

Q. 1

# Hash File

## Hash files

Assume we have a file with 250,000 records, each of them of size 300 bytes, 75 of which for the key field. Each block is of size 1024 bytes. A block pointer is 4 bytes.

- If we use a hash organization with **1200** buckets, how many blocks do we need for the bucket directory?
- How many blocks do we need for the buckets, assuming a uniform distribution of records into buckets?
- Still assuming a uniform distribution of records, what is the average number of accesses needed to find a record in the file?
- How many buckets should we create to have instead an average number of accesses less or equal 10, still assuming a uniform distribution of records in the buckets?

a) 1200 pointer corresponds bucket

$$\text{No. of Pointer} \times \text{Block} (\text{Ptrck}) = \left\lceil \frac{\text{Block Size}}{\text{Pointer size}} \right\rceil = \frac{1024}{4} = 256$$

↑  
App. per df.

$$\text{No. Block} \times \text{Bucket dir} = \frac{\lceil \text{No. Bucket} \rceil}{\text{Ptrck}} = \frac{1200}{256} = 5$$

↓  
App per acc.

$$\text{Record} \times \text{Block} = \frac{\text{Block Size} - \text{pointer size}}{\text{Record Size}} = \frac{1024 - 4}{300} = 3$$

↓  
per df

$$\text{Record} \times \text{Bucket} = \frac{N_{\text{Record}}}{\text{No. Block}} = \frac{250000}{1200} = 208$$

↓  
per access

$$\text{Block per Bucket} = \frac{\text{Recr} \times \text{Bucket}}{\text{Record} \times \text{Block}} = \frac{208}{3} = 70$$

$$\text{Tot Block} = \text{Block per Bucket} \cdot N \text{ Blocks} = 70 \cdot 1200 = 84000$$

c) n° medio blocchi da controllare per trans record

$$\text{Avg Time} = \frac{\text{Block per Bucket}}{2} = 35 \quad \text{per record}$$

d) Voglia che Avg Time passi da 35 a  $\leq 10$

Questo parametro modificabile è n° bucket (1200) anziché due  
due refre quindi tutti i conti all'indietro:

$$\frac{\text{Block} \times \text{Bucket}}{2} \leq 10 \Rightarrow \text{Block} \times \text{Bucket} \leq 20$$

$$\Rightarrow \frac{\text{Record} \times \text{Bucket}}{\text{Record} \times \text{Block}} \leq 20 \Rightarrow$$

Record  $\times$  Block  $\rightarrow$  non modificabile

$$\frac{\text{Record} \times \text{Bucket}}{3} \leq 20 \Rightarrow \text{Record} \times \text{Bucket} \leq 60$$

$$\frac{N^{\circ} \text{ record}}{N^{\circ} \text{ Bucket}} \leq 60 \quad \rightarrow \text{non modificabile}$$

$$N^{\circ} \text{ Bucket} \cdot 60 \geq N^{\circ} \text{ record}$$

$$N^{\circ} \text{ Bucket} \geq \frac{N^{\circ} \text{ record}}{60}$$

$$N^{\circ} \text{ Bucket} \geq \frac{250000}{60} =$$

$$N^{\circ} \text{ Bucket} \geq 4167$$

Q 2

Assume we have a file with 780,000 records, each of them of size 250 bytes. Each block is of size 1024 bytes. A block pointer is 4 bytes. We are using a hash organization with 2500 buckets.

- How many blocks do we need to use for the bucket directory and for the buckets, assuming a uniform distribution of records?
- How many blocks do we need to use, assuming that 30% of records are uniformly distributed in 1000 buckets, and the remaining 70% is uniformly distributed in the remaining 1500 buckets?

b) File A = 30% of 780,000 = 234,000

No Bucket = 1000

File B = 70% of 780,000 = 546,000

No Bucket = 1500

R.f.c.c. of stress, called, data, calculation, total Block

Q 3

ISAM

Assume we have a file of **150,000** records, each of them of size **250** bytes, **50** of which are used for the key field. Each block is of size **1024** bytes. A block pointer is of size **4** bytes.

- a) If we use an ISAM organization, we assume that blocks are filled at most at **70%** of their capacity, how many blocks do we need for the index file?  
 b) If we use binary search, what is the maximum number of accesses to blocks to find a record in the file for case a)?

a)

$$\text{Block Size} = 1024$$

$$\text{Space Usable (max 70\%)} = \frac{1024 \cdot 70}{100} \overset{\substack{\text{per difete} \\ \uparrow}}{=} 716$$

$$\text{Record} \times \text{Block} = \frac{\text{Block Size}}{\text{Record Size}} = \frac{716}{250} = 2$$

$$\text{Tot Block} = \frac{N^{\circ} \text{ Record}}{\text{Record} \times \text{Block}} = \frac{150000}{2} = 75000$$

$$\text{Record per Block} = \frac{\text{Block size}}{\text{key size} + \text{Pointer size}} \overset{\substack{\text{per dif} \\ \uparrow}}{=} \frac{716}{54} = 13$$

$$N^{\circ} \text{ Block Index} = \frac{\text{Tot Block}}{\text{Record} \times \text{Block}} = \frac{75000}{13} = 5770$$



$$b) \text{ Search Time} = \lceil \log_2 (\text{N}^\circ \text{Block}) \rceil + 1 = 13 + 1 = 14$$