

# TD3 - Optimisation de requêtes

- Parcours séquentiel d'une table R :

$$P(R) = \frac{\text{card}(R)}{T_{\text{page}} / \text{largeur}(R)}, \quad T_{\text{page}} = \text{taille d'une page en octets}$$

nombre de n-uplets de R dans une page.

$$\text{coût}(R) = P(R) \times c \quad \text{où} \quad \begin{cases} c < 1 & \text{si pages contigües (= 0,27 Oracle)} \\ c = 1 & \text{sinon} \end{cases}$$

lecture séquentielle (Table Access FULL)

- Sélection  $\sigma$  :
  - si E expression composé :  $\text{coût}(\sigma_{p(A)}(E)) = \text{coût}(E)$  ←
  - si T table et A pas indexé :  $\text{coût}(\sigma_{p(A)}(T)) = P(T)$

- Sélection index non plasant :

$$\text{coût}(\sigma_{p(A)}(R)) = C_{\text{index}} + C_{\text{rowid}} + \text{card}(\sigma_{p(A)}(R)) \times \frac{CF}{\text{card}(R)}$$

où :  $C_{\text{index}} = \begin{cases} 0 & \text{si l'index tient en mémoire} \\ \text{hauteur de l'arbre} - 1 & \text{sinon} \end{cases}$

$$C_{\text{rowid}} = \begin{cases} 0 & \text{index unique scan} \\ \left\lceil \frac{\text{card}(\sigma_{p(A)}(R))}{\text{card}(R)} \times \text{nb pages} \right\rceil & \text{index range scan} \end{cases}$$

$$CF = \frac{\text{card}(R)}{N/P} \quad \text{où } N \text{ rowid font référence à } P \text{ pages à l'ine}$$

↳ clustering factor

cas général :  $\text{coût}(\sigma_{p(A)}(R)) = \text{card}(\sigma_{p(A)}(R))$  ←

- Sélection index plasant :

$$\begin{aligned} \text{coût}(\sigma_{p(A)}(R)) &= \lceil P(R) \times SF(p(A)) \rceil \\ &= P(R) \times \frac{\text{card}(\sigma_{p(A)}(R))}{\text{card}(R)} \quad \leftarrow \end{aligned}$$

- Jointure par boucles imbriquées

→ index :  $\text{coût}(R \bowtie_{n.a=s.a} S) = \text{coût}(R) + \text{card}(R) \times \text{coût}(\sigma_{a=r}(S))$

avec  $\text{coût}(\sigma_{a=r}(S)) = \begin{cases} 1 & \text{si } a \text{ est une clé de } S \\ C_{\text{rowid}} + \text{card}(\sigma_{a=r}(S)) \times \frac{CF}{\text{card}(S)} & \text{si} \\ \text{index non plasant} \\ \lceil P(S) \times SF(a=r) \rceil & \text{si index plasant} \end{cases}$

→ table sans index :  $\text{coût}(R \bowtie S) = \text{coût}(R) + P(R) \times P(S)$

→ si S est une expression :  $\text{coût}(R \bowtie S) = \text{coût}(S) + P(S) + \text{coût}(R) + P(R) \times P(S)$



- Jointure par hachage :  
 $\text{coût}(R \bowtie S) = \text{coût}(S) + \text{coût}(R)$

hypothèse  $P(R) \geq P(S)$

cas hachage externe (S ne tient pas en mémoire)  
 $\text{coût}(R \bowtie S) = 2 \lceil \log_k (P(S)) \rceil (P(R) + P(S)) + P(R) + P(S)$

avec taille mémoire =  $k+1$  pages

- Tri externe (merge-sort)

- E table :  $\text{coût}(\text{tri}(E)) = \text{coût}(E) + 2 \times P(E) \times (s-1)$

-  $\text{coût}(\text{tri}(R)) = 2 \cdot P(R) \cdot (s-1) + P(R)$  si on ne matérialise pas le résultat (on lit seulement)

-  $\text{coût}(\text{tri}(R)) = 2 \cdot P(R) \cdot s$  coût total

avec  $s = \lceil \log_k (P(R)) \rceil$ , s tq  $k^s \geq P(R)$

- cardinalité d'une sélection :

$$\text{card}(\sigma_{p(A)}(R)) = \text{SF}(p(A)) \times \text{card}(R)$$

→ si égalité :  $\text{SF}(p(A)) = \frac{\text{nb val}}{D(R, A)} \leftarrow \begin{matrix} \text{nombre valeurs distinctes} \\ = \text{IT}_A(R) \end{matrix}$

→ si inégalité :  $\text{SF}(p(A)) = \frac{\text{longueur du segment}}{L(A)} = \frac{\text{longueur du segment}}{\text{max}(A) - \text{min}(A)}$

où borne :  $A \in [\text{min}(A), \text{max}(A)]$

$$\text{longueur du segment} = \begin{cases} r - \text{min}(A) & \text{si } A < r \\ \text{max}(A) - r & \text{si } A > r \\ r_2 - r_1 & \text{si } A \geq r_1 \text{ and } A \leq r_2 \\ \text{max}(A) - r_2 + r_1 - \text{min}(A) & \text{si } A < r_1 \text{ or } A > r_2 \end{cases}$$

- sélection sur plusieurs attributs

-  $\text{SF}(p(A) \text{ and } p(B)) = \text{SF}(p(A)) \times \text{SF}(p(B))$

-  $\text{SF}(p(A) \text{ or } p(B)) = \text{SF}(p(A)) + \text{SF}(p(B)) - \text{SF}(p(A)) \times \text{SF}(p(B))$

-  $\text{SF}(\text{not } p(A)) = 1 - \text{SF}(p(A))$

il faut remonter

- cardinalité sur les opérations :

-  $\text{card}(\text{IT}_A(R)) = \text{card}(R)$

-  $\text{card}(R \times S) = \text{card}(R) \times \text{card}(S)$

-  $\text{card}(R \cup S) \in [\text{max}(\text{card}(R), \text{card}(S)), \text{card}(R) + \text{card}(S)]$

-  $\text{card}(R - S) \in [0, \text{card}(R)]$

- cardinalité jointure :

table :  $\text{card}(R \bowtie_A S) = \text{card}(S) \times \text{card}(R) \times \frac{1}{D(\text{entité}, A)}$

entité : table avec primary key = 1 si jointure naturelle

entre 2 sélections :

$$\text{card}(\sigma_{p(A)}(R) \bowtie_A \sigma_{p(B)}(S)) = \text{SF}(p(A)) \times \text{SF}(p(B)) \times \text{card}(R \bowtie_A S)$$