

Basi di Dati

Esercizio su costi di accesso su join

Esercizio completo costi di accesso (join)

Sono date le seguenti relazioni:

- ▶ **BUS (CB, MODELLO, TARGA, ANNO,...)**
con 100 n-ple in 50 pagine, **CB** è la chiave e **MODELLO** ha 20 valori diversi;
- ▶ **PERCORSI (CP, LUOGO_P, LUOGO_A,....)**
con 500 n-ple in 100 pagine, **CP** è la chiave
LUOGO_P e **LUOGO_A** hanno entrambi 100 valori diversi;
- ▶ **CORSE (CC, CB, CP, T_IN, T_OUT,...)**
con 10000 n-ple in 1000 pagine, **CC** è la chiave ,
il numero di valori in **T_IN** e **T_OUT** è ignoto.



Esercizio completo costi di accesso (join)

Scrivere un join che risolva l'interrogazione:

“selezionare tutti i bus di modello =Z partiti prima di un tempo TX e rientrati prima di un tempo TY, dove LUOGO_P=A oppure LUOGO_A=B”:

```
SELECT *, --, --  
FROM BUS B, CORSE C, PERCORSI P  
WHERE   B.MODELLO = 'Z'  
AND     C.T_OUT < TX AND C.T_IN < TY  
AND     (P.LUOGO_P = 'A' OR P.LUOGO_A = 'B')  
AND     B.CB = C.CB AND P.CP = C.CP
```



Esercizio completo costi di accesso (join)

Per le sequenze: **BUS \Rightarrow CORSE \Rightarrow PERCORSI** e
 PERCORSI \Rightarrow CORSE \Rightarrow BUS

- ▶ Si stabilisca la **migliore disposizione di indici** per il metodo nested-loops considerando le relazioni non ordinate.
- ▶ Si proponga un **solo indice per relazione**.
- ▶ Si confronti il costo di accesso con il costo di accesso che si avrebbe in **assenza di indici**.
- ▶ Si assuma il numero di pagine di ogni indice uguale al **10%** del numero di pagine della relazione.



Esercizio completo costi di accesso (join)

CASO: BUS \Rightarrow CORSE \Rightarrow PERCORSI

a) BUS ESTERNA

$$C_{\text{BUS}}(\text{seq}) = 50$$

$$F(\text{modello}) = 1/20$$

$$\begin{aligned} C_{\text{BUS}}(\text{modello}) &= \lceil 5 \times 1/20 \rceil + \lceil 100/20 \rceil = \\ &= 1 + \lceil 5 \rceil = 6 \end{aligned}$$

$$E_{\text{BUS}} = \lceil 100 \times 1/20 \rceil = 5$$



Esercizio completo costi di accesso (join)

CASO: BUS \Rightarrow CORSE \Rightarrow PERCORSI

b) CORSE INTERNA

$$C_{\text{CORSE}}(\text{seq}) = 1000$$

$$F(t_{\text{out}}) = 1/3 \quad F(t_{\text{in}}) = 1/3 \quad F(\text{cb}) = 1/100$$

$$\begin{aligned} C_{\text{CORSE}}(\text{cb}) &= \lceil 100 \times 1/100 \rceil + \lceil 10000/100 \rceil \\ &= 1 + \lceil 100 \rceil = 1 + 100 = 101 \end{aligned}$$

$$\begin{aligned} C_{\text{BUS-CORSE}} &= C_{\text{BUS}}(\text{modello}) + E_1 \times C_{\text{CORSE}}(\text{cb}) = \\ &= 6 + 5 \times 101 = 511 \end{aligned}$$

$$E_{\text{BUS-CORSE}} = \lceil 100 \times 10000 \times 1/20 \times 1/3 \times 1/3 \times 1/100 \rceil = 56$$



Esercizio completo costi di accesso (join)

CASO: BUS \Rightarrow CORSE \Rightarrow PERCORSI

c) PERCORSI

$$C_{\text{PERCORSI}}(\text{seq}) = 100 \quad F(\text{cp}) = 1/500$$

$$C_{\text{PERCORSI}}(\text{cp}) = \lceil 1/500 \times 10 \rceil + \lceil 500/500 \rceil = 2$$

$$\begin{aligned} C_{\text{B-C-P}} &= C_{\text{BUS-CORSE}} + E_{\text{BUS-CORSE}} \times C_{\text{PERCORSI}}(\text{cp}) = \\ &= 511 + 56 \times 2 = 623 \end{aligned}$$



Esercizio completo costi di accesso (join)

CASO : PERCORSI \Rightarrow CORSE \Rightarrow BUS

d) PERCORSI ESTERNA

$$C_{\text{PERCORSI}}(\text{seq}) = 100$$

$$F(\text{luogo}_p) = 1/100$$

$$F(\text{luogo}_a) = 1/100$$

$$\begin{aligned} F(\text{luogo}_p = \langle \text{val} \rangle \text{ or } \text{luogo}_a = \langle \text{val} \rangle) &= \\ &= F(\text{luogo}_p) + F(\text{luogo}_a) - F(\text{luogo}_p) \times F(\text{luogo}_a) = \\ &= 1/100 + 1/100 - 1/100 \times 1/100 = 199/10000 \end{aligned}$$

$$E_{\text{PERCORSI}} = \lceil 500 \times 199/10000 \rceil = 10$$



Esercizio completo costi di accesso (join)

CASO : PERCORSI \Rightarrow CORSE \Rightarrow BUS

e) CORSE INTERNA

$$C_{\text{CORSE}}(\text{seq}) = 1000$$

$$F(t_{\text{out}}) = 1/3 \quad F(t_{\text{in}}) = 1/3 \quad F(\text{cp}) = 1/500$$

$$\begin{aligned} C_{\text{CORSE}}(\text{cp}) &= \lceil 1/500 \times 100 \rceil + \lceil 10000/500 \rceil = \\ &= 1 + \lceil 20 \rceil = 21 \end{aligned}$$

$$\begin{aligned} C_{\text{PERCORSI-CORSE}} &= C_{\text{PERCORSI}}(\text{seq}) + E_{\text{PERCORSI}} \times C_{\text{CORSE}}(\text{cp}) = \\ &= 100 + 10 \times 21 = 310 \end{aligned}$$

$$\begin{aligned} E_{\text{PERCORSI-CORSE}} &= \\ &= \lceil 10000 \times 500 \times 199/10000 \times 1/3 \times 1/3 \times 1/500 \rceil = 23 \end{aligned}$$



Esercizio completo costi di accesso (join)

CASO : PERCORSI \Rightarrow CORSE \Rightarrow BUS

f) BUS

$$C_{\text{BUS}}(\text{seq}) = 50,$$

$$F(\text{modello}) = 1/20 \quad , \quad C_{\text{BUS}}(\text{modello}) = 6$$

$$F(\text{cb}) = 1/100 \quad C_{\text{BUS}}(\text{cb}) = \lceil 1/100 \times 5 \rceil + \lceil 1 \rceil = 2$$

$$\begin{aligned} C_{\text{P-C-B}} &= C_{\text{PERCORSI-CORSE}} + E_{\text{PERCORSI-CORSE}} \times C_{\text{BUS}}(\text{cb}) = \\ &= 310 + 23 \times 2 = 356 \end{aligned}$$

Risulta migliore PERCORSI \Rightarrow CORSE \Rightarrow BUS :

PER LA RELAZIONE PERCORSI: SCAN. SEQ.

PER LA RELAZIONE CORSE: INDICE SU CP,

PER LA RELAZIONE BUS: INDICE SU CB,



Esercizio completo costi di accesso (join)

CASO : PERCORSI \Rightarrow CORSE \Rightarrow BUS

Senza indici avremmo avuto

nel caso PERCORSI \Rightarrow CORSE \Rightarrow BUS :

$$\begin{aligned} C_{P-C-B} &= C_{\text{PERCORSI}}(\text{seq}) + E_{\text{PERCORSI}} \times C_{\text{CORSE}}(\text{seq}) + \\ &\quad + E_{\text{PERCORSI-CORSE}} \times C_{\text{BUS}}(\text{seq}) = \\ &= 100 + 10 \times 1000 + 23 \times 50 = 11250 \end{aligned}$$

nel caso BUS \Rightarrow CORSE \Rightarrow PERCORSI :

$$\begin{aligned} C_{B-C-P} &= C_{\text{BUS}}(\text{seq}) + E_{\text{BUS}} \times C_{\text{CORSE}}(\text{seq}) + \\ &\quad + E_{\text{BUS-CORSE}} \times C_{\text{PERCORSI}}(\text{seq}) = \\ &= 50 + 5 \times 1000 + 56 \times 100 = 10650 \end{aligned}$$

