

COMPITO DI SISTEMI INFORMATIVI

13 luglio 2010 (Tot. 16) Tempo: 2h

Esercizio 1 (punti 3)

Si consideri il seguente log:

- | | |
|----------------------|----------------------|
| 1. U(T2,O1,B1,A1) | 14. CK(T4,T5,T6) |
| 2. U(T1,O2,B2,A2) | 15. D(T6,O5,B15) |
| 3. C(T2) | 16. C(T5) |
| 4. B(T3) | 17. B(T7) |
| 5. U(T3,O2,B4,A4) | 18. U(T7,O3,B18,A18) |
| 6. B(T4) | 19. C(T6) |
| 7. D(T4,O2,B7) | 20. D(T7,O4,B20) |
| 8. B(T5) | 21. B(T8) |
| 9. I(T5,O3,A9) | 22. I(T8,O6,A22) |
| 10. C(T1) | 23. U(T4,O6,B23,A23) |
| 11. C(T3) | 24. U(T8,O7,B24,A24) |
| 12. B(T6) | 25. D(T7,O8,B25) |
| 13. U(T6,O4,B13,A13) | 26. U(T4,O7,B26,A26) |

si mostrino le operazioni di recovery da effettuare supponendo che il guasto avvenga subito dopo l'ultimo record del log.

Esercizio 2 (punti 4)

Dato il seguente schedule:

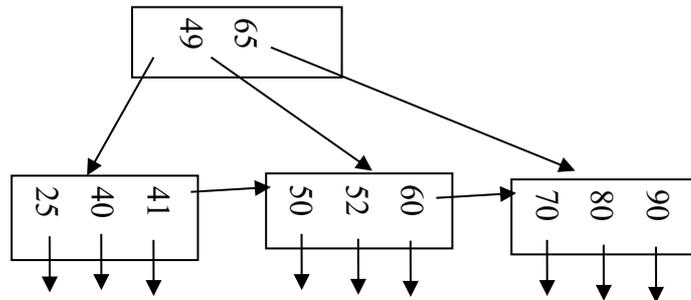
$r2(y), r1(y), r3(x), w3(x), w1(z), w3(y), w2(w), r2(u), r1(x)$

si indichi se è view-equivalente allo schedule seriale

$r2(y), w2(w), r2(u), r3(x), w3(x), w3(y), r1(y), w1(z), r1(x),$

Esercizio 3 (punti 3)

Sia dato il seguente B+ tree di ordine 3.



Si mostri l'albero risultante dall'inserimento della chiave 75. Si mostri l'albero risultante dall'inserimento della chiave 30 nell'albero ottenuto in precedenza.

Esercizio 4 (punti 6)

Siano date le relazioni

Lettore(CodiceLet, Nome, Cognome, Indirizzo, DataDiNascita, ISBNLibroPreferito)

Libro(ISBN, Titolo, Autore, AnnoPubbli)

la query

```
SELECT Le.*, Li.*
```

```
FROM Lettore AS Le, Libro AS Li
```

```
WHERE Le ISBNLibroPreferito =Li.ISBN AND Li.AnnoPubbli >=1950
```

e i parametri:

buffer di memoria centrale disponibili: $M=2000000$

dimensione del buffer: $B=2000$ bytes

numero di tuple: $T(\text{Lettore})=1000000$, $T(\text{Libro})=2000000$,

dimensione delle tuple: $S(\text{Lettore})=5000$ bytes, $S(\text{Libro})=3000$, $S(\text{ISBN})=100$,

$S(\text{ISBNLibroPreferito})=100$

$V(\text{Lettore}, \text{ISBNLibroPreferito})=1000$

range di Libro.AnnoPubbli = [1900,1999]

Si calcoli il costo minimo di ciascuna delle seguenti sequenze:

$\sigma_{\text{AnnoPubbli} \geq 1950} (\text{Libro} \bowtie \text{Lettore})$

$(\sigma_{\text{AnnoPubbli} \geq 1950} \text{Libro}) \bowtie \text{Lettore}$

Si indichi quale delle sequenze ha costo inferiore.

Si supponga che i record siano impaccati nei buffer e che tutto lo spazio nei buffer sia occupato dai record.

Si supponga di avere indici primari su Libro.AnnoPubbli e su Lettore. ISBNLibroPreferito.

SOLUZIONE

Esercizio 1

1. U(T2,O1,B1,A1)
2. U(T1,O2,B2,A2)
3. C(T2)
4. B(T3)
5. U(T3,O2,B4,A4)
6. B(T4)
7. D(T4,O2,B7)
8. B(T5)
9. I(T5,O3,A9)
10. C(T1)
11. C(T3)
12. B(T6)
13. U(T6,O4,B13,A13)
14. CK(T4,T5,T6)
15. D(T6,O5,B15)
16. C(T5)
17. B(T7)
18. U(T7,O3,B18,A18)
19. C(T6)
20. D(T7,O4,B20)
21. B(T8)
22. I(T8,O6,A22)
23. U(T4,O6,B23,A23)
24. U(T8,O7,B24,A24)
25. D(T7,O8,B25)
26. U(T4,O7,B26,A26)

- 14 UNDO={T4,T5,T6} REDO={}
16 UNDO={T4,T6} REDO={T5}
17 UNDO={T4,T6,T7} REDO={T5}
19 UNDO={T4,T7} REDO={T5, T6}
21 UNDO={T4,T7,T8} REDO={T5, T6}

UNDO

- 26 O7=B26
25 I(O8,B25)
24 O7=B24
23 O6=B23
22 D(O6)
20 I(O4,B20)
18 O3=B18
7 I(O2,B7)

REDO

- 9 I(O3,A9)
13 O4=A13
15 D(O5)

Esercizio 2

$r_2(y), r_1(y), r_3(x), w_3(x), w_1(z), w_3(y), w_2(w), r_2(u), r_1(x)$
"legge-da" = $\{(r_1(x), w_3(x))\}$

scritture finali = $\{(t_2, w), (t_3, y), (t_1, z), (t_3, x)\}$

Schedule seriale

$r_2(y), w_2(w), r_2(u), r_3(x), w_3(x), w_3(y), r_1(y), w_1(z), r_1(x),$

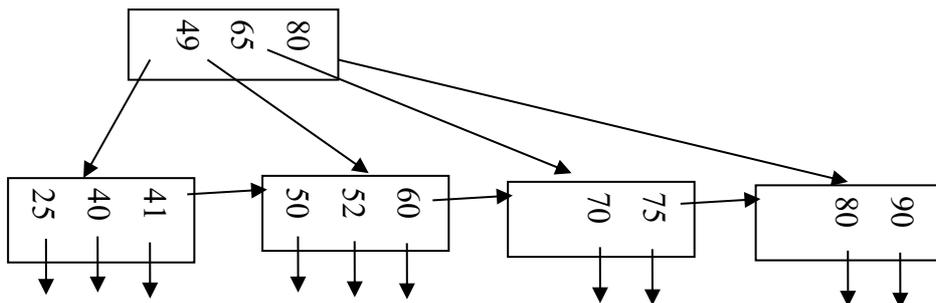
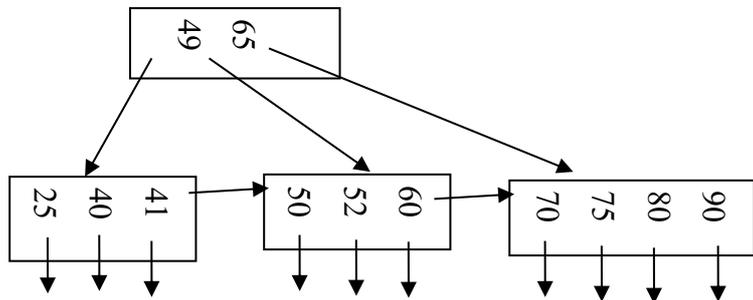
"legge-da" = $\{(r_1(x), w_3(x)), (r_1(y), w_3(y))\}$

scritture finali = $\{(t_2, w), (t_3, y), (t_1, z), (t_3, x)\}$

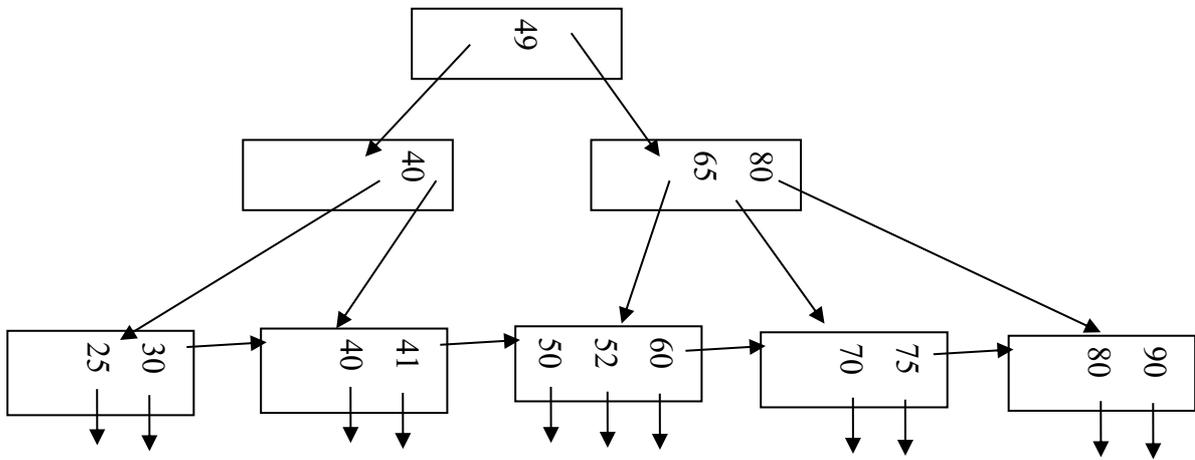
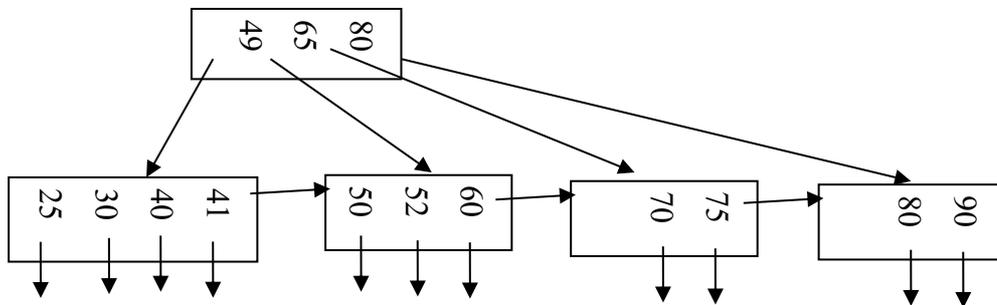
I due schedule non sono view-equivalenti.

Esercizio 3

Inserimento di 75



Inserimento di 30



Esercizio 4

buffer di memoria centrale disponibili: $M=2000000$

dimensione del buffer: $B=2000$ bytes

numero di tuple: $T(\text{Lettore})=1000000$, $T(\text{Libro})=2000000$,

dimensione delle tuple: $S(\text{Lettore})=5000$ bytes, $S(\text{Libro})=3000$, $S(\text{ISBN})=100$

$V(\text{Lettore, ISBNLibroPreferito})=1000$

range di $\text{Libro.AnnoPubbli}=[1900,1999]$

Calcoliamo innanzitutto il numero di blocchi occupati da ciascuna relazione:

$$B(\text{Lettore})=T(\text{Lettore}) \cdot S(\text{Lettore})/B = 1000000 \cdot 5000/2000=2500000$$

$$B(\text{Libro}) = 2000000 \cdot 3000/2000=3000000$$

Prima sequenza: $(\sigma_{\text{AnnoPubbli} \geq 1950} \text{ Libro}) \triangleright \triangleleft \text{ Lettore}$

Sia $X = (\sigma_{\text{AnnoPubbli} \geq 1950} \text{ Libro})$

C'è un indice secondario su Libro.AnnoPubbli

$$F = (1999 - 1950 + 1) / (1999 - 1900 + 1) = 0.50$$

$$\text{Costo}(X) = 3 + \lceil f * B(\text{Libro}) \rceil = 3 + 0.5 * 3000000 = 1500003$$

$$B(X) = 0.5 * 2000000 = 1500000$$

X sta in memoria centrale, quindi si usa il join a un passo

$$\text{Costo}(X \triangleright \triangleleft \text{ Lettore}) = B(X) + B(\text{Lettore}) = 1500000 + 2500000 = 4000000$$

$$\text{Costo totale} = 1500003 + 4000000 = 5500003$$

Seconda sequenza: $\sigma_{\text{AnnoPubbli} \geq 1950} (\text{ Libro } \triangleright \triangleleft \text{ Lettore})$

Sia $Z = \text{ Libro } \triangleright \triangleleft \text{ Lettore}$

Né Libro né Utente stanno in memoria centrale, quindi si prova con il join nested-loop basato sui blocchi, con l'hash join ibrido e con il join con indice usando l'indice primario su Lettore.

ISBNLibroPreferito

$$\text{CostoJNL}(Z) = B(\text{Lettore}) + B(\text{Lettore}) * B(\text{Libro}) / M = 2500000 + 2500000 * 3000000 / 2000000 = 6250000$$

$$\text{CostoHJI}(Z) = (3 - 2M / B(\text{Lettore})) (B(\text{Lettore}) + B(\text{Libro})) = (3 - 2 * 2000000 / 2500000) * (2500000 + 3000000) = 7700000$$

$$\text{CostoIJ}(Z) = B(\text{Libro}) + T(\text{Libro}) \lceil B(\text{Lettore}) / V(\text{Lettore}, \text{ISBNLibroPreferito}) \rceil = 3000000 + 2000000 * \lceil 2500000 / 1000 \rceil = 3000000 + 2000000 * 2500 = 5003000000$$

$$T(Z) = T(\text{Libro}) * T(\text{Lettore}) / \max \{ V(\text{Lettore}, \text{ISBNLibroPreferito}), V(\text{Libro}, \text{ISBN}) \} = 2500000 * 3000000 / 2000000 = 3750000$$

$$S(Z) = S(\text{Lettore}) + S(\text{Libro}) - S(\text{ISBN}) = 5000 + 3000 - 100 = 7900$$

$$B(Z) = 3750000 * 7050 / 2000 = 13218750$$

Dato che non ci sono indici su Z:

$$\text{Costo} (\sigma_{\text{AnnoPubbli} \geq 1950} Z) = B(Z) = 13218750$$

$$\text{Costo totale} = 6250000 + 13218750 = 19468750$$

La sequenza di costo minore è la prima