

Pràctica 2

Anna Marín Nuño

Monday, 10th of March 2025 (7:19)

Problema 1

Donat un sistema amb tres instal·lacions que anomenarem A, B i C.

La instal·lació A fa 5.000 lectures a dades pròpies així com 200 a dades de B i 100 de C.

La instal·lació B realitza 1.000 consultes diàries a dades pròpies, 200 a dades de A i 300 a dades de C. A més, realitza també 200 actualitzacions de les seves dades.

Finalment, la C llegeix i actualitza les seves dades 500 vegades al dia i fa 500 consultes a dades de B.

Considerant una matriu de costos com

	Local	Remota
Lectura	1	3
Escriptura	2	5

Decideix un esquema de localització de les dades per a minimitzar els costos de les comunicacions, raonant totes les decisions que preneu.

- Instal·lació A: dades A
- Instal·lació B: dades A i B
- Instal·lació C: dades C

On posar les dades A:

- Només a A:
 $A = 5.000$
 $B = 200 \cdot 3 = 600$
 $cost = 5.600$
- Replicat A i B
 $A = 5.000$
 $B = 200$
 $cost = 5.200$

On posar les dades B:

- Només a A:
 $A : 200$
 $B : 1.000 \cdot 3 + 200 \cdot 5 = 4.000$
 $C : 500 \cdot 3 = 1.500$
 $cost = 5.700$
- Només a B:
 $A : 200 \cdot 3 = 600$
 $B : 1.000 + 200 \cdot 2 = 1.400$
 $C : 500 \cdot 3 = 1.500$
 $cost = 3.500$
- Només a C:
 $A : 200 \cdot 3 = 600$
 $B : 1.000 \cdot 3 + 200 \cdot 5 = 4.000$
 $C : 500$
 $cost = 5.100$

- Replicat a A i B:

$$A : 200$$

$$B : 1.000 + 200 \cdot 2 + 200 \cdot 5 = 2.400$$

$$C : 500 \cdot 3 = 1.500$$

$$cost = 4.100$$

- Replicat a A i C:

$$A : 200$$

$$B : 1.000 \cdot 3 + 200 \cdot 5 \cdot 2 = 5.000$$

$$C : 500$$

$$cost = 5.700$$

- Replicat a B i C:

$$A : 200 \cdot 3 = 600$$

$$B : 1.000 + 200 \cdot 2 + 200 \cdot 5 = 2.400$$

$$C : 500$$

$$cost = 3.500$$

- Totalment replicat:

$$A : 200$$

$$B : 1.000 + 2 \cdot 5 \cdot 200 + 200 \cdot 2 = 3.400$$

$$C : 500$$

$$cost = 4.100$$

Convé guardar les dades de B només a B

On posar les dades de C:

- Només a A:

$$A : 100$$

$$B : 300 \cdot 3 = 900$$

$$C : 500 \cdot 3 + 500 \cdot 5 = 4.000$$

$$cost = 5.000$$

- Només a B:

$$A : 100 \cdot 3 = 300$$

$$B : 300$$

$$C : 500 \cdot 3 + 500 \cdot 5 = 4.000$$

$$cost = 4.600$$

- Només a C:

$$A : 100 \cdot 3 = 300$$

$$B : 300 \cdot 3 = 900$$

$$C : 500 + 500 \cdot 2 = 1.500$$

$$cost = 1.700$$

- Replicat a A i B:

$$A : 100$$

$$B : 300$$

$$C : 500 \cdot 3 + 2 \cdot 5 \cdot 500 = 6.500$$

$$cost = 6.900$$

- Replicat a A i C:

$$A : 100$$

$$B : 300 \cdot 3 = 900$$

$$C : 500 + 500 \cdot 2 + 500 \cdot 5 = 4.000$$

$$cost = 5.000$$

- Replicat B i C:

$$A : 100 \cdot 3 = 300$$

$$B : 300$$

$$C : 500 + 500 \cdot 2 + 500 \cdot 5 = 4.000$$

$$cost = 4.600$$

- Totalment replicat:

$$A : 100$$

$$B : 300$$

$$C : 500 + 500 \cdot 2 + 500 \cdot 5 \cdot 2 = 6.500$$

$$cost = 6.800$$

Problema 2

Donat el següent **esquema** de BD:

Alumne(id_al, nom, adreça, telèfon, compte_corrent)

Assignatura(id_as, nom, crèdits, curs)

Grup(id_gr, assignatura_id, horari)

Matricula(id_ma, alumne_id, grup_id, data)

I la següent **consulta**:

```
SELECT al.nom, al.compte_corrent
FROM Alumne al, Assignatura as, Matricula ma, Grup gr
WHERE as.curs = 2 AND ma.data = 2024
      AND ma.alumne_id = al.id_al
      AND ma.grup_id = gr.id_gr
      AND gr.assignatura_id = as.id_as
```

Considereu la següent fragmentació de la base de dades:

- **Alumne:**
 - Alumne1: Alumne(id_al ≤ 200)
 - Alumne2: Alumne(id_al > 200)[id_al, nom]
 - Alumne3: Alumne(id_al > 200)[id_al, adreça, telèfon, compte_corrent]
 - **Assignatura:** sense fragmentar
 - **Grup:** sense fragmentar
 - **Matricula:** sense fragmentar
- a) Indiqueu el **tipus de fragmentació** de la relació Alumne i **raoneu la seva correctesa** (completesa, reconstruïbilitat i disjuntivitat). Dibuixeu, si s'escau, el seu **arbre de fragmentació**.
- b) Suposant que teniu **transparència de localització**, però **heu perdut la de fragmentació**, re-escriviu la consulta del problema 1, **raonant** el procés per adaptar-la.

a) Tipus de fragmentació i anàlisi de la seva correctesa

La relació Alumne presenta una fragmentació mixta (híbrida):

Fragmentació horitzontal inicial basada en el predicat $id_al \leq 200$:

Alumne1: Conté tots els registres on $id_al \leq 200$ amb tots els atributs

I la resta de registres on $id_al > 200$ es distribueixen entre Alumne2 i Alumne3

Fragmentació vertical per als registres amb $id_al > 200$:

Alumne2: Conté [id_al, nom]

Alumne3: Conté [id_al, adreça, telèfon, compte_corrent]

Anàlisi formal de correctesa:

Completesa:

Tots els registres estan representats: $Alumne = \sigma_{id_al \leq 200}(Alumne) \cup \sigma_{id_al > 200}(Alumne)$

Tots els atributs estan presents: $att(Alumne) = att(Alumne1)$ per a $id_al \leq 200$

$att(Alumne2) \cup att(Alumne3) \setminus id_al = att(Alumne) \setminus id_al$ per a $id_al > 200$

Reconstruïbilitat:

La relació original es pot reconstruir amb: $Alumne = Alumne1 \cup (Alumne2 \bowtie_{id_{al}} Alumne3)$

Els registres amb $id_{al} > 200$ es reconstrueixen unint Alumne2 i Alumne3 per id_{al}

Disjuntivitat:

Horitzontalment: $Alumne1 \cap \sigma_{id_{al} > 200}(Alumne) = \emptyset$

Verticalment: $(att(Alumne2) \cap att(Alumne3)) \setminus id_{al} = \emptyset$

b)

```
-- Part 1: Alumnes amb id_al ≤ 200 (utilitzant Alumne1)
SELECT al1.nom, al1.compte_corrent
FROM Alumne1 al1, Assignatura as, Matricula ma, Grup gr
WHERE as.curs = 2 AND ma.data = 2024
AND ma.alumne_id = al1.id_al
AND ma.grup_id = gr.id_gr
AND gr.assignatura_id = as.id_as

UNION

-- Part 2: Alumnes amb id_al > 200 (combinant Alumne2 i Alumne3)
SELECT al2.nom, al3.compte_corrent
FROM Alumne2 al2, Alumne3 al3, Assignatura as, Matricula ma, Grup gr
WHERE al2.id_al = al3.id_al -- Unió entre Alumne2 i Alumne3
AND as.curs = 2 AND ma.data = 2024
AND ma.alumne_id = al2.id_al
AND ma.grup_id = gr.id_gr
AND gr.assignatura_id = as.id_as
```