

# Basi di dati, primo modulo

## 15 giugno 2001

Tempo a disposizione: un'ora e trenta minuti. Libri chiusi.

### Domanda 1 (40%)

Si consideri una relazione IMPIEGATO(Matricola,Cognome,Nome,DataNascita) con un numero di ennuple pari a  $N$  molto variabile nel tempo e una dimensione di ciascuna ennupla (a lunghezza fissa) pari a  $L$  byte, di cui  $K$  per la chiave.

Supporre di avere a disposizione un DBMS che permetta strutture fisiche disordinate (heap), ordinate (con indice primario sparso) e hash e che preveda la possibilità di definire indici secondari e operi su un sistema operativo che utilizza blocchi di dimensione  $B$  e con puntatori ai blocchi di  $P$  caratteri.

Siano le seguenti le operazioni principali:

1. ricerca sul numero di matricola con frequenza  $f_1$
2. ricerca sul cognome (o sua sottostringa iniziale) con frequenza  $f_2$

Individuare una possibile organizzazione fisica per tale relazione e calcolare (approssimativamente) il numero di accessi a memoria secondaria (nell'unità di tempo) supponendo  $N = 10.000.000$ ,  $L = 100$ ,  $K = 5$ ,  $B = 500$ ,  $P = 4$ ,  $f_1 = 100$ ,  $f_2 = 1.000$ .

### Domanda 2 (15%)

Spiegare perché si utilizza il metodo di Lamport nell'ambito del controllo di concorrenza basato su timestamp in contesti distribuiti.

### Domanda 3 (30%)

Descrivere i principali algoritmi per l'esecuzione dell'operatore di join, discutendone anche la complessità (distinguendo casi diversi a seconda delle strutture fisiche definite sulle relazioni).

### Domanda 4 (15%)

Indicare quali fra le seguenti affermazioni sono vere per i data warehouse:

1. Le operazioni sono complesse e non predefinite
2. Le proprietà "acide" non sono rilevanti
3. I dati sono, istante per istante, una sintesi dei dati operativi
4. Le strutture fisiche cercano di conciliare le esigenze degli aggiornamenti e quelle delle interrogazioni
5. Le attività supportate sono quelle quotidiane (ad esempio la vendita e le attività di sportello)
6. Gli utenti sono prevalentemente di livello più alto nella gerarchia aziendale rispetto a quelli che utilizzano i sistemi OLTP