

الجمهورية الجزائرية الديمقراطية الشعبية

République Algérienne Démocratique et Populaire

وزارة التعليم العالي و البحث العلمي

*Ministère de l'Enseignement Supérieur et de la Recherche
Scientifique*

*Ecole Supérieure d'Informatique
De Sidi Bel Abbès*



مدرسة العليا للإعلام الإلكتروني
بسيدي بلعباس

ORGANISATION DES FICHIERS

Méthodes Index

Présenté par Dr.Nabil KESKES.

Année 2017-2018.

PLAN

- Introduction.
- Définition
- Index Dense et non Dense
- Index primaire a un niveau
- Index primaire a plusieurs niveaux
- Conclusion

Ecole Supérieure d'Informatique
De Sidi Bel Abbès



مدرسة العليا للإعلام الآلي
بمبدي بلعباس

Objectifs du Chapitre

L'objectif de notre cours consiste donc à répondre aux deux questions suivantes :

Comment définir :

- ✓ l'organisation du fichier : séparation des champs, des articles, ..
- ✓ la méthode d'accès : comment localiser un article sur le fichier ?

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Ecole Supérieure d'Informatique
De Sidi Bel Abbès



مدرسة العليا للإعلام الآلي
بمسيسي بلعباس

Ca revient donc à étudier les opérations de base :

- recherche
- insertion
- suppression
- requête à intervalle

sous différentes formes d'accès qui sont :

- ✓ les méthodes d'index
- ✓ l'accès multi-clés

Ecole Supérieure d'Informatique
De Sidi Bel Abbés



مدرسة العليا للإعلام الآلي
بمسيسي بلعباس

Introduction

Si l'opération de recherche est très fréquente pour une application donnée, la meilleure alternative parmi les structures simples de fichiers est le choix d'un tableau ordonné avec un format fixe des articles. Par conséquent, la méthode se trouve limitée à des fichiers dont **la taille devienne volumineuse**

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Ecole Supérieure d'Informatique
De Sidi Bel Abbés



مدرسة العليا للإعلام الآلي
بمسيدى بلعباس



les opérations d'accès (recherche, insertion, ...) deviennent très inefficaces.



Les méthodes d'index permettent d'améliorer, dans une certaine mesure, les performances en gérant une structure auxiliaire (table d'index) accélérant la recherche.

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Ecole Supérieure d'Informatique
De Sidi Bel Abbés



مدرسة العليا للإعلام الآلي
بمبيدي بلعباس

En l'absence d'un index, seules solutions :

- ✓ Parcours séquentiel (complexité linéaire)
- ✓ Recherche par dichotomie si fichier trié (complexité logarithmique)

Avec un index :

Parcours de l'index, puis accès direct à l'enregistrement Mais attention : mises à jour plus coûteuses !!

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Ecole Supérieur d'Informatique
De Sidi Bel Abbés



مدرسة العليا للإعلام الآلي
بمسيسي بلعباس

Definition

They are based on ***indexing field(s)*** that are used to construct the ***index***.

Un index est (généralement) une table ordonnée contenant les couples <clé, adr > utilisée pour accélérer la recherche des enregistrements d'un fichier.

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي و البحث العلمي
*Ministère de l'Enseignement Supérieur et de la Recherche
Scientifique*

*Ecole Supérieure d'Informatique
De Sidi Bel Abbés*



مدرسة العليا للإعلام الآلي
بمسيسي بلعباس

Remarque

✓ Pour ces méthodes, deux fichiers sont utilisés : *le fichier d'index* et *le fichier au données*. En règle générale, le fichier de données n'est pas trié. Il peut être un tableau ou une liste linéaire chaînée de blocs sur le disque. Le fichier d'index est supposé en mémoire et est trié selon les clés des articles. L'index peut être à un ou plusieurs niveaux comme nous allons le décrire par la suite. L'organisation interne du fichier est quelconque.

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Ecole Supérieure d'Informatique
De Sidi Bel Abbés



مدرسة العليا للإعلام الآلي
بمبيدي بعباس

Exemple 1:

L'utilisation des index est basée sur l'observation suivante: pour trouver un livre dans une bibliothèque, au lieu d'examiner un par un chaque livre (ce qui correspond à une [recherche séquentielle](#)), il est plus rapide de consulter le catalogue où ils sont classés par thème, auteur et titre¹. Chaque entrée d'un index comporte une valeur extraite des données et un pointeur sur son emplacement d'origine. Un enregistrement peut être ainsi facilement retrouvé en recherchant sa localisation dans l'index².



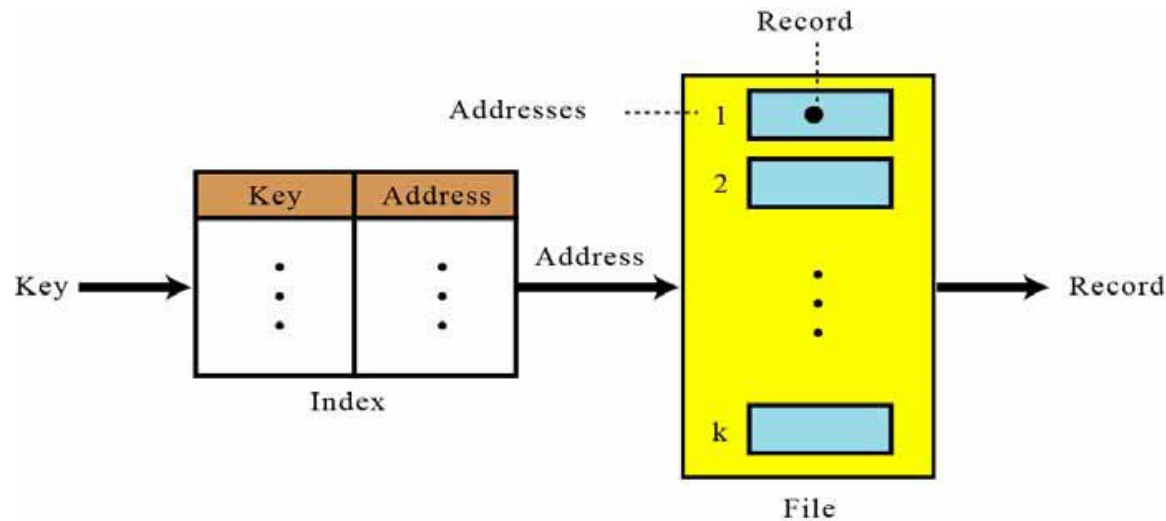
*Ecole Supérieur d'Informatique
 De Sidi Bel Abbés*

مدرسة العليا للإعلام الآلي
 بسيسي بلعباس

Exemple 2:

INDEXED FILES

To access a record in a file randomly, we need to know the address of the record.



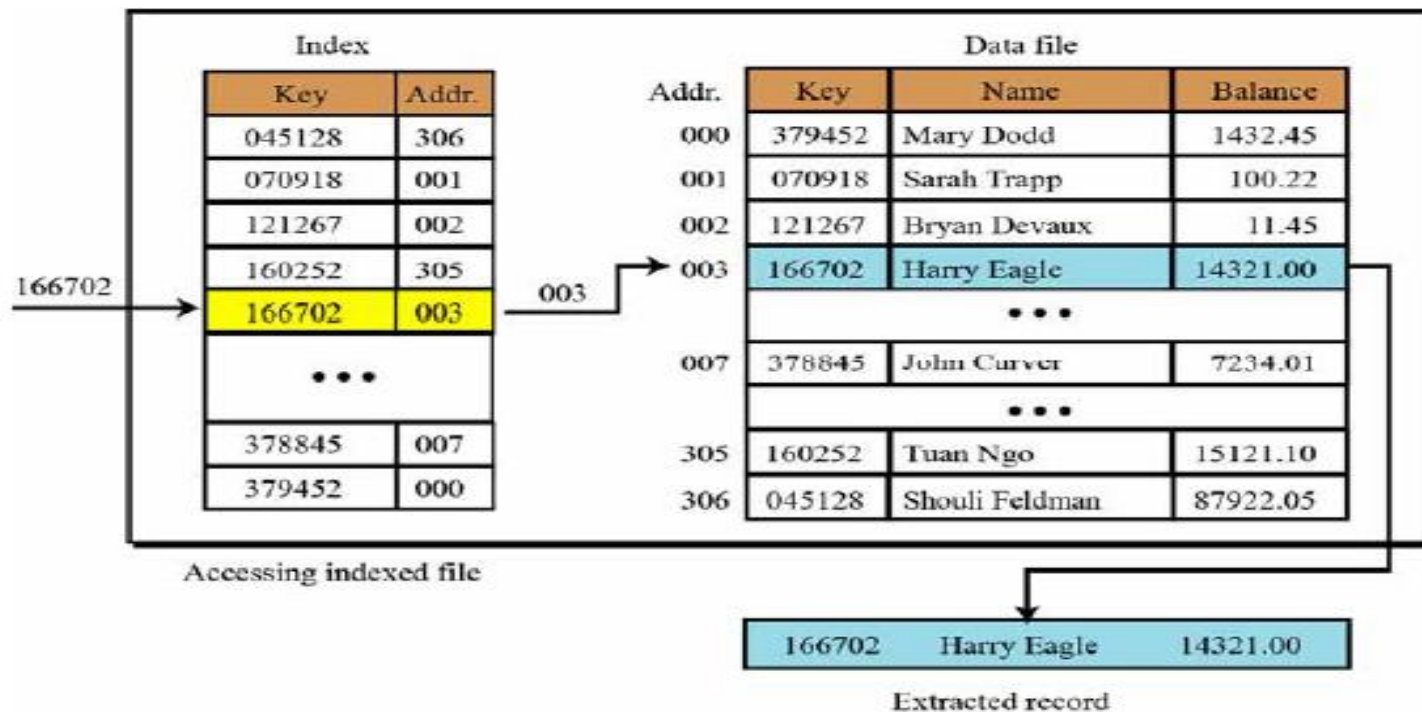
الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
 وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique



Ecole Supérieur d'Informatique
De Sidi Bel Abbés

مدرسة العليا للإعلام الآلي
 بسيسي بلعباس

Exemple 3:





Avantages des méthodes d'index

Même si l'index n'est pas en mémoire, les méthodes d'index ont les avantages suivants :

- ✓ permettent la recherche dichotomique même pour les articles de longueur variable.
- ✓ faire la recherche dichotomique sur le fichier d'index est beaucoup plus rapide que sur le fichier de données lui-même.
- ✓ L'effacement des articles par des "flag" se fait sans accès disque.



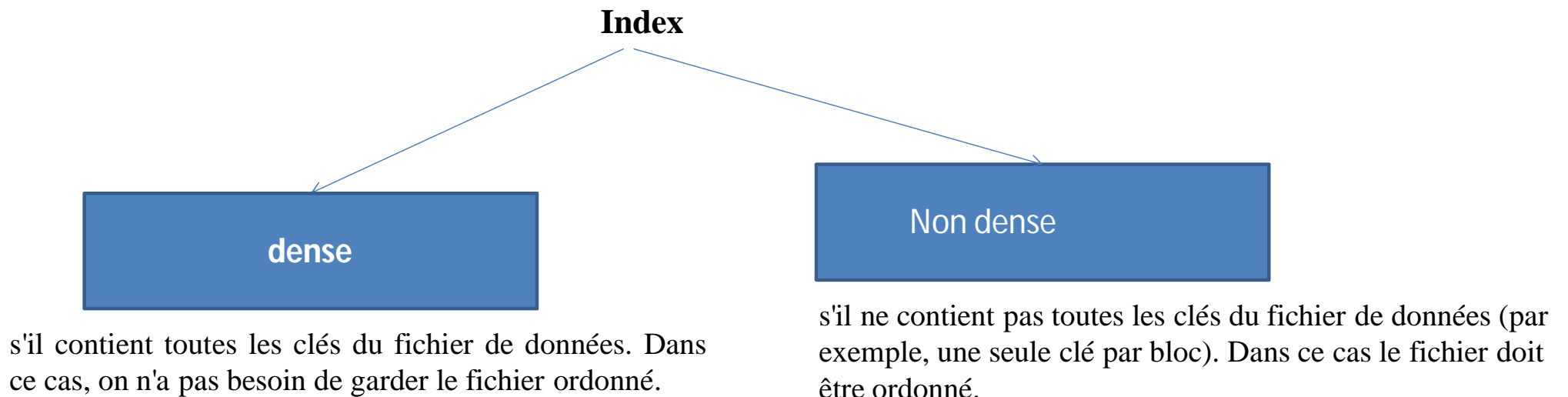
Inconvénients des méthodes d'index

Si l'index est large pour être contenu en mémoire

- ✓ recherche binaire chère
- ✓ réarrangement coûteux à cause des décalages) on a recours à d'autres techniques : Arbre, hachage



Typologies selon la dense du clé



Ecole Supérieur d'Informatique
De Sidi Bel Abbés



مدرسة العليا للإعلام الآلي
بمبيدي بلعباس

Index dense :

Comporte un enregistrement pour *chaque valeur de la clé* de tri du fichier indexé.

Avec le temps, ce type d'index peut occuper beaucoup d'espace mémoire.

Mais..ils ont un temps de recherche très court.

الجمهورية الجزائرية الديمقراطية الشعبية

République Algérienne Démocratique et Populaire

وزارة التعليم العالي و البحث العلمي

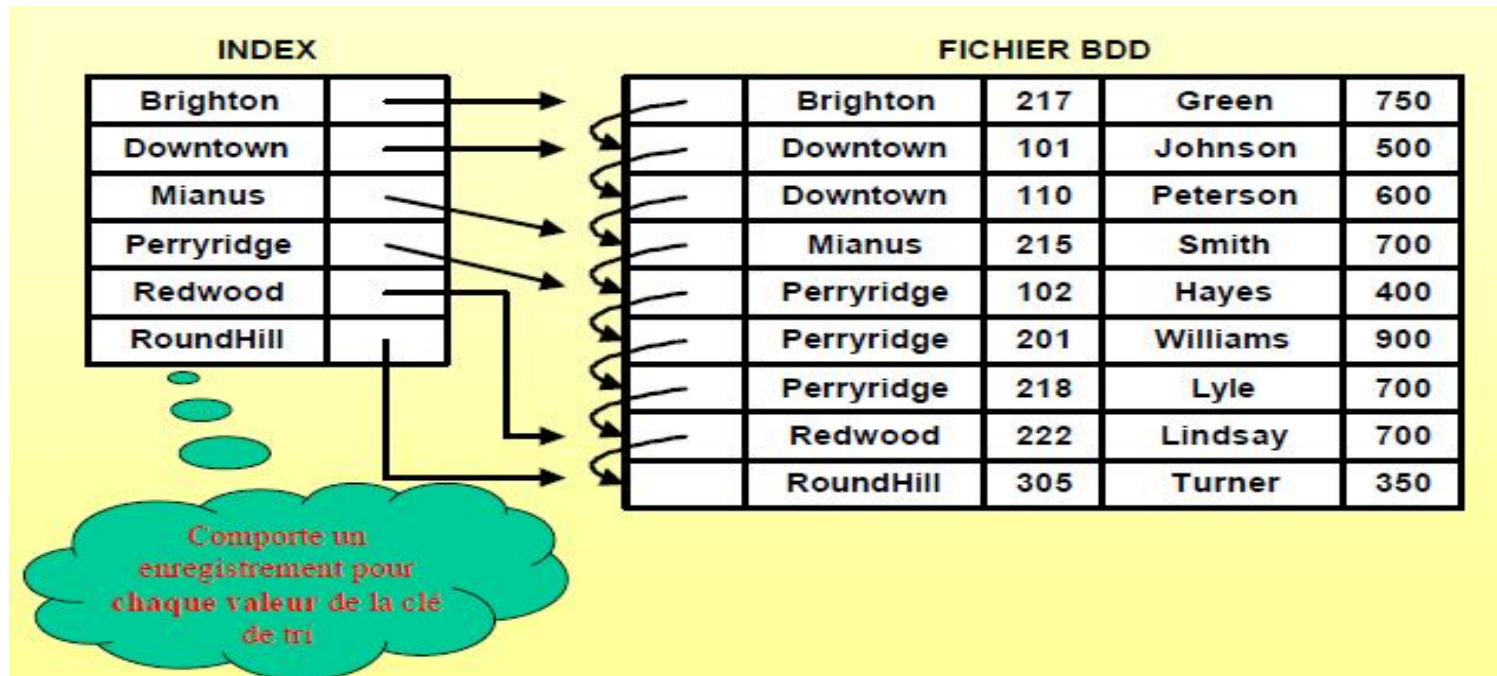
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Ecole Supérieure d'Informatique

De Sidi Bel Abbès



مدرسة العليا للإعلام الآلي
بسيدي بلعباس



Exemple d'index dense

Ecole Supérieure d'Informatique
De Sidi Bel Abbés



مدرسة العليا للإعلام الآلي
بسيدي بلعباس

Index non dense

Comporte des enregistrements pour certaines valeurs de la clé de tri du fichier indexé.

Les index non dense ont des avantages :

- ✓ ils occupent **moins d'espace**;
- ✓ ils imposent **moins de servitudes lors d'insertions et d'effacements**.

الجمهورية الجزائرية الديمقراطية الشعبية

République Algérienne Démocratique et Populaire

وزارة التعليم العالي و البحث العلمي

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

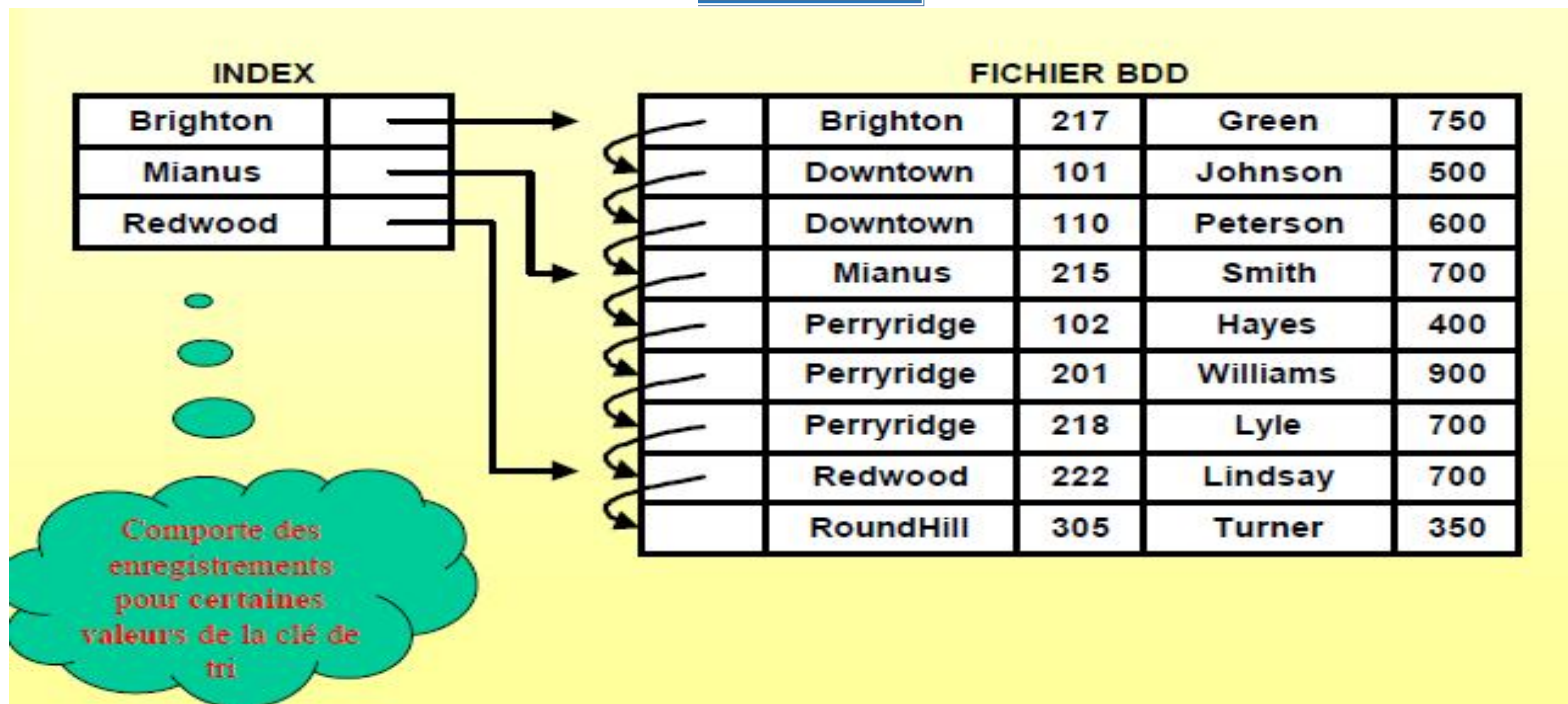


Ecole Supérieur d'Informatique

De Sidi Bel Abbés

مدرسة العليا للإعلام الآلي

بمسيسي بلعباس



Exemple d'index non dense



Typologies selon la nature du clé

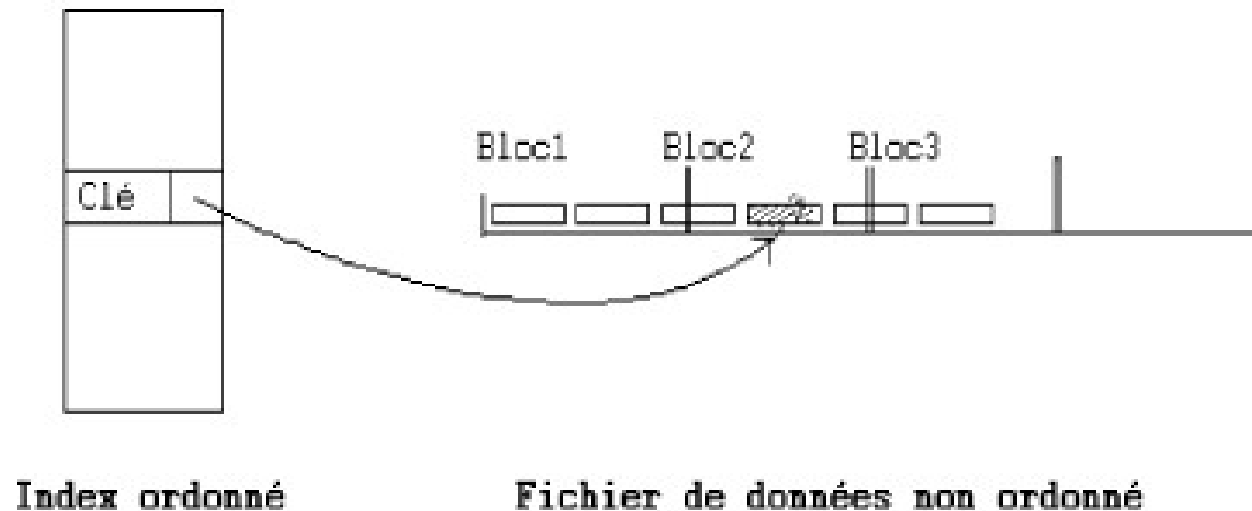
Index primaire

Si le champ clé ne contient de valeurs en double, l'index est alors « primaire »

Primary Index: An index ordered in the same way as the data file, which is sequentially ordered according to a key. (The indexing field is equal to this Key.)

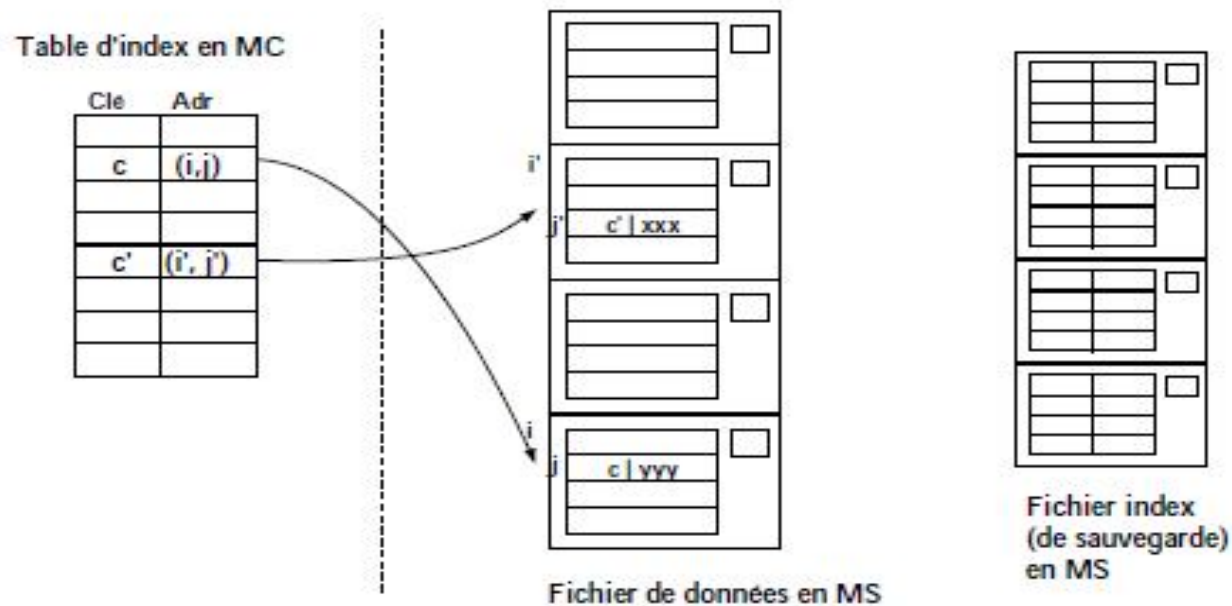
index primaire à un niveau

L'index est un ensemble de couples (clé, adresse) rangés entièrement en mémoire centrale. Le fichier est un ensemble de blocs (dans l'exemple les blocs sont contigus) sur le disque. Le bloc contient un ensemble d'articles. les articles peuvent être à cheval sur deux blocs logiquement consécutifs. C'est le cas de la figure suivante :



index primaire à un niveau

Les fichiers de données et d'index peuvent être de n'importe quelle structure (blocs contigus, blocs chaînés, ...). De même que les enregistrements peuvent être à format fixe ou variable (avec ou sans chevauchement).



Operations de base

Pour mettre au point une méthode d'index, les opérations suivantes sont nécessaires :

- ✓ créer un fichier d'index et de données
- ✓ charger le fichier d'index en mémoire avant utilisation
- ✓ rechercher un article de clé donnée,
- ✓ insérer un article,
- ✓ supprimer un article,
- ✓ modifier un article.

rechercher un article de clé donnée

Pour rechercher un article de clé donnée dans le fichier, on commence par faire une recherche dichotomique sur l'index.

✓ Si la clé est trouvée, le bloc contenant l'article est ramené en mémoire centrale afin de récupérer l'article.

✓ Si ce dernier est à cheval sur deux bloc, le deuxième bloc est ramené pour récupérer le reste de l'article.

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Ecole Supérieure d'Informatique
De Sidi Bel Abbès



مدرسة العليا للإعلام الآلي
بمبدي بلعباس

✓insérer un article

Une insertion d'article est réalisée comme suit :

- ✓ si la clé n'est pas trouvée dans l'index, elle est y insérée et peut donc entraîner des décalages.
- ✓ l'article est inséré en fin de fichier.

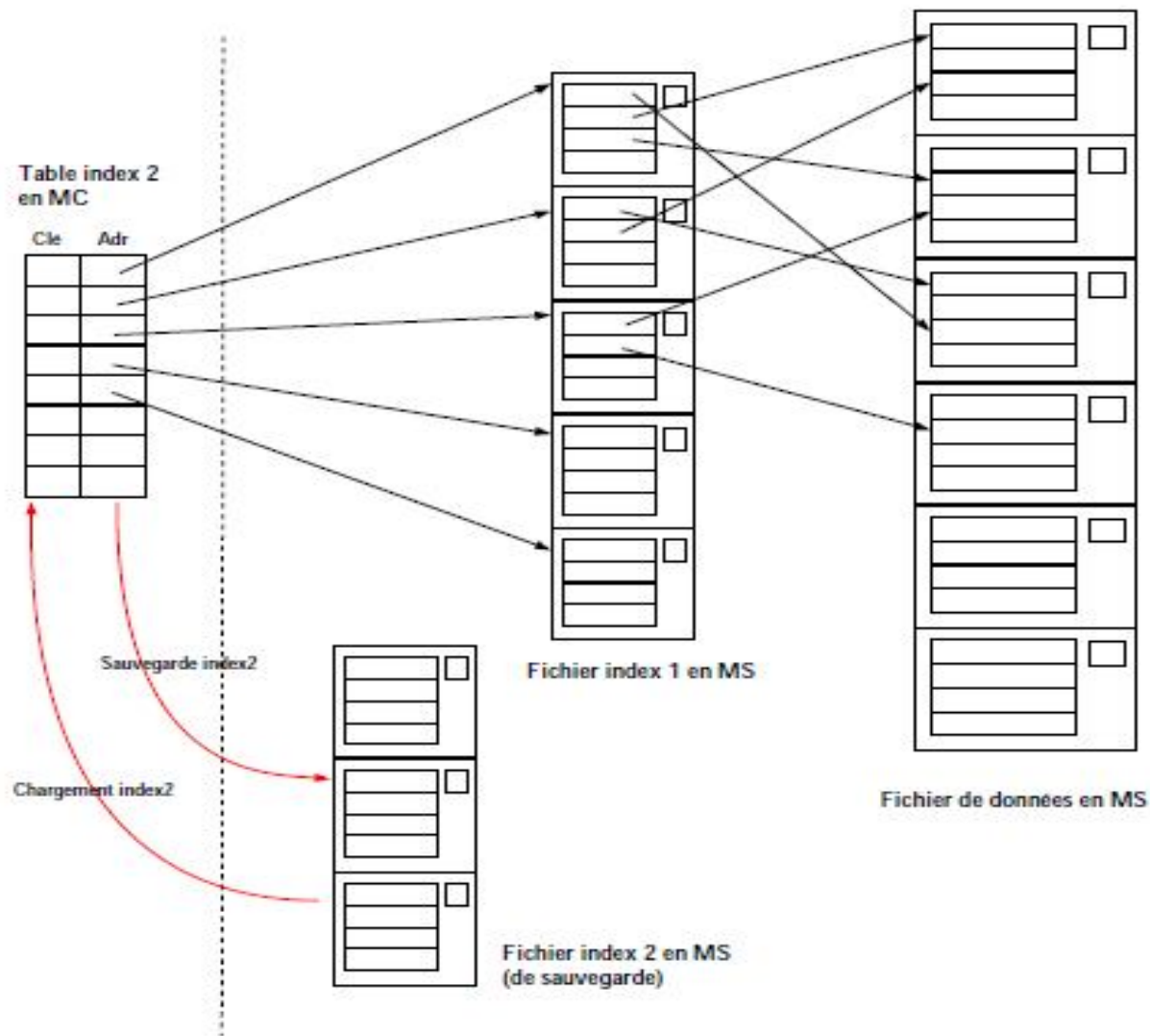
Supprimer un article

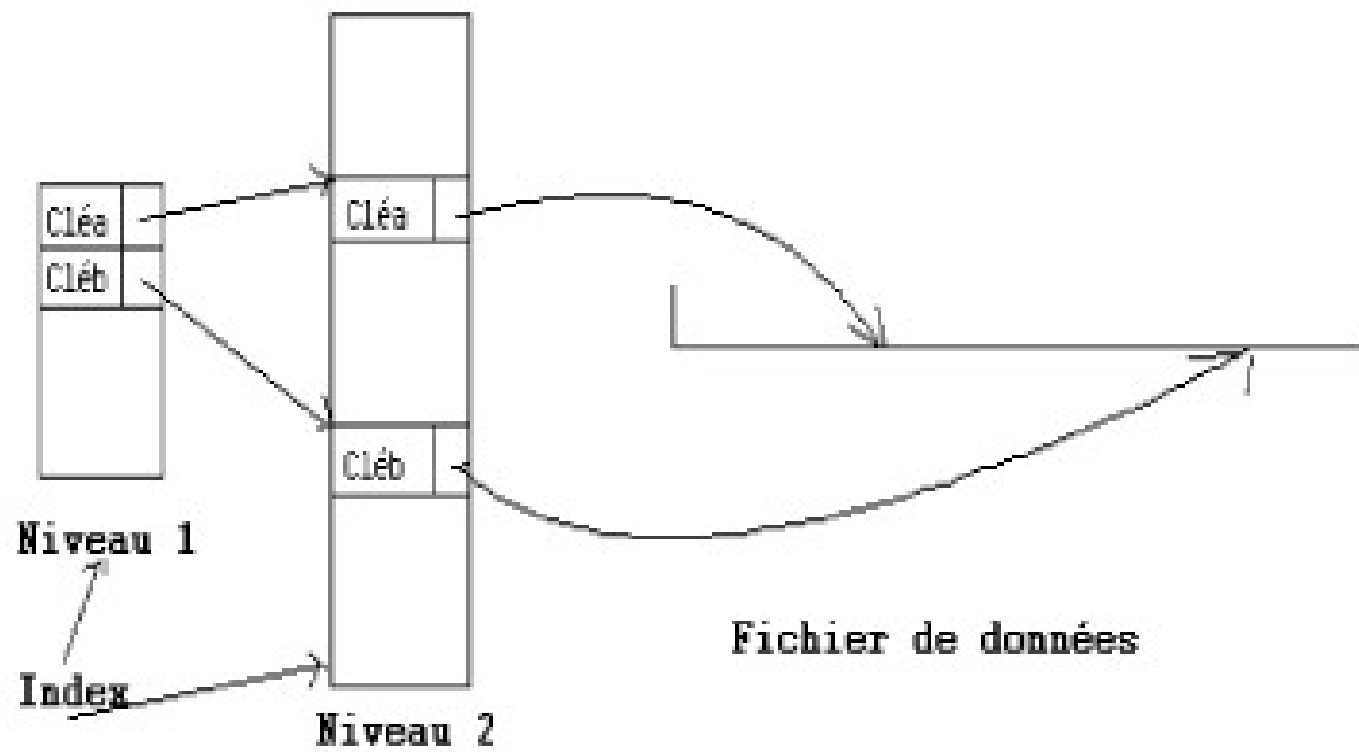
- ✓ Une suppression est généralement logique. Elle est donc extrêmement rapide.
- ✓ Une réorganisation est donc à prévoir afin d'éliminer physiquement les articles supprimés. Généralement, on construit un autre fichier.

index primaire à plusieurs niveaux

✓ Si l'index est trop grand pour résider en MC, on construit un deuxième index sur le fichier index (ordonné). Dans ce cas, on choisira une seule clé pour chaque bloc du fichier index (index non dense) pour construire le deuxième index.

✓ Si le deuxième index est encore trop grand pour résider en MC, on le stocke sur disque (deuxième fichier index) et on construit un troisième index en choisissant une clé par bloc du deuxième fichier index. On peut répéter ce procédé au tant que nécessaire.





rechercher un article de clé donnée

Pour rechercher un article de clé donnée, on commence par rechercher sa clé dans l'index du niveau 1. un intervalle est alors sélectionné. La recherche continue dans l'index de niveau 2 uniquement dans cet intervalle. Si la clé est trouvée, on procède de la même façon que dans le cas avec un seul index. Il est clair que les deux recherches dichotomiques sur les deux petits vecteurs (index de niveau 1 et une partie de l'index de niveau 2) sont beaucoup plus rapides qu'une seule recherche dichotomique sur un seul grand vecteur (index de niveau 2).

الجمهورية الجزائرية الديمقراطية الشعبية

République Algérienne Démocratique et Populaire

وزارة التعليم العالي و البحث العلمي

*Ministère de l'Enseignement Supérieur et de la Recherche
Scientifique*

*Ecole Supérieure d'Informatique
De Sidi Bel Abbés*



مدرسة العليا للإعلام الآلي
بسبيدي بلعباس

Conclusion

- ✓ Les méthodes d'index permettent d'améliorer, dans une certaine mesure, les performances en gérant une structure auxiliaire (table d'index) accélérant la recherche
- ✓ Les méthodes d'index comme celles présentées dans ce chapitre, sont dédiées aux fichiers statiques (c.-à-d. dans les cas où le nombre d'insertions et de suppressions est relativement faible)

Références

Sitographie

<https://sites.google.com/a/esi.dz/hidouci/competences-professionnelles/algo2>

<https://www.lri.fr/~benzaken/documents/slindex.pdf>

<http://www.nhu.edu.tw/~chun/CS-ch13-File%20Structures.pdf>

<http://www.cs.cityu.edu.hk/~helena/cs34622000B/Files.pdf>

<http://zegour.esi.dz/Publication/Livre2/Partie-sdf/Part2-ss.htm>