

COMPITO DI SISTEMI INFORMATIVI

9 gennaio 2006 (Tot. 16) Tempo: 2h

Esercizio 1 (punti 3)

Si consideri il seguente log:

- | | |
|-------------------|--------------------|
| 1. B(T1) | 14. U(T2,O5,B6,A6) |
| 2. B(T3) | 15. B(T5) |
| 3. I(T3,O1,A1) | 16. U(T5,O6,B7,A7) |
| 4. U(T1,O2,B1,A2) | 17. U(T4,O6,B8,A8) |
| 5. D(T3,O3,B2) | 18. C(T4) |
| 6. I(T3,O4,A3) | 19. B(T6) |
| 7. B(T4) | 20. D(T6,O6,B9) |
| 8. U(T4,O4,B3,A4) | 21. C(T1) |
| 9. D(T4,O4,B4) | 22. I(T2,O7,A9) |
| 10. CK(T1,T3,T4) | 23. C(T2) |
| 11. I(T3,O5,A5) | 24. I(T5,O8,A10) |
| 12. D(T1,O1,B5) | 25. I(T6,O9,A11) |
| 13. B(T2) | 26. D(T5,O7,B10) |

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

Esercizio 2 (punti 4)

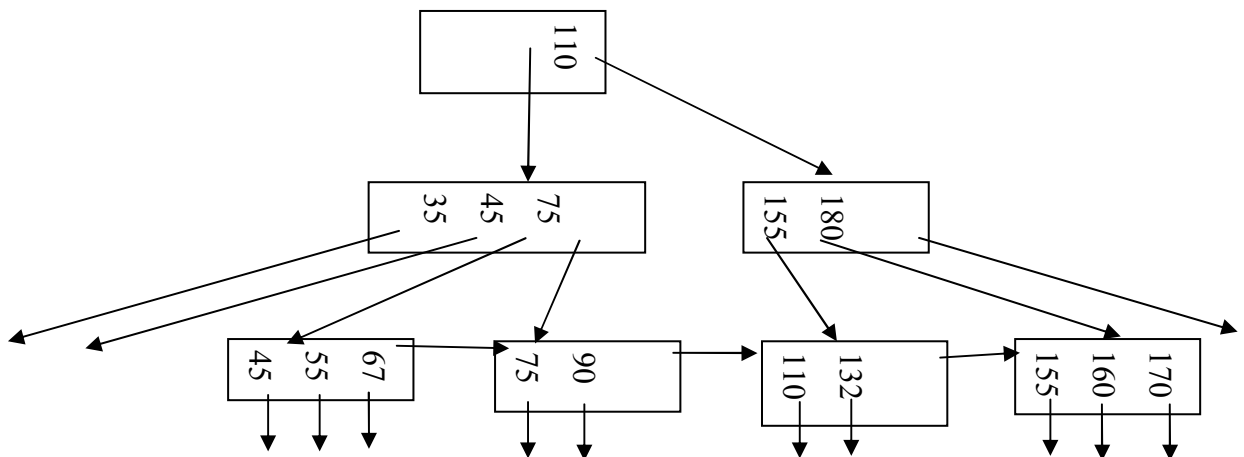
Dato il seguente schedule:

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

si dica se è conflict serializzabile.

Esercizio 3 (punti 3)

Sia dato il seguente B+ tree di ordine 4.



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

Esercizio 4 (punti 6)

Siano date le relazioni

Impiegato(CodImp, Nome, Cognome, DataDiNascita, Dipartimento, Stipendio)

Consulente(CodCons, Nome, Cognome, Dipartimento, Stipendio, Indirizzo)

la query

```
SELECT I.*, C.*
```

```
FROM Impiegato AS I, Consulente AS C
```

```
WHERE I.Dipartimento=C.Dipartimento AND I.Nome="Francesco"
```

e i parametri:

buffer di memoria centrale disponibili: $M=100$

dimensione del buffer: $B=500$ bytes

numero di tuple: $T(\text{Impiegato})=250.000$, $T(\text{Consulente})=800$,

dimensione delle tuple: $S(\text{Impiegato})=160$ bytes, $S(\text{Consulente})=250$, $S(\text{Dipartimento})=50$

numero di valori: $V(\text{Impiegato}, \text{Dipartimento})=2.000$, $V(\text{Consulente}, \text{Dipartimento})=500$,

$V(\text{Impiegato}, \text{Nome})=400$

Si stabilisca qual'è l'ordine migliore con cui eseguire le operazioni supponendo di utilizzare l'Hash Join Ibrido. In particolare, occorrerà calcolare il costo delle due sequenze:

$(\sigma_{\text{Impiegato.Nome}='Francesco'} \text{Impiegato}) \bowtie \text{Consulente}$ e

$\sigma_{\text{Impiegato.Nome}='Francesco'}(\text{Impiegato} \bowtie \text{Consulente})$.

Si supponga che i record siano impaccati nei buffer e che tutto lo spazio nei buffer sia occupato dai record. Si supponga inoltre che ci sia un indice secondario su Impiegato.Nome. Si supponga che non vengano costruiti indici su Impiegato \bowtie Consulente.

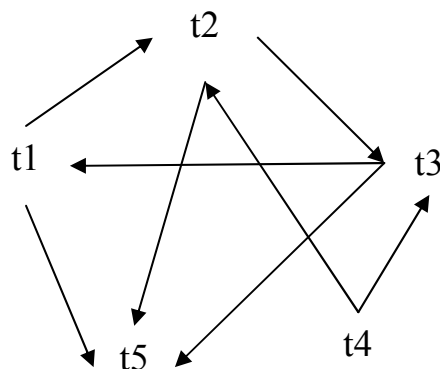
SOLUZIONE

Esercizio 1

- 27. B(T1)
 - 28. B(T3)
 - 29. I(T3,O1,A1)
 - 30. U(T1,O2,B1,A2)
 - 31. D(T3,O3,B2)
 - 32. I(T3,O4,A3)
 - 33. B(T4)
 - 34. U(T4,O4,B3,A4)
 - 35. D(T4,O4,B4)
 - 36. CK(T1,T3,T4)
 - 37. I(T3,O5,A5)
 - 38. D(T1,O1,B5)
 - 39. B(T2)
 - 40. U(T2,O5,B6,A6)
 - 41. B(T5)
 - 42. U(T5,O6,B7,A7)
 - 43. U(T4,O6,B8,A8)
 - 44. C(T4)
 - 45. B(T6)
 - 46. D(T6,O6,B9)
 - 47. C(T1)
 - 48. I(T2,O7,A9)
 - 49. C(T2)
 - 50. I(T5,O8,A10)
 - 51. I(T6,O9,A11)
 - 52. D(T5,O7,B10)
 - 10 UNDO={T1,T3,T4} REDO={}
 - 13 UNDO={T1,T2,T3,T4} REDO={}
 - 15 UNDO={T1,T2,T3,T4,T5} REDO={}
 - 18 UNDO={T1,T2,T3,T5} REDO={T4}
 - 19 UNDO={T1,T2,T3,T5,T6} REDO={T4}
 - 21 UNDO={T2,T3,T5,T6} REDO={T1,T4}
 - 23 UNDO={T3,T5,T6} REDO={T1,T2,T4}
- UNDO
- 26 I(O7,B10)
 - 25 D(O9)
 - 24 D(O8)
 - 20 I(O6,B9)
 - 16 O6=B7
 - 11 D(O5)
 - 6 D(O4)
 - 5 I(O3,B2)
 - 3 D(O1)
- REDO
- 4 O2=A2
 - 8 O4=A4
 - 9 D(O4)
 - 12 D(O1)
 - 14 O5=A6
 - 17 O6=A8
 - 22 I(O7,A9)

Esercizio 2

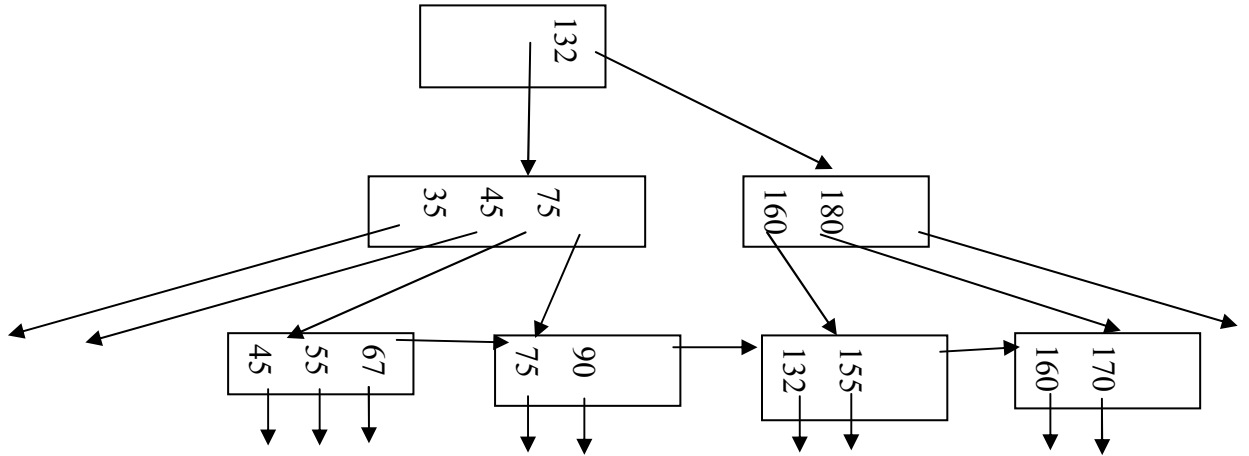
r1(x) w3(z) w4(y) w2(x) r2(y) r3(x) w5(x) r3(y) r1(z)



C'è un ciclo quindi non è conflict serializzabile.

Esercizio 3

Cancellazione di 110



La radice può cambiare o rimanere uguale, l'albero e' ugualmente corretto.

Esercizio 4

buffer di memoria centrale disponibili: $M=100$

dimensione del buffer: $B=500$ bytes

numero di tuple: $T(\text{Impiegato})=250.000$, $T(\text{Consulente})=800$,

dimensione delle tuple: $S(\text{Impiegato})=160$ bytes, $S(\text{Consulente})=250$, $S(\text{Dipartimento})=50$

numero di valori: $V(\text{Impiegato},\text{Dipartimento})=2.000$, $V(\text{Consulente},\text{Dipartimento})=500$,

$V(\text{Impiegato},\text{Nome})=400$

Calcoliamo innanzitutto il numero di blocchi occupati da ciascuna relazione:

$B(\text{Impiegato})=T(\text{Impiegato}) \cdot S(\text{Impiegato})/B = 250.000 \cdot 160/500 = 80.000$

$B(\text{Consulente})=800 \cdot 250/500 = 400$

Prima sequenza: $(\sigma_{\text{Impiegato.Nome}='Francesco'} \text{Impiegato}) \bowtie \text{Consulente}$

Sia $X = \sigma_{\text{Impiegato.Nome}='Francesco'} \text{Impiegato}$

$\text{Costo}(X) = 3 + \lceil T(\text{Impiegato})/V(\text{Impiegato},\text{Nome}) \rceil = 3 + \lceil 250.000/400 \rceil = 3 + 625 = 628.$

$T(X) = 250.000/400 = 625$

$B(X) = 625 \cdot 160/500 = 200$

$\text{Costo}(X \bowtie \text{Consulente}) = (3 - 2 \cdot 100/200) \cdot (200 + 400) = 1.200$

$\text{Costo}((\sigma_{\text{Impiegato.Nome}='Francesco'} \text{Impiegato}) \bowtie \text{Consulente}) = 628 + 1.200 = 1.828$

Seconda sequenza: $\sigma_{\text{Impiegato.Nome}='Francesco'}(\text{Impiegato} \bowtie \text{Consulente})$

Sia $Y = \text{Impiegato} \bowtie \text{Consulente}$

$\text{Costo}(Y) = (3 - 2 \cdot 100/400) \cdot (80.000 + 400) = 201.000$

$S(Y) = S(\text{Impiegato}) + S(\text{Consulente}) - S(\text{Dipartimento}) = 160 + 250 - 50 = 360$

$T(Y) = T(\text{Impiegato}) \cdot T(\text{Consulente}) / \max(V(\text{Impiegato},\text{Dipartimento}), V(\text{Consulente},\text{Dipartimento}))$
 $= 250.000 \cdot 800 / 2.000 = 100.000$

$B(Y) = 100.000 \cdot 360/500 = 72.000$

$\text{Costo}(\sigma_{\text{Impiegato.Nome}='Francesco'} Y) = B(Y) = 72.000$

$\text{Costo}(\sigma_{\text{Impiegato.Nome}='Francesco'}(\text{Impiegato} \bowtie \text{Consulente})) = 72.000 + 201.000 = 273.000$

La sequenza migliore per realizzare l'operazione è la prima.