

COMPITO DI SISTEMI INFORMATIVI
18 gennaio 2016 (Tot. 16) Tempo: 2h, Compito A

Esercizio 1 (punti 3)

Si consideri il seguente log:

- | | |
|----------------------|----------------------|
| 1. B(T1) | 15. B(T5) |
| 2. U(T1,O1,B2,A2) | 16. U(T5,O6,B16,A16) |
| 3. B(T2) | 17. I(T5,O7,A17) |
| 4. D(T2,O1,B4) | 18. C(T4) |
| 5. B(T3) | 19. D(T2,O8,B19) |
| 6. B(T4) | 20. U(T3,O7,B20,A20) |
| 7. I(T4,O2, A7) | 21. C(T3) |
| 8. U(T4,O3,B8,A8) | 22. C(T5) |
| 9. D(T4,O2,B9) | 23. B(T6) |
| 10. U(T3,O4,B10,A10) | 24. I(T6,O10,A24) |
| 11. C(T1) | 25. D(T6,O7,B25) |
| 12. I(T3,O5,A12) | 26. D(T6,O6,B26) |
| 13. CK(T2,T3,T4) | 27. U(T6,O9,B27,A27) |
| 14. D(T3,O4,B14) | |

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

Esercizio 2 (punti 4)

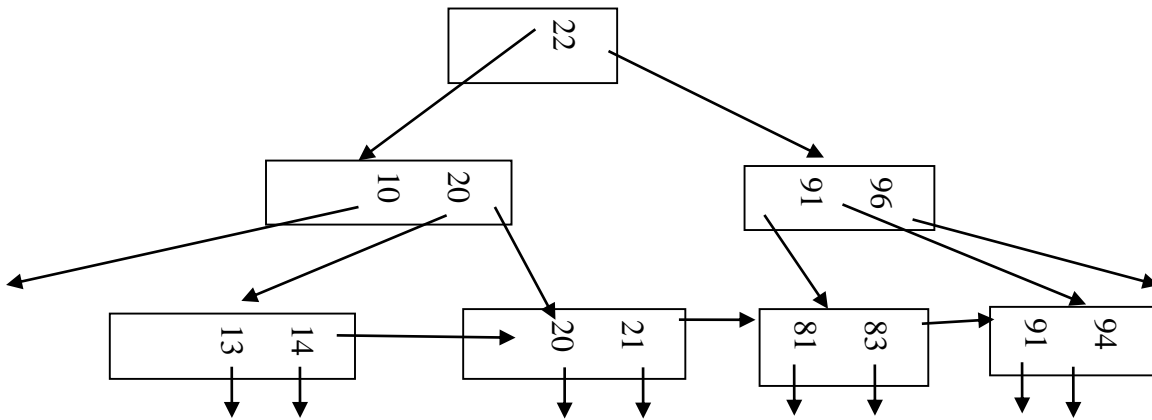
Dato il seguente schedule:

w1(x) r2(y) r2(x) w3(y) r4(y) w2(x) w5(y) w3(z) r3(y) r5(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 4.



Si mostri l'albero risultante dalla cancellazione della chiave 21

Esercizio 4 (punti 6)

Siano date le seguenti relazioni

Assicuratore(CodAss, Nome, Cognome, Via, Città, CAP, Filiale, Telefono)

Cliente(CodCli, Nome, Cognome, Via, Città, CAP, Filiale, Telefono, CodAss, Portafoglio)

la query

```
SELECT A.*, C.*
```

```
FROM Assicuratore AS A, Cliente AS C
```

```
WHERE A.Filiale = C.Filiale AND C.Portafoglio >40000
```

e i parametri:

buffer di memoria centrale disponibili: $M=500.000$

dimensione del buffer: $B=20.000$ bytes

numero di tuple: $T(\text{Cliente})=10.000.000$, $T(\text{Assicuratore})=200.000$,

dimensione delle tuple: $S(\text{Cliente})=10.000$, $S(\text{Assicuratore})=10.000$, $S(\text{Filiale})=50$

numero di valori $V(\text{Cliente}, \text{Filiale})=60.000$, $V(\text{Assicuratore}, \text{Filiale})=70.000$

$\text{Min}(\text{Cliente}, \text{Portafoglio})=1.000$ $\text{Max}(\text{Cliente}, \text{Portafoglio})=100.000$

Si calcoli il costo minimo di ciascuna delle seguenti sequenze:

$(\sigma_{\text{Portafoglio}>40000} \text{ Cliente}) \triangleright \triangleleft \text{Assicuratore}$

$\sigma_{\text{Portafoglio}>40000} (\text{Cliente} \triangleright \triangleleft \text{Assicuratore})$

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 secondario su Cliente.Filiale e uno primario su Cliente.Portafoglio.

SOLUZIONE

Esercizio 1

1. B(T1)
2. U(T1,O1,B2,A2)
3. B(T2)
4. D(T2,O1,B4)
5. B(T3)
6. B(T4)
7. I(T4,O2, A7)
8. U(T4,O3,B8,A8)
9. D(T4,O2,B9)
10. U(T3,O4,B10,A10)
11. C(T1)
12. I(T3,O5,A12)
13. CK(T2,T3,T4)
14. D(T3,O4,B14)
15. B(T5)
16. U(T5,O6,B16,A16)
17. I(T5,O7,A17)
18. C(T4)
19. D(T2,O8,B19)
20. U(T3,O7,B20,A20)
21. C(T3)
22. C(T5)
23. B(T6)
24. I(T6,O10,A24)
25. D(T6,O7,B25)
26. D(T6,O6,B26)
27. U(T6,O9,B27,A27)

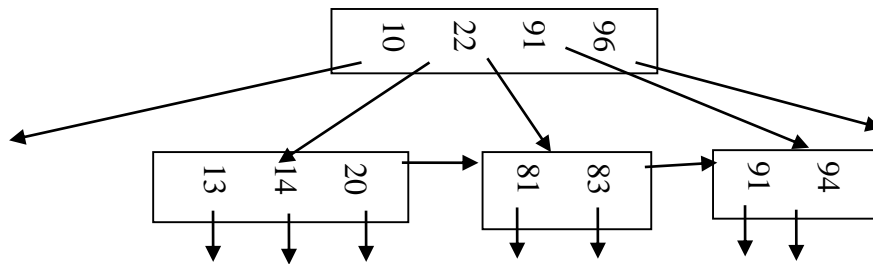
13 UNDO={ T2,T3,T4 } REDO={ }
15 UNDO={ T2,T3,T4,T5 } REDO={ }
18 UNDO={ T2,T3,T5 } REDO={ T4 }
21 UNDO={ T2,T5 } REDO={ T3,T4 }
22 UNDO={ T2 } REDO={ T3,T4,T5 }
23 UNDO={ T2,T6 } REDO={ T3,T4,T5 }

UNDO

27 O9=B27
26 I(O6,B26)
25 I(O7,B25)
24 D(O10)
19 I(O8,B19)
4 I(O1,B4)

REDO

7 I(O2,A7)
8 O3=A8
9 D(O2)
10 O4=A10
12 I(O5,A12)
14 D(O4)
16 O6=A16
17 I(O7,A17)
20 O7=A20



Esercizio 4

buffer di memoria centrale disponibili: $M=500000$

dimensione del buffer: $B=20000$ bytes

numero di tuple: $T(\text{Cliente})=10000000$, $T(\text{Assicuratore})=200000$,

dimensione delle tuple: $S(\text{Cliente})=10000$, $S(\text{Assicuratore})=10000$, $S(\text{Filiale})=50$

numero di valori $V(\text{Cliente}, \text{Filiale})=60000$, $V(\text{Assicuratore}, \text{Filiale})=70000$

$\text{Min}(\text{Cliente}, \text{Portafoglio})=1000$ $\text{Max}(\text{Cliente}, \text{Portafoglio})=100000$

Calcoliamo innanzitutto il numero di blocchi occupati da ciascuna relazione:

$B(\text{Cliente})=T(\text{Cliente}) \cdot S(\text{Cliente})/B = 10000000 \cdot 10000/20000 = 5000000$

$B(\text{Assicuratore})=T(\text{Assicuratore}) \cdot S(\text{Assicuratore})/B = 200000 \cdot 10000/20000 = 100000$

$(\sigma_{\text{Portafoglio} > 40000} \text{ Cliente}) \triangleright \triangleleft \text{Assicuratore}$

$\sigma_{\text{Portafoglio} > 40000} (\text{Cliente} \triangleright \triangleleft \text{Assicuratore})$

Prima sequenza: $(\sigma_{\text{Portafoglio} > 40000} \text{ Cliente}) \triangleright \triangleleft \text{Assicuratore}$

Sia $X = (\sigma_{\text{Portafoglio} > 40000} \text{ Cliente})$

C'è un indice primario su $\text{Cliente.Portafoglio}$

$f = (100000 - 40000) / (100000 - 1000 + 1) = 0.606$

$\text{Costo}(X) = 3 + \lceil f \cdot B(\text{Cliente}) \rceil = 3 + \lceil 0.606 \cdot 5000000 \rceil = 3030003$

$B(X) = 0.606 \cdot 5000000 = 3030000$

Assicuratore sta in memoria centrale quindi si usa il join a un passo.

$\text{CostoJ1P}(X \triangleright \triangleleft \text{Assicuratore}) = B(\text{Assicuratore}) + B(X) = 100000 + 3030000 = 3130000$

Costo totale = $3030003 + 3130000 = 6.160.003$

Seconda sequenza: $\sigma_{\text{Portafoglio} > 40000} (\text{Cliente} \triangleright \triangleleft \text{Assicuratore})$

Sia $Z = \text{Cliente} \triangleright \triangleleft \text{Assicuratore}$

Assicuratore sta in memoria centrale quindi si usa il join a un passo.

$\text{CostoJ1P}(Z) = B(\text{Assicuratore}) + B(\text{Cliente}) = 100000 + 5000000 = 5100000$

$$T(Z) = T(\text{Assicuratore}) * T(\text{Cliente}) / \max\{V(\text{Assicuratore, Filiale}), V(\text{Assicuratore, Filiale})\} = 200000 * 10000000 / 70000 = 28571429$$

$$S(Z) = S(\text{Cliente}) + S(\text{Assicuratore}) - S(\text{Filiale}) = 10000 + 10000 - 50 = 19950$$

$$B(Z) = 28571429 * 19950 / 20000 = 28500000$$

Dato che non ci sono indici su Z:

$$\text{Costo } \sigma_{\text{Portafoglio} > 40000} (Z) = B(Z) = 28500000$$

$$\text{Costo totale} = 5100000 + 28500000 = 33.600.000$$

La sequenza di costo minore è la prima