

COMPITO DI SISTEMI INFORMATIVI/BASI DI DATI II

24 gennaio 2008 (Tot. 16) Tempo: 2h

Esercizio 1 (punti 3)

Si consideri il seguente log:

- | | |
|----------------------|----------------------|
| 1. D(T1,O1,B1) | 16. U(T4,O4,B16,A16) |
| 2. B(T2) | 17. D(T2,O8,B17) |
| 3. I(T2,O2,A3) | 18. C(T4) |
| 4. U(T1,O3,B4,A4) | 19. B(T5) |
| 5. B(T3) | 20. U(T5,O4,B20,A20) |
| 6. I(T3,O4,A6) | 21. B(T6) |
| 7. D(T2,O5,B7) | 22. I(T6,O9,A22) |
| 8. U(T3,O4,B8,A8) | 23. U(T5,O9,B23,A23) |
| 9. C(T1) | 24. U(T6,O6,B24,A24) |
| 10. B(T4) | 25. U(T2,O2,B25,A25) |
| 11. U(T4,O2,B11,A11) | 26. D(T3,O2,B26) |
| 12. I(T2,O6,A12) | 27. C(T3) |
| 13. U(T3,O6,B13,A13) | 28. I(T2,O10,A28) |
| 14. D(T2,O7,B14) | 29. C(T6) |
| 15. CK(T2,T3,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)

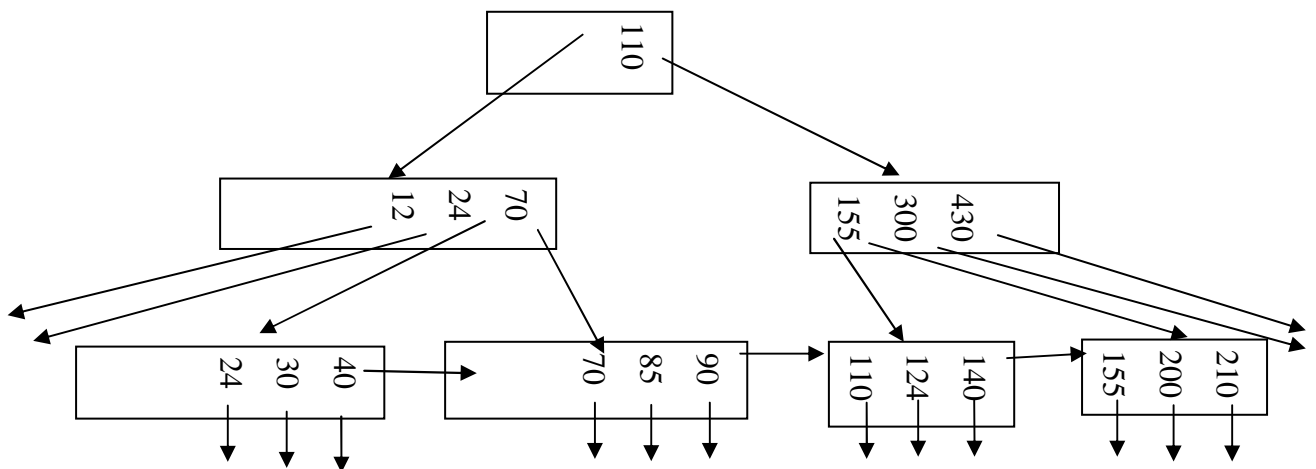
Dato il seguente schedule:

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

si indichi se è conflict serializzabile.

Esercizio 3 (punti 3)

Sia dato il seguente B+ tree di ordine 3.



Si mostri come si modifica l'albero nel caso di inserimento della chiave 76

Esercizio 4 (punti 6)

Siano date le relazioni

Prof(CodiceProf, Nome, Cognome, Titolo, Dipartimento, Università, Indirizzo)

Corso(CodiceCorso, Nome, Docente, Università, Trimestre, Ore)

la query

```
SELECT P.*, C.*
FROM Corso AS C, Professore AS P
WHERE P.CodiceProf=C.Docente AND C.Ore>30
```

e i parametri:

buffer di memoria centrale disponibili: $M=5000$

dimensione del buffer: $B=4.000$ bytes

numero di tuple: $T(\text{Prof})= 80000$, $T(\text{Corso})= 16000$,

dimensione delle tuple: $S(\text{Prof})=2000$ bytes, $S(\text{Corso})=4000$,

$S(\text{CodiceProf})=200$

numero di valori: $V(\text{Corso}, \text{Docente})=8000$,

Si calcoli il costo minimo di ciascuna delle seguenti sequenze:

$(\sigma_{\text{Ore}>30} \text{ Corso}) \triangleright \triangleleft \text{Prof}$

$\sigma_{\text{Ore}>30} (\text{Corso} \triangleright \triangleleft \text{Prof})$

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

Si supponga di avere soltanto un indice primario su Prof.CodiceProf.

SOLUZIONE

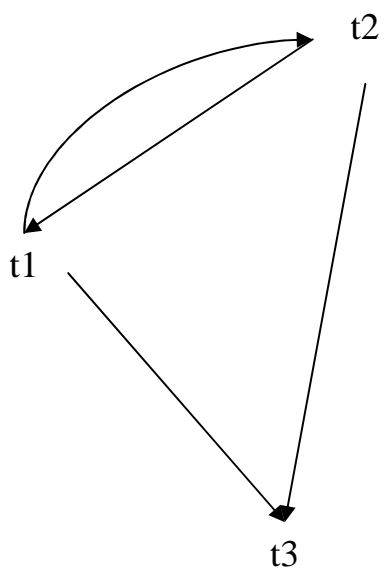
Esercizio 1

1. D(T1,O1,B1)
2. B(T2)
3. I(T2,O2,A3)
4. U(T1,O3,B4,A4)
5. B(T3)
6. I(T3,O4,A6)
7. D(T2,O5,B7)
8. U(T3,O4,B8,A8)
9. C(T1)
10. B(T4)
11. U(T4,O2,B11,A11)
12. I(T2,O6,A12)
13. U(T3,O6,B13,A13)
14. D(T2,O7,B14)
15. CK(T2,T3,T4)
16. U(T4,O4,B16,A16)
17. D(T2,O8,B17)
18. C(T4)
19. B(T5)
20. U(T5,O4,B20,A20)
21. B(T6)
22. I(T6,O9,A22)
23. U(T5,O9,B23,A23)
24. U(T6,O6,B24,A24)
25. U(T2,O2,B25,A25)
26. D(T3,O2,B26)
27. C(T3)
28. I(T2,O10,A28)
29. C(T6)

15 UNDO={T2,T3,T4 } REDO={}
18 UNDO={T2,T3 } REDO={T4}
19 UNDO={T2,T3,T5 } REDO={T4}
21 UNDO={T2,T3,T5,T6 } REDO={T4}
27 UNDO={T2, T5,T6 } REDO={T3,T4}
29 UNDO={T2, T5 } REDO={T3,T4,T6}
UNDO
28 D(O10)
25 O2=B25
23 O9=B23
20 O4=B20
17 I(O8,B17)
14 I(O7,B14)
12 D(O6)
7 I(O5,B7)
3 D(O2)
REDO
6 I(O4,A6)
8 O4=A8
11 O2=A11
13 O6=A13
16 O4=A16
22 I(O9,A22)
24 O6=A24
26 D(O2)

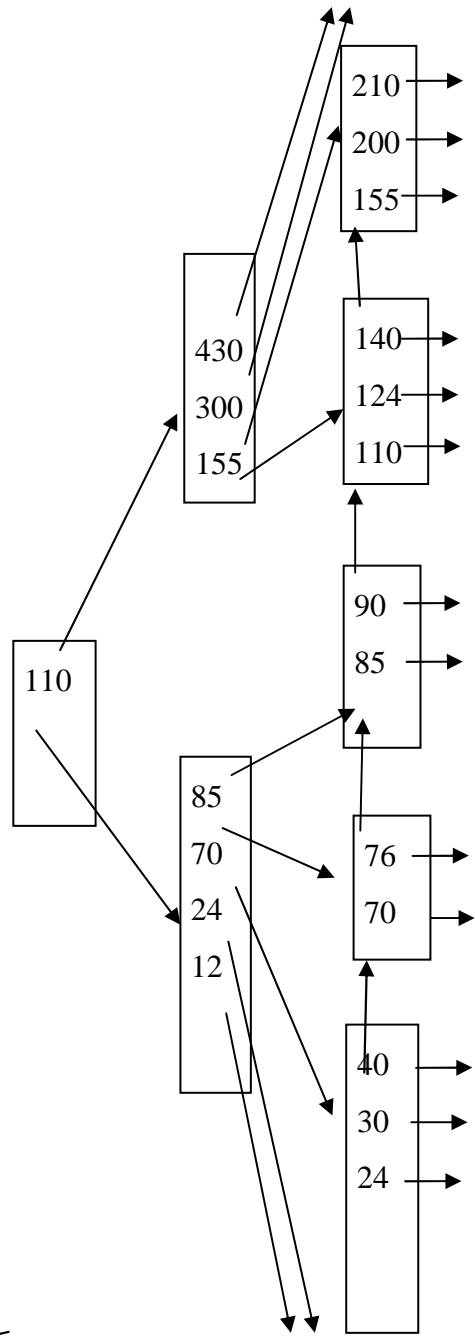
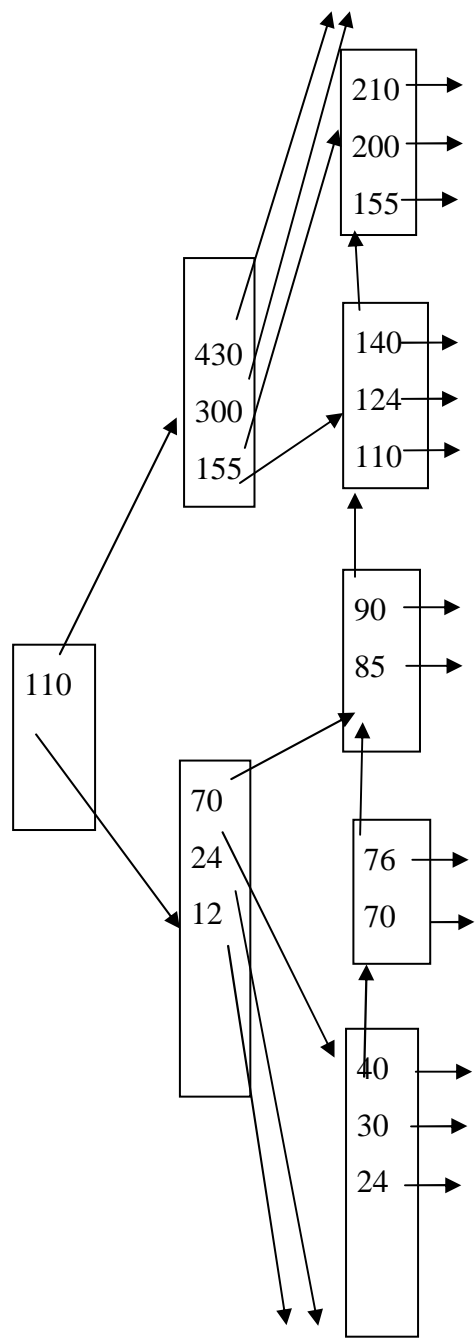
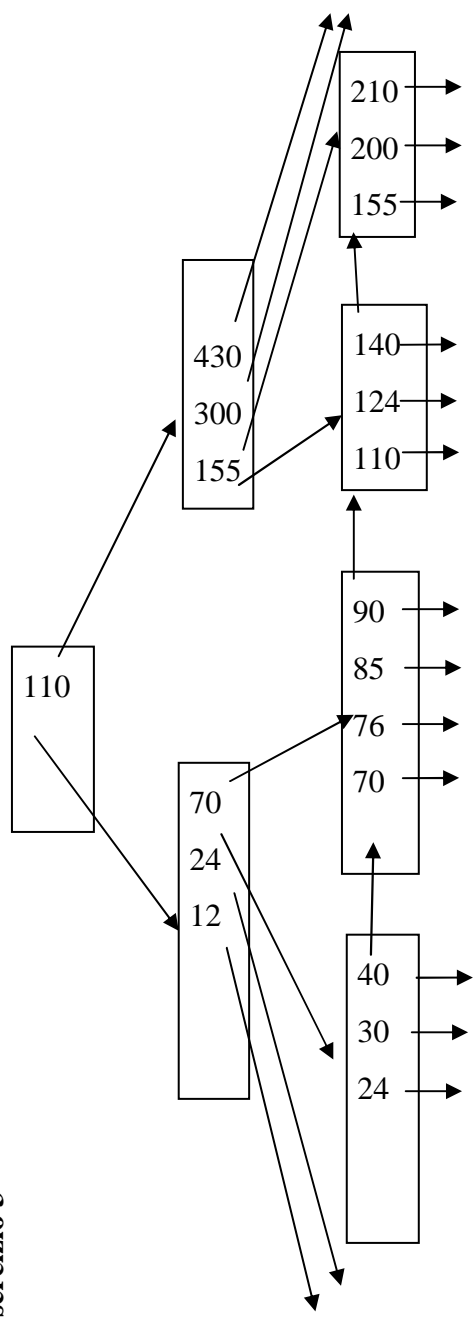
Esercizio 2

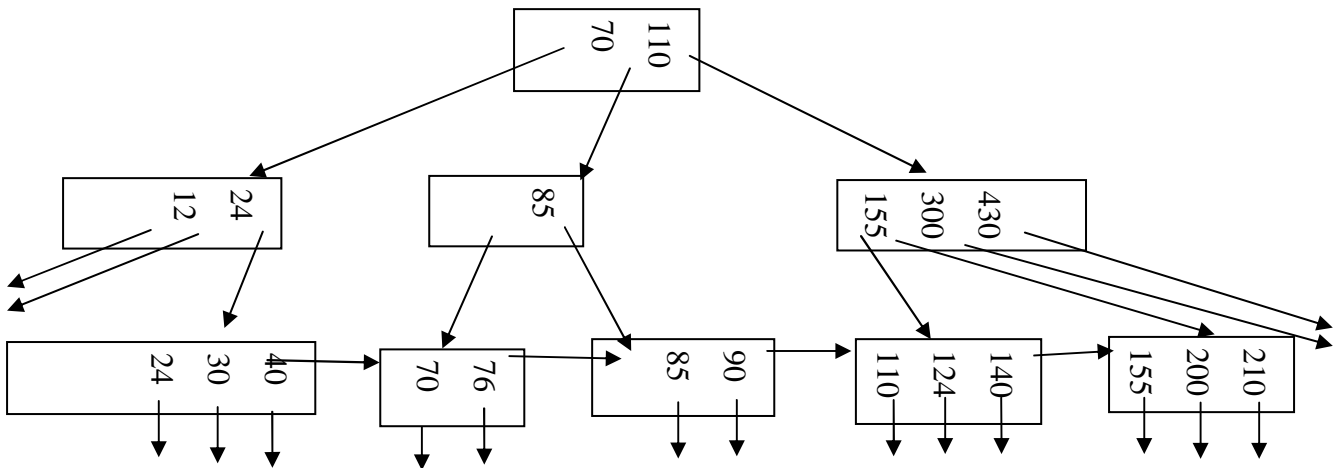
r1(x) r2(x) w1(x) w1(w) w3(y) r2(x) w2(x) r3(w) w2(z) r2(w) r3(x)



Lo schedale non è conflict serializzabile.

Esercizio 3





Esercizio 4

buffer di memoria centrale disponibili: $M=5000$

dimensione del buffer: $B=4.000$ bytes

numero di tuple: $T(\text{Prof})= 80000$, $T(\text{Corso})= 16000$,

dimensione delle tuple: $S(\text{Prof})=2000$ bytes, $S(\text{Corso})=4000$,
 $S(\text{CodiceProf})=200$

numero di valori: $V(\text{Corso},\text{Docente})=8000$,

Si calcoli il costo minimo di ciascuna delle seguenti sequenze:

$(\sigma_{\text{Ore}>30} \text{Corso}) \triangleright \triangleleft \text{Prof}$

$\sigma_{\text{Ore}>30} (\text{Corso} \triangleright \triangleleft \text{Prof})$

Calcoliamo innanzitutto il numero di blocchi occupati da ciascuna relazione:

$B(\text{Corso})=T(\text{Corso}) * S(\text{Corso}) / B = 16000 * 4000 / 4000 = 16000$

$B(\text{Prof}) = 80000 * 2000 / 4000 = 40000$

Prima sequenza: $(\sigma_{\text{Ore}>30} \text{Corso}) \triangleright \triangleleft \text{Prof}$

Sia $X = \sigma_{\text{Ore}>30} \text{Corso}$

Non ci sono indici su Ore quindi

$\text{Costo}(X) = B(\text{Corso}) = 16000$

$T(X) = B(\text{Corso}) / 3 = 16000 / 3 = 5333$

$B(X) = 5333 * 4000 / 4000 = 5333$

Né X né Prof 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 visto che c'è un indice su CodiceProf.

$\text{CostoJNL}(X \triangleright \triangleleft \text{Prof}) = B(X) + B(X) * B(\text{Prof}) / M = 5333 + 5333 * 40000 / 5000 = 47997$

$\text{CostoHJI}(X \triangleright \triangleleft \text{Prof}) = (3 - 2M/B(X)) * (B(X) + B(\text{Prof})) = (3 - 2 * 5000 / 5333) * (5333 + 40000) = 50994$

$\text{CostoIJ}(X \triangleright \triangleleft \text{Prof}) = B(X) + T(X) * \lceil B(\text{Prof}) / V(\text{Prod}, \text{CodiceProf}) \rceil = 5333 + 5333 * \lceil 40000 / 80000 \rceil = 5333 + 5333 * 1 = 10666$

Costo totale = $16000 + 10666 = 26666$

Seconda sequenza: $\sigma_{Ore > 30}(\text{Corso} \bowtie \text{Prof})$

Sia $Z = \text{Corso} \bowtie \text{Prof}$

Né Corso né Prof 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 visto che c'è un indice su CodiceProf.

$\text{CostoJNL}(Z) = B(\text{Corso}) + B(\text{Corso}) * B(\text{Prof}) / M = 16000 + 16000 * 40000 / 5000 = 144000$

$\text{CostoHJI}(Z) = (3 - 2M/B(\text{Corso})) * (B(\text{Prof}) + B(\text{Corso})) = (3 - 2 * 5000 / 16000) * (16000 + 40000) = 133000$

$\text{CostoIJ}(Z) = B(\text{Corso}) + T(\text{Corso}) \lceil B(\text{Prof}) / V(\text{Prod}, \text{CodiceProf}) \rceil = 16000 + 16000 * 1 = 32000$

$T(Z) = T(\text{Corso}) * T(\text{Prof}) / \max\{V(\text{Corso}, \text{Docente}), V(\text{Prof}, \text{CodiceProf})\} =$

$80000 * 16000 / \max\{8000, 80000\} = 16000$

$S(Z) = S(\text{Corso}) + S(\text{Prof}) - S(\text{CodiceProf}) = 4000 + 2000 - 200 = 5800$

$B(Z) = 16000 * 5800 / 4000 = 23200$

Dato che non ci sono indici su Z:

$\text{Costo}(\sigma_{Ore > 30} Z) = B(Z) = 23200$

$\text{Costo totale} = 32000 + 23200 = 55200$