de índice está asociada con una clave de búsqueda concreta. Al igual que en el catálogo de una biblioteca, un índice almacena de manera ordenada los valores de las claves de búsqueda, y asocia a cada clave los registros que contienen esa clave de búsqueda.

Los registros en el archivo indexado pueden estar a su vez almacenados siguiendo un orden, semejante a como los libros están ordenados en una biblioteca por algún atributo. Un archivo puede tener varios índices según diferentes claves de búsqueda.

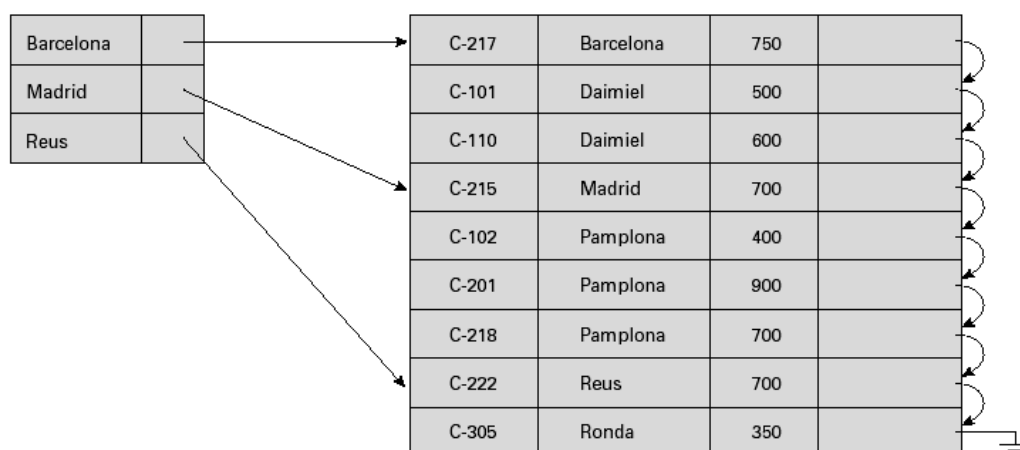
Si el archivo que contiene los registros está ordenado secuencialmente, el índice cuya clave de búsqueda especifica el orden secuencial del archivo es el índice primario. (El término índice primario se emplea algunas veces para hacer alusión a un índice según una clave primaria. Sin embargo, tal uso no es normal y debería evitarse.) Los índices primarios también se llaman índices con agrupación (clustering indices.) La clave de búsqueda de un índice primario es normalmente la clave primaria, aunque no es así necesariamente.

Los índices cuyas claves de búsqueda especifican un orden diferente del orden secuencial del archivo se llaman índices secundarios o índices sin agrupación (non clustering indices).

3.1 Índice primario

En este apartado se asume que todos los archivos están ordenados secuencialmente según alguna clave de búsqueda.

Estos archivos con índice primario según una clave de búsqueda se llaman archivos secuenciales indexados. Representan uno de los esquemas de índices más antiguos usados por los sistemas de bases de datos. Se emplean en aquellas aplicaciones que demandan un procesamiento secuencial del archivo completo así como un acceso directo a sus registros.

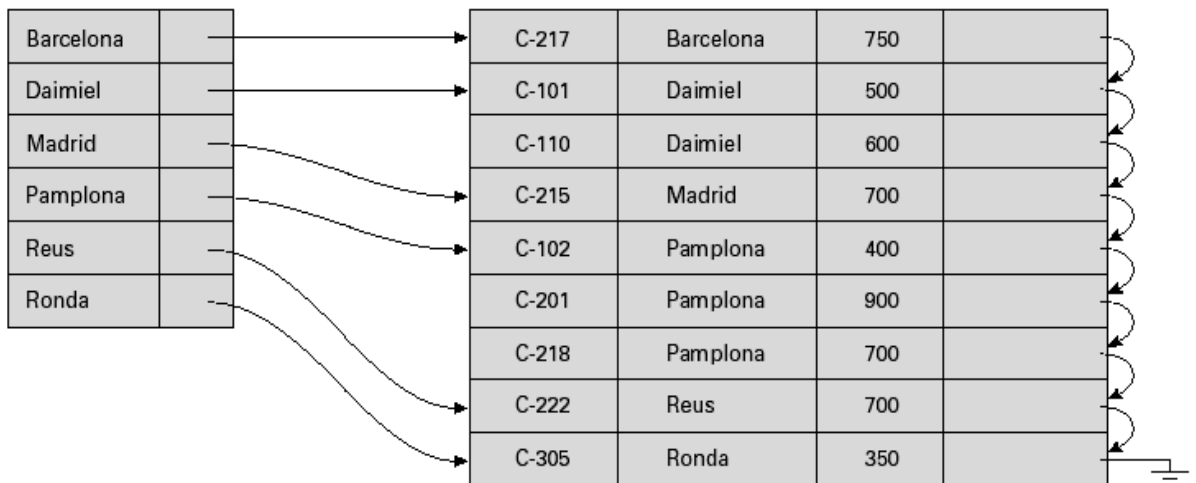


3.1.1 Índices densos y dispersos

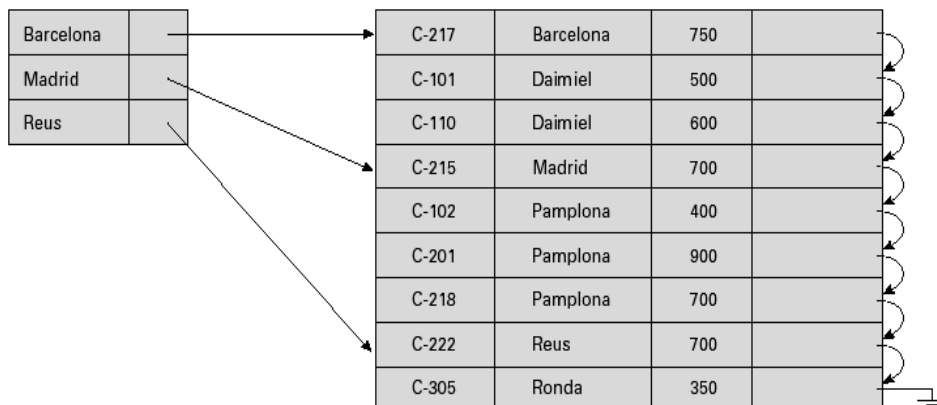
Un registro índice o entrada del índice consiste en un valor de la clave de búsqueda y punteros a uno o más registros con ese valor de la clave de búsqueda. El puntero a un registro consiste en el identificador de un bloque de disco y un desplazamiento en el bloque de disco para identificar el registro dentro del bloque. Hay dos clases de índices ordenados que se pueden emplear:

- Índice denso. Aparece un registro índice por cada valor de la clave de búsqueda en el archivo. El registro índice contiene el valor de la clave y un puntero al primer registro con ese valor de la clave de búsqueda. El resto de registros con el mismo valor de la clave de búsqueda se almacenan consecutivamente después del primer registro, dado

que, ya que el índice es primario, los registros se ordenan sobre la misma clave de búsqueda.



- Índice disperso. Sólo se crea un registro índice para algunos de los valores. Al igual que en los índices densos, cada registro índice contiene un valor de la clave de búsqueda y un puntero al primer registro con ese valor de la clave. Para localizar un registro se busca la entrada del índice con el valor más grande que sea menor o igual que el valor que se está buscando. Se empieza por el registro apuntado por esa entrada del índice y se continúa con los punteros del archivo hasta encontrar el registro deseado.

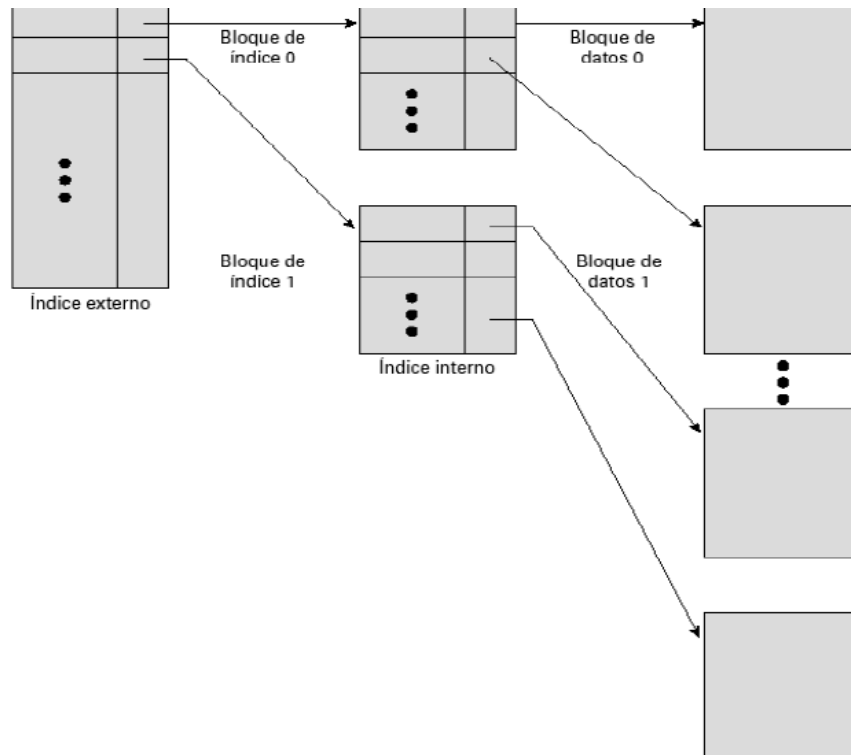


3.1.2 Índices multinivel

Incluso si se usan índices dispersos, el propio índice podría ser demasiado grande para un procesamiento eficiente. En la práctica no es excesivo tener un archivo con 100.000 registros, con 10 registros almacenados en cada bloque. Si tenemos un registro índice por cada bloque, el índice tendrá 10.000 registros. Como los registros índices son más pequeños que los registros de datos, podemos suponer que caben 100 registros índices en un bloque. Por tanto, el índice ocuparía 100 bloques. Estos índices de mayor tamaño se almacenan como archivos secuenciales en disco.

Para reducir el número de accesos a disco para los registros del índice, tratar de mantener el índice primario sobre disco como un archivo secuencial y construir un índice disperso en él.

- índice externo – un índice disperso de índice primario
- índice interno – el archivo del índice primario



3.1.3 Actualización del índice

Sin importar el tipo de índice que se esté usando, los índices se deben actualizar siempre que se inserte o borre un registro del archivo.

3.2 Índices secundarios

Los índices secundarios deben ser densos, con una entrada en el índice por cada valor de la clave de búsqueda, y un puntero a cada registro del archivo. Un índice primario puede ser disperso, almacenando sólo algunos de los valores de la clave de búsqueda, ya que siempre es posible encontrar registros con valores de la clave de búsqueda intermedios mediante un acceso secuencial a parte del archivo, como se describió antes. Si un índice secundario almacena sólo algunos de los valores de la clave de búsqueda, los registros con los valores de la clave de búsqueda intermedios pueden estar en cualquier lugar del archivo y, en general, no se pueden encontrar sin explorar el archivo completo.

Un índice secundario sobre una clave candidata es como un índice denso primario, excepto en que los registros apuntados por los sucesivos valores del índice no están almacenados secuencialmente.

Por lo general, los índices secundarios están estructurados de manera diferente a como lo están los índices primarios. Si la clave de búsqueda de un índice primario no es una clave

candidata, es suficiente si el valor de cada entrada en el índice apunta al primer registro con ese valor en la clave de búsqueda, ya que los otros registros podrían ser alcanzados por una búsqueda secuencial del archivo.

En cambio, si la clave de búsqueda de un índice secundario no es una clave candidata, no sería suficiente apuntar sólo al primer registro de cada valor de la clave.

El resto de registros con el mismo valor de la clave de búsqueda podrían estar en cualquier otro sitio del archivo, ya que los registros están ordenados según la clave de búsqueda del índice primario, en vez de la clave de búsqueda del índice secundario. Por tanto, un índice secundario debe contener punteros a todos los registros.

