

PARCIALITO DE PROCESAMIENTO DE CONSULTAS

EJERCICIO 1 - Dado el siguiente esquema sobre profesores de una facultad:

- profesores(legajo, nombre, apellido, genero, titulo)

Se pide:

1. Analice cuál es el método de acceso más eficiente para resolver la siguiente consulta y estime su costo (en bloques accedidos):

$$\sigma_{titulo="LICENCIADO" \wedge genero="M"}(profesores)$$

2. Estime la cantidad de tuplas devueltas por dicha selección.

Información que se posee:

- $n(\text{profesores}) = 500$ y $F(\text{profesores}) = 25$.
- $V(\text{titulo}, \text{profesores}) = 10$ y $V(\text{genero}, \text{profesores}) = 2$.
- La tabla profesores tiene un índice llamado prtít por título con $\text{Height}(\text{prtít}) = 2$. Este índice **no** es de clustering.
- Los valores de las columnas titulo y genero se almacenan siempre en mayúsculas.

EJERCICIO 2 - Dado el siguiente esquema que registra cuando un usuario le da “Me gusta” a una publicación:

- megusta(id_usuario, id_publicación, fecha_hora)

Se busca encontrar pares de usuarios a los que les guste una misma publicación. Se pide:

1. Analice cuál es el método de acceso más eficiente para resolver la siguiente consulta y estime su costo (en bloques accedidos):

$$megusta \bowtie_{id_usuario \neq id_usuario' \wedge id_publicacion = id_publicacion'} megusta'$$

2. Estime la cantidad de tuplas devueltas por dicha selección.

Información que se posee:

- $n(\text{megusta}) = 100,000,000$ y $F(\text{megusta}) = 1,000$
- $V(\text{id_usuario}, \text{megusta}) = 50,000$ y $V(\text{id_publicación}, \text{megusta}) = 10,000,000$
- No se cuenta con índices y se dispone de $M = 1,001$ bloques de memoria disponibles.