

## COMPITO DI SISTEMI INFORMATIVI

13 giugno 2014 (Tot. 16) Tempo: 2h

### Esercizio 1 (punti 3)

Si consideri il seguente log:

- |                      |                      |
|----------------------|----------------------|
| 1. B(T1)             | 15. U(T5,O6,B15,A15) |
| 2. I(T1,O1,A2)       | 16. U(T3,O6,B16,A16) |
| 3. B(T2)             | 17. I(T4,O7,A17)     |
| 4. U(T2,O1,B4,A4)    | 18. C(T4)            |
| 5. B(T3)             | 19. D(T1,O2,B19)     |
| 6. D(T3,O1,B6)       | 20. U(T5,O7,B20,A20) |
| 7. U(T1,O2,B7,A7)    | 21. C(T1)            |
| 8. I(T2,O3,A8)       | 22. B(T6)            |
| 9. B(T4)             | 23. U(T6,O7,B23,A23) |
| 10. U(T4,O4,B10,A10) | 24. I(T5,O8,A24)     |
| 11. C(T2)            | 25. D(T6,O7,B25)     |
| 12. I(T3,O5,A12)     | 26. D(T5,O6,B26)     |
| 13. CK(T1,T3,T4)     | 27. U(T3,O9,B27,A27) |
| 14. B(T5)            |                      |

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:

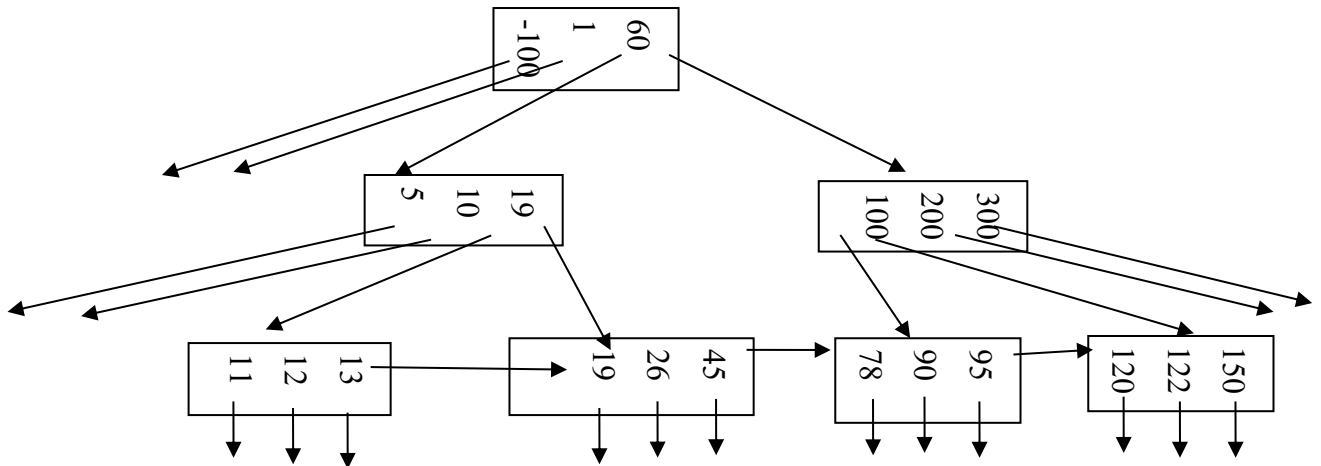
$r1(x), r2(y), r1(y), r3(x), w3(x), w1(z), w3(y), w2(y), r2(u)$

si indichi se è view-equivalente allo schedule seriale

$r2(y), w2(y), r2(u), r1(x), r1(y), w1(z), r3(x), w3(x), w3(y)$

### Esercizio 3 (punti 3)

Sia dato il seguente B+ tree di ordine 3.



Si mostri l'albero risultante dall'inserimento della chiave 91

### 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 < 1985
```

e i parametri:

buffer di memoria centrale disponibili:  $M=40000$

dimensione del buffer:  $B=20000$  bytes

numero di tuple:  $T(\text{Studente})=1000000$ ,  $T(\text{Laurea})=50000$ ,

dimensione delle tuple:  $S(\text{Studente})=10000$  bytes,  $S(\text{Laurea})=20000$ ,  $S(\text{CodiceLaurea})=100$

$V(\text{Studente}, \text{CodiceLaurea})=40000$

range di Studente.AnnoDiNascita=[1980,1989]

Si calcoli il costo minimo di ciascuna delle seguenti sequenze:

$(\sigma_{\text{AnnoDiNascita} < 1985} \text{ Studente}) \triangleright \triangleleft \text{Laurea}$

$\sigma_{\text{AnnoDiNascita} < 1985} (\text{Studente} \triangleright \triangleleft \text{Laurea})$

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 indici secondari su Studente.AnnoDiNascita e su Laurea.CodiceLaurea.

## SOLUZIONE

### Esercizio 1

1. B(T1)
2. I(T1,O1,A2)
3. B(T2)
4. U(T2,O1,B4,A4)
5. B(T3)
6. D(T3,O1,B6)
7. U(T1,O2,B7,A7)
8. I(T2,O3,A8)
9. B(T4)
10. U(T4,O4,B10,A10)
11. C(T2)
12. I(T3,O5,A12)
13. CK(T1,T3,T4)
14. B(T5)
15. U(T5,O6,B15,A15)
16. U(T3,O6,B16,A16)
17. I(T4,O7,A17)
18. C(T4)
19. D(T1,O2,B19)
20. U(T5,O7,B20,A20)
21. C(T1)
22. B(T6)
23. U(T6,O7,B23,A23)
24. I(T5,O8,A24)
25. D(T6,O7,B25)
26. D(T5,O6,B26)
27. U(T3,O9,B27,A27)

- 13 UNDO={T1,T3,T4 } REDO={}  
14 UNDO={T1,T3,T4,T5 } REDO={}  
18 UNDO={T1,T3,T5 } REDO={T4}  
21 UNDO={ T3,T5 } REDO={ T1,T4}  
22 UNDO={ T3,T5,T6 } REDO={ T1,T4}

#### UNDO

- 27 O9=B27  
26 I(O6,B26)  
25 I(O7,B25)  
24 D(O8)  
23 O7=B23  
20 O7=B20  
16 O6=B16  
15 O6=B15  
12 D(O5)  
6 I(O1,B6)

#### REDO

- 2 I(O1,A2)  
7 O2=A7  
10 O4=A10  
17 I(O7,A17)  
19 D(O2)

### Esercizio 2

S1=r1(x),r2(y),r1(y),r3(x),w3(x),w1(z),w3(y),w2(y),r2(u)  
 S2=r2(y),w2(y),r2(u),r1(x),r1(y),w1(z),r3(x),w3(x), w3(y)

S1:

"legge-da"={ }

scritture finali={(t2,y),(t1,z),(t3,x) }

S1:

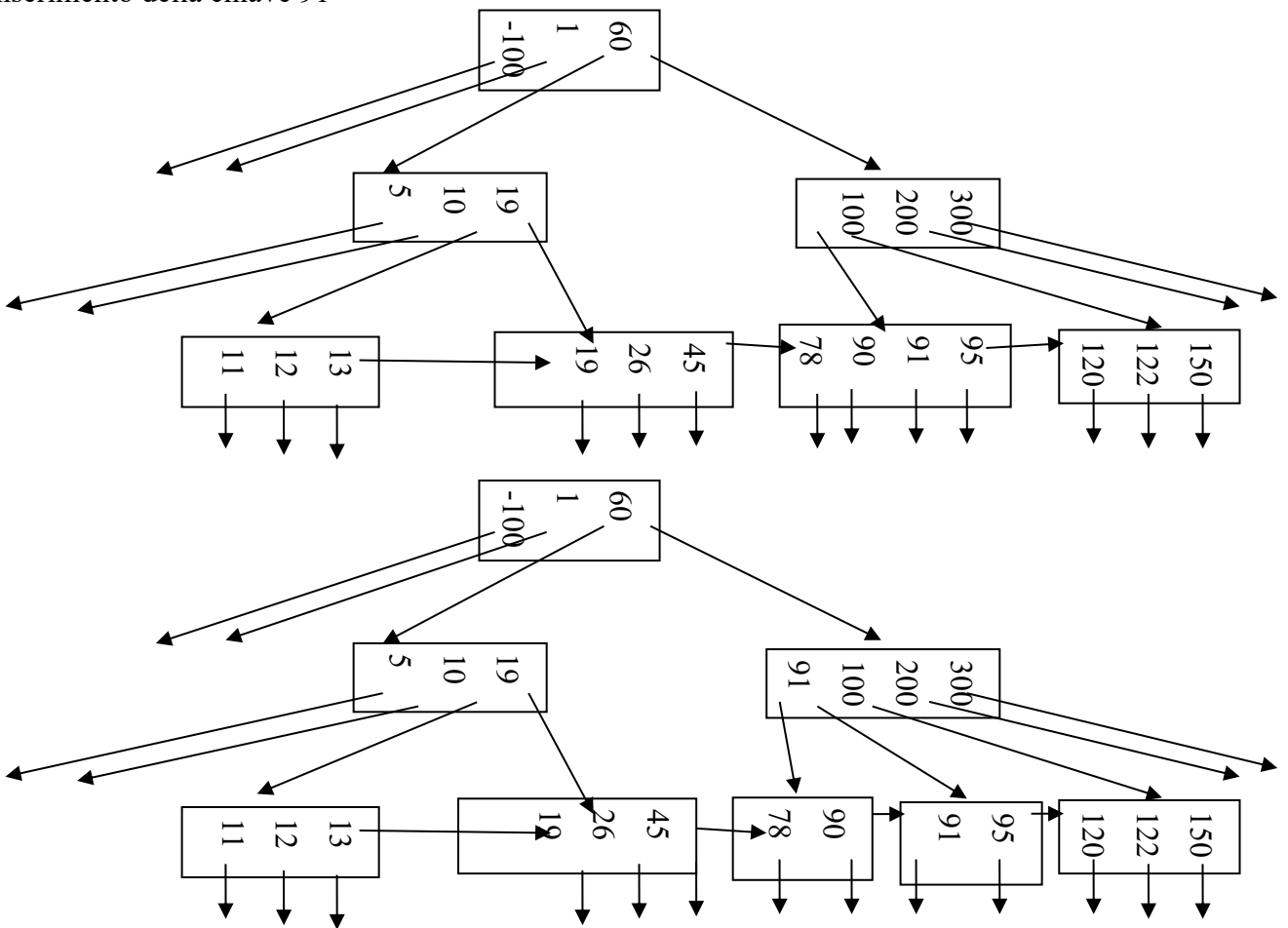
"legge-da"={ (r1(y),w2(y)) }

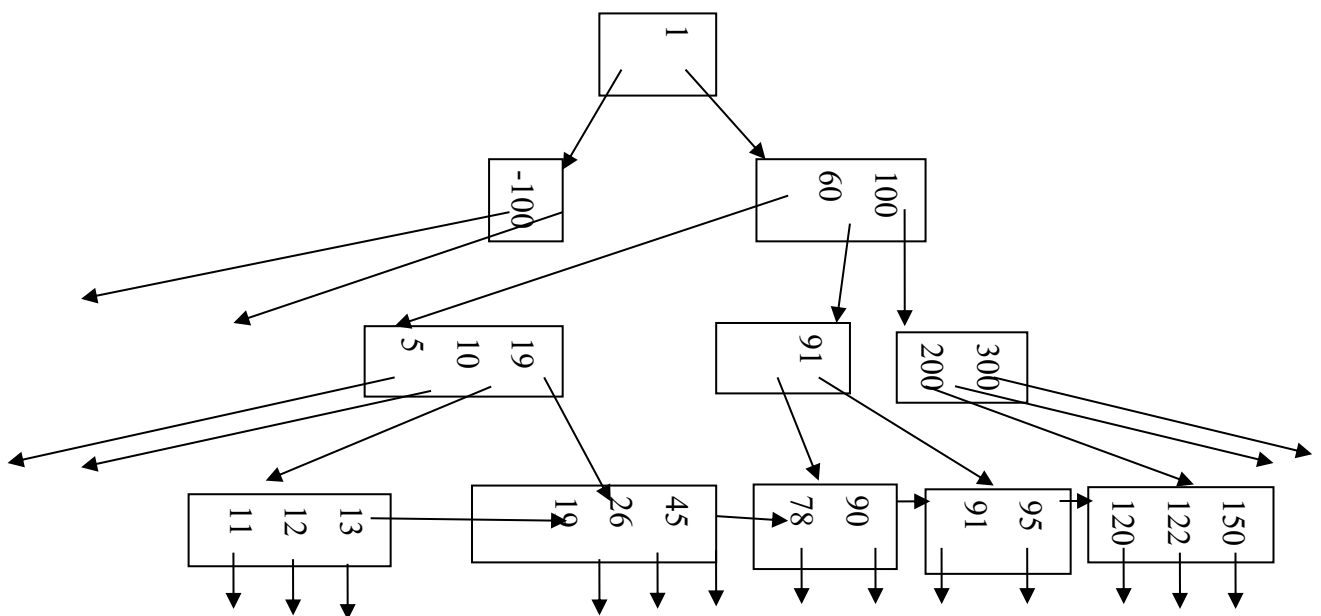
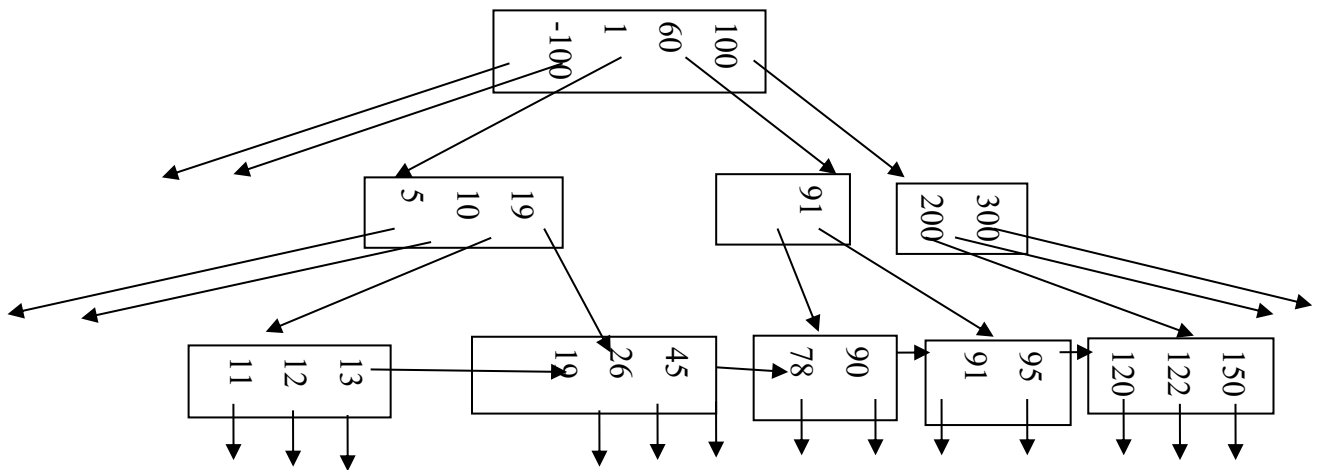
scritture finali={(t3,y),(t1,z),(t3,x) }

I due schedule non sono view equivalenti

### Esercizio 3

Inserimento della chiave 91





#### Esercizio 4

buffer di memoria centrale disponibili:  $M=40000$

dimensione del buffer:  $B=20000$  bytes

numero di tuple:  $T(\text{Studente})= 1000000$ ,  $T(\text{Laurea})= 50000$ ,

dimensione delle tuple:  $S(\text{Studente})=10000$  bytes,  $S(\text{Laurea})=20000$ ,  $S(\text{CodiceLaurea})=100$

$V(\text{Studente}, \text{CodiceLaurea})= 40000$

range di  $\text{Studente.AnnoDiNascita}=[1980,1989]$

Calcoliamo innanzitutto il numero di blocchi occupati da ciascuna relazione:

$$B(\text{Studente})=T(\text{Studente}) \cdot S(\text{Studente}) / B = 1000000 \cdot 10000 / 20000 = 500000$$

$$B(\text{Laurea}) = 50000 \cdot 20000 / 20000 = 50000$$

Prima sequenza:  $(\sigma_{\text{AnnoDiNascita} < 1985} \text{ Studente}) \triangleright \triangleleft \text{Laurea}$

Sia  $X = (\sigma_{\text{AnnoDiNascita} < 1985} \text{ Studente})$

C'è un indice secondario su Studente. AnnoDiNascita quindi si prova usando l'indice e con la scansione

$$f = (1985 - 1980) / (1989 - 1980 + 1) = 0.5$$

$$\text{CostoIndice}(X) = 3 + f * T(\text{Studente}) = 3 + 0.5 * 1000000 = 3 + 500000 = 500003$$

$$\text{CostoScansione}(X) = B(\text{Studente}) = 500000$$

$$B(X) = 0.5 * 500000 = 250000$$

$$T(X) = 0.5 * 1000000 = 500000$$

Costa meno la scansione, si usa quella.

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 secondario su Laurea.CodiceLaurea.

$$\text{CostoJNL}(Z) = B(\text{Laurea}) + B(X) * B(\text{Laurea}) / M = 50000 + 250000 * 50000 / 40000 = 362500$$

$$\text{CostoHJI}(Z) = (3 - 2M/B(\text{Laurea})) * (B(X) + B(\text{Laurea})) = (3 - 2 * 40000 / 50000) * (250000 + 50000) = 420000$$

$$\text{CostoIJ}(Z) = B(X) + T(X) * 1 = 250000 + 500000 * 1 = 750000$$

$$\text{Costo totale} = 500000 + 362500 = 862500$$

Seconda sequenza:  $\sigma_{\text{AnnoDiNascita} < 1985}(\text{Studente} \bowtie \triangleleft \text{Laurea})$

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

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 secondario su Laurea.CodiceLaurea

$$\text{CostoJNL}(Z) = B(\text{Laurea}) + B(\text{Studente}) * B(\text{Laurea}) / M = 50000 + 50000 * 500000 / 40000 = 675000$$

$$\text{CostoHJI}(Z) = (3 - 2M/B(\text{Laurea})) * (B(\text{Studente}) + B(\text{Laurea})) = (3 - 2 * 40000 / 50000) * (500000 + 50000) = 770000$$

$$\text{CostoIJ}(Z) = B(\text{Studente}) + T(\text{Studente}) * 1 = 500000 + 1000000 * 1 = 1500000$$

$$T(Z) = T(\text{Laurea}) * T(\text{Studente}) / \max\{V(\text{Laurea}, \text{CodiceLaurea}), V(\text{Studente}, \text{CodiceLaurea})\} = 50000 * 1000000 / 50000 = 1000000$$

$$S(Z) = S(\text{Laurea}) + S(\text{Studente}) - S(\text{CodiceLaurea}) = 20000 + 10000 - 100 = 29900$$

$$B(Z) = 1000000 * 29900 / 20000 = 1495000$$

Dato che non ci sono indici su Z:

$$\text{Costo } \sigma_{\text{AnnoDiNascita} < 1985}(Z) = B(Z) = 1495000$$

Costo totale =  $675000 + 1495000 = 2170000$

La sequenza di costo minore è la prima