

COMPITO DI SISTEMI INFORMATIVI

21 luglio 2009 – Compito (Tot. 16) Tempo: 2h

Esercizio 1 (punti 3)

Si consideri il seguente log:

- | | |
|----------------------|----------------------|
| 1. I(T1,O1,A1) | 15. B(T5) |
| 2. B(T2) | 16. I(T5,O7,A16) |
| 3. U(T2,O1,B3,A3) | 17. C(T2) |
| 4. B(T3) | 18. U(T3,O7,B18,A18) |
| 5. U(T3,O2,B5,A5) | 19. I(T4,O8,A19) |
| 6. B(T4) | 20. B(T6) |
| 7. D(T4,O3,B7) | 21. U(T6,O8,B21,A21) |
| 8. C(T1) | 22. U(T5,O4,B22,A22) |
| 9. U(T2,O1,B9,A9) | 23. U(T3,O8,B23,A23) |
| 10. I(T4,O4,A10) | 24. I(T6,O9,A24) |
| 11. U(T3,O4,B11,A11) | 25. B(T7) |
| 12. D(T2,O5,B12) | 26. U(T7,O6,B26,A26) |
| 13. CK(T2,T3,T4) | 27. C(T5) |
| 14. U(T4,O6,B14,A14) | |

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:

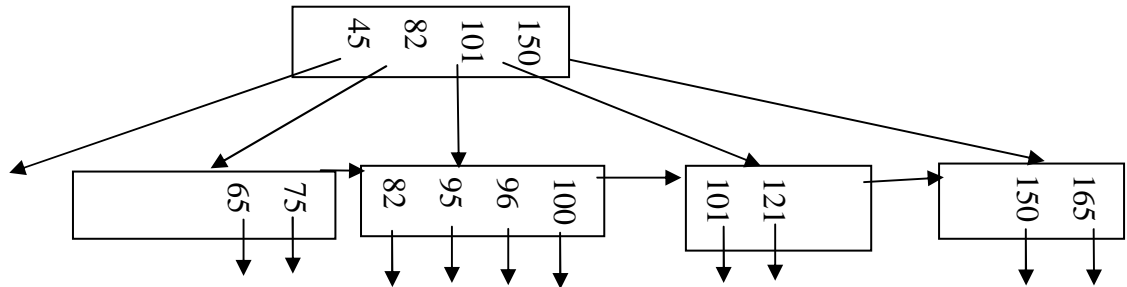
w1(x) r2(y) r2(x) w3(y) r4(y) w2(x) w5(y) w3(z) r5(y) r3(z)

si indichi se è view equivalente al seguente schedule seriale

w1(x) r2(y) r2(x) w2(x) w3(y) w3(z) r3(z) r4(y) w5(y) r5(y)

Esercizio 3 (punti 3)

Sia dato il seguente B+ tree di ordine 4.



Si mostri come si modifica l'albero nel caso di inserimento della chiave 97.

Esercizio 4 (punti 6)

Sia data la seguente relazione

Agenzia(CodAg, Nome, Via, Città, CAP, Telefono, Fax)

Cliente(CodiceCli, Nome, Cognome, Via, Città, CAP, Telefono, CodAg, Età)

la query

```
SELECT A.*, C.*
```

```
FROM Agenzia AS A, Cliente AS C
```

```
WHERE A.Città = C.Città AND C.Età >35
```

e i parametri:

buffer di memoria centrale disponibili: $M=1000000$

dimensione del buffer: $B=2000$ bytes

numero di tuple: $T(\text{Cliente})=1000000$, $T(\text{Agenzia})=20000$,

dimensione delle tuple: $S(\text{Cliente})=4000$, $S(\text{Agenzia})=4000$, $S(\text{Città})=100$

numero di valori $V(\text{Cliente}, \text{Città})=500000$, $V(\text{Agenzia}, \text{Città})=1000$

$\text{Min}(\text{Cliente}, \text{Età})=18$ $\text{Max}(\text{Cliente}, \text{Età})=55$

Si calcoli il costo minimo necessario ad eseguire la query.

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

Si supponga che ci sia un indice primario su Cliente.Età

SOLUZIONE

Esercizio 1

1. I(T1,O1,A1)
2. B(T2)
3. U(T2,O1,B3,A3)
4. B(T3)
5. U(T3,O2,B5,A5)
6. B(T4)
7. D(T4,O3,B7)
8. C(T1)
9. U(T2,O1,B9,A9)
10. I(T4,O4,A10)
11. U(T3,O4,B11,A11)
12. D(T2,O5,B12)
13. CK(T2,T3,T4)
14. U(T4,O6,B14,A14)
15. B(T5)
16. I(T5,O7,A16)
17. C(T2)
18. U(T3,O7,B18,A18)
19. I(T4,O8,A19)
20. B(T6)
21. U(T6,O8,B21,A21)
22. U(T5,O4,B22,A22)
23. U(T3,O8,B23,A23)
24. I(T6,O9,A24)
25. B(T7)
26. U(T7,O6,B26,A26)
27. C(T5)

13 UNDO={T2,T3,T4} REDO={}
15 UNDO={T2,T3,T4,T5} REDO={}
17 UNDO={T3,T4,T5} REDO={T2}
20 UNDO={T3,T4,T5,T6} REDO={T2}
25 UNDO={T3,T4,T5,T6,T7} REDO={T2}
27 UNDO={T3,T4,T6,T7} REDO={T2,T5}

UNDO

26 O6=B26
24 D(O9)
23 O8=B23
21 O8=B21
19 D(O8)
18 O7=B18
14 O6=B14
11 O4=B11
10 D(O4)
7 I(O2,B7)
5 O2=B5

REDO

3 O1=A3
9 O1=A9
12 D(I5)
16 I(O7,A16)
22 O4=A22

Esercizio 2

w1(x) r2(y) r2(x) w3(y) r4(y) w2(x) w5(y) w3(z) r5(y) r3(z)

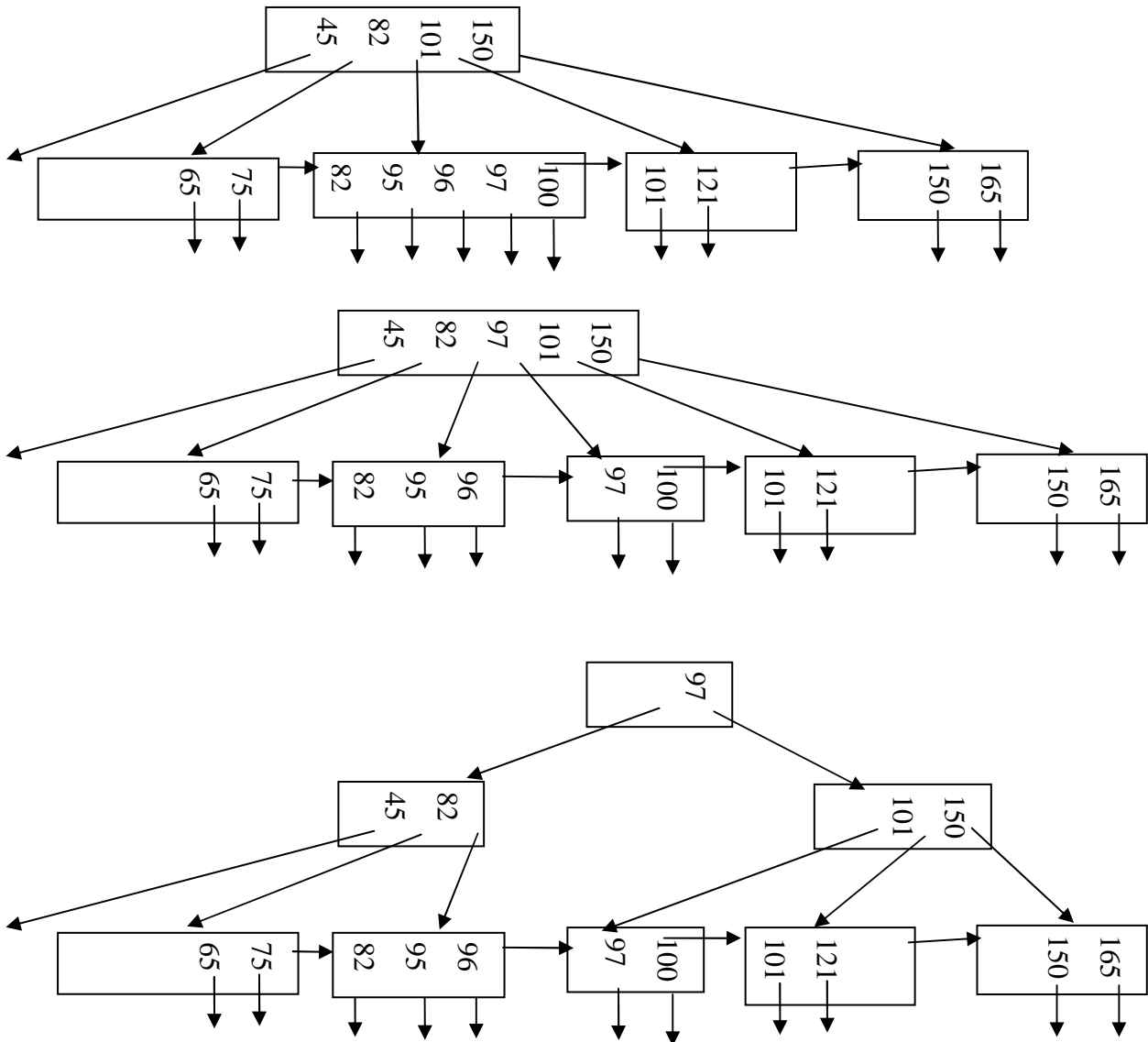
"legge-da"={ (r2(x),w1(x)),(r4(y),w3(y)),(r5(y),w5(y)),(r3(z),w3(z)) }
scritture finali={ (t2,x),(t5,y),(t3,z) }

w1(x) r2(y) r2(x) w2(x) w3(y) w3(z) r3(z) r4(y) w5(y) r5(y)

"legge-da"={ (r2(x),w1(x)),(r4(y),w3(y)),(r5(y),w5(y)),(r3(z),w3(z)) }
scritture finali={ (t2,x),(t5,y),(t3,z) }

I due schedule sono view-equivalenti

Esercizio 3



Esercizio 4

Parametri:

buffer di memoria centrale disponibili: $M=1000000$

dimensione del buffer: $B=2000$ bytes

numero di tuple: $T(\text{Cliente})=1000000$, $T(\text{Agenzia})=20000$,

dimensione delle tuple: $S(\text{Cliente})=4000$, $S(\text{Agenzia})=4000$

numero di valori $V(\text{Cliente}, \text{Città})=500000$, $V(\text{Agenzia}, \text{Città})=1000$

$\text{Min}(\text{Cliente}, \text{Età})=18$, $\text{Max}(\text{Cliente}, \text{Età})=55$

Sequenze

$(\sigma_{\text{Età} > 35} \text{ Cliente}) \triangleright \triangleleft \text{Agenzia}$

$\sigma_{\text{Età} > 35} (\text{Cliente} \triangleright \triangleleft \text{Agenzia})$

Calcoliamo innanzitutto il numero di blocchi occupati dalle relazioni:

$B(\text{Cliente})=T(\text{Cliente}) \cdot S(\text{Cliente})/B = 1000000 \cdot 4000/2000 = 2000000$

$B(\text{Agenzia})=T(\text{Agenzia}) \cdot S(\text{Agenzia})/B = 20000 \cdot 4000/2000 = 40000$

Prima sequenza: $(\sigma_{\text{DataNascita} > "1/1/1960"} \text{ Cliente}) \triangleright \triangleleft \text{Agenzia}$

Sia $X = \sigma_{\text{Età} > 35} \text{ Cliente}$

C'è un indice primario su Età

$$F = (55-35)/(55-18+1) = 0.526$$

$$\text{Costo}(X) = 3 + \lceil f * B(\text{Cliente}) \rceil = 3 + \lceil 0.526 * 2000000 \rceil = 1052003$$

$$T(X) = 0.526 * 1000000 = 526000$$

$$B(X) = 526000 * 4000 / 2000 = 1052000$$

Agenzia sta in memoria centrale quindi si usa il join a un passo:

$$\text{Costo}(X \triangleright \triangleleft \text{Agenzia}) = B(X) + B(\text{Agenzia}) = 1052000 + 40000 = 1092000$$

$$\text{Costo totale} = 1052003 + 1092000 = 2144003$$

Seconda sequenza: $\sigma_{\text{Età} > 35} (\text{Cliente} \triangleright \triangleleft \text{Agenzia})$

Sia $Z = \text{Cliente} \triangleright \triangleleft \text{Agenzia}$

Agenzia sta in memoria centrale quindi si usa il join a un passo:

$$\text{Costo}(Z) = B(\text{Cliente}) + B(\text{Agenzia}) = 2000000 + 40000 = 2040000$$

$$T(Z) = T(\text{Cliente}) * T(\text{Agenzia}) / \max\{V(\text{Agenzia, Città}), V(\text{Cliente, Città})\} = 1000000 * 20000 / 500000 = 40000$$

$$S(Z) = S(\text{Cliente}) + S(\text{Agenzia}) - S(\text{CodiceAg}) = 4000 + 4000 - 100 = 7900$$

$$B(Z) = 40000 * 7900 / 2000 = 158000$$

Dato che non ci sono indici su Z:

$$\text{Costo: } (\sigma_{\text{Età} > 35} Z) = B(Z) = 158000$$

$$\text{Costo totale} = 2040000 + 158000 = 2198000$$

Il costo minimo è 2144003