

COMPITO DI SISTEMI INFORMATIVI

1 febbraio 2016 (Tot. 16) Tempo: 2h

Esercizio 1 (punti 3)

Si consideri il seguente log:

- | | |
|----------------------|----------------------|
| 1. D(T1,O1,B1) | 15. B(T5) |
| 2. U(T1,O2,B2,A2) | 16. I(T5,O6,A16) |
| 3. B(T2) | 17. U(T5,O7,B17,A17) |
| 4. U(T2,O2,B4,A4) | 18. D(T2,O8,B18) |
| 5. B(T3) | 19. C(T2) |
| 6. B(T4) | 20. U(T3,O7,B20,A20) |
| 7. D(T4,O2,B7) | 21. C(T3) |
| 8. U(T3,O3,B8,A8) | 22. C(T4) |
| 9. I(T4,O4,A9) | 23. B(T6) |
| 10. U(T2,O4,B10,A10) | 24. D(T6,O10,B24) |
| 11. C(T1) | 25. D(T5,O7,B25) |
| 12. I(T3,O5,A12) | 26. I(T6,O6,A26) |
| 13. D(T2,O5,B13) | 27. U(T5,O9,B27,A27) |
| 14. 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)

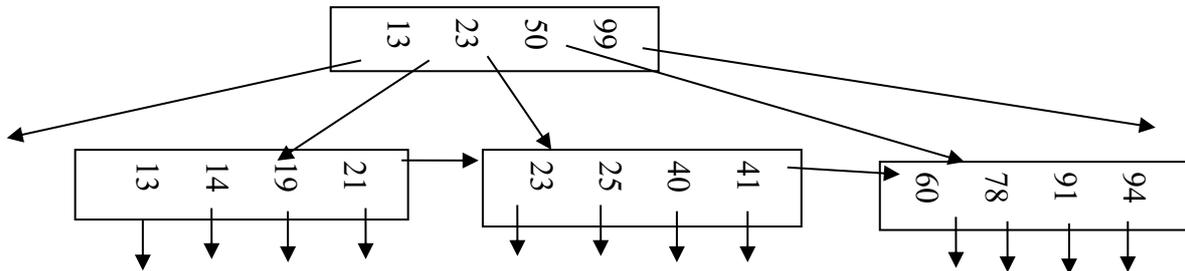
Dato il seguente schedule:

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

si indichi se è conflict serializzabile.

Esercizio 3 (punti 3)

Sia dato il seguente B+ tree di ordine 4.



Si mostri l'albero risultante dall'inserimento della chiave 24

Esercizio 4 (punti 6)

Siano date le seguenti relazioni

Studente(CodStu, Nome, Cognome, Via, Città, CAP, Università, Telefono)

Professore(CodPro, Nome, Cognome, Via, Città, CAP, Università, Telefono, OreLezione)

la query

```
SELECT S.*, P.*
```

```
FROM Professore AS P, Studenti AS S
```

```
WHERE P.Univerità= S.Univerità AND P.OreLezione>60
```

e i parametri:

buffer di memoria centrale disponibili: $M=500.000$

dimensione del buffer: $B=20.000$ bytes

numero di tuple: $T(\text{Studente})= 5.000.000$, $T(\text{Professore})= 2.000.000$,

dimensione delle tuple: $S(\text{Studente})=10.000$, $S(\text{Professore})=10.000$, $S(\text{Università})=50$

numero di valori $V(\text{Studente}, \text{Università})=500$, $V(\text{Professore}, \text{Università})=400$

$\text{Min}(\text{Professore}, \text{OreLezione})=0$ $\text{Max}(\text{Professore}, \text{OreLezione})=150$

Si calcoli il costo minimo di ciascuna delle seguenti sequenze:

$(\sigma_{\text{OreLezione}>60} \text{Professore}) \triangleright \triangleleft \text{Studente}$

$\sigma_{\text{OreLezione}>60} (\text{Professore} \triangleright \triangleleft \text{Studente})$

Si indichi quale delle sequenze ha costo inferiore.

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 primario su Studente.Università e uno secondario su Professore.OreLezione.

SOLUZIONE

Esercizio 1

1. D(T1,O1,B1)
2. U(T1,O2,B2,A2)
3. B(T2)
4. U(T2,O2,B4,A4)
5. B(T3)
6. B(T4)
7. D(T4,O2,B7)
8. U(T3,O3,B8,A8)
9. I(T4,O4,A9)
10. U(T2,O4,B10,A10)
11. C(T1)
12. I(T3,O5,A12)
13. D(T2,O5,B13)
14. CK(T2,T3,T4)
15. B(T5)
16. I(T5,O6,A16)
17. U(T5,O7,B17,A17)
18. D(T2,O8,B18)
19. C(T2)
20. U(T3,O7,B20,A20)
21. C(T3)
22. C(T4)
23. B(T6)
24. D(T6,O10,B24)
25. D(T5,O7,B25)
26. I(T6,O6,A26)
27. U(T5,O9,B27,A27)

- 14 UNDO={T2,T3,T4 } REDO={}
15 UNDO={T2,T3,T4,T5 } REDO={}
19 UNDO={T3,T4,T5 } REDO={T2}
21 UNDO={ T4,T5 } REDO={ T2,T3}
22 UNDO={ T5 } REDO={ T2,T3,T4}
23 UNDO={ T5,T6 } REDO={T2,T3,T4}

UNDO

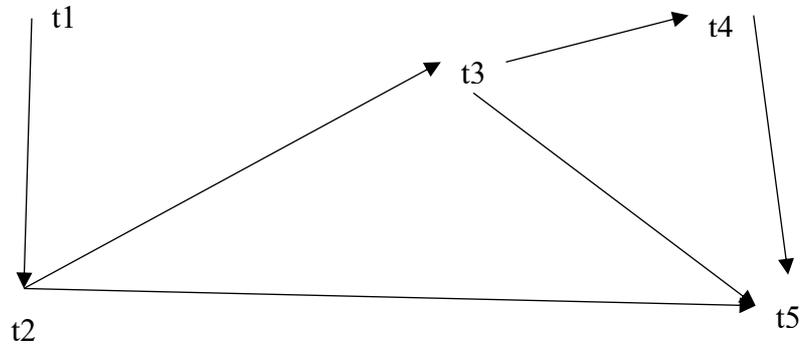
- 27 O9=B27
26 D(O6)
25 I(O7,B25)
24 I(O10,B24)
17 O7=B17
16 D(O6)

REDO

- 4 O2=A4
7 D(O2)
8 O3=A8
9 I(O4,A9)
10 O4=A10
12 I(O5,A12)
13 D(O5)
18 D(O8)
20 O7=A20

Esercizio 2

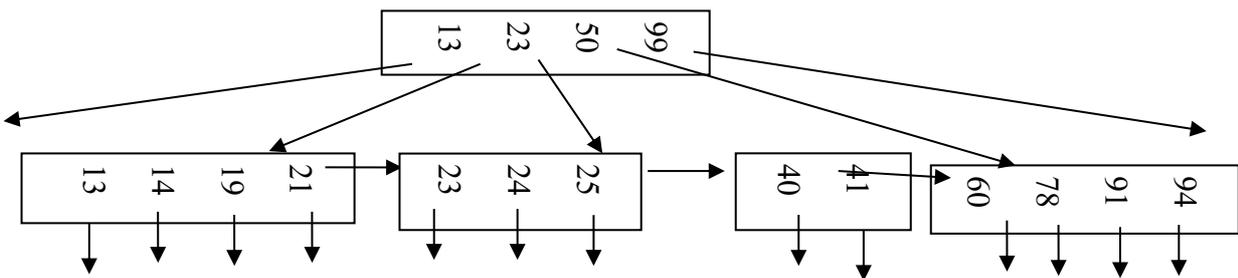
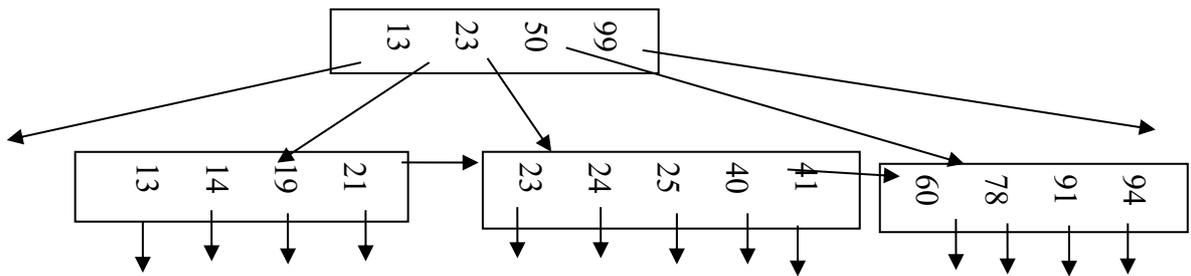
w1(x) r2(y) r2(x) w3(y) r4(y) w2(x) w5(y) w3(z) r3(z) r5(z)

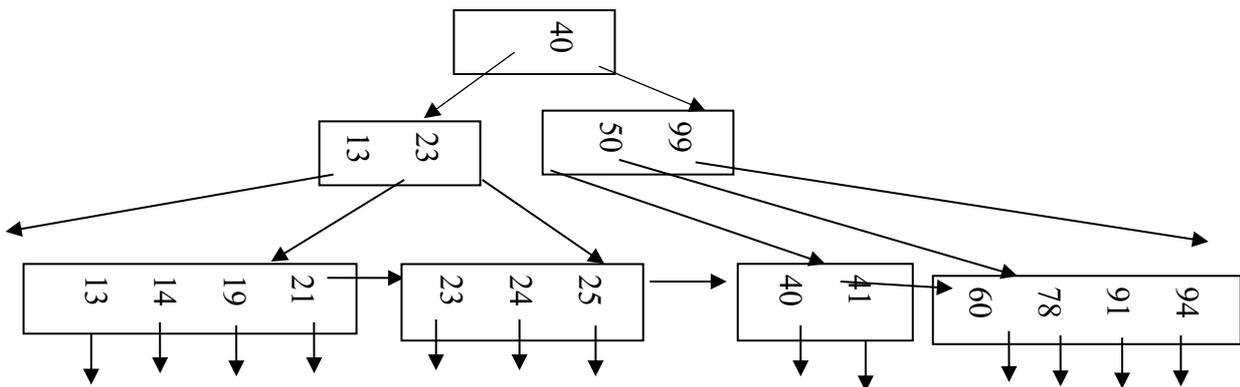
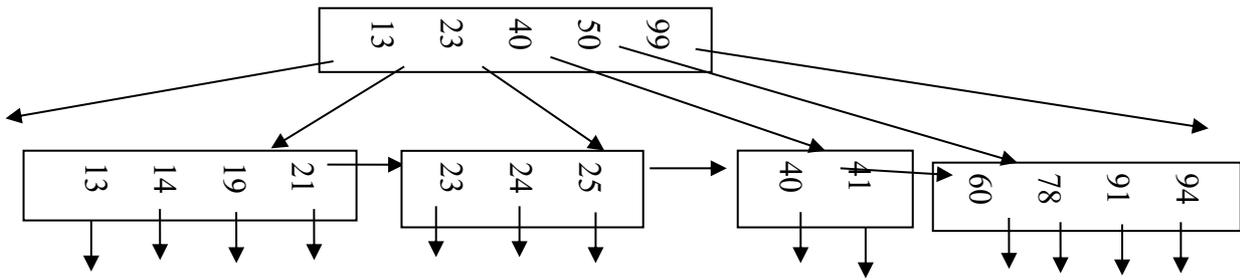


Lo schedule è conflict serializzabile.

Esercizio 3

Inserimento della chiave 24





Esercizio 4

buffer di memoria centrale disponibili: $M=500.000$

dimensione del buffer: $B=20.000$ bytes

numero di tuple: $T(\text{Studente})= 5.000.000$, $T(\text{Professore})= 2.000.000$,

dimensione delle tuple: $S(\text{Studente})=10.000$, $S(\text{Professore})=10.000$, $S(\text{Università})=50$

numero di valori $V(\text{Studente}, \text{Università})=500$, $V(\text{Professore}, \text{Università})=400$

$\text{Min}(\text{Professore}, \text{OreLezione})=0$ $\text{Max}(\text{Professore}, \text{OreLezione})=150$

Calcoliamo innanzitutto il numero di blocchi occupati da ciascuna relazione:

$$B(\text{Studente})=T(\text{Studente}) \cdot S(\text{Studente})/B = 5000000 \cdot 10000/20000=2500000$$

$$B(\text{Professore})=T(\text{Professore}) \cdot S(\text{Professore})/B = 2000000 \cdot 10000/20000=1000000$$

$(\sigma_{\text{OreLezione}>60}) \text{ Professore} \triangleright \triangleleft \text{ Studente}$

$\sigma_{\text{OreLezione}>60} (\text{Studente} \triangleright \triangleleft \text{Professore})$

Prima sequenza: $(\sigma_{\text{OreLezione}>60} \text{ Studente}) \triangleright \triangleleft \text{Professore}$

Sia $X = (\sigma_{\text{OreLezione}>60} \text{ Professore})$

C'è un indice secondario su $\text{Professore.OreLezione}$

$$f=(150-60)/(150-0+1)= 0.596$$

$$\text{CostoIndice}(X) = 3+ f \cdot T(\text{Professore}) = 3+0.596 \cdot 2000000=1192003$$

$$\text{CostoScansione}(X) = B(\text{Professore}) = 1000000$$

$$B(X) = 0.596 \cdot 1000000=596000$$

$$T(X) = 0.596 * 2000000 = 1192000$$

Né X né Studente 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 usando l'indice primario su Studente.Università.

$$\text{CostoJNL}(X \triangleright \triangleleft \text{Studente}) = B(X) + B(X) * B(\text{Studente}) / M \\ = 596000 + 596000 * 2500000 / 500000 = 3576000$$

$$\text{CostoHJI}(X \triangleright \triangleleft \text{Studente}) = (3 - 2M/B(X))(B(X) + B(\text{Studente})) = (3 - 2 * 500000 / 596000) * (596000 + 2500000) = 4093369$$

$$\text{CostoIJ}(X \triangleright \triangleleft \text{Studente}) = B(X) + T(X) \lceil B(\text{Studente}) / V(\text{Studente}, \text{Università}) \rceil \\ = 596000 + 1192000 * \lceil 2500000 / 500 \rceil = 596000 + 1192000 * 5000 = 5960596000$$

$$\text{Costo totale} = 1000000 + 3576000 = 4576000$$

Seconda sequenza: $\sigma_{\text{OreLezione} > 60}(\text{Studente} \triangleright \triangleleft \text{Professore})$

Sia $Z = \text{Studente} \triangleright \triangleleft \text{Professore}$

Né Professore né Studente 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 usando l'indice primario su Studente.Università.

$$\text{CostoJNL}(\text{Professore} \triangleright \triangleleft \text{Studente}) = B(\text{Professore}) + B(\text{Professore}) * B(\text{Studente}) / M \\ = 1000000 + 1000000 * 2500000 / 500000 = 6000000$$

$$\text{CostoHJI}(\text{Professore} \triangleright \triangleleft \text{Studente}) = (3 - 2M/B(\text{Professore}))(B(\text{Professore}) + B(\text{Studente})) = (3 - 2 * 500000 / 1000000) * (1000000 + 2500000) = 7000000$$

$$\text{CostoIJ}(\text{Professore} \triangleright \triangleleft \text{Studente}) = B(\text{Professore}) + T(\text{Professore}) \lceil B(\text{Studente}) / V(\text{Studente}, \text{Università}) \rceil \\ = 1000000 + 2000000 * \lceil 2500000 / 500 \rceil = 1000000 + 2000000 * 500 = 1001000000$$

$$T(Z) = T(\text{Professore}) * T(\text{Studente}) / \max\{V(\text{Professore}, \text{Università}), V(\text{Studente}, \text{Università})\} = 2000000 * 5000000 / 500 = 20000000000$$

$$S(Z) = S(\text{Studente}) + S(\text{Professore}) - S(\text{Università}) = 10000 + 10000 - 50 = 19950$$

$$B(Z) = 20000000000 * 19950 / 20000 = 19950000000$$

Dato che non ci sono indici su Z:

$$\text{Costo } \sigma_{\text{OreLezione} > 60}(Z) = B(Z) = 19950000000$$

$$\text{Costo totale} = 6000000 + 19950000000 = 19.956.000.000$$

La sequenza di costo minore è la prima