

# Capítulo 4

Ejemplo de ejercicio de ejecución de plan de  
evaluación con materialización

# Caso de estudio

- Sea la consulta:

$\Pi$  legajo, profe.nombre ( $\sigma$  materia.nombre = “Intro a los Algoritmos” (profe legajo  $\bowtie$  a\_cargo materia))

- La tabla materia tiene 100 registros, ocupando 20 bloques.
- La tabla profe tiene 2.000 registros, ocupando 500 bloques.
- Evaluar el costo de la consulta en transferencias de bloques.

# Árbol binario de ejecución

$\Pi_{\text{legajo, profe.nombre}}$

$\sigma_{\text{materia.nombre} == \text{"intro a los algoritmos"}}$

|

legajo ⋈ a cargo

|

Profe

|

materia

# Factor de selectividad

- $r = \text{profe}_{\text{legajo}} \bowtie_{a\_cargo} \text{materia}$
- $S = \sigma_{\text{materia.nombre} == \text{"intro a los algoritmos"}} r$
- Calculamos los factores de selectividad.
- **Resultado:** Si B clave foránea en s referenciando r:  $fs(r.A == s.B, r, s) = 1 / |r|$
- Como *a\_cargo* clave foránea en *materia* referenciando a *profe* se tiene:  
$$fs(\text{profe.legajo} = \text{materia.a\_cargo}, \text{profe}, \text{materia}) = 1 / |\text{profe}| = 1/2000$$
- **Resultado:**  $fs(A == c, r) = 1/V(A, r)$ , donde  $V(A, r)$  número de distintos valores que aparecen en  $r$  para  $A$ . Aquí se asume uniformidad.
- $fs(\text{nombre\_materia} == \text{'intro a los algoritmos'}, s) = 1 / V(\text{materia.nombre}, s) = 1/100$

# Operadores físicos

- **Reunión selectiva**

- Algoritmo de reunión por mezcla.
- Asumimos que profe y materia ordenados según *legajo, a cargo*.
- Los otros algoritmos son peores.

- **Selección**

- Como opera sobre resultado intermedio, no usa índices.
- Consideramos algoritmo de búsqueda lineal.

- **Proyección**

- Requiere recorrer todos los registros y realizar una proyección en cada uno.
- Se recorren todos los bloques de la tabla.

# Tamaño en bloques de las tablas

- Según el enunciado materia tiene 20 bloques y profe tiene 500 bloques.

# Tamaño de los resultados intermedios

- Sea  $r = \text{profe}_{\text{legajo}} \bowtie_{\text{a cargo}} \text{materia}$
- $|r| = |\text{profe}| * |\text{materia}| * \text{fs}(\text{profe.legajo} == \text{materia.a_cargo}, r) = 2000 * 100 * 1/2000 = 100$
- Hay 500 bloques de profe y hay 2000 profes; entonces en un bloque entran: 4 materias por bloque.
- Hay 20 bloques de materia y hay 100 materias; entonces en un bloque entran 5 materias.
- ¿Cuántas tuplas de la reunión selectiva entran por bloque?
- Juntar dos tuplas aproximadamente duplica el tamaño de una de las dos tuplas. Entonces estimamos que van a entrar 2 tuplas de  $r$  por bloque.
- $B_r = 100 / 2 = 50$ . O sea, necesitamos 50 bloques para la reunión selectiva.
- Sea  $s = \sigma_{\text{materia.nombre} == \text{"intro a los algoritmos"}} r$
- $|s| = |r| * \text{fs}(\text{materia.nombre} == \text{'intro\_algoritmos'}, r) = 100 * 1/100 = 1$
- $B_s = 1$ . Para guardar una tupla basta con un bloque.

# Costo de los operadores físicos

- **Reunión selectiva**

- Costo en cantidad de transferencia de bloques sin contar ordenación:
- $B_{profe} + B_{materia} = 500 + 20 = 520$
- Por lo tanto, el costo es 520 transferencias de bloques.

- **Selección**

- Costo  $B_r = 50$  transferencias de bloques
- $B_r$  denota el número de bloques conteniendo registros de la tabla resultado de la reunión selectiva.

- **Proyección**

- Costo  $B_s$  transferencias de bloques donde  $s$  es la tabla resultado de la operación de selección.
- $B_s$  denota el número de bloques conteniendo registros de la tabla  $s$
- $B_s = 1$



# Sumar los costos totales

- Costo total = costo de operaciones + costo de materialización
- $= (520 + 50 + 1) + (50 + 1) = 622$  transferencias de bloques