

COMPITO DI SISTEMI INFORMATIVI

14 settembre 2011 (Tot. 16) Tempo: 2h

Esercizio 1 (punti 3)

Si consideri il seguente log:

- | | |
|---------------------|----------------------|
| 1. U(T1,O1,B1,A1) | 13. B(T4) |
| 2. I(T1,O2,A2) | 14. U(T4,O2,B14,A14) |
| 3. B(T2) | 15. D(T3,O2,B15) |
| 4. U(T2,O2,B4,A4) | 16. B(T5) |
| 5. B(T3) | 17. U(T5,O3,B17,A17) |
| 6. I(T3,O3,A6) | 18. I(T4,O5,A18) |
| 7. C(T1) | 19. B(T6) |
| 8. U(T2,O3,B8,A8) | 20. C(T3) |
| 9. I(T3,O4,A9) | 21. C(T4) |
| 10. U(T2,O4,B10,A1) | 22. I(T5,O6,A23) |
| 11. D(T3,O1,B11) | 23. D(T6,O3,B24) |
| 12. CK(T2,T3) | |

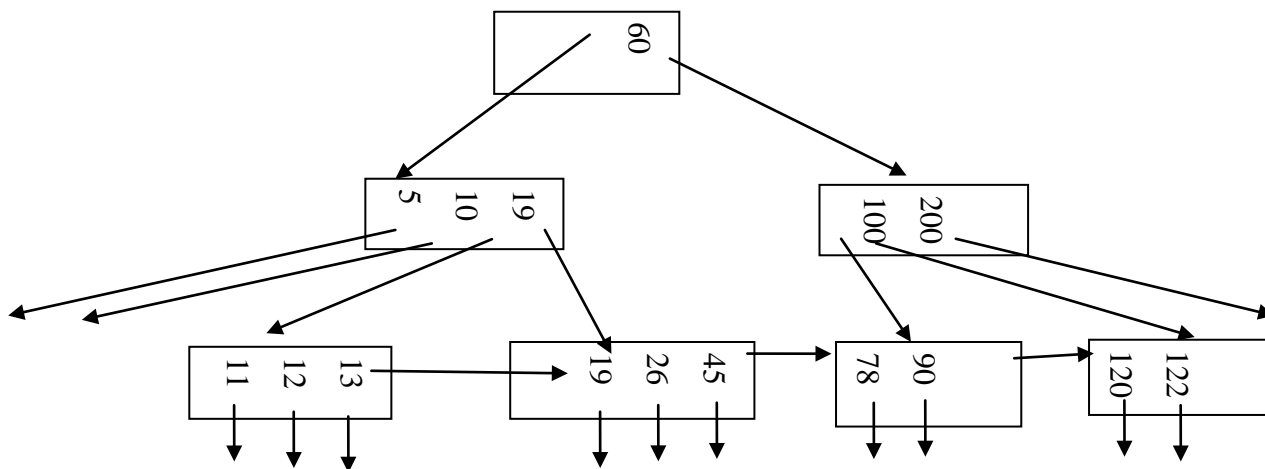
si mostrino le operazioni di recovery da effettuare supponendo che il guasto avvenga subito dopo l'ultimo record del log.

Esercizio 2 (punti 4)

$r1(x) w2(x) w3(y) r1(x) r2(z) r3(y) r4(y) r3(x) w2(z)$
si indichi se è view-equivalente allo schedule seriale
 $w2(x) r2(z) w2(z) r1(x) r1(x) w3(y) r3(Y) r3(x) r4(y)$

Esercizio 3 (punti 3)

Sia dato il seguente B+ tree di ordine 3.



Si mostri l'albero risultante dall'inserimento della chiave 23

Esercizio 4 (punti 6)

Siano date le relazioni

Studente(CodiceStu, Nome, Cognome, Indirizzo, AnnoDiNascita, CodiceLaurea)

Laurea(CodiceLaurea, NomeLaurea, Durata, Università)

la query

```
SELECT U.*, S.*
```

```
FROM Studente AS S, Laurea AS L
```

```
WHERE S.CodiceLaurea = L.CodiceLaurea AND S.AnnoDiNascita > 1978
```

e i parametri:

buffer di memoria centrale disponibili: $M=2000$

dimensione del buffer: $B=40000$ bytes

numero di tuple: $T(\text{Studente})=8000000$, $T(\text{Laurea})=20000$,

dimensione delle tuple: $S(\text{Studente})=4000$ bytes, $S(\text{Laurea})=5000$, $S(\text{CodiceLaurea})=50$

$V(\text{Studente}, \text{CodiceLaurea})=20000$

range di $\text{Studente.AnnoDiNascita}=[1970, 1990]$

Si calcoli il costo minimo di ciascuna delle seguenti sequenze:

$(\sigma_{\text{AnnoDiNascita} > 1978} \text{ Studente}) \triangleright \triangleleft \text{Laurea}$

$\sigma_{\text{AnnoDiNascita} > 1978} (\text{Studente} \triangleright \triangleleft \text{Laurea})$

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 $\text{Studente.AnnoDiNascita}$, uno primario su $\text{Laurea.CodiceLaurea}$ e uno secondario su $\text{Studente.CodiceLaurea}$

SOLUZIONE

Esercizio 1

1. U(T1,O1,B1,A1)
2. I(T1,O2,A2)
3. B(T2)
4. U(T2,O2,B4,A4)
5. B(T3)
6. I(T3,O3,A6)
7. C(T1)
8. U(T2,O3,B8,A8)
9. I(T3,O4,A9)
10. U(T2,O4,B10,A1)
11. D(T3,O1,B11)
12. CK(T2,T3)
13. B(T4)
14. U(T4,O2,B14,A14)
15. D(T3,O2,B15)
16. B(T5)
17. U(T5,O3,B17,A17)
18. I(T4,O5,A18)
19. B(T6)
20. C(T3)
21. C(T4)
22. I(T5,O6,A22)
23. D(T6,O3,B23)

- 12 UNDO={T2,T3} REDO={}
13 UNDO={T2,T3,T4} REDO={}
16 UNDO={T2,T3,T4,T5} REDO={}
19 UNDO={T2,T3,T4,T5,T6} REDO={}
20 UNDO={T2, T4,T5,T6} REDO={ T3}
21 UNDO={T2,T5,T6} REDO={ T3, T4}

UNDO

- 23 I(O3,B23)
22 D(O6)
17 O3=B17
10 O4=B10
8 O3=B8
4 O2=B4

REDO

- 6 I(O3,A6)
9 I(O4,A9)
11 D(O1)
14 O2=A14
15 D(O2)
18 I(O5,A18)

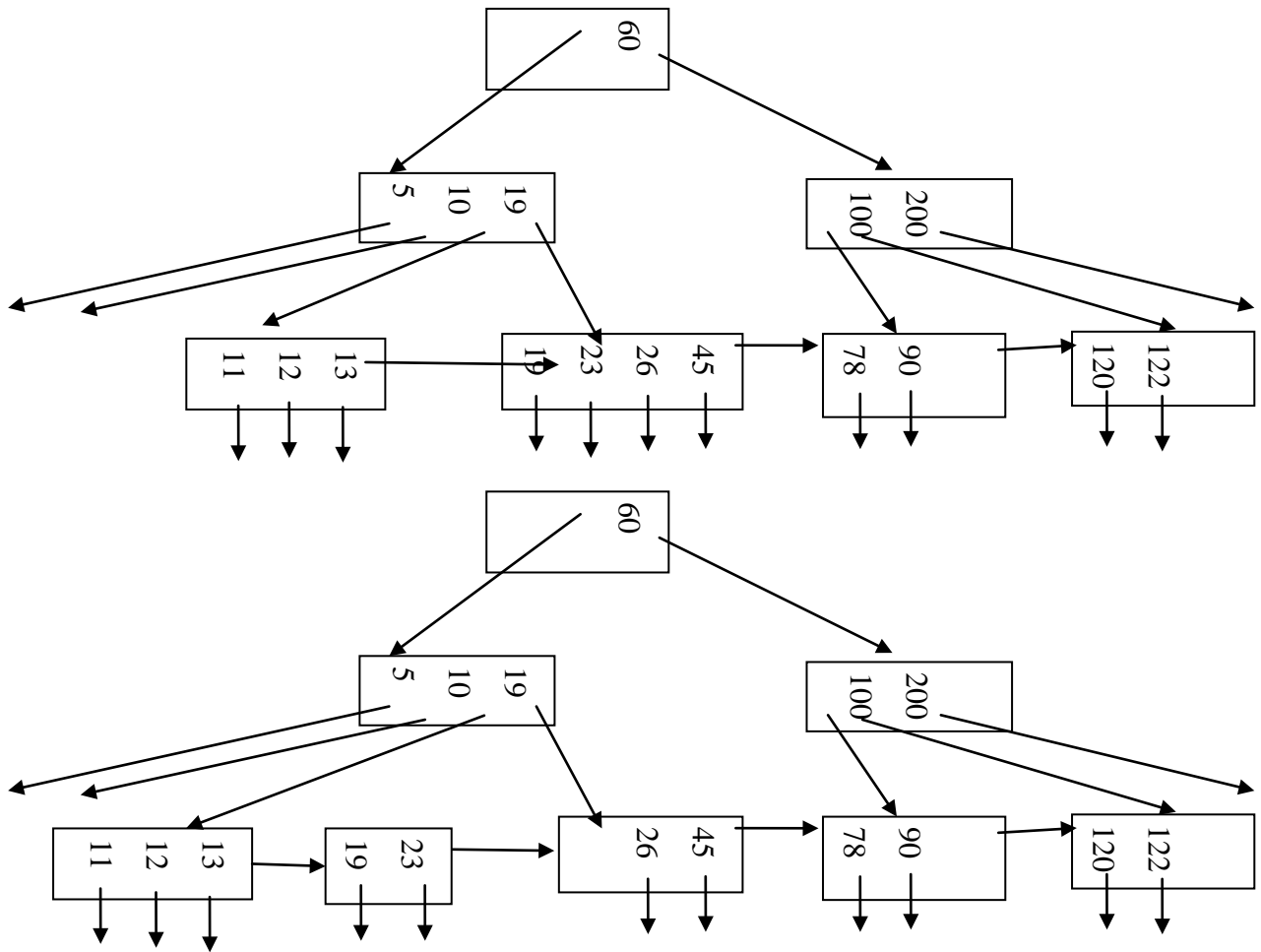
Esercizio 2

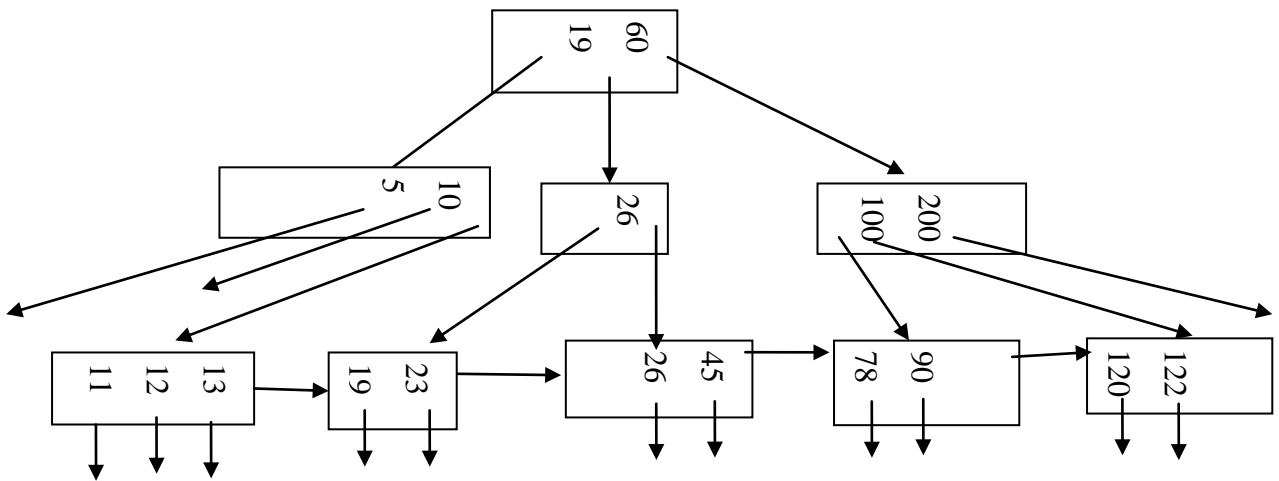
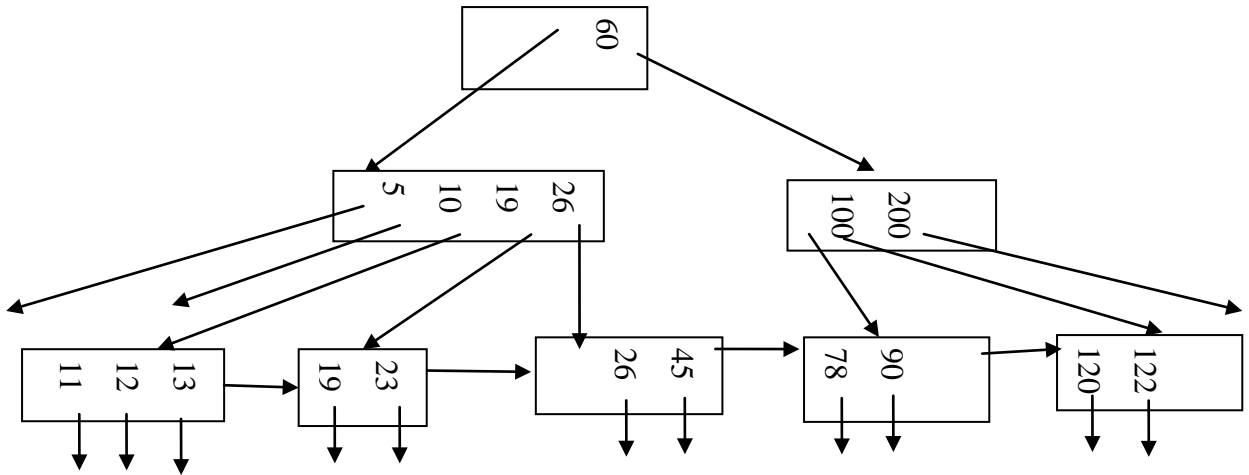
r1(x) w2(x) w3(y) r1(x) r2(z) r3(y) r4(y) r3(x) w2(z)
 si indichi se è view-equivalente allo schedule seriale
 w2(x) r2(z) w2(z) r1(x) r1(x) w3(y) r3(y) r3(x) r4(y)

"legge-da"={ (r1(x),w2(x)),(r3(y),w3(y)),(r4(y),w3(y)),(r3(x),w2(x)) }
 scritture finali={ (t2,z),(t3,y),(t2,x) }
 per entrambi gli schedule

Esercizio 3

Inserimento della chiave 23





Esercizio 4

buffer di memoria centrale disponibili: $M=2000$

dimensione del buffer: $B=40000$ bytes

numero di tuple: $T(\text{Studente})= 8000000$, $T(\text{Laurea})= 20000$,

dimensione delle tuple: $S(\text{Studente})=4000$ bytes, $S(\text{Laurea})=5000$, $S(\text{CodiceLaurea})=50$

$V(\text{Studente}, \text{CodiceLaurea})= 20000$

range di $\text{Studente.AnnoDiNascita}=[1970,1990]$

Calcoliamo innanzitutto il numero di blocchi occupati da ciascuna relazione:

$$B(\text{Studente})=T(\text{Studente}) \cdot S(\text{Studente})/B = 8000000 \cdot 4000 / 40000 = 800000$$

$$B(\text{Laurea}) = 20000 \cdot 5000 / 40000 = 2500$$

Prima sequenza: $(\sigma_{\text{AnnoDiNascita} > 1978} \text{ Studente}) \triangleright \triangleleft \text{Laurea}$

Sia $X = (\sigma_{\text{AnnoDiNascita} > 1978} \text{ Studente})$

C'è un indice primario su $\text{Studente.AnnoDiNascita}$

$$f = (1990 - 1978) / (1990 - 1970 + 1) = 0.57143$$

$$\text{Costo}(X) = 3 + \lceil f \cdot B(\text{Studente}) \rceil = 3 + \lceil 0.57143 \cdot 800000 \rceil = 3 + 457144 = 457147$$

$$B(X) = 0.57143 * 800000 = 457144$$

$$T(X) = 0.57143 * 800000 = 457144$$

Né Laurea né X 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 Laurea.CodiceLaurea.

$$\text{CostoJNL}(Z) = B(\text{Laurea}) + B(X) * B(\text{Laurea}) / M = 2500 + 457144 * 2500 / 2000 = 573930$$

$$\text{CostoHJI}(Z) = (3 - 2M/B(\text{Laurea})) * (B(X) + B(\text{Laurea})) = (3 - 2 * 2000 / 2500) * (457144 + 2500) = 643502$$

$$\text{CostoIJ}(Z) = B(X) + T(X) * 1 = 457144 + 457144 * 1 = 5028584$$

$$\text{Costo totale} = 457147 + 573930 = 1031077$$

Seconda sequenza: $\sigma_{\text{AnnoDiNascita} > 1978}$ (Studente $\triangleright \triangleleft$ Laurea)

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

Né Laurea 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 Laurea.CodiceLaurea e secondario su Studente.CodiceLaurea

$$\text{CostoJNL}(Z) = B(\text{Laurea}) + B(\text{Studente}) * B(\text{Laurea}) / M = 2500 + 800000 * 2500 / 2000 = 1002500$$

$$\text{CostoHJI}(Z) = (3 - 2M/B(\text{Laurea})) * (B(\text{Studente}) + B(\text{Laurea})) = (3 - 2 * 2000 / 2500) * (2500 + 800000) = 1123500$$

$$\text{CostoIJ1}(Z) = B(\text{Studente}) + T(\text{Studente}) * 1 = 800000 + 800000 * 1 = 8800000$$

$$\text{CostoIJ2}(Z) = B(\text{Laurea}) + T(\text{Laurea}) * T(\text{Studente}) / V(\text{Studente}, \text{CodiceLaurea}) = 2500 + 20000 * 800000 / 20000 = 8002500$$

$$T(Z) = T(\text{Laurea}) * T(\text{Studente}) / \max\{V(\text{Laurea}, \text{CodiceLaurea}), V(\text{Studente}, \text{CodiceLaurea})\} = 20000 * 800000 / 20000 = 8000000$$

$$S(Z) = S(\text{Laurea}) + S(\text{Studente}) - S(\text{CodiceLaurea}) = 5000 + 4000 - 50 = 8950$$

$$B(Z) = 8000000 * 8950 / 40000 = 1790000$$

Dato che non ci sono indici su Z:

$$\text{Costo } \sigma_{\text{DataDiNascita} > 1978}(Z) = B(Z) = 1790000$$

$$\text{Costo totale} = 1002500 + 1790000 = 2792500$$

La sequenza di costo minore è la prima