

Esercizi sulla organizzazione fisica dei dati

1.

Supponiamo di avere un file di 250.000 record. Ogni record occupa 300 byte, di cui 75 per il campo chiave. Ogni blocco contiene 1024 byte. Un puntatore a blocco occupa 4 byte.

- a) Se usiamo una organizzazione hash con 1200 bucket, quanti blocchi occorrono per la bucket directory ?
- b) Quanti blocchi occorrono per i bucket, assumendo una distribuzione uniforme dei record nei bucket?
- c) Assumendo ancora che tutti i bucket contengano il numero medio di record, qual è il numero medio di accessi a blocco per ricercare un record che sia presente nel file?
- d) Quanti bucket dovremmo creare per avere invece un numero medio di accessi a blocco inferiore o al massimo uguale a 10, assumendo comunque una distribuzione uniforme dei record nei bucket?

2.

Supponiamo di avere un file di 150.000 records. Ogni record occupa 250 byte, di cui 50 per il campo chiave. Ogni blocco contiene 1024 byte. Un puntatore a blocco occupa 4 byte.

- a) Se usiamo un indice ISAM sparso, e assumiamo che i record non siano puntati e che il fattore di utilizzo dei blocchi del file sia 0,7 (cioè i blocchi non sono completamente pieni, ma pieni al 70%), quanti blocchi dobbiamo utilizzare per l'indice ?
- b) Se usiamo un indice ISAM sparso, e assumiamo che i record siano puntati e che i blocchi del file siano pieni, quanti blocchi dobbiamo utilizzare per l'indice ?
- c) Se utilizziamo la ricerca binaria, quale è il numero massimo di accessi a blocco per ricercare un record presente nel file nei casi a) e b), supponendo nel caso b) di non avere liste di overflow ?

3.

Supponiamo di avere un file di 170.000 record. Ogni record occupa 200 byte, di cui 20 per il campo chiave. Ogni blocco contiene 1024 byte. Un puntatore a blocco occupa 4 byte.

- a) Se usiamo un B-tree e assumiamo che sia i blocchi indice che i blocchi del file sono pieni al minimo, quanti blocchi vengono usati per il livello foglia (main file) e quanti per l'indice, considerando tutti i livelli non foglia ?
- b) Se il B-tree è usato invece come indice secondario, con i blocchi indice dei livelli intermedi pieni al massimo e i blocchi del livello foglia (indice denso) pieni al 60%, quanti blocchi vengono usati per i blocchi indice, compreso il livello foglia (indice denso), sapendo che per l'indice secondario il numero di chiavi NK è 2500, che occupano ognuna LK = 10 byte e che un puntatore a record PR occupa 3 byte?