

AMPLI_P1_2021

miércoles, 29 de marzo de 2023 19:51

1. (4 punts) 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
```

I la següent informació al diccionari de dades:

- Empleat:
 - Cardinalitat: 25.000
 - Factor de bloqueig (b) = 50
 - Un 20% dels empleats tenen un sou més gran o igual que 30.000
 - Índex primari per NIF (b₁ = 200)
 - Índex secundari per sou (b₁ = 100)
- Sucursal:
 - Cardinalitat: 3.000
 - Factor de bloqueig (b) = 25
 - Índex primari per codi_sucursal (b₁ = 200)
- Ciutat:
 - Cardinalitat: 1.000
 - Factor de bloqueig (b) = 10
 - Índex primari per codi_ciutat (b₁ = 200)

Calcula el cost de l'estratègia d'execució següent:

- 1) Seleccionar Empleat pel seu sou
- 2) Fer el join entre el resultat de 1 i Sucursal
- 3) Fer el join entre el resultat de 2 i Ciutat

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).

Recorda els algorismes a considerar:

- | | |
|--------------------------------|-----------------------|
| Selecció: | Combinació (join): |
| • Índex primari o clúster | • Bucles imbricats |
| • Índex secundari (no clúster) | • Sort-merge |
| • Intersecció d'índexs | • Merge-join d'índexs |
| • Seqüencial | • Bucle amb índex |

2^o=1; 2¹=2; 2²=4; 2³=8; 2⁴=16; 2⁵=32; 2⁶=64; 2⁷=128; 2⁸=256; 2⁹=512; 2¹⁰=1024;

$$① \quad B_{\text{empleat}} = \frac{25000}{50} = 500$$

$$B_{\text{sucursal}} = \frac{3000}{25} = 120$$

$$B_{\text{ciutat}} = \frac{1000}{10} = 100$$

1) Seleccionar Empleat pel sou

- índex primari? No

- índex secundari? Si

$$\text{Cost} = \text{select}(\text{sou} \geq 30000) * \text{Card}(\text{Empleat}) = 0,2 \cdot 25000 = 5000$$

- intersecció d'índexs? No

- seqüencial?

$$\text{cost} = \frac{\text{Card}(\text{Empleat})}{b_{\text{empleat}}} = \frac{25000}{50} = \boxed{500}$$

Guardar

$$\text{cost} = \frac{\text{select}(\text{sou} \geq 30000) * \text{Card}(\text{Empleat})}{b_{\text{empleat}}} = \frac{0,2 * 25000}{50} = \boxed{100}$$

2) Join de 1 i Sucursal

$$B_1 = 100$$

$$B_{\text{sucursal}} = 120$$

- Bucles imbricats

$$\text{cost} = B_1 + B_1 * B_{\text{sucursal}} = 100 + 100 * 120 = 12100$$

- Sort-merge

$$\text{cost} = B_{\text{sucursal}} + B_1 + B_1 * \log_2(B_1) = 120 + 100 + 100 \log_2(100) = 120 + 100 + 100 \cdot 7 = \boxed{920}$$

- Merge-join índex

No

- Bucle amb índex

$$\text{Cost} = B_1 + \text{Card}(B_1) + \text{Card}(\text{Join}(\text{Sucursal}, B_1)) = 100 + 5000 + 5000 = 10100$$

Guardar

$$\frac{1}{B_{\text{join}}} = \frac{1}{50} + \frac{1}{25} = 16$$

$$\text{cost} = \frac{\text{Card}(\text{join})}{b_{\text{join}}} = \frac{5000}{16} = \boxed{313}$$

3) Join de 2 amb Ciutat

$$B_2 = 313$$

$$B_{\text{ciutat}} = 100$$

- Bucles - imbricats

$$\text{cost} = B_{\text{ciutat}} + B_{\text{ciutat}} \cdot B_2 = 100 + 100 \cdot 313 = 31400$$

- Sort-merge

$$\text{cost} = B_{\text{ciutat}} + B_2 + B_2 \log_2(B_2) = 100 + 313 + 313 \log_2(313) = 100 + 313 + 313 \cdot 9 = \boxed{3230}$$

- Merge join per índex: No

- Bucle amb índex:

$$\text{Cost} = B_2 + \text{ind}(B_2) + (\text{ind}(\text{Join}(B_2, \text{Ciutat}))) = 313 + 5000 + 5000 = 10313$$

$$\text{Total} \Rightarrow 500 + 100 + 920 + 313 + 3230 = \boxed{5063}$$

2. (2 punts) Donat el següent esquema de fragmentació de les relacions del problema 1:

Ciutat-Lleida = Ciutat(provincia = 'Lleida')
 Altres-Ciutats = Ciutat - Ciutat-Lleida
 Sucursals-Lleida = Sucursal \times Ciutat-Lleida
 Altres-Sucursals = Sucursal \times Altres-Ciutats
 Empleat-Bàsic = Empleat[NIF, nom, suc_assignada]
 Empleat-Nòmies = Empleat[NIF, categoria, sou]

- a) **Raona** la correctesa de la fragmentació (completitud, reconstruïbilitat i disjuntivitat)
 b) Tenint en compte que tenim **transparència de localització**, però **hem perdut la de fragmentació**, expressa en SQL (i pseudocodi si cal) la consulta del problema 1.

a) Completitud:

- Ciutat: És complet, reconstruïble i disjunt (al fn la resta no coincideix a triple)
- Sucursal: És completa, reconstruïble i disjunt
- Empleat: És completa (a les projeccions no ens deixem cap cap), reconstruïble (fn join entre els dos fragments) és fàcilment disjunt perquè no hi ha cap repetició de PK

b) (SELECT eb.nom, cl.ciutat
 FROM Empleat-bàsic eb, Ciutat-Lleida cl, Sucursals-Lleida sl
 WHERE eb.sou >= 30000
 AND eb.sucursal_assignada = sl.codi_sucursal
 AND sl.codi_ciutat = cl.codi_ciutat)
 UNION
 (SELECT eb.nom, cl.ciutat
 FROM Empleat-bàsic eb, Altres-Lleida al, Altres-Sucursals as
 WHERE eb.sou >= 30000
 AND eb.sucursal_assignada = as.codi_sucursal
 AND as.codi_ciutat = al.codi_ciutat)

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

transparència de localització i de fragmentació

- la consulta aquesta es de transparència total
- si no tens de fragmentació, utilitzes els fragments de dalt
- si no tens de localització, et diria com a info que els fragments estan a x lloc i hauries de fer un select (al SELECT e.nom at x, FROM Empleat e at x

3. (2 punts) Considera l'esquema global consensuat per la base de dades heterogènia:

Treballador(NIF, nom, oficina_assignada)
 Oficina(codi_oficina, adreça, nom_ciutat)
 Sou(NIF, sou)

On una de les bases de dades components té l'esquema (local) descrit al problema 1.

a) Expressa en DATALOG el **mapping** d'aquest esquema local al global

b) Expressa en DATALOG la **consulta** següent sobre l'esquema global

SELECT t.nom, o.adreça, s.sou
 FROM Treballador t, Oficina o, Sou s
 WHERE t.oficina_assignada = o.codi_oficina
 AND o.nom_ciutat = 'Barcelona'
 AND t.NIF = s.NIF

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

c) En funció dels mappings d'a), expressa en DATALOG la **consulta** de b) per la base de dades component segons el seu **esquema local**

a) Treballador(n1, n0, o0) ← Empleat(n1, n0, -, -, o0)

Oficina(c0, ad, nc) ← Sucursal(c0, -, ad, c0, -),
 Ciutat(cc, nc, -)

Sou(n1, sou) ← Empleat(n1, -, -, sou, -)

b) q(n0, ad, sou) ← Treballador(n1, n0, o0),

Oficina(o0, ad, "Barcelona"),
 Sou(n1, sou)

q(n0, ad, sou, nc) nc = "Barcelona"

c) q(n0, ad, sou) ← Empleat(-, n0, -, sou, o0),

c) $q(\text{no}, \text{ad}, \text{sou}) \leftarrow \text{Enpket}(-, \text{no}, -, \text{sou}, \text{sa}),$
 $\text{Suansel}(\text{sa}, -, \text{ad}, -, -)$

$q(\text{no}, \text{ad}, \text{sou}, \text{nc})$ $\text{nc} = \text{"Barcelona"}$