

**CUARTA CONSULTA – INDICES**

**Presentado a:**

Julio Cesar Florez Baez

**Presentado por:**

Johan Esteban Castaño Martinez - 20191020029

Jhony Alejandro Caro Umbariba - 20191020055

Samuel Andrés Romero Bueno - 20191020127

**Equipo Número 1**

Facultad de Ingeniería.

Ciencias de la Computación II.

2 de Octubre de 2022.

**INDICE**

[**1.**](#_heading=h.30j0zll) **Índices:** 2

[**2.**](#_heading=h.1fob9te) **Tema 2:** 2

[**3.**](#_heading=h.3znysh7) **Tema 3:** 3

[**3.1.**](#_heading=h.2et92p0) **Subtema 3.1:** 3

[**3.2.**](#_heading=h.tyjcwt) **Subtema 3.2:** 3

[**3.3.**](#_heading=h.3dy6vkm) **Subtema 3.3:** 3

1. **Índices:**
   1. Primera definición:

El método de búsqueda secuencial con índices trabaja con bloques y con archivos de índices. En el archivo de índices se almacenan las claves que hacen referencia a cada bloque y la dirección de los bloques en el archivo, La búsqueda de un elemento comienza recorriendo el archivo de índices, comparando las claves allí almacenadas con la clave del elemento en cuestión.

Una vez que se determina el bloque en el cual se puede encontrar el registro buscado, se continúa la búsqueda ahora recorriendo secuencialmente dicho bloque.

La desventaja de este método es que requiere más espacio de memoria, ya que se trabaja con dos archivos: el principal, en el cual se almacenan los registros, y el de índices. Una forma de acelerar el proceso de búsqueda consiste en mantener en memoria principal el archivo de índices.[[1]](#footnote-1)

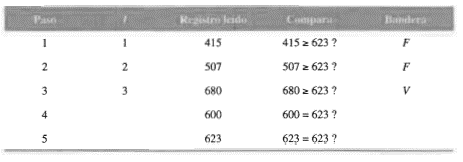


Imagen 1. Ejemplo de índices tomado de: “ (Guardati & Cairo)”

* 1. Segunda definición:
  2. Tercera definición:

1. **De un solo nivel:**
   1. Primera definición:

Los índices de un solo nivel son ficheros ordenados, sus registros tienen dos campos:

* Campo de indexación: coincide con uno de los campos del fichero de datos.
* Dirección del registro que corresponde al valor del campo de indexación.

Al ser ficheros ordenados se pueden realizar búsquedas binarias.[[2]](#footnote-2)

* 1. Segunda definición:
  2. Tercera definición:
     1. **Primarios**
        1. Primera definición:

En un archivo ordenado secuencialmente, es el indice cuya clave de búsqueda especifica el orden secuencial del archivo.

**Entradas**: registros de longitud fija. valor de la clave del primer/último registro del bloque y puntero a dicho bloque.

**Campo de indexación**: campo clave de ordenación del fichero de datos.

**Índice no denso.**

**Búsqueda binaria sobre el índice**: visita menos bloques de disco.

**Problema**: son ficheros ordenados (opciones: fichero de desbordamiento desordenado; lista enlazada de registros de desbordamiento).

**Importante**: sobre un fichero ordenado por clave sólo puede definirse un índice primario.[[3]](#footnote-3)

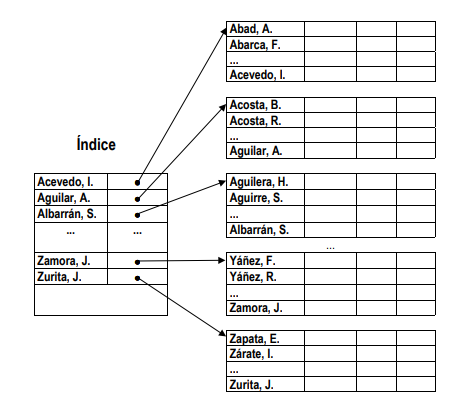


Imagen 1. Ejemplo de indice pr tomado de: <http://www3.uji.es/~mmarques/f47/teoria/tema2b.pdf>

* + - 1. Segunda definición:
      2. Tercera definición:
    1. **Secundarios**
       1. Primera definición:

Es un índice cuya clave de búsqueda especifica un orden distinto del orden secuencial del archivo.

**Campo de indexación**: cualquier campo que no sea el campo de ordenación.

* **Si es un campo clave**: índice denso.
* **Si es un campo no clave**: hay varias opciones.

**Importante**: pueden definirse varios índices secundarios sobre un mismo fichero. Los índices densos proporcionan un ordenamiento lógico de los registros según el campo de indexación.[[4]](#footnote-4)

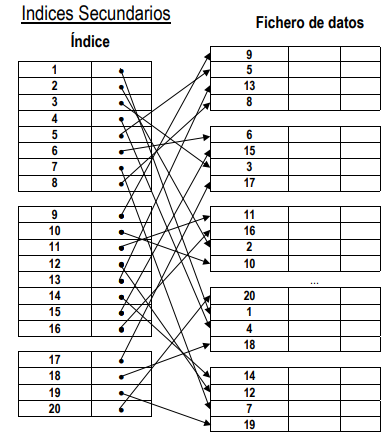


Imagen 1. Ejemplo de indice secundario tomado de: “http://www3.uji.es/~mmarques/f47/teoria/tema2b.pdf”

**Índice secundario sobre un campo no clave**:

* Una entrada por cada registro: índice denso.
* Registros de longitud variable: índice no denso y el campo de la dirección contiene una lista de punteros.
* Registros de longitud fija: índice no denso con un nivel extra de indirección para manejar punteros múltiples. [[5]](#footnote-5)

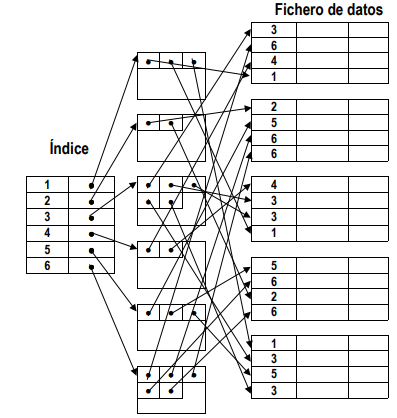


Imagen 1. Ejemplo de indice secundario tomado de: “http://www3.uji.es/~mmarques/f47/teoria/tema2b.pdf”

* + - 1. Segunda definición:
      2. Tercera definición:
    1. **Agrupado**
       1. Primera definición:

**Entradas**: registros de longitud fija. Una por cada valor distinto del campo de agrupamiento. El puntero apunta al primer bloque que contiene un registro con dicho valor.

**Campo de indexación**: campo no clave de ordenación del fichero de datos (campo de agrupamiento).

**Índice no denso**

**Búsqueda binaria sobre el índice**: visita menos bloques de disco.

**Problema**: son ficheros ordenados. (opción: reservar un bloque entero para cada valor distinto del campo de agrupamiento).

**Importante**: sobre un fichero ordenado por un campo no clave sólo puede definirse un índice de agrupamiento.[[6]](#footnote-6)

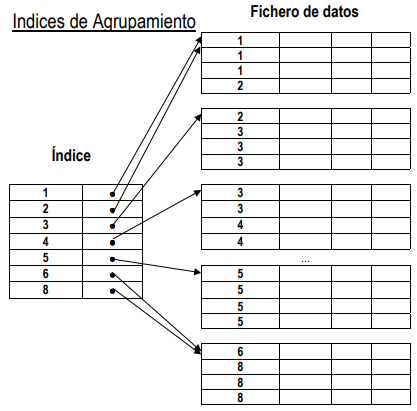


Imagen 1. Ejemplo de indice agrupado tomado de: “http://www3.uji.es/~mmarques/f47/teoria/tema2b.pdf”

Aquí se ha reservado un bloque para cada valor distinto del campo de agrupamiento. Se van añadiendo y enlazando bloques conforme sea necesario[[7]](#footnote-7)

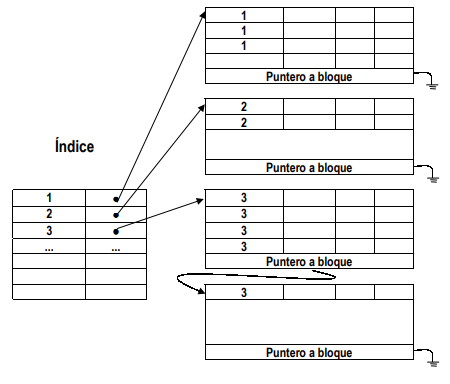


Imagen 1. Ejemplo de indice agrupado tomado de: “http://www3.uji.es/~mmarques/f47/teoria/tema2b.pdf”

* + - 1. Segunda definición:
      2. Tercera definición:

1. **Multinivel:**
   1. Primera definición:

**Objetivo**: reducir más que con la búsqueda binaria el trozo de índice en donde seguir buscando.

**Primer nivel**: fichero ordenado con entradas de tamaño fijo y un valor distinto del campo de indexación en cada una.

**Siguientes niveles**: índices primarios sobre el nivel anterior.

Número de registros por bloque: **r**

* primer nivel entradas
* segundo nivel entradas
* tercer nivel entradas ...

Se necesita un nivel más si el anterior ocupa más de un bloque.

Un índice multinivel con i1 entradas en el primer nivel tiene niveles.

Reducen el número de accesos a bloque al hacer búsquedas, pero son ficheros ordenados.[[8]](#footnote-8)

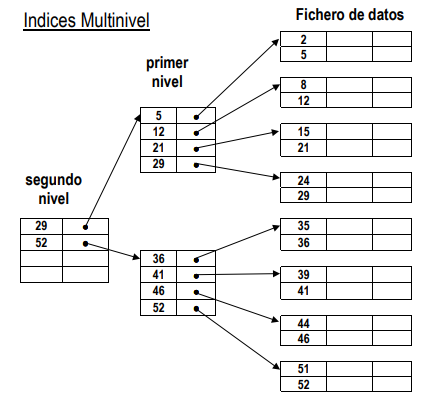


Imagen 2. Ejemplo de índice multinivel tomado de: “http://www3.uji.es/~mmarques/f47/teoria/tema2b.pdf”

* 1. Segunda definición:
  2. Tercera definición:
     1. **Multinivel con base en Primario**
        1. Primera definición:

Los intervalos se agrupan en áreas de control (puede ser o no un cilindro).

Dentro de los intervalos de control se pueden dejar espacios libres al final de los mismos y en un área puede haber intervalos completamente vacíos.

El tamaño de un área de control suele estar definido por el sistema, lo que se permite es definir el número de intervalos que se quiere que estén vacíos. Podemos definir la longitud de los intervalos de control.

**área de Datos + área de índices = Clúster**

En el área de índices se tendrá un árbol B+. Cada nodo del árbol será un intervalo de control. En las hojas se encuentran todas las claves y los nodos de las hojas están enlazados por punteros. Las hojas forman lo que se llama conjunto de secuencias. Los elementos de dicho conjunto son los nodos con entradas que serán: como clave la mayor contenida en un intervalo de control del área de datos, y un puntero a ese intervalo del área de datos.

Cuando hay varios intervalos vacíos habrá nodos que lo indiquen y un puntero a esos intervalos.

El acceso directo se hace con la búsqueda en el árbol. El acceso secuencial se hace empleando los punteros horizontales que enlazan las hojas del índice.

Veamos un esquema que representa cómo sería el árbol del área de índices, y cómo estaría unido al área de datos.

Los registros de datos pueden ser de longitud fija o variable, y al principio de cada intervalo hay unos caracteres de control que indican el nivel de ocupación (interesa al hacer un recorrido secuencial).

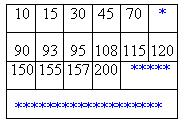
Al eliminar un registro los que estén a su derecha se moverán a la izquierda dejando siempre los espacios libres al final del intervalo. Y si algún intervalo quedará vacío aparecerá en el conjunto de secuencias como una entrada de vacío (esto se llama reclamación dinámica de espacio libre).

Este proceso evita la necesidad de tener zonas de excedentes.

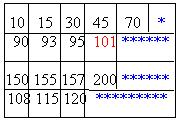
**Procesos de Partición de Intervalos y de áreas.**

En una inserción si hay espacio al final no hay problema. Si el intervalo estuviera completo lo que se hace es partir en dos el intervalo de control, pasando a ocupar una de las mitades alguno de los intervalos libres que queden. A continuación, podemos ver un esquema que nos sirve de ejemplo de una situación en la que esto ocurriría:[[9]](#footnote-9)

En un área de control los intervalos



Esto provoca que sea posible que en un área de control los intervalos no estén en orden de clave, pero lo que siempre estará ordenado es el conjunto de secuencias, por eso en el recorrido secuencial se emplean las hojas.



Lo que queremos decir con esto queda expresado en el siguiente esquema:

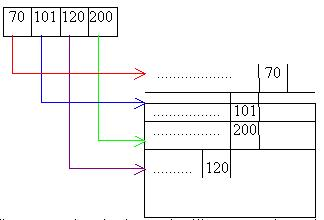


Imagen 2. Ejemplo de indice multinivel primario tomado de: “https://www.um.es/docencia/barzana/IAGP/Iagp7.html”

* + - 1. Segunda definición:
      2. Tercera definición:
    1. **Multinivel con base en Secundario:** 
       1. Primera definición:

Este método usa un fichero de datos secuencial y un índice secuencial.

Divide el espacio del soporte en tres zonas: *área de Datos, área de índices* y *área de Desborde*, las cuales se subdividen en otras según la estructura de los soportes. Los datos se organizan en pistas (que es la unidad de transferencia con la memoria principal) y éstas en cilindros lógicos.

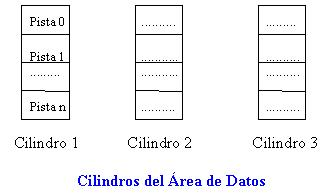


Imagen 2. Cilindros de Datos de: “https://www.um.es/docencia/barzana/IAGP/Iagp7.html”

La pista 0 de todos los cilindros se reserva para crear los llamados índices de pistas y alguna más para los excedentes del cilindro (al final).

Cuando se llena una pista se pasa a la siguiente pista libre de ese mismo cilindro (se va rellenando cilindro a cilindro). Al rellenar una pista se crea en el índice de pista una entrada con la clave de mayor orden de esa pista y un puntero a esa pista.

Al llenar un cilindro, en el área de índices se crea una entrada en el índice de cilindros con la clave de mayor orden y un puntero al cilindro.

Puede existir un tercer índice, el índice maestro, muy pequeño que apunta al índice del cilindro.

La mejora que obtenemos con este método es que al poder llevar una pista entera a memoria principal se trabaja más rápido; si al hacer una inserción excede el tamaño de la pista el/los registro/s excedente/s va/n a las pistas del área de excedentes del cilindro.

Pueden almacenarse en una zona (un cilindro o más) exclusiva para ellos. Otra forma sería reservar pistas para los registros excedentes al final de cada cilindro. Por último una tercera forma consiste en una mezcla de las dos anteriores, es decir tener pistas al final de los cilindros y una zona exclusiva.

Esta 3ª forma es la más utilizada, ya que la 1ª presenta el inconveniente de tener que hacer movimientos de las cabezas del disco para acceder a los excedentes, y la 2ª, aunque no tiene este problema tiene otros dos inconvenientes: que se puede agotar el espacio reservado o bien que por miedo a que esto ocurra se desaprovecha mucho espacio en el soporte.

Para localizarlos según que técnica empleemos tardaremos mucho (búsqueda secuencial, índice de pistas para los Excedentes). La técnica más empleada consiste en que en el índice de pistas cada entrada sean en realidad dos entradas, una para los registros almacenados normalmente, y otra para los Excedentarios. Por tanto cada entrada estará compuesta por una entrada N que será un puntero a la pista y como clave la mayor de la pista, y una entrada O que tiene un puntero a la menor entrada correspondiente a esa pista que esté en el área de Excedentes y como clave la mayor de dichos Excedentes.[[10]](#footnote-10)

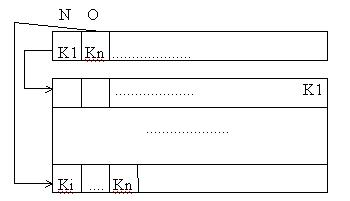


Imagen 2. Ejemplo de índices Secuenciales Multinivel de: “<https://www.um.es/docencia/barzana/IAGP/Iagp7.html>”

* + - 1. Segunda definición:
      2. Tercera definición:

1. **Forma de Búsqueda:**
   * 1. No densas o Dispersas:
        1. Primera definición:

Sólo se crea un registro índice para algunos de los valores de la clave de búsqueda.

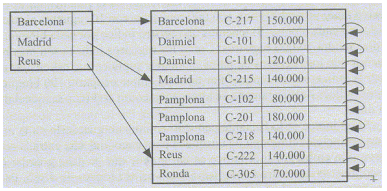


Imagen 2. Ejemplo de forma de búsqueda dispersa de: “http://gssi.det.uvigo.es/users/mramos/public\_html/bd/tema6-4.pdf”

Generalmente es más rápido localizar un registro con Índice Denso que con Disperso.[[11]](#footnote-11)

* + - 1. Segunda definición:
      2. Tercera definición:
    1. Densas:
       1. Primera definición:

Aparece un registro índice para cada valor de la clave de búsqueda en el archivo.

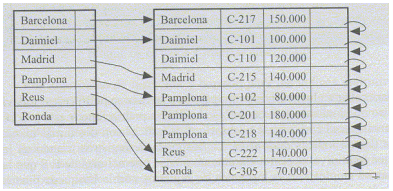


Imagen 2. Ejemplo de forma de búsqueda densa tomado de: “http://gssi.det.uvigo.es/users/mramos/public\_html/bd/tema6-4.pdf”

Los índices dispersos utilizan menos espacio, y tienen un mantenimiento menor para las inserciones y borrados.

Un buen compromiso entre tiempo de acceso y espacio adicional requerido es tener un índice disperso con una entrada del índice por cada bloque. [[12]](#footnote-12)

* + - 1. Segunda definición:
      2. Tercera definición:

**Bibliografía**

Guardati, S., & Cairo, O. (s.f.). *Estructua de datos.* Ciudad de Mexico: MC GRAW HILL TERCERA EDICIÓN.

Universidad de Murcia. (30 de Septiembre de 2022). *Organización de Ficheros y Métodos de Acceso*. Obtenido de https://www.um.es/docencia/barzana/IAGP/Iagp7.html

Universidade de Vigo. (01 de Octubre de 2022). *Bases de Datos - Indexación y Asociación*. Obtenido de http://gssi.det.uvigo.es/users/mramos/public\_html/bd/tema6-4.pdf

Universitat Jaume I. (30 de Septiembre de 2022). *Organizaciones de ficheros y estructuras de acceso*. Obtenido de http://www3.uji.es/~mmarques/f47/teoria/tema2b.pdf

1. (Guardati & Cairo, Estructua de datos) [↑](#footnote-ref-1)
2. (Universitat Jaume I, Organizaciones de ficheros y estructuras de acceso) [↑](#footnote-ref-2)
3. (Universitat Jaume I, Organizaciones de ficheros y estructuras de acceso) [↑](#footnote-ref-3)
4. (Universitat Jaume I, Organizaciones de ficheros y estructuras de acceso) [↑](#footnote-ref-4)
5. (Universitat Jaume I, Organizaciones de ficheros y estructuras de acceso) [↑](#footnote-ref-5)
6. (Universitat Jaume I, Organizaciones de ficheros y estructuras de acceso) [↑](#footnote-ref-6)
7. (Universitat Jaume I, Organizaciones de ficheros y estructuras de acceso) [↑](#footnote-ref-7)
8. (Universitat Jaume I, Organizaciones de ficheros y estructuras de acceso) [↑](#footnote-ref-8)
9. (Universidad de Murcia, Organización de Ficheros y Métodos de Acceso) [↑](#footnote-ref-9)
10. (Universidad de Murcia, Organización de Ficheros y Métodos de Acceso) [↑](#footnote-ref-10)
11. (Universidad de Vigo, Indexación y Asociación) [↑](#footnote-ref-11)
12. (Universidad de Vigo, Indexación y Asociación) [↑](#footnote-ref-12)