Étude bibliographique : Le R-tree

# Sources

Guttman, A. (1984). ["R-Trees: A Dynamic Index Structure for Spatial Searching"](Gut84.Rtree.pdf). Proceedings of the 1984 ACM SIGMOD international conference on Management of data - SIGMOD '84. p. 47.

<https://en.wikipedia.org/wiki/R-tree>

<https://technet.microsoft.com/en-us/library/ms190969%28v=sql.105%29.aspx> (Data pages)

<https://technet.microsoft.com/en-us/library/ms179316%28v=sql.105%29.aspx> (Filegroups)

# Résumé

**Un R-tree ne stocke pas le vide.**

# Arbre binaire

<https://en.wikipedia.org/wiki/Binary_tree>

Un arbre qui a au plus deux enfants (gauche-droite)

## Définition

(nœud | (arbre binaire, nœud, arbre binaire))

# Arbres binaires de recherche

# B-tree

# R-tree

Englober des objets dans leur boîte englobante minimale : une recherche n’intersectant pas la boîte ne peut pas intersecter les objets contenus.

## R-tree et stockage

Le R-tree est conçu pour indexer des pages. Un gros fichier est divisé en pages ; à chaque accès au disque dur, histoire de ne pas déplacer la tête de lecture pour pas grand-chose, on lit ou on écrit une page entière. Par exemple, dans MS SQL Server, une page fait 8 ko. Les pages doivent être le plus remplies possible pour que les accès disques soient rentables ; mais remplir beaucoup les pages signifie que (?).

Dans une base de données comme MS SQL Server, il y a un fichier principal qui pointe vers tous les fichiers secondaires où sont stockées les données. Ces fichiers secondaires contiennent les pages évoquées plus haut. Une page est toujours associée à un unique objet (table), et est constitué d’un en-tête, suivie des lignes de données (sauf certaines données de longueur variable), et finie par un pointeur vers le début de chaque ligne. Tous ces fichiers sont stockés dans un système FAT ou NTFS.

## Notations (Wikipedia)

M := Nombre maximum d’entrée dans une page

## Propriétés

Équilibré

De Recherche

Remplissage des pages garanti (conseil 30–40 %)