

Donat el següent esquema de BD:

Empleat(NIF, nom, categoria, sou, sucursal_assignada)
Sucursal(codi_sucursal, nom, adreça, codi_ciutat, nif_cap)
Ciutat(codi_ciutat, nom_ciutat, província)

I la següent consulta:

```
SELECT e.nom, c.nom_ciutat
FROM Empleat e, Ciutat c, Sucursal s
WHERE e.sou >= 30.000
AND e.sucursal_assignada = s.codi_sucursal
AND s.codi_ciutat = c.codi_ciutat
```

select empleat i sucursal i ciutat

Compara els costos de les següents estratègies d'execució i indica quina és la millor

Estratègia A:

- 1) Fer el join entre el Empleat i Sucursal (tot seleccionant Empleat)
- 2) Fer el join entre el resultat de 1 i Ciutat

Estratègia B:

- 1) Seleccionar Empleat
- 2) Fer el join entre Sucursal i Ciutat
- 3) Fer el join entre el resultats de 1 i de 2

Per cada estratègia, justifica l'aplicabilitat o no de cada algoritme; què representen els diferents elements als càlculs que feu; quin és el cost de guardar els resultats intermedis i si estan ordenats o no (i per quin atribut).

$2^0=1; 2^1=2; 2^2=4; 2^3=8; 2^4=16; 2^5=32; 2^6=64; 2^7=128; 2^8=256; 2^9=512; 2^{10}=1.024;$
 $2^{11}=2.048; 2^{12}=4.096; 2^{13}=8.192; 2^{14}=16.384; 2^{15}=32.768; 2^{16}=65.536$

• GUARDAR EL JOIN:

$$\text{card}(\text{Empleat} \bowtie \text{Sucursal}) = \text{card}(\text{Empleat}) \cdot \frac{\text{card}(\text{Sucursal})}{\text{card}(\text{Empleat})} = 30.000 \cdot \frac{3.000}{30.000} = 3.000$$

selecció

$$\frac{1}{b} = \frac{1}{50} + \frac{1}{30} = \frac{1}{(1/50 + 1/30)} = 18,75 \approx 19$$

$$\text{cost} \rightarrow \frac{3.000}{19} = 157,89 \approx 158$$

ordenat per codi de sucursal

I la següent informació al diccionari de dades:

- Empleat:
 - Cardinalitat: 30.000
 - Factor de bloqueig (b) = 30
 - Un 10% dels empleats tenen un sou més gran o igual que 30.000
 - Índex primari per NIF ($b_1 = 250$)
 - Índex secundari per sou ($b_1 = 250$)
 - Sucursal:
 - Cardinalitat: 3.000
 - Factor de bloqueig (b) = 50
 - Índex primari per codi_sucursal ($b_1 = 250$)
 - Ciutat:
 - Cardinalitat: 1.000
 - Factor de bloqueig (b) = 100
 - Índex primari per codi_ciutat ($b_1 = 250$)
- Combinació (join):**
- Bucles imbricats
 - Merge-join
 - Merge-join d'índexos
 - Bucle amb índex
- Selecció:**
- Índex primari (o clúster)
 - Índex secundari (o no clúster)
 - Intersecció d'índexos
 - Seqüencial

A 1 Join empleat i sucursal (tot seleccionant empleat)

• Bucle imbricat:

Blocks = card blocking factor

$$\text{Bucle Sucursal: } \frac{3.000}{50} = 60 \text{ (extrem)}$$

$$\text{Bucle Empleat: } \frac{30.000}{30} = 1.000 \text{ (intern)}$$

$$60 + 60 \cdot 1.000 = 60.060$$

• MERGE-JOIN:

Sucursal ordenada per codi de sucursal ✓
empleat ordenat per NIF ✗

→ ordenar empleat: $1.000 \cdot (\log_2 1.000) = 1.000 \cdot 10 = 10.000$

→ Ajustar: $10.000 + 60 = 10.060$

11060

• MERGE-JOIN INDEXOS:

No es aplicable, cal que les dues relacions tinguin índexos per al join dels atributs. ✗

• Bucle amb índex

$$1.000 + 0,1 \cdot 30.000 + 0,1 \cdot 30.000 = 7.000$$

2. JOIN R2 i Ciutat:

* Bucle imbricat:

Blocks

$$\text{Bucle } R_2 \rightarrow 167$$

$$\text{Bucle Ciutat} \rightarrow \frac{1.000}{100} = 10$$

$$10 + 10 \cdot 167 = 1.680$$

* Merge-Join: Ciutat ordenat per codi-ciutat
R1 no ordenat

→ ordenar R1: $167 \cdot \log_2(167) = 167 \cdot 8 = 1.336$

→ Ajustar: $167 + 10 = 177$

1513

$$B_{R1} + \text{card}(R1) + \text{card}(\text{join}(Ciutat, R1))$$

* Merge-Join indexos: R1 no té índex

* Bucle amb índex: $167 + 3.000 + 3.000 = 6.167$

• ESTRATÈGIA A → $7.000 + 167 + 1.513 = 8.680$

B 1 Seleccionar empleat:

* Índex primari: ✗

* Intersecció de índexos: ✗

$$sl(\text{empleat}) \cdot \text{card}(\text{emp})$$

* Índex secundari: $\frac{0,1 \cdot 30.000}{30.000} \cdot 30.000 = 3.000$

* Seqüencial: $B = \frac{\text{card}(\text{empleat})}{b} = \frac{30.000}{30} = 1.000$

• GUARDAR SELECCIÓ

→ $\text{card}(\text{Empleat}') = 30.000 \cdot 0,1 = 3.000 \rightarrow \text{cost} = B_f = \frac{3.000}{30} = 100$

ordenat per NIF

$$\hookrightarrow \text{card}(\text{Empleat'}) = 30000 \cdot 0.1 = 3000 \rightarrow \text{Cost} = B_F = \frac{3000}{30} = \boxed{100} \text{ ordenat per NIF}$$

2. join sucursal i ciutat:

* Bucle imbricat

$$\begin{aligned} B_{\text{sucursal}} &\rightarrow \frac{3000}{50} = 60 \\ B_{\text{ciutat}} &\rightarrow \frac{1000}{100} = 10 \end{aligned} \rightarrow 10 + 10 \cdot 60 = 610$$

* Merge-Join: Ciutat ordenat per codi ciutat
Sucursal no ordenat

$$\begin{aligned} \rightarrow \text{ordenar Sucursal: } 60 \cdot \log_2(60) &= 60 \cdot 6 = 360 \\ \rightarrow \text{Ajustar: } 60 + 10 &= 70 \end{aligned} \rightarrow \boxed{430}$$

* Merge-index: No es pot

$$* \text{Bucle amb index: } 60 + 3000 + 3000 = 6060$$

• Guardar els joins

$$\text{card}(\text{Sucursal} \bowtie \text{Ciutat}) = \text{card}(\text{Sucursal}) = 3000$$

$$1/b = 1/100 + 1/50 \rightarrow \frac{1}{(\frac{1}{100} + \frac{1}{50})} = 100/3 = 33.\bar{3} \approx 33 // \quad \text{Cost} = \frac{3000}{33} = 90.\bar{90} \approx \boxed{91}$$

ordenat per codi-ciutat

3. fer el join entre els resultats 1 i 2

* Bucle imbricat

$$\text{Bucle R1: } \frac{3000}{33} \approx 91$$

$$91 + 1000 \cdot 91 = 91091$$

$$\text{Bucle empleat: } \frac{30000}{30} = 1000$$

* Merge-Join: R1 no ordenat
empleat' ordenat per NIF

$$\text{ordenar empleat': } 100 \cdot \log_2(100) = 100 \cdot 7 = 700$$

$$\text{ordenar R1: } 91 \cdot \log_2(91) = 91 \cdot 6.5 = 637 \rightarrow \boxed{1529}$$

$$\rightarrow \text{Ajustar: } 91 + 100 = 191$$

* Merge-Join index: R1 no es té

* Bucle amb index:

$$B_S + \text{card}(S) + \text{card}(\text{join}(R, S))$$

$$91 + 3000 + 30000 = 33091$$

$$\bullet \text{Cost estratègia B} \Rightarrow 1000 + 100 + 430 + 91 + 1528 = \boxed{3149}$$