

Introduction aux S. G. B. D.

1

1. L 'origine

2

Les concepts de base

- Donnée : représentation d'une information
- Type de donnée : définit la nature du codage et les opérations autorisées sur une donnée
- Donnée simple : non décomposable
- Donnée complexe : composée de données simples ou complexes
- ex 1 : entier : donnée simple
- ex 2 : chaîne de caractère : simple ou complexe
- ex 2 : donnée multimédia : simple ou complexe

3

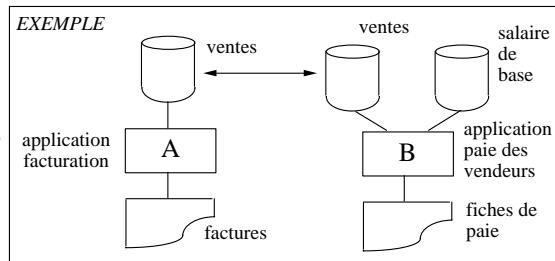
Les données en mémoire

- Mémoire centrale :
 - RAM, volatile
- Mémoire secondaire :
 - disque magnétique, stockage permanent
- Mémoire tertiaire :
 - bande, disque optique, redondance pour sauvegarde

4

Approche traditionnelle

- Approche « fichier »
- Approche par application
- Pour une application, on définit :
 - les entrées
 - les sorties
 - les fichiers
 - les supports physiques
 - le langage
 - le responsable, ...



5

Approche « fichier » - inconvénients

- Redondance des données
 - multiplication des tâches de saisie, mise à jour
 - multiplication des programmes de contrôle
 - augmentation du volume => supports physiques
 - incohérence => fiabilité
- Impossibilité de partager les données
 - pas d'accès simultané
- Difficultés d'accès à l'information
 - technique : comment ?
 - pratique : où trouver les données ?

6

2. La gestion de fichiers

Gestion de fichiers : notions de base (1)

- Fichier : ensemble de données semblables physiquement regroupées sur un support
 - ex : données concernant tous les clients d'une entreprise
 - un fichier est composé d'articles
- Article : unité d'entrée, de traitement ou de sortie des informations
 - ex : les données relatives à un client
 - un article est composé de rubriques

Gestion de fichiers : notions de base (2)

- Rubrique ou champ : niveau élémentaire de découpage de l'information
 - ex : le nom d'un client
 - si l'une des rubriques (ou plusieurs) permet l'identification d'un seul article dans le fichier, c'est une clé

N° client	Nom	Adresse
-----------	-----	---------

CLE

9

Opérations de gestion de fichiers

- Créer / détruire un fichier
- Ouvrir / fermer un fichier
- Insérer / modifier / supprimer un article
- Rechercher un article ayant une valeur donnée pour un de ses champs (ou plusieurs champs)

10

Autres notions

- Clé d'accès : champ sur lequel s'effectue la recherche dans un fichier
 - peut être discriminante ou non
- Chemin d'accès :
 - structure de données accélérant la recherche des articles sur une valeur de clé d'accès
- Page mémoire :
 - unité de lecture (et écriture) sur le support mémoire
- Organisation d'un fichier :
 - mode de stockage de l'article dans les pages mémoire

11

Organisation séquentielle

- Les articles sont stockés les uns à la suite des autres sans trous.
- Les articles sont généralement triés selon un ordre déterminé sur des rubriques.
 - simple => peu coûteux
 - peu performant
 - accès séquentiel uniquement

12

Organisation relative

adresse de
l'article de = $i \text{ Div } \frac{\text{taille page}}{\text{taille article}} + 1$
clé i

- Articles stockés séquentiellement et de longueur fixe
- Recherche par clé discriminante = numéro d'ordre de l'article dans le fichier
- Inconvénients :
 - taille fixe
 - nécessité d'une clé discriminante continue

13

Organisation hachée ou aléatoire

valeur de clé $\xrightarrow[\text{de hachage } h]{\text{fonction}}$ n° paquet

- Découper un fichier en paquets de faible taille et répartir les articles dans les paquets en fonction de leur clé
- Si la fonction de hachage est non injective, la recherche est plus compliquée
- Il est difficile de faire des recherches autres que sur l'égalité d'une valeur de clé

14

Organisation triée ou arborescente

Fichier index (valeur-clé, adresse-page)

plus petite valeur de clé des
articles de cette page

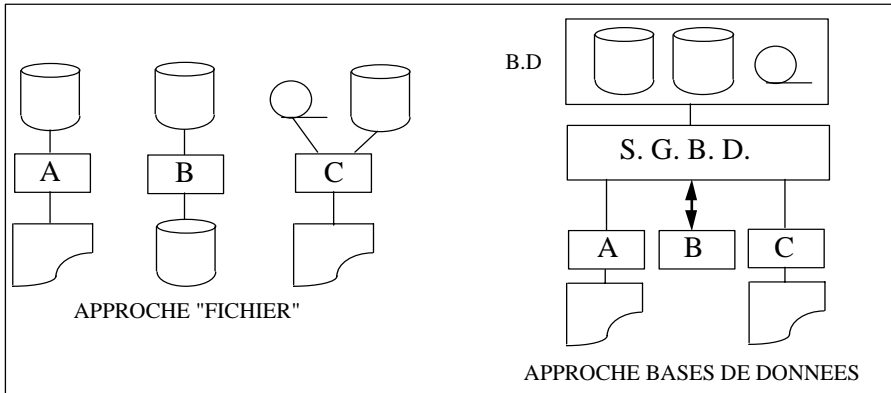
trié sur valeur-clé

- Fichier trié selon une clé d'accès
- Organisation arborescente des données
 - pour faciliter les insertions et les suppressions
- Index primaire (plaçant) = l'ordre des clés est l'ordre des articles
- Index secondaire (non plaçant) = ré-ordonnancement logique des articles du fichier

15

3. Les SGBD - Définition, principes et architecture

Approche base de données



- Regroupement de l'ensemble des données
- Accès aux données géré et contrôlé par le S.G.B.D.

17

Base de données et SGBD

- Base de données :
 - Ensemble de données :
 - fortement structurées
 - persistantes
 - dont la structure est définie dans un schéma
 - gérées par un système de gestion de base de données
- Système de gestion de base de données
 - logiciel spécialisé pour la gestion de base de données

18

Objectifs des bases de données

- Gestion centralisée : non-redondance
- Contrôle d'accès
- Partage des données
- Représentation abstraite des données
 - faciliter l'accès et faciliter l'évolution et la maintenance
- Performances : comparables ou meilleures

19

Qu'est ce qu'une base de données ?

- Collection de données reliées
- Stockées sans redondance inutile
- Utilisables pour de multiples applications
- La structure de stockage doit être indépendante des applications utilisant ces données
- La manipulation des données se fait suivant une approche commune et contrôlée :
 - extraction
 - ajout
 - modification
 - suppression
- Les données sont structurées pour permettre le développement cohérent des applications futures

20

Fonctions essentielles du SGBD

- Description : à partir d'une représentation abstraite des données, assure la description de la structure physique des données
- Création : offre un langage d'introduction de données nouvelles dans la base, contrôle la validité des données
- Mise à jour : assure la maintenance des données
- Extraction : permet l'interrogation interactive ou via des programmes

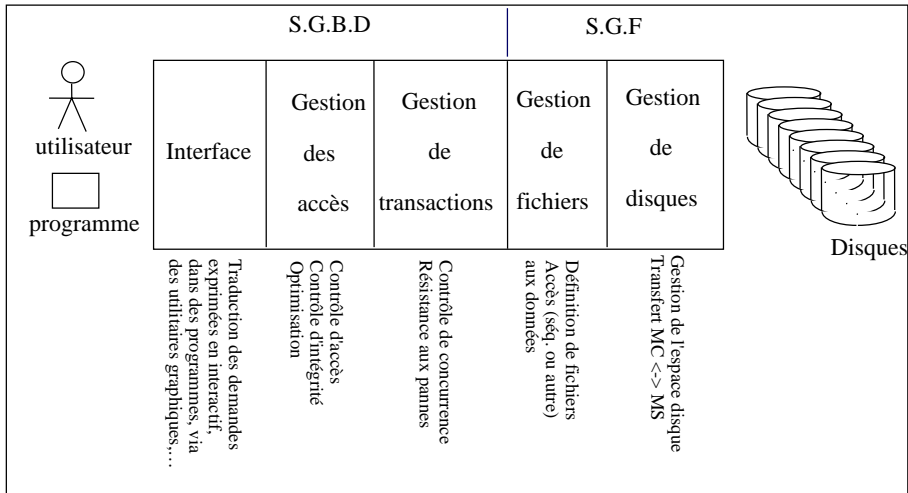
21

SGBD - Fonctions induites

- Optimisation de l'accès => performances
- Contrôle de l'intégrité et de la cohérence
 - Contrôle individuel des données : type, valeur, ...
 - Contrôle des relations entre les données
- Contrôle du partage des données (concurrency)
- Contrôle de l'accès :
 - Confidentialité et sécurité des données
 - Prise de conscience de la valeur de l'information
- Résistance aux pannes :
 - Sauvegardes, journalisation et points de reprise
- Faciliter l'accès aux données : interfaces variées
 - Programmes, graphiques ..., langage naturel

22

Architecture d'un SGBD



23

Principe d'indépendance physique

- La manipulation des données doit se faire indépendamment de leur structure de stockage
 - représentation abstraite des données
 - structure logique ou conceptuelle ° structure physique
 - On peut modifier la structure physique sans que les utilisateurs soient affectés et sans modification des programmes

Exemple :

Structure logique

Employé : Nom
 Adresse
 Fonction
 Enfants

Prénom
 Age

Structure physique

- 1 seul fichier
- ou bien
- 1 fichier Employé et 1 fichier Enfants et des pointeurs entre les deux

24

Principe d'indépendance logique

- Une application travaille sur une représentation logique des seules données qu'elle manipule
 - Le SGBD assure la correspondance avec les données réelles
 - Les programmes ne doivent pas être revus lors de la modification de données qu'ils n'utilisent pas.
 - Exemple : Les données d'un hôpital : médecins, malades, chambres, ...
 - Application n° 1 : Suivi des malades (Nom, N° SS, N° Chambre, Médecin)
 - Application n° 2 : Données d'un médecin (Nom, N°SS, N°Chambre, Thérapie)
 - Application n° 3 : Gestion du personnel : les médecins (Nom, Grade, Spécialité, Salaire)

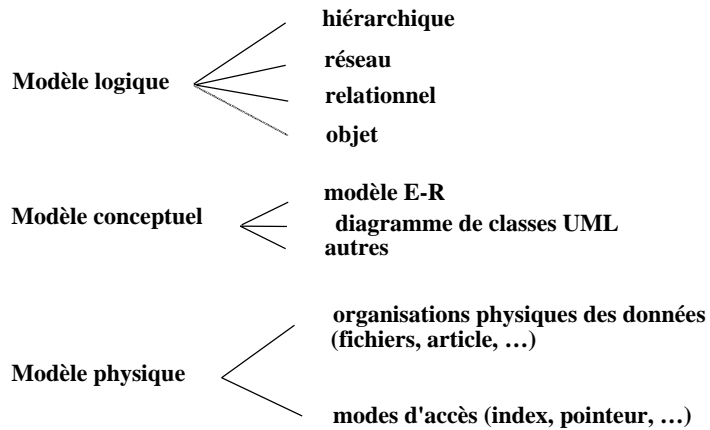
25

Niveaux de présentation (ANSI/SPARC)

- Niveau interne dépendant du SGBD
 - Organisation physique des données
- Niveau conceptuel indépendant du SGBD
 - Vue abstraite des données
 - On y distingue le niveau logique, dépendant du type de SGBD (hiérarchique, réseau, relationnel, objet) du niveau conceptuel
 - plus proche de l'utilisateur
- Niveau externe
 - Vue partielle des données, sous-ensemble du niveau conceptuel, pour une application ou un groupe d'utilisateurs

26

Modèle de données : outil de représentation des données



27

4. Le fonctionnement

Les langages

- LDD Langage de Définition des Données
- LMD Langage de Manipulation des Données
- Langage Hôte Langage de programmation enrichi d'instructions de manipulation de données
- L4G Langage de 4ème Génération :
 - intégration conviviale d'instructions du LMD + instructions procédurales + utilitaires de mise en forme de résultats + ...

29

Schéma de la base de données

- Structure des données définie
 - avec le LDD
 - ou avec un modèle

30

Interface BD-programme

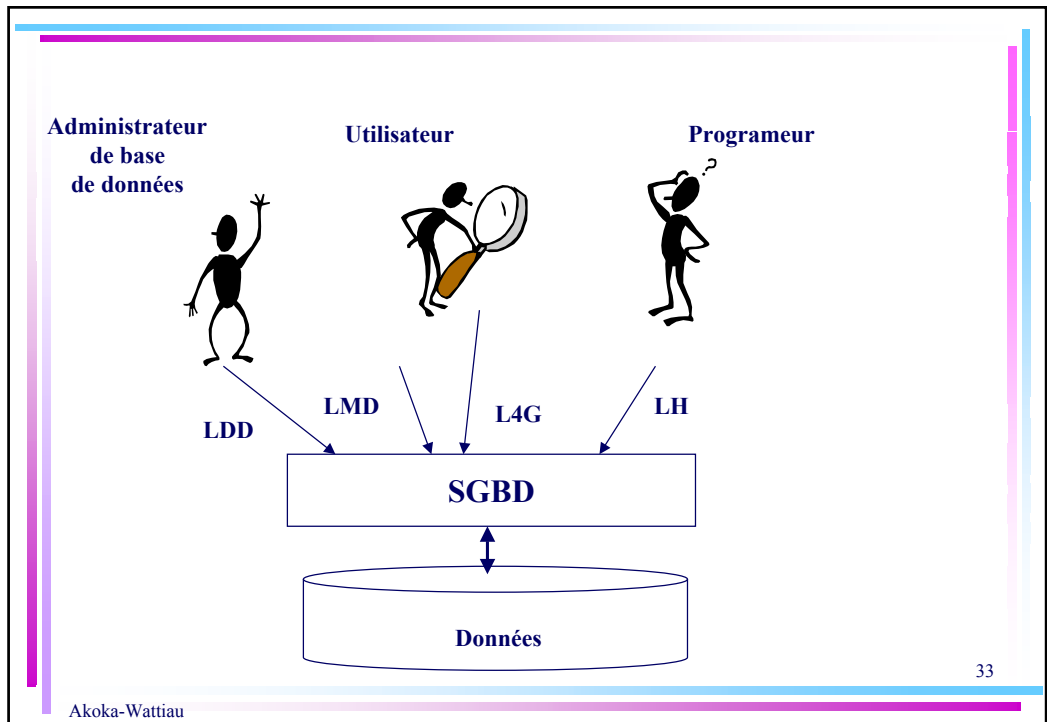
- Précompilateur
- A.P.I. (Application Programming Interface)
 - ex : JDBC Java DataBase Connectivity
 - une requête d'accès à la base = une chaîne de caractère, paramètre d'une instruction

31

Les acteurs

- Concepteur de Base de Données :
 - décide de l'information à introduire, de sa structure
- Administrateur de Bases de Données :
 - décide des autorisations d'accès, des modifications d'organisation physique, des procédures de sauvegarde,...
 - tout ce qui n'est pas automatisable et qui n'est pas du ressort de l'utilisateur
- Utilisateur :
 - accès à la B.D. pour interrogation et mise à jour
- Analyste système et programmeur
 - détermine les besoins des utilisateurs et programme les accès

32



5. Typologie des SGBD

Akoka - Wattiau

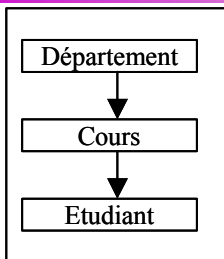
34

Les standards et quelques produits

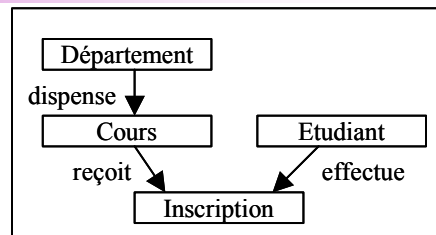
Relationnel-
objet

Modèles	Produits	Normes	Apparition
Hiérarchique	IMS/DL1		Années 1960
	System 2000		
Réseau	IDMS	CODASYL ANSI/SPARC	Années 1970
	IDS		
	Socrate		
	Total		
	MDBS		
Relationnel	Oracle	SQL	Années 1980
	DB2		
	Ingres		
	Informix		
	Sybase		
	SQL Server		
	mySQL		
Objet	Versant	ODMG	Années 1990 35
	GemStone		
	ObjectStore		

Akoka-Wattiau



hiérarchique



réseau

Département	code dépt	nom dépt
	M	Math
	I	Informatique

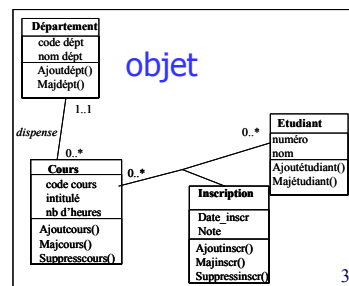
Cours	code cours	intitulé	nb d'heures	code dépt
	A12	algèbre	100	M
	G11	géométrie	50	M
	A15	analyse	120	M
	150	programmation	100	I

Etudiant	numéro	nom
	14	Jean
	18	Jacques
	25	Pierre
	27	Paul

relationnel

Inscription	code cours	numéro	date inscription	note
	A12	14	12/12/2002	12
	A12	18	10/10/2002	10
	G11	27	10/10/2002	18
	G11	25	10/11/2002	9
	G11	14	08/09/2002	18

La scolarité



36

Akoka-Wattiau

L 'organisation des BD et des SGBD

- BD Centralisée
 - données et SGBD sont sur le même ordinateur
- BD répartie ou distribuée
 - les données sont sur plusieurs ordinateurs
- BD parallèle
 - utilise le parallélisme des processeurs, de la mémoire vive et/ou de la mémoire secondaire
 - pour améliorer performance et fiabilité

37

Choix d 'une architecture (1)

- Architecture centralisée
 - mainframe / terminaux
- Architecture client/serveur
 - deux ordinateurs se répartissent les tâches de :
 - gestion des données
 - logique applicative
 - présentation des résultats
- Middleware :
 - logiciel de communication entre le programme client et le serveur

38

Choix d'une architecture (2)

- Architecture 3 tier
 - serveur de données
 - serveur d'applications
 - un poste client (« léger ») : navigateur web
- Architecture n tier
 - répartition de la logique applicative sur plusieurs serveurs

39

Entrepôts de données

- Base de données à vocation décisionnelle
- Même logiciel ou logiciel différent
- OLTP (On Line Transactional Processing) vs. OLAP (On Line Analytical Processing)
- Règles de mise à jour différentes

40

BD et ERP

- ERP - Enterprise Resource Planning
 - SAP, Oracle, PeopleSoft
- Une BD + modules applicatifs dédiés
 - à un secteur d 'activité
 - et/ou à un domaine de gestion
- Utilisent un moteur SGBD
- mais certains modules (ex : intégrité) peuvent être contournés et remplacés par du code spécifique