

1)

Considerar la siguiente consulta a ser procesada usando el procedimiento de dos fases de materialización.

$\Pi_{DNI} ((\sigma_{\text{antigüedad} > 5} \text{ bibliotecario}) \bowtie (\sigma_{\text{nombreBib} = \text{'FaMaf'}} \text{ trabajaEn}))$

Asumir la siguiente información:

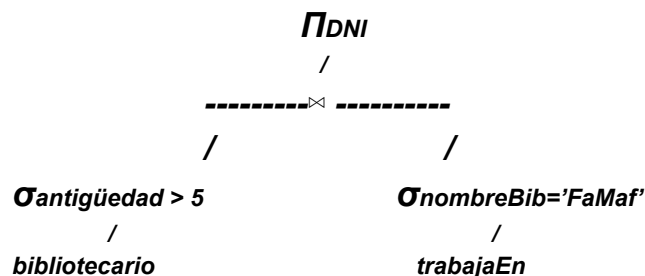
tabla	bloques	registros	Índice
<i>bibliotecario</i> (<u>DNI</u> , antigüedad)	3	600	antigüedad secundario
<i>trabajaEn</i> (<u>DNI</u> , <u>nombreBib</u>)	32	2048	nombreBib primario DNI secundario

Asumir que los bibliotecarios están equi-distribuidos por antigüedad y que las bibliotecas están equi-distribuidas en trabajaEn. Asumir que hay 35 antigüedades (de 1 a 35) en bibliotecario, que hay 200 bibliotecas en trabajaEn y que en trabajaEn hay 1024 valores de DNI. Además, asuma que en memoria entran 6 bloques máximo.

Se pide:

1. Calcular factor de selectividad para selecciones y reunión natural.
2. Indicar el tamaño de los resultados intermedios en bloques.
3. Indicar el algoritmo más eficiente a utilizar para cada operador de la consulta.
Considerar los algoritmos del apunte de procesamiento de consultas en forma de filminas.
4. Calcular el costo del algoritmo escogido para reunión natural en cantidad de transferencias de bloques.

Primero realizo el árbol binario:



Luego calculo los factores de selectividad (1)

$$Fs(\text{nombreBib}='FaMaF', \text{trabajaEn}) = 1/(V(\text{nombreBib}, \text{trabajaEn})) \\ = 1/200$$

$$Fs(\text{antigüedad} > '5', \text{bibliotecario}) = (\max(\text{antigüedad}, \text{bibliotecario}) - 5) / (\max(\text{antigüedad}, \text{bibliotecario}) - \min(\text{antigüedad}, \text{bibliotecario})) \\ = (35-5)/(35-1) = 0,88$$

$$Fs(\text{bibliotecario.DNI} = \text{trabajaEn.DNI}, \text{bibliotecario}, \text{trabajaEn}) \\ = 1/\max(V(\text{DNI}, \text{bibliotecario}), V(\text{DNI}, \text{trabajaEn})) \\ = 1/\max(600, 1024) \\ = 1/1024$$

Luego para conseguir el tamaño de los resultados intermedios en bloques (2)

$$|\sigma_{\text{antigüedad} > 5} \text{ bibliotecario}| = |\text{bibliotecario}| * 0,88 \\ = 528$$

Como entran 200 registros por bloque, requiero de los 3 bloques

$$|\sigma_{\text{nombreBib}='FaMaF'} \text{ trabajaEn}| = |\text{trabajaEn}| * 1/200 \\ = 10,24$$

Como entran 64 registros por bloque, requiero de un solo bloque.

En cuanto a los algoritmos de operadores físicos, es conveniente optar por:

- **En la selección de antigüedad el algoritmo para selección con índice secundario con comparación**
- **En la selección de nombreBib el algoritmo para selección con índice primario e igualdad en clave candidata**
- **Y finalmente en la reunión el algoritmo de merge-sort join**

No me dió el tiempo para realizar los algoritmos