

Structures de données avancées :

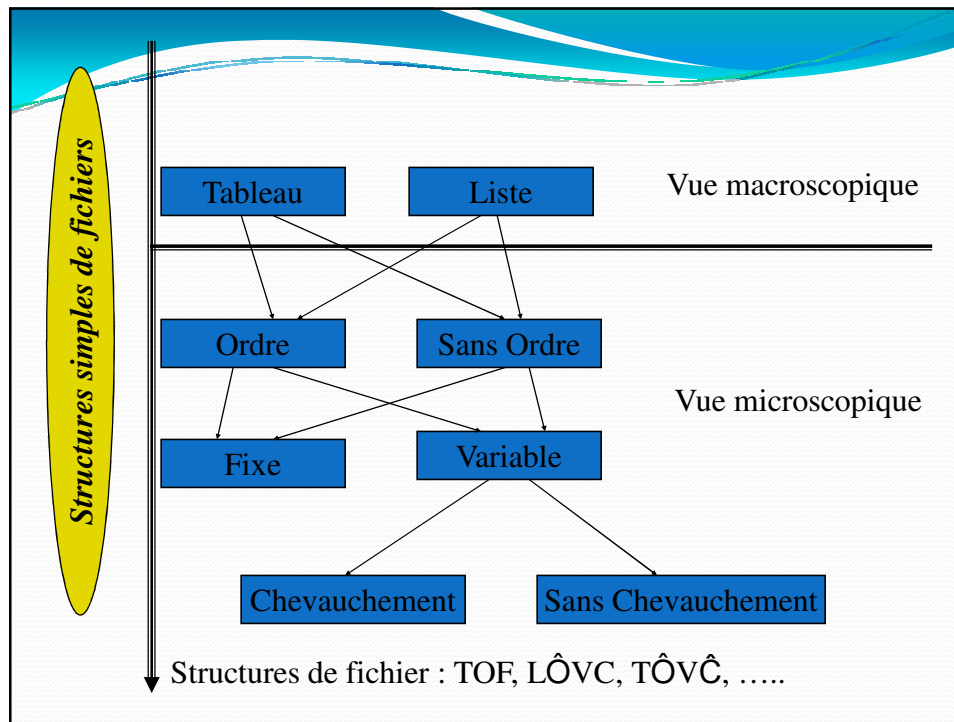
Principales structures de fichiers

Pr ZEGOUR DJAMEL EDDINE
Ecole Supérieure d'Informatique (ESI)
www.zegour.univ.dz
email: d_zegour@esi.dz

Les principales structures de fichiers

Structures simples de fichiers

- Fichier = Tableau de blocs
- Fichier = Liste linéaire chaînée de blocs
- Remarques
 - ✓ L'organisation du fichier est quelconque (format fixe ou variable)
 - ✓ Si fichier ordonné : chargement initial à 60-70%.
==> économiser les décalages intra-blocs.
 - ✓ limitées à des petits fichiers. (< 6 accès disque)



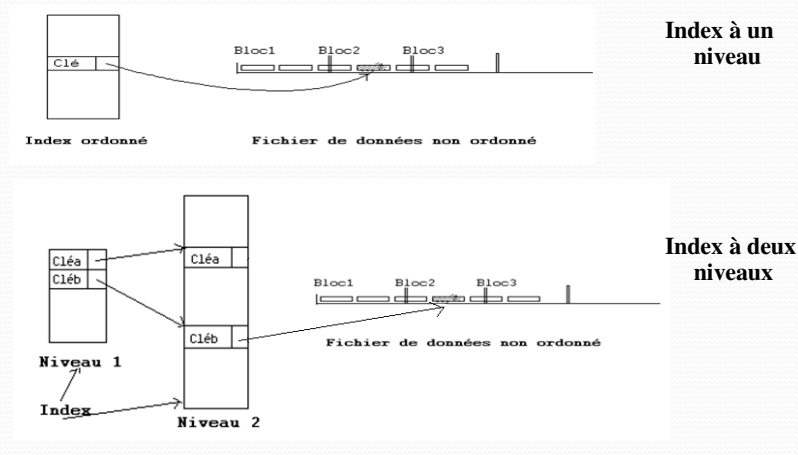
Les principales structures de fichiers

Les méthodes simples d'index (Variante 1)

- 2 fichiers sont utilisés : fichier d'index et fichier de données.
- En règle générale, le fichier de données n'est pas trié.
- Fichier de données = Structure simple de fichier non ordonné (tableau ou une liste linéaire chaînée).
- Fichier d'index = Tableau ordonné (selon les clés) en mémoire. (recherche dichotomique en mémoire).

Les principales structures de fichiers

Les méthodes simples d'index (Variante 1)



Les principales structures de fichiers

Les méthodes simples d'index (Variante 1)

- Opérations
 - ❑ Recherche :
 - a) Recherche dichotomique sur l'index.
 - b) Si clé trouvée, ramener le bloc en mémoire centrale
 - c) Si article à cheval sur deux blocs, ramener le deuxième bloc
 - ❑ Insertion
 - a) Si clé non trouvée dans l'index, il y a insertion avec décalages.
 - b) l'article est inséré en fin de fichier.

Les principales structures de fichiers

Les méthodes simples d'index (Variante 1)

❑ Suppression

Généralement logique ==> extrêmement rapide.

❑ Réorganisation

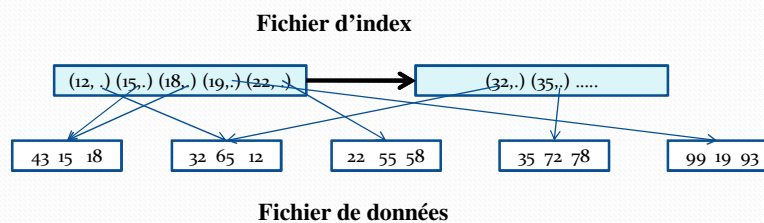
Éliminer physiquement les articles supprimés. Généralement, on construit un autre fichier.

Les principales structures de fichiers

Les méthodes simples d'index (Variante 2)

Fichier de données non ordonné.

Fichier d'index ordonné maintenu sur le disque (contient toutes les clés) [*Dense index*]

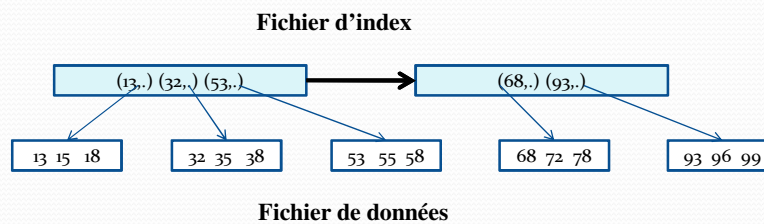


Les principales structures de fichiers

Les méthodes simples d'index (Variante 3)

Fichier de données ordonné.

Fichier d'index ordonné maintenu sur le disque (ne contient pas toutes les clés) [*Sparse index*]



Les principales structures de fichiers

Accès séquentiel indexé

Définition

- pour localiser un article on utilise un index.
- pour avoir le suivant, on ne repasse pas par l'index.

Les principales structures de fichiers

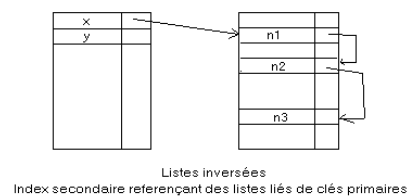
Accès multicritères

- **Concepts**
Un article = (cle1, cle2 ...cléd)
Une clé primaire et d-1 clés secondaires
Récupérer l'article par plusieurs clés
- **Exemples de requêtes :**
Livres parus en 1973,
Livres parus en 1973 de l'auteur X donné.
- **Solution**
Créer des indexes secondaires pour chaque attribut
Passer toujours par l'index primaire.

Les principales structures de fichiers

Accès multicritères

- **Organisation d'un index secondaire**
"listes inversées". (clé secondaire ==> clé primaire ==> article)



Les principales structures de fichiers

Accès multicritères

- Algorithme qui Récupère les articles possédant l'attribut X et l'attribut Y.
 - ❑ Il existe donc des indexes sur X et Y
 - ❑ L' index sur X donne l'ensemble des clés primaires avec l'attribut X.
 - ❑ L'index sur Y donne l'ensemble des clés primaires avec l'attribut Y.
 - ❑ L'intersection donne l'ensemble des clés primaires ayant l'attribut X et Y.
 - ❑ Pour chaque élément de cet ensemble, on fait alors un accès disque.

Les principales structures de fichiers

Accès multicritères

- Suppression :
 ===> Ne pas toucher aux indexes secondaires
 (car les clés secondaires référencent les clés primaires)
- Recherche :
 C'est dans l'index primaire qu'on s'aperçoit que l'article à été supprimé.

Les principales structures de fichiers

Les méthodes d'arbres

- Si le fichier est très volumineux, on ne peut garder l'index en RAM.
- Garder l'index en mémoire secondaire coûte chère car :
 - 1) la recherche dichotomique exige beaucoup d'accès disque
 - 2) il faut maintenir l'index ordonné
- ➔ Trouver une meilleure façon de faire les insertions et les suppressions d'articles avec seulement une réorganisation locale.

Les principales structures de fichiers

Les méthodes d'arbres

- Utilisation d'un arbre de recherche m-aire (ARM) comme méthode d'accès.
- ❑ Un ARM est surtout utilisé pour le stockage en mémoire externe.
- ❑ Un nœud de l'arbre est alors un bloc sur le disque.
- ❑ A un moment donné seule un bloc est ramené en mémoire.
- ❑ Nécessite des blocs de très grande taille (ordre de 200).

Les principales structures de fichiers

Les méthodes d'arbres

- Implémentation des ARM
 - Deux façons de ranger les articles :
 - (i) dans les nœuds avec les clés, (fichier de données = ARM)
 - (ii) séparément. pointeur additionnel vers l'information. (fichier = index vers un fichier de données)
 - Si le fichier est grand : utilisation de (ii)

Les principales structures de fichiers

Les méthodes d'arbres

- Utilisation des ARM
 - Avantages
 - ✓ Le problème de taille ne se pose plus
 - ✓ Fichier ordonné
 - Inconvénients
 - ✓ Destinés pour les fichiers statiques
 - ✓ Déséquilibre de l' ARM
 - Pas de solution → Réorganisation périodique