

Supponiamo di avere un file di 3.000.000 record. I record sono a lunghezza fissa e ogni record occupa 250 byte. Ogni blocco contiene 2048 byte. Un puntatore a blocco occupa 5 byte.
a) Usando una organizzazione hash, con una funzione che distribuisce uniformemente 1/3 dei record in 250 bucket e i rimanenti 2/3 in altri 250 bucket, quanti blocchi dobbiamo utilizzare complessivamente per la bucket directory e per i bucket?
b) Calcolare il numero medio di accessi a blocco per la ricerca in ognuna delle due parti dell'organizzazione hash.

ho complessivamente 500 bucket, che occupano $500 \cdot 5 = 2500$ byte, necessitiamo quindi di **2 blocchi**.

Un blocco può contenere $\lfloor \frac{2048}{250} \rfloor = 8$ record. Abbiamo $3000000 \cdot \frac{1}{3} = 1000000$ di record in 250 bucket, in ogni

bucket ci sono $\frac{1000000}{250} = 4000$ record, e si necessita di $\frac{10000}{8} = 500$ blocchi. Abbiamo poi $3000000 \cdot \frac{2}{3} = 2000000$

di record in 250 bucket, in ogni bucket ci sono $\frac{2000000}{250} = 8000$ record, e per contenerli servono **1000 blocchi**.

Quando si fa una ricerca, gli accessi medi sono $\frac{1500}{2} = 250$ per una parte, e $\frac{1000}{2}$ per l'altra.

2)
È dato un file di 1.750.000 record. Ogni record occupa 130 byte, di cui 35 per la chiave. Un puntatore a blocco occupa 5 byte. Un blocco di memoria contiene 2048 byte. Utilizziamo un indice ISAM, e supponiamo che ogni blocco sia per i dati che per l'indice sia utilizzato al massimo.
- Quanti blocchi di memoria occupa il file principale?
- Quanti blocchi di memoria occupa il file indice?
- Quale è il numero massimo di accessi per ricercare un record, utilizzando la ricerca binaria sul file indice?

In un blocco entrano $\lfloor \frac{2048}{130} \rfloor = 15$ record, il file principale necessita di $\frac{1750000}{15} = 116667$ blocchi. Una

coppia key-pointer pesa 40 byte, in un blocco ne entrano $\lfloor \frac{2048}{40} \rfloor = 51$, quindi necessito di

$\frac{116667}{51} = 2288$ blocchi. Il costo massimo di una ricerca è $\lceil \log_2 2288 \rceil + 1 = 13$ accessi.

6)
Supponiamo di avere un file di 2.750.000 record. Ogni record occupa 256 byte, di cui 35 per il campo chiave. Ogni blocco contiene 4096 byte. Un puntatore a blocco occupa 5 byte. Usiamo una organizzazione B*-tree. La differenza con il B-tree classico è che ogni blocco del solo livello foglia punta al blocco successivo. Qual è il numero minimo di blocchi che dobbiamo utilizzare complessivamente per il file indice (totale di blocchi di tutti i livelli) e per il file principale (livello foglia)? Qual è il costo di una ricerca in questo caso?

NUMERO MINIMO DI BLOCCHI \Rightarrow RIEMPIRI AL MASSIMO

In un blocco foglia, entrano $\lfloor \frac{4096-5}{256} \rfloor = 15$ record, il file principale conterrà $\frac{2750000}{15} = 183334$ blocchi.

In un blocco del file indice, entrano $\lfloor \frac{4096-5}{35+5} \rfloor = 102$ coppie chiave-puntatore, quindi 103 puntatori.

LIV1: $\frac{183334}{103} = 1780$ blocchi LIV2: $\frac{1780}{103} = 18$ LIV3: $\frac{18}{103} = 1 \Rightarrow$ RADICE IndexFile = $1780 + 18 + 1 = 1799$ blocchi.

Se la radice non fosse in Ram, necessiteremo di **4 accessi**.