

Basi di Dati 2

Formulario

by Palla

Legenda

#blocchi = N
#record = L
Profondità dell'indice = I
dim puntatore = p
dim chiave = a
dim Rlocco = B
dim Record = R
#record d'interesse = n
Livelli = Liv

Nested Loop

Costo senza indici

$$N1 + N1*N2$$

Costo con indici

Albero

$$N1 + L1*I2$$

Hash

$N1 + L1 * 1, \dots \rightarrow$ poco più di 1 (uno) per la questione delle collisioni

Marge Scan / Marge Join

Costo

$$N1 + N2$$

Con i file disordinati

Con indice poco selettivo

$$L1 + L2$$

Con indice molto selettivo

Si visitano solo le foglie dell'albero

Hash Join

Costo

$$2*N1 + 2*N2 + L1 + L2 \approx L1 + L2$$

Costo su una sola tabella

$$2N1 + L1 + N2 \approx L1 + N2$$

B+Tree

Costo

Sparso con ricerca ad intervalli

$$Liv + b$$

$$b = n / (B/R)$$

$$Liv = \log_{F'} (L / (B/R))$$

$$F' = 2/3 F$$

$$F = B / (a + p)$$

Indice secondario denso

$$\text{Liv} + n(1 + 1/F')$$

$$\text{Liv} = \log_{F'} L$$

$$\# \text{foglie dell'indice da visitare} = n/F'$$

DIMENSIONE DELL'INDICE

- L numero di record nel file
 - B dimensione dei blocchi
 - R lunghezza dei record (fissa)
 - K lunghezza del campo pseudochiave
 - P lunghezza degli indirizzi (ai blocchi)
- N. di blocchi per il file (circa): $NF = L / (B/R)$
N. di blocchi per un indice denso: $ND = L / (B/(KP))$
N. di blocchi per un indice sparso: $NS = NF / (B/(K+P))$