

COMPITO DI SISTEMI INFORMATIVI/BASI DI DATI II

26 giugno 2006 (Tot. 16) Tempo: 2h

Esercizio 1 (punti 3)

Si consideri il seguente log:

- | | |
|--------------------|----------------------|
| 1. B(T1) | 15. D(T4,O3,B7) |
| 2. D(T1,O1,B1) | 16. B(T5) |
| 3. B(T2) | 17. I(T5,O5,A7) |
| 4. U(T1,O2,B2,A1) | 18. U(T4,O5,B8,A8) |
| 5. U(T2,O2,B3,A2) | 19. U(T2,O4,B9) |
| 6. I(T1,O3,A3) | 20. I(T3,O6,A9) |
| 7. B(T3) | 21. C(T3) |
| 8. D(T3,O2,B4) | 22. B(T6) |
| 9. U(T2,O3,B5,A4) | 23. D(T2,O4,B10) |
| 10. C(T1) | 24. I(T5,O7,A10) |
| 11. I(T3,O4,A5) | 25. C(T5) |
| 12. U(T2,O4,B6,A6) | 26. U(T6,O6,B11,A11) |
| 13. B(T4) | 27. D(T2,O5,B12) |
| 14. CK(T2,T3,T4) | 28. C(T4) |

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

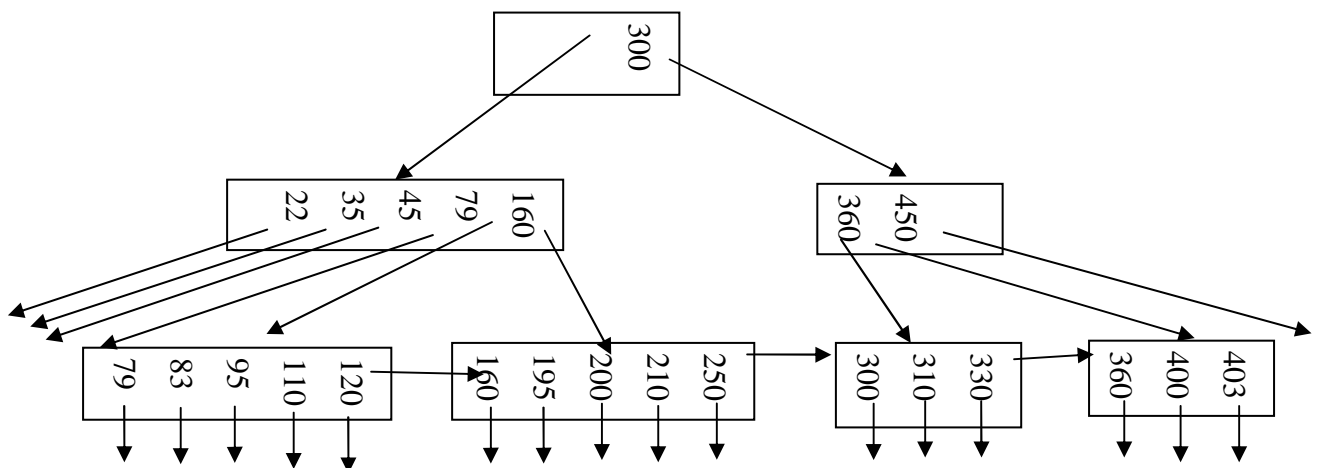
Esercizio 2 (punti 4)

Dato il seguente schedule:

r1(x) w3(z) w4(y) w2(x) r2(y) r3(x) w5(x) r3(y) r1(z)
si indichi se è view-equivalente allo schedule seriale
r1(x) r1(z) w3(z) r3(x) r3(y) w4(y) w2(x) r2(y) w5(x)

Esercizio 3 (punti 3)

Sia dato il seguente B+ tree di ordine 5.



Si mostri come si modifica l'albero nel caso di inserimento della chiave 205.

Esercizio 4 (punti 6)

Siano date le relazioni

Cliente(CodCli, Nome, Cognome, DataDiNascita, Città)

Venditore(CodVend, Nome, Cognome, Città)

Filiale(CodFil, Nome, Città)

la query

```
SELECT C.*, V.*, F.*
```

```
FROM Cliente AS C, Venditore AS V, Filiale AS F
```

```
WHERE C.Città=V.Città AND V.Città=F.Città
```

e i parametri:

buffer di memoria centrale disponibili: $M=100$

dimensione del buffer: $B=500$ bytes

numero di tuple: $T(\text{Cliente})=40.000$, $T(\text{Venditore})=10.000$, $T(\text{Filiale})=1.000$

dimensione delle tuple: $S(\text{Cliente})=100$ bytes, $S(\text{Venditore})=125$, $S(\text{Filiale})=250$, $S(\text{Città})=25$

numero di valori: $V(\text{Cliente}, \text{Città})=1.000$, $V(\text{Venditore}, \text{Città})=800$, $V(\text{Filiale}, \text{Città})=500$

Si stabilisca qual'è l'ordine migliore con cui eseguire i singoli join supponendo di utilizzare sempre l'Hash Join Ibrido. In particolare, occorrerà calcolare il costo delle tre sequenze:

$(\text{Cliente} \triangleright \langle \text{Venditore} \rangle) \triangleright \langle \text{Filiale} \rangle$, $(\text{Cliente} \triangleright \langle \text{Filiale} \rangle) \triangleright \langle \text{Venditore} \rangle$ e

$(\text{Venditore} \triangleright \langle \text{Filiale} \rangle) \triangleright \langle \text{Cliente} \rangle$.

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

SOLUZIONE

Esercizio 1

1. B(T1)
 2. D(T1,O1,B1)
 3. B(T2)
 4. U(T1,O2,B2,A1)
 5. U(T2,O2,B3,A2)
 6. I(T1,O3,A3)
 7. B(T3)
 8. D(T3,O2,B4)
 9. U(T2,O3,B5,A4)
 10. C(T1)
 11. I(T3,O4,A5)
 12. U(T2,O4,B6,A6)
 13. B(T4)
 14. CK(T2,T3,T4)
 15. D(T4,O3,B7)
 16. B(T5)
 17. I(T5,O5,A7)
 18. U(T4,O5,B8,A8)
 19. U(T2,O4,B9)
 20. I(T3,O6,A9)
 21. C(T3)
 22. B(T6)
 23. D(T2,O4,B10)
 24. I(T5,O7,A10)
 25. C(T5)
 26. U(T6,O6,B11,A11)
 27. D(T2,O5,B12)
 28. C(T4)
- 14 UNDO={T2,T3,T4} REDO={}
- 16 UNDO={T2,T3,T4,T5} REDO={}
- 21 UNDO={T2,T4,T5} REDO={T3}
- 22 UNDO={T2,T4,T5,T6} REDO={T3}
- 25 UNDO={T2,T4,T6} REDO={T3,T5}
- 28 UNDO={T2,T6} REDO={T3,T4,T5}
- UNDO
- 27 I(O5,B12)
- 26 O6=B11
- 23 I(O4,B10)
- 19 O4=B9
- 12 O4=B6
- 9 O3=B5
- 5 O2=B3
- REDO
- 8 D(O2)
- 11 I(O4,A5)
- 15 D(O3)
- 17 I(O5,A7)
- 18 O5=A8
- 20 I(O6,A9)
- 24 I(O7,A10)

Esercizio 2

$r1(x) w3(z) w4(y) w2(x) r2(y) r3(x) w5(x) r3(y) r1(z)$

"legge-da" = $\{(t1,t3),(t2,t4),(t3,t4),(t3,t2)\}$

scritture finali = $\{(t5,x),(t4,y),(t3,z)\}$

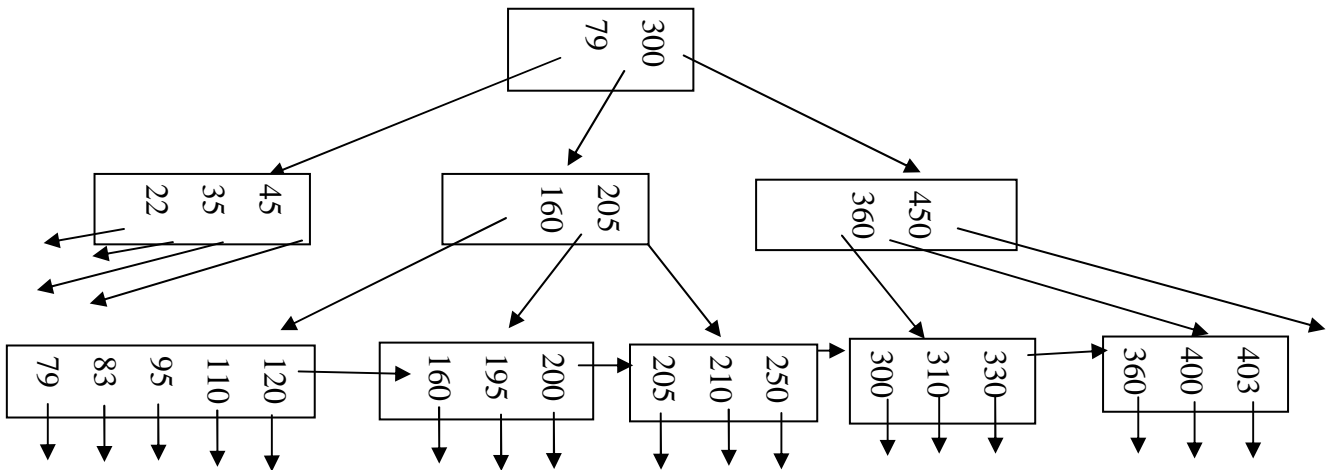
$r1(x) r1(z) w3(z) r3(x) r3(y) w4(y) w2(x) r2(y) w5(x)$

"legge-da" = $\{(t2,t4)\}$

scritture finali = $\{(t5,x),(t4,y),(t3,z)\}$

I due schedule non sono view-equivalenti

Esercizio 3



Esercizio 4

buffer di memoria centrale disponibili: $M=100$

dimensione del buffer: $B=500$ bytes

numero di tuple: $T(\text{Cliente})=40.000$, $T(\text{Venditore})=10.000$, $T(\text{Filiale})=1.000$

dimensione delle tuple: $S(\text{Cliente})=100$ bytes, $S(\text{Venditore})=125$, $S(\text{Filiale})=250$, $S(\text{Città})=25$

numero di valori: $V(\text{Cliente},\text{Città})=1.000$, $V(\text{Venditore},\text{Città})=800$, $V(\text{Filiale},\text{Città})=500$

Calcoliamo innanzitutto il numero di blocchi occupati da ciascuna relazione:

$$B(\text{Cliente})=T(\text{Cliente}) \cdot S(\text{Cliente})/B=40.000 \cdot 100/500=8.000$$

$$B(\text{Venditore})=10.000 \cdot 125/500=2.500$$

$$B(\text{Filiale})=1.000 \cdot 250/500=500$$

Prima sequenza: $(\text{Cliente} \triangleright \langle \text{Venditore} \rangle) \triangleright \langle \text{Filiale} \rangle$

$$\text{Costo}(\text{Cliente} \triangleright \langle \text{Venditore} \rangle)=(3-2 \cdot 100/2.500) \cdot (2.500+8.000)=30.660$$

Sia $X=\text{Cliente} \triangleright \langle \text{Venditore} \rangle$

$$S(X)=S(\text{Cliente})+S(\text{Venditore})-S(\text{Città})=100+125-25=200$$

$$T(X)=T(\text{Cliente}) \cdot T(\text{Venditore})/\max(V(\text{Cliente},\text{Città}),V(\text{Venditore},\text{Città}))=40.000 \cdot 10.000/1.000=400.000$$

$$B(X)=400.000 \cdot 200/500=160.000$$

$$\text{Costo}(X \triangleright \langle \text{Filiale} \rangle)=(3-2 \cdot 100/500) \cdot (500+160.000)=417.300$$

$$\text{Costo}((\text{Cliente} \triangleright \langle \text{Venditore} \rangle) \triangleright \langle \text{Filiale} \rangle)=30.660+417.300=447.960$$

Seconda sequenza: $(\text{Cliente} \triangleright \langle \text{Filiale} \rangle) \triangleright \langle \text{Venditore} \rangle$

$$\text{Costo}(\text{Cliente} \triangleright \langle \text{Filiale} \rangle)=(3-2 \cdot 100/500) \cdot (8.000+500)=22.100$$

Sia $Y=\text{Cliente} \triangleright \langle \text{Filiale} \rangle$

$$S(Y)=100+250-25=325$$

$$T(Y)=40.000 \cdot 1.000/1.000=40.000$$

$$B(Y)=40.000 \cdot 325/500=26.000$$

$$\text{Costo}(Y \triangleright \langle \text{Venditore} \rangle)=(3-2 \cdot 100/2.500) \cdot (26.000+2.500)=83.220$$

$$\text{Costo}((\text{Cliente} \triangleright \langle \text{Filiale} \rangle) \triangleright \langle \text{Venditore} \rangle)=83.220+22.100=105.320$$

Terza sequenza: $(\text{Venditore} \triangleright \langle \text{Filiale} \rangle) \triangleright \langle \text{Cliente} \rangle$

$$\text{Costo}(\text{Venditore} \triangleright \langle \text{Filiale} \rangle)=(3-2 \cdot 100/500) \cdot (2.500+500)=7.800$$

Sia $Z=\text{Venditore} \triangleright \langle \text{Filiale} \rangle$

$$S(Z)=125+250-25=350$$

$$T(Z)=10.000*1.000/800=12.500$$

$$B(Z)=12.500*350/500=8.750$$

$$\text{Costo}(Z \triangleright \triangleleft \text{Cliente})=(3-2*100/8.000)(8.750+8.000)=49.831$$

$$\text{Costo}((\text{Venditore} \triangleright \triangleleft \text{Filiale}) \triangleright \triangleleft \text{Cliente})=7.800+49.831=57.631$$

La sequenza migliore per realizzare il join è la terza sequenza.