

Cost Estimation Formulas

Estimated cost of an indexed selection $\sigma_{\theta}(X)$:

- $f B(X)$ if the index is clustered or
- $f T(X)$ if the index is unclustered,
- where f is the selectivity of the condition θ

Estimated cost of $X \text{ JOIN } Y$ when $M (\geq 1)$ extra memory blocks are available:

- Nested loop join: $B(X) + B(X) B(Y) / M$
- Sorted merge join:
 - $B(X) + B(Y)$ if $\max(B(X), B(Y)) \leq M$
 - $4 B(X) + 4 B(Y)$ otherwise
- Hash join $B(X) + B(Y)$ if $\min(B(X), B(Y)) \leq M$
- Indexed join using an appropriate index on Y
 - $B(X) + f T(X) B(Y)$ if the index is clustered or
 - $B(X) + f T(X) T(Y)$ if the index is unclustered,
 - where f is the selectivity of the condition $A = c$

Estimated selectivity of conditions:

- For $A = c$, the selectivity is $1 / (\# \text{ distinct values of } A)$
- For $A < c$, the selectivity is $(c - \text{lowest value of } A) / (\text{highest} - \text{lowest value of } A)$
- For $c < A < d$, the selectivity is $(d - c) / (\text{highest} - \text{lowest value of } A)$