

COMPITO DI SISTEMI INFORMATIVI/BASI DI DATI II

15 gennaio 2007 (Tot. 16) Tempo: 2h

Esercizio 1 (punti 3)

Si consideri il seguente log:

1. I(T1,O1,A1)
2. D(T2,O2,B2)
3. B(T3)
4. U(T3,O1,B4,A4)
5. C(T1)
6. I(T3,O3,A6)
7. C(T2)
8. B(T4)
9. U(T4,O4,B9,A9)
10. D(T3,O5,B10)
11. B(T5)
12. U(T5,O3,B12,A12)
13. I(T4,O6,A13)
14. I(T5,O7,A14)
15. U(T3,O7,B15,A15)
16. CK(T3,T4,T5)
17. D(T5,O4,B17)
18. U(T4,O6,B18,A18)
19. B(T6)
20. I(T6,O8,A20)
21. C(T5)
22. U(T6,O6,B22,A22)
23. B(T7)
24. U(T7,O3,B24,A24)
25. D(T6,O3,B25)
26. I(T3,O9,A26)
27. C(T3)
28. U(T4,O9,B28,A28)
29. I(T6,O10,A29)
30. C(T6)

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

Esercizio 2 (punti 4)

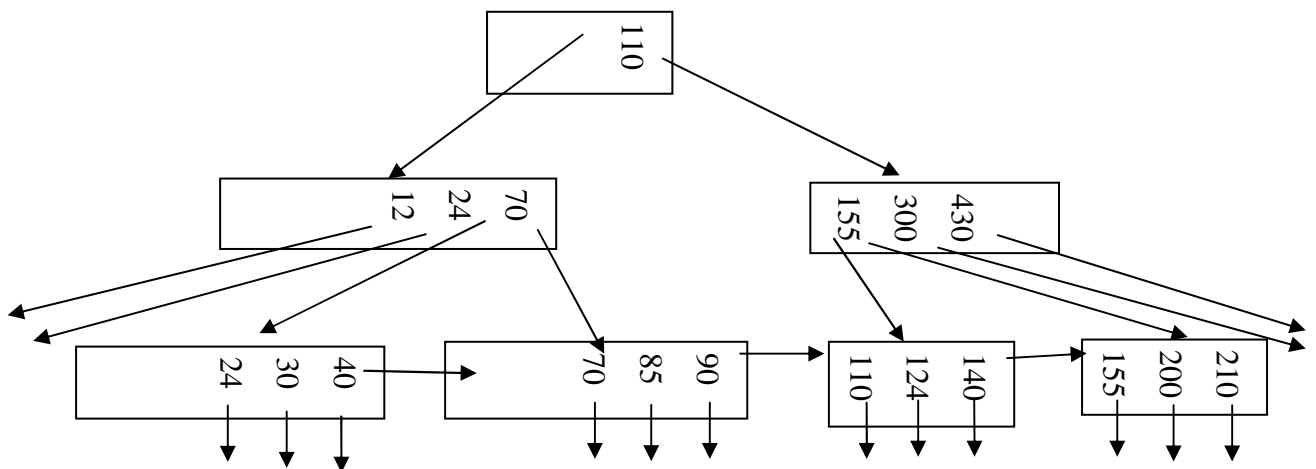
Dato il seguente schedule:

r1(x) r2(x) r3(y) r1(y) r2(z) r3(z) w3(y) r1(x) w2(x) w1(x) w2(z)

si indichi se rispetta il two-phase locking. Nel caso lo rispetti, si indichi una sequenza di acquisizione e rilascio di lock compatibile con lo schedule e che rispetti il two-phase locking. Nel caso non lo rispetti, si indichi almeno uno dei conflitti.

Esercizio 3 (punti 3)

Sia dato il seguente B+ tree di ordine 6.



Si mostri come si modifica l'albero nel caso di cancellazione della chiave 110

Esercizio 4 (punti 6)

Siano date le relazioni

Vendita(CodVendita,Data,Negoziio, CodArticolo, Quantità)

Articolo(CodArticolo,Descrizione,Prezzo)

la query

SELECT V.*, A.*

FROM Vendita AS V, Articolo AS A

WHERE V.CodArticolo=A.CodArticolo AND A.Prezzo=10 AND V.Data='1/1/2000'

e i parametri:

buffer di memoria centrale disponibili: M=2.000

dimensione del buffer: B=4.000 bytes

numero di tuple: T(Vendita)=200.000, T(Articolo)=15.000,

dimensione delle tuple: S(Vendita)=1000 bytes, S(Articolo)=2000,

S(CodArticolo)=100

numero di valori: V(Vendita,CodArticolo)=8000 V(Vendita,Data)=60

V(Articolo,Prezzo)=10

Si calcoli il costo minimo di ciascuna delle seguenti sequenze:

$(\sigma_{Prezzo=10}Articolo) \triangleright \triangleleft (\sigma_{Data='1/1/2000'}Vendita)$

$\sigma_{Data='1/1/2000' \wedge Prezzo=10}(Articolo) \triangleright \triangleleft Vendita)$

Si supponga che i record siano impaccati nei buffer e che tutto lo spazio nei buffer sia occupato dai record. Si supponga di avere un indice secondario su Articolo.Prezzo e un indice primario su Vendita.Data. CodArticolo è chiave primario di Articolo.

SOLUZIONE

Esercizio 1

1. I(T1,O1,A1)
2. D(T2,O2,B2)
3. B(T3)
4. U(T3,O1,B4,A4)
5. C(T1)
6. I(T3,O3,A6)
7. C(T2)
8. B(T4)
9. U(T4,O4,B9,A9)
10. D(T3,O5,B10)
11. B(T5)
12. U(T5,O3,B12,A12)
13. I(T4,O6,A13)
14. I(T5,O7,A14)
15. U(T3,O7,B15,A15)
16. CK(T3,T4,T5)
17. D(T5,O4,B17)
18. U(T4,O6,B18,A18)
19. B(T6)
20. I(T6,O8,A20)
21. C(T5)
22. U(T6,O6,B22,A22)
23. B(T7)
24. U(T7,O3,B24,A24)
25. D(T6,O3,B25)
26. I(T3,O9,A26)
27. C(T3)
28. U(T4,O9,B28,A28)
29. I(T6,O10,A29)
30. C(T6)

16 UNDO={T3,T4,T5} REDO={}
19 UNDO={T3,T4,T5,T6} REDO={}
21 UNDO={T3,T4,T6} REDO={T5}
23 UNDO={T3,T4,T6,T7} REDO={T5}
27 UNDO={T4,T6,T7} REDO={T3,T5}
30 UNDO={T4,T7} REDO={T3,T5,T6}

UNDO

28 O9=B28
24 O3=B24
18 O6=B18
13 D(O6)
9 O4=B9

REDO

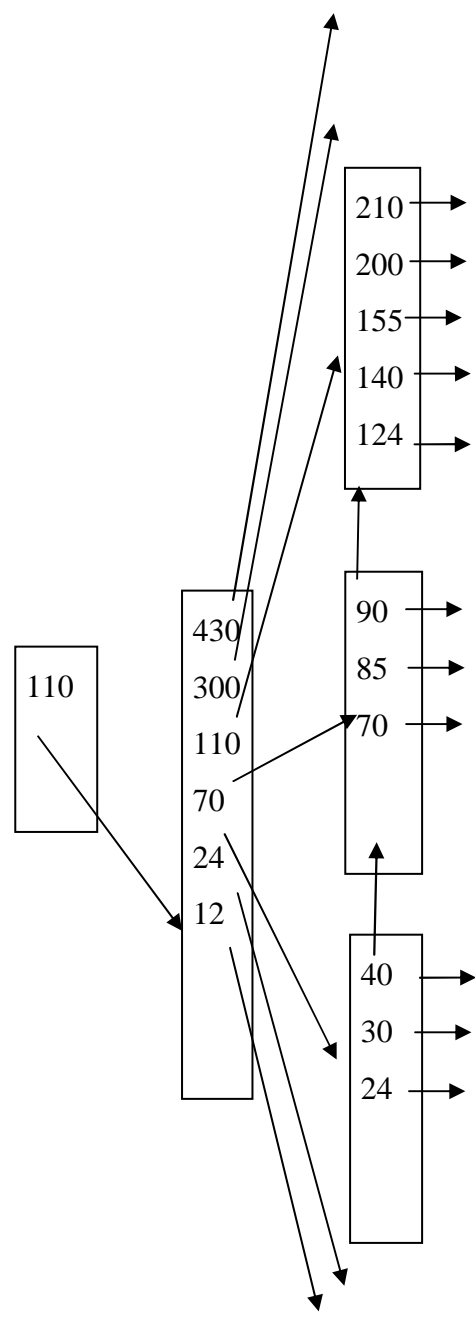
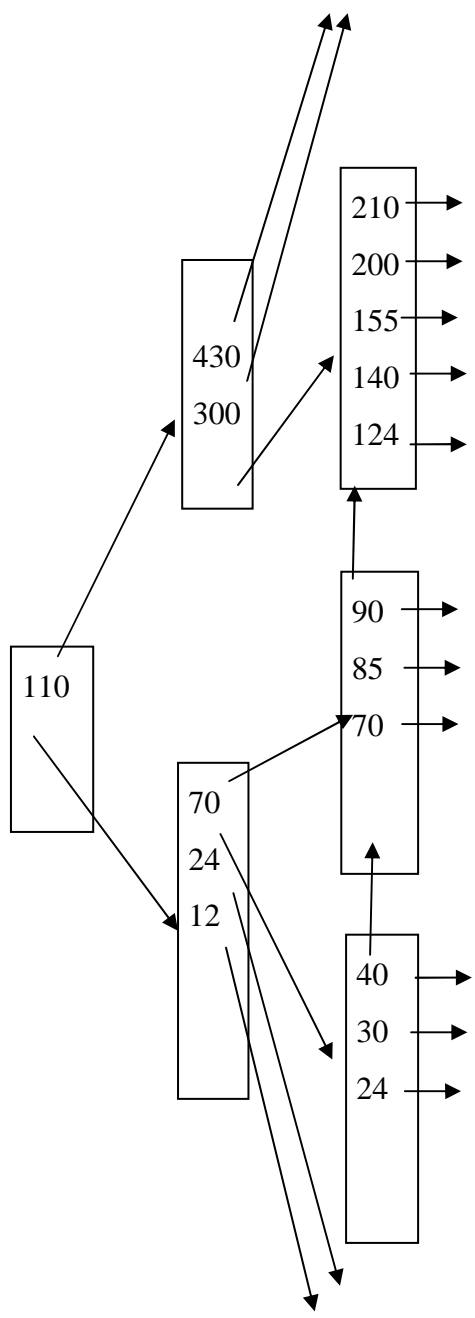
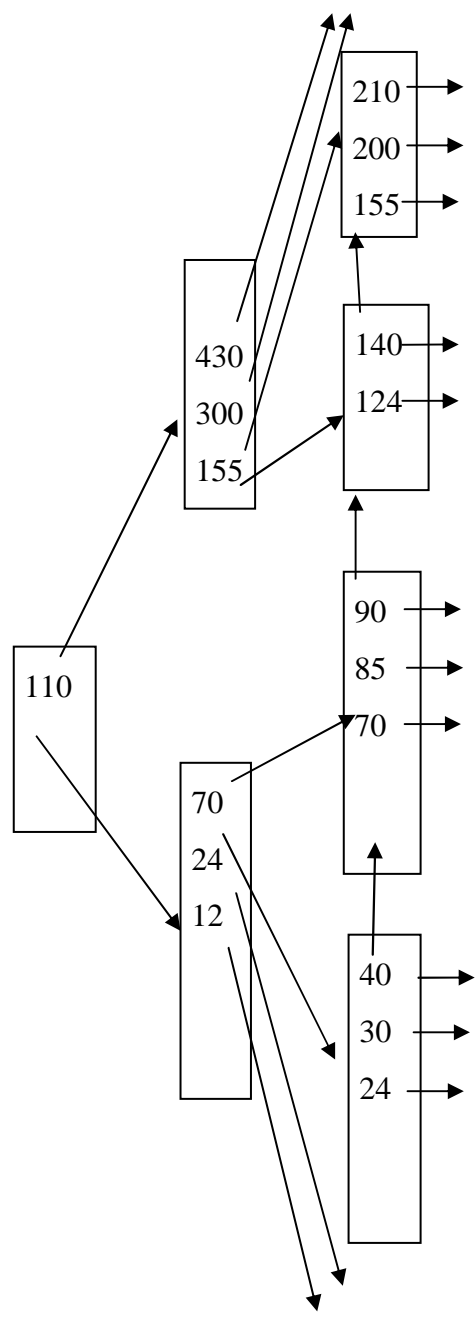
4 O1=A4
6 I(O3,A6)
10 D(O5)
12 O3=A12
14 I(O7,A14)
15 O7=A15
17 D(O4)
20 I(O8,A20)
22 O6=A22
25 D(O3)
26 I(O9,A26)
29 I(O10,A29)

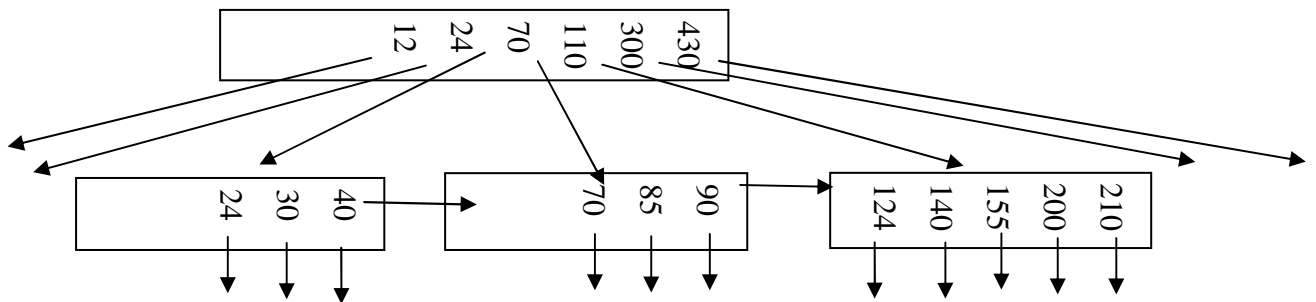
Esercizio 2

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

Lo schedale non rispetta il two-phase locking perche' t1 deve rilasciare il lock su y affinché t3 possa effettuare $w_3(y)$ ma poi deve acquisire un lock su x per effettuare $w_1(x)$ e non poteva acquisirlo prima perché t2 deve scrivere su x dopo $w_3(y)$

Esercizio 3





Esercizio 4

buffer di memoria centrale disponibili: $M=2.000$
 dimensione del buffer: $B=4.000$ bytes
 numero di tuple: $T(\text{Vendita})=200.000$, $T(\text{Articolo})=15.000$,
 dimensione delle tuple: $S(\text{Vendita})=1000$ bytes, $S(\text{Articolo})=2000$,
 $S(\text{CodArticolo})=100$
 $V(\text{Vendita}, \text{CodArticolo})=8000$ $V(\text{Vendita}, \text{Data})=60$
 $V(\text{Articolo}, \text{Prezzo})=10$

Calcoliamo innanzitutto il numero di blocchi occupati da ciascuna relazione:

$B(\text{Vendita})=T(\text{Vendita}) \cdot S(\text{Vendita})/B=200000 \cdot 1000/4000=50000$
 $B(\text{Articolo})=15000 \cdot 2000/4000=7500$

Prima sequenza: $(\sigma_{\text{Prezzo}=10} \text{Articolo}) \bowtie (\sigma_{\text{Data}='1/1/2000'} \text{Vendita})$

Sia $X = \sigma_{\text{Prezzo}=10} \text{Articolo}$

Utilizzando l'indice secondario su Prezzo

$\text{Costo}(X) = 3 + T(\text{Articolo})/V(\text{Articolo}, \text{Prezzo}) = 3 + 15000/10 = 1503$

$T(X) = T(\text{Articolo})/V(\text{Articolo}, \text{Prezzo}) = 15000/10 = 1500$

$B(X) = 1500 \cdot 2000/4000 = 750$

Sia $Y = \sigma_{\text{Data}='1/1/2000'} \text{Vendita}$

Utilizzando l'indice primario su Data

$\text{Costo}(Y) = 3 + \lceil B(\text{Vendita})/V(\text{Vendita}, \text{Data}) \rceil = 3 + \lceil 50000/60 \rceil = 3 + \lceil 833.3333 \rceil = 3 + 834 = 837$

$T(Y) = T(\text{Vendita})/V(\text{Vendita}, \text{Data}) = 200000/60 = 3333$

$B(Y) = 3333 \cdot 1000/4000 = 833$

X e Y stanno interamente in memoria centrale quindi posso usare il join a un passo

$\text{Costo}(X \bowtie Y) = B(X) + B(Y) = 750 + 833 = 1583$

$\text{Costo}((\sigma_{\text{Prezzo}=10} \text{Articolo}) \bowtie (\sigma_{\text{Data}='1/1/2000'} \text{Vendita})) = 1503 + 837 + 1583 = 3923$

Seconda sequenza: $\sigma_{\text{Data}='1/1/2000 \text{ AND } \text{Prezzo}=10}(\text{Articolo} \bowtie \text{Vendita})$

Sia $Z = \text{Articolo} \bowtie \text{Vendita}$

Né Articolo né Vendita stanno in memoria centrale, quindi provo con il join nested-loop basato sui blocchi e con l'hash join ibrido

Join nested-loop basato sui blocchi:

$$\text{Costo}(Z) = B(\text{Articolo}) + B(\text{Articolo})B(\text{Vendita}) / (M-1) = 7500 + 7500 * 50000 / 1999 = 195094$$

Hash join ibrido: si può fare perchè $2000 > \sqrt{7500} = 87$

$$\text{Costo}(Z) = (3 - 2 * 2000 / 7500) * (7500 + 50000) = 141833$$

Il migliore è l'hash-join ibrido basato sui blocchi.

$$T(Z) = T(\text{Articolo}) * T(\text{Vendita}) / \max\{V(\text{Articolo}, \text{CodArticolo}), V(\text{Vendita}, \text{CodArticolo})\} = 200000 * 15000 / 15000 = 200000$$

$$S(Z) = S(\text{Vendita}) + S(\text{Articolo}) - S(\text{CodArticolo}) = 1000 + 2000 - 100 = 2900$$

$$B(Z) = 200000 * 2900 / 4000 = 145000$$

Dato che non ci sono indici su Z:

$$\text{Costo}(\sigma_{\text{Data}='1/1/2000 \text{ AND } \text{Prezzo}=10} Z) = B(Z) = 145000$$

$$\text{Costo}(\sigma_{\text{Data}='1/1/2000 \text{ AND } \text{Prezzo}=10}(\text{Articolo} \bowtie \text{Vendita})) = 141833 + 145000 = 286833$$