

Kostenabschätzung

Terminologie	
m	Anzahl der Seitenrahmen im Datenbank-Puffer
b_R, b_S	Anzahl der Seiten (im Hintergrundspeicher) für die Relationen R bzw. S
M_R, M_S	Anzahl der Tupel von R bzw. S
p_R, p_S	Anzahl der Tupel pro Seite (von R bzw. S)
Selektion	
ohne Index	b_R
key A - B+ Index	≤ 4
key A - Hash Index (statisch)	1
key A - Hash Index (erweiterbar)	2
Index nur mit TID's	3
ungeballter Index	$M_R * sel_{A=C}$
geballter Index	$b_R * sel_{A=C}$
Sortierung	
$2 * b_R * (1 + \lceil \log_{m-1}(\lceil b_R / m \rceil) \rceil)$	
Nested Loop Join	
Simple	$b_R + M_R * b_S$
Pagewise	$b_R + b_R * b_S$
Block	$b_R + k + \lceil b_R / (m-k-1) \rceil * (b_S - k)$
Block (k=1)	$b_R + 1 + \lceil b_R / (m-2) \rceil * (b_S - 1)$
Index (1 qual. Tupel in S)	$b_R + c * M_R$ mit $c \in [1,5]$
Index (n qual. Tupel in S) geballter Index	$b_R + M_R * (c + b_S * sel_{RS})$ mit $c \in [1,5]$
Index (n qual. Tupel in S) ungeballter index	$b_R + M_R * (c + M_S * sel_{RS})$ mit $c \in [1,5]$
Sort Merge Join	
$Sort_R + Sort_S + b_R + b_S$	
Hash Join	
$3 * (b_R + b_S)$	
Hybrid Hash Join	
$(3 - 2 * k/n) * (b_R + b_S)$ Idealfall: eine der Relationen passt zur Gänze in den Puffer: $k=n \Rightarrow k/n=1$ $\Rightarrow b_R + b_S$	

Projektion: mit anderem Schritt kombiniert $\Rightarrow 0$

Duplikatelimination, Gruppierung,

Mengenoperationen: mittels Sortierung oder Hashing implementiert