

Basi di dati II, primo modulo

Esercizi di autovalutazione — 4 marzo 2011

Note

- Gli esercizi sono facoltativi ma è estremamente importante svolgerli (e anche discuterli ad esempio sul forum), perché le domande proposte nei compiti di esame possono essere molto simili.

Domanda 1 Si consideri un disco con una singola faccia che contenga $T=20.000$ tracce e con una velocità di rotazione di $v=10.000$ giri al minuto. Ogni traccia contiene $N=240$ blocchi e ogni blocco contiene $L=2KB$. Rispondere alle seguenti domande mostrando formula e valore numerico numero (N.B. non servono calcolatrici, i risultati sono semplici, approssimare $1GB=1000MB=1.000.000KB$).

1. Qual è la capacità del disco (in GB)?
2. Qual è la massima velocità di trasferimento (in MB al secondo)?

Domanda 2 Considerare un disco con una velocità di rotazione di $v = 10.000$ giri al minuto e un tempo medio di posizionamento della testina (tempo di seek) $t_S = 5$ msec. Ogni traccia contiene $N = 240$ blocchi e ogni blocco contiene $B = 4$ KB. Rispondere alle seguenti domande mostrando formula e valore numerico numero (N.B. non servono calcolatrici, i risultati sono semplici, approssimare $1 MB = 1000 KB$).

1. Qual è tempo medio di latenza (attesa dovuta alla rotazione) t_L ?
2. Qual è la massima velocità di trasferimento (in MB al secondo)?
3. Qual è il tempo minimo t_B necessario per leggere un blocco?
4. Qual è il tempo medio necessario per leggere un blocco?
5. Qual è il tempo medio necessario per leggere $n = 40$ blocchi contigui?

Domanda 3 Installare un DBMS commerciale o di pubblico dominio, ad esempio uno fra DB2, Oracle, SQLServer, Postgres, Derby, MySQL (sono più o meno tutti scaricabili in rete).

Domanda 4 Installare il DBMS didattico SimpleDB (<http://www.cs.bc.edu/sciore/simpledb/>).

Domanda 5 Si supponga di avere un file con numero di record $N=1.000.000$ di $L=100$ byte ciascuno con un sistema che utilizzi blocchi di dimensione $D=4Kbyte$. Quanti accessi a memoria secondaria sono necessari per trovare un record con un certo valore su un campo?

Domanda 6 Si supponga di avere un file definito su vari campi fra cui un identificatore che viene costruito concatenando un numero d'ordine relativo alla giornata con la data corrente (ad esempio, 1223022011 è il dodicesimo identificatore generato il 23 febbraio 2011. Spiegare perché, su questo campo, una funzione hash basata sul resto della divisione sarebbe estremamente inefficace).