

Chapitre 01

Généralités sur les bases de données

Généralités sur les bases de données

- Qu'est-ce donc qu'une base de données ?
- Que peut-on attendre d'un système de gestion de bases de données ?
- Que peut-on faire avec une base de données ?

Des données ? Est ce important pour vous ?

- Des relevés de banques, de cartes de crédit
- Des carnets d'adresses
- La consommation de téléphone
- Des inscriptions à des clubs, associations,
- Des papiers utiles
- Des horaires et disponibilités de transport
- Des programmes de télé

Des données ? Est ce important pour vous ?

- Assurer l'accès aux données
- Assurer la sécurité de ces données
 - Confidentialité
 - Authentification
 - Signature digitale
 - Intégrité
- Le tout efficacement, rapidement, partout etc...

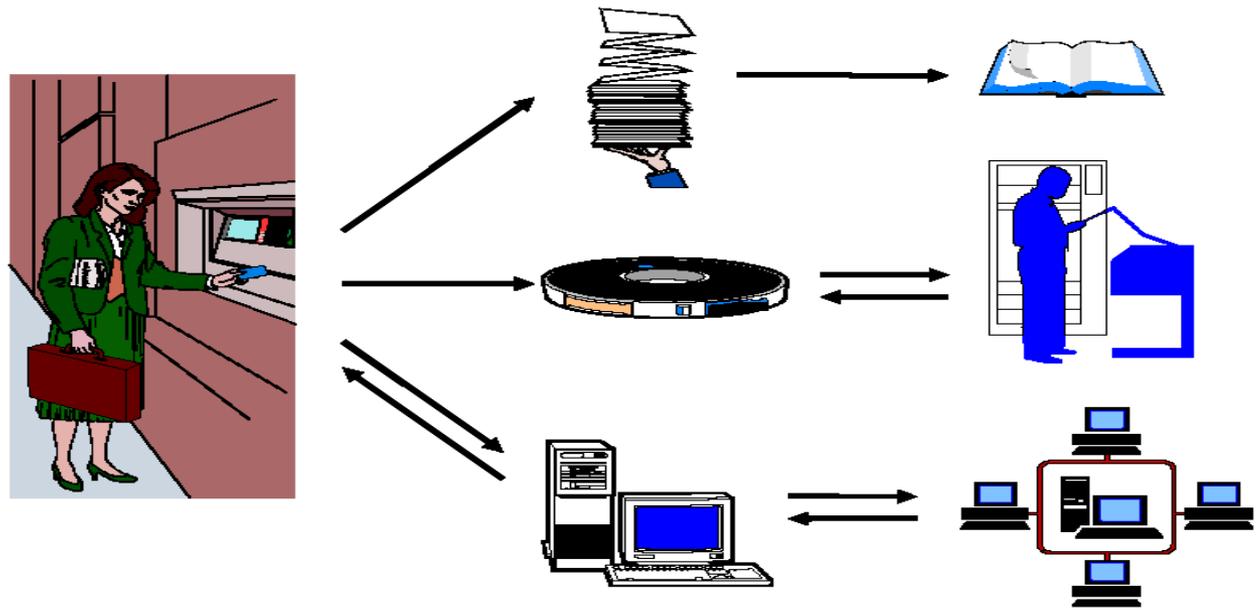
Des données ? Est ce important pour vous ?

- C'est important pour vous...
- C'est impératif pour les entreprises !

Le stockage / l'accès

Data and Files - Storage

◆ How is it stored?



Un point dans le temps

- 1950-1960
 - Des fichiers séquentiels, du ‘batch’
- 1960 – 1970
 - Le début des bases de données hiérarchiques
- 1970 – 1980
 - La naissance du modèle relationnel
- Début des années 90
 - Sql, l’aide à la décision
- Fin des années 90
 - Croissance du volume des données, Internet, modèle

Systemes de fichiers

Caractéristiques



Comptabilité



Chirurgie

Consultations



Psychiatrie



Problèmes

Format des fichiers



Dupont

Symptomes : y
Turlututu : sqj
Symptomes : y
Turlututu : sdd
Analyses : xxx

Dupond

Turlututusqjsk
Symptom: yyyy
Analyses xxxx

Turlututudhjsd
Analyses :xx



Duhpon

Symptomes : yy
Analyses : xxxx

Symptomes : yy

Duipont

Turlututu : sq

Symptomyyyy
Analysesxxxx

Turlututudhjsd



Caractéristiques

Plusieurs applications

- plusieurs formats
- plusieurs langages

Problèmes

- Difficultés de gestion

Redondance (données)



Dupont
Symptomes : y
Turlututu : sqj
Symptomes : y
Turlututu : sdd
Analyses : xxx

Dupond
Turlututu : sqj
Symptom: yyyy
Analyses xxxx
Turlututudhjsd
Analyses :xx



Duhpon
Symptomes : yy
Analyses : xxxx
Symptomes : yy

Duipont
Turlututu : sqj
Symptomyyyy
Analysesxxxx
Turlututudhjsd



Caractéristiques

Plusieurs applications

- plusieurs formats
- plusieurs langages

Redondance de données

Problèmes

- Difficultés de gestion
- Incohérence des données

Interrogations



Caractéristiques

Plusieurs applications

- plusieurs formats
- plusieurs langages

Redondance de données

Pas de facilité d'interrogation

- Question ⇒ développement

Problèmes

- Difficultés de gestion
- Incohérence des données
- Coûts élevés
- Maintenance difficile



Pannes ???



Caractéristiques

Plusieurs applications

- plusieurs formats
- plusieurs langages

Redondance de données

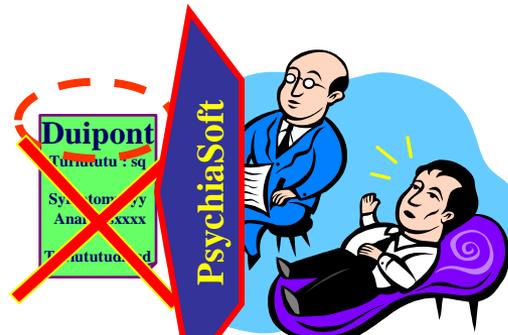
Pas de facilité d'interrogation

- Question ⇒ développement

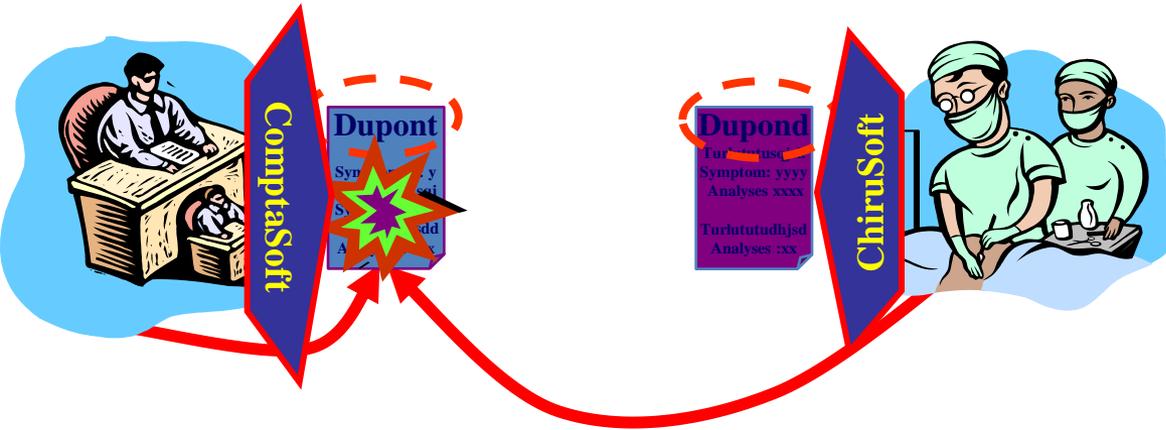
Redondance de code

Problèmes

- Difficultés de gestion
- Incohérence des données
- Coûts élevés
- Maintenance difficile
- Gestion de pannes ???



Partage de données



Caractéristiques

Plusieurs applications

- plusieurs formats
- plusieurs langages

Redondance de données

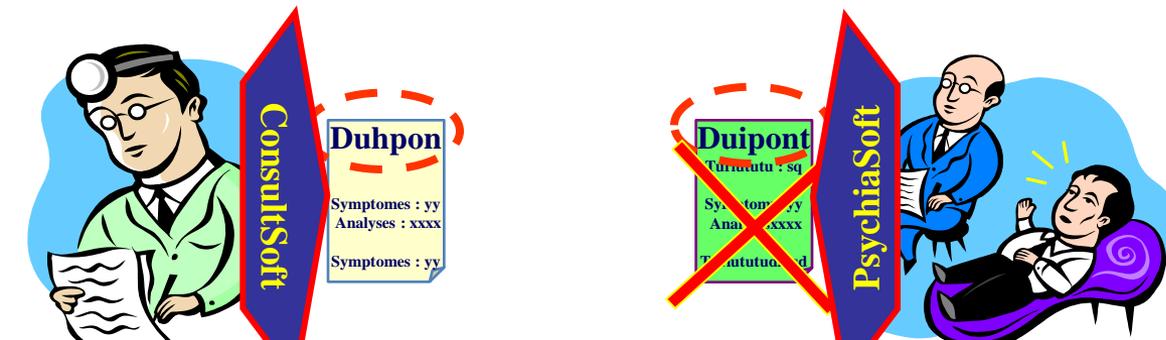
Pas de facilité d'interrogation

- Question ⇒ développement

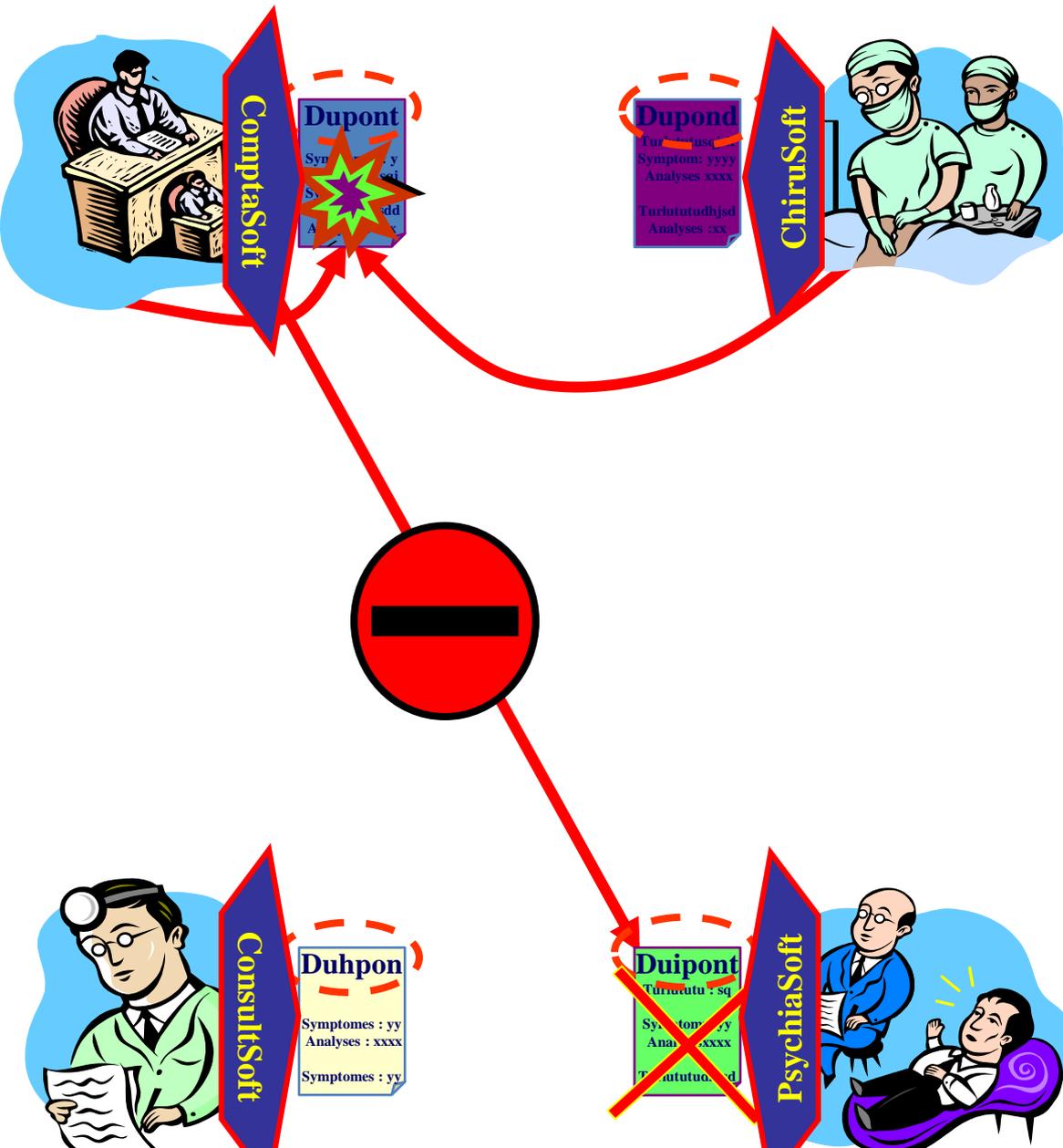
Redondance de code

Problèmes

- Difficultés de gestion
- Incohérence des données
- Coûts élevés
- Maintenance difficile
- Gestion de pannes ???
- Partage des données ???



Confidentialité



Caractéristiques

Plusieurs applications

- plusieurs formats
- plusieurs langages

Redondance de données

Pas de facilité d'interrogation

- Question ⇒ développement

Redondance de code

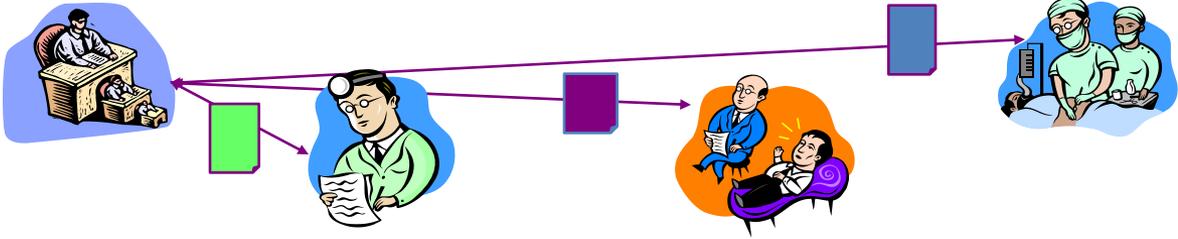
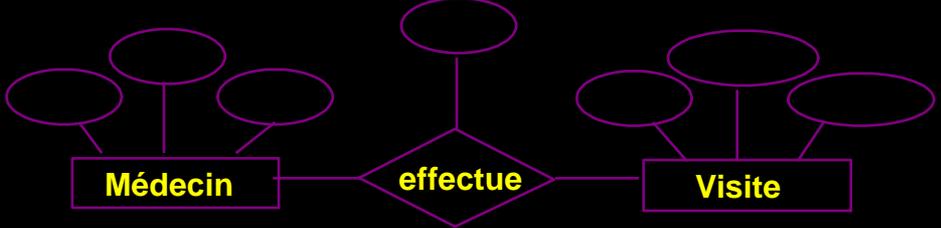
Problèmes

- Difficultés de gestion
- Incohérence des données
- Coûts élevés
- Maintenance difficile
- Gestion de pannes ???
- Partage des données ???
- Confidentialité ???

L'approche "Bases de données"

- **Modélisation des données**
 - Eliminer la **redondance** de données
 - **Centraliser** et **organiser** correctement les données
 - Plusieurs niveaux de modélisation
 - Outils de conception
- **Logiciel «Système de Gestion de Bases de Données»**
 - **Factorisation** des modules de contrôle des applications
 - Interrogation, cohérence, partage, gestion de pannes, etc...
 - Administration facilitées des données

Modélisation du réel

| | | | | | |
|--------------------------|---|--|--------------------|--------------|------------|
| <p>Réel</p> |  | | | | |
| <p>Modèle conceptuel</p> | <ul style="list-style-type: none"> • Indépendant du modèle de données • Indépendant du SGBD |  | | | |
| <p>Modèle logique</p> | <ul style="list-style-type: none"> • Dépendant du modèle de données • Indépendant du SGBD | <p>CodasyI</p> | <p>Relationnel</p> | <p>Objet</p> | <p>XML</p> |
| <p>Modèle Physique</p> | <ul style="list-style-type: none"> • Dépendant du modèle de données • Dépendant du SGBD | <ul style="list-style-type: none"> • Organisation physique des données • Structures de stockage des données • Structures accélératrices (index) | | | |

Notions de bases

Définition intuitive d'une BD (1)

- Définition intuitive : on peut considérer une **Base de Données** (BD) comme une grande quantité de données (ou ensemble d'informations), centralisées ou non, servant pour les besoins d'une ou plusieurs applications, interrogeables et modifiables par un groupe d'utilisateurs travaillant en parallèle.
- Exemples d'application
 - Système Socrate : SNCF
 - Annuaire électronique
 - Catalogue électronique d'une bibliothèque

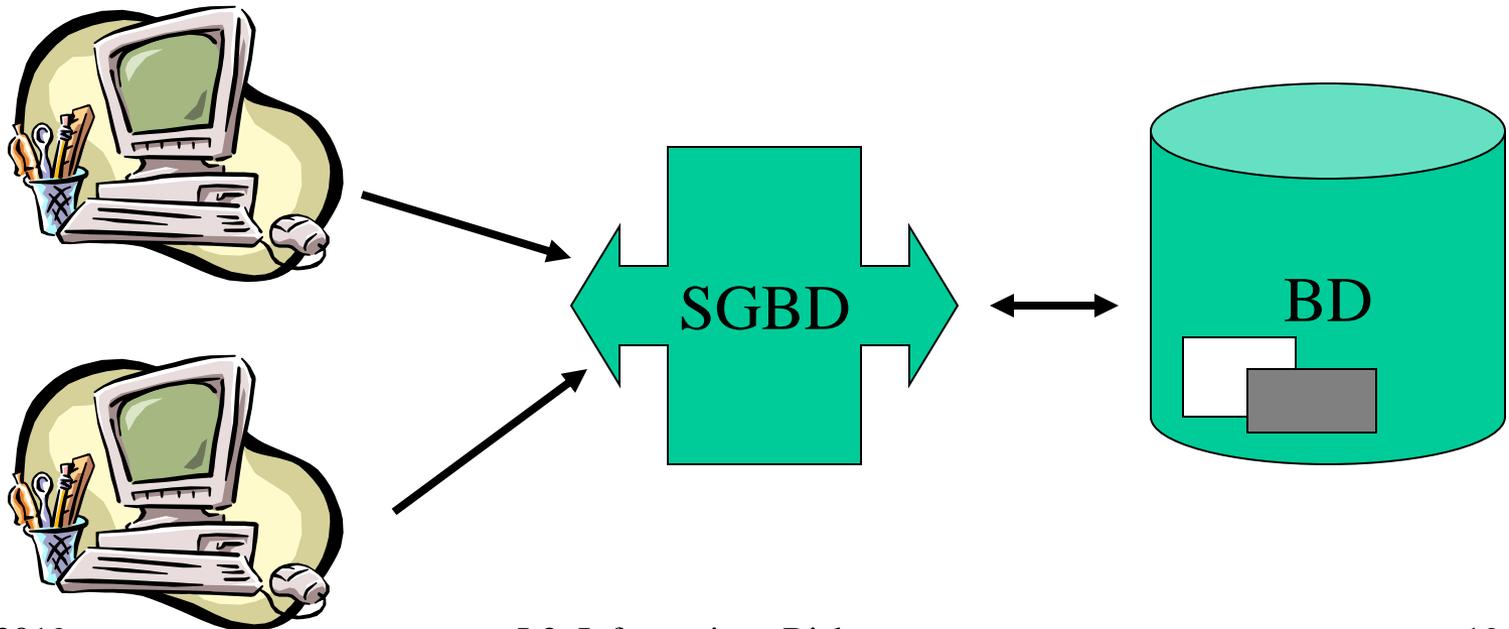
Notions de bases

Définition intuitive d'une BD (2)

- **Définition** ; une base de données est un ensemble structuré de données (1) enregistrées sur des supports accessibles par l'ordinateur (2) pour satisfaire simultanément plusieurs utilisateurs (3) de manière sélective (4) en un temps opportun (5).
 - (1) : Organisation et description de données
 - (2) : Stockage sur disque
 - (3) : Partage des données
 - (4) : Confidentialité
 - (5) : Performance

SGBD (1)

Définition : Le logiciel qui permet d'interagir avec une BD est Système de Gestion de Base de Données (SGBD)

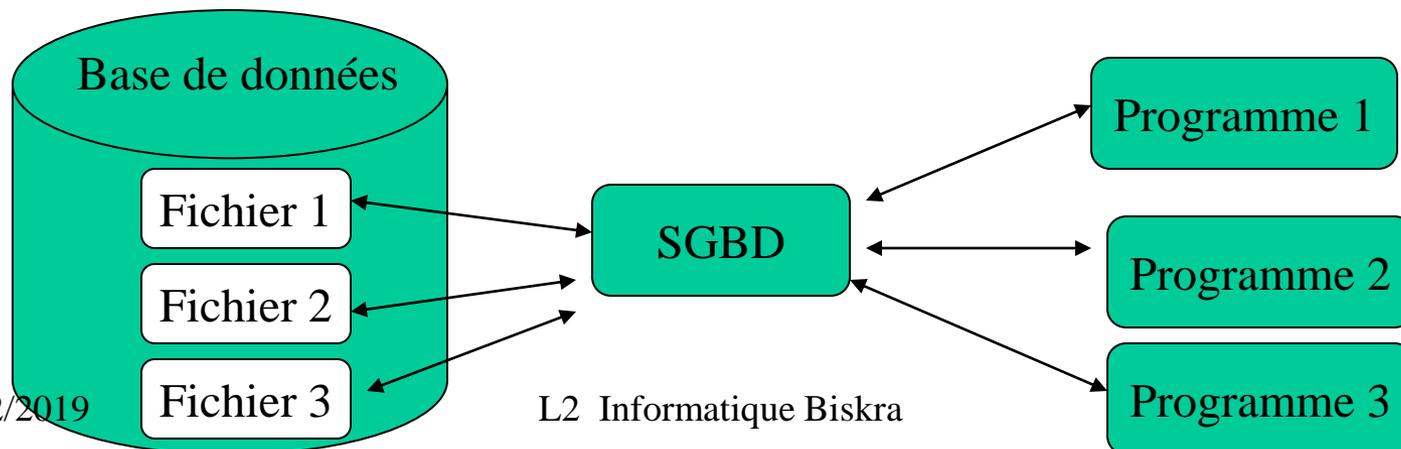


SGBD (2)

Un SGBD est un intermédiaire entre les utilisateurs et les fichiers physiques

Un SGBD facilite

- la gestion de données, avec une représentation simple sous forme de table par exemple
- la manipulation de données. On peut insérer, modifier les données et les structures sans modifier les programmes qui manipulent la base de données



Objectifs des SGBD (1)

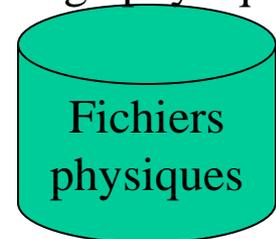
Faciliter la représentation et la description de données

- Indépendance physique (1) : Plus besoin de travailler directement sur les fichiers physiques (tels qu'ils sont enregistrés sur disque). Un SGBD nous permet de décrire les données et les liens entre elles d'une façon logique sans se soucier du comment cela va se faire physiquement dans les fichiers. On parle alors d'image logique de la base de données, (ou aussi description logique ou conceptuelle ou encore de schéma logique). Ce schéma est décrit dans un modèle de données par exemple le modèles de tables, appelé le modèle relationnel.

Image logique



Image physique



Objectifs des SGBD (2)

- Indépendance physique (2) : La manipulation des données doit être facilitée en travaillant directement sur le schéma logique. On peut insérer, supprimer, modifier des données directement sur l'image logique. Le SGBD va s'occuper de faire le travail sur les fichiers physiques.
- Indépendance logique : Un même ensemble de données peut être vu différemment par des utilisateurs différents. Toutes ces visions personnelles des données doivent être intégrées dans une vision globale.
- Manipulations des données par des non informaticiens. Il faut pouvoir accéder aux données sans savoir programmer ce qui signifie des langages « quasi naturels ».
- Efficacité des accès aux données : Ces langages doivent permettre d'obtenir des réponses aux interrogations en un temps « raisonnable ». Il doivent donc être optimisés et, entre autres, il faut un mécanisme permettant de minimiser le nombre d'accès disques. Tout ceci, bien sur, de façon complètement transparente pour l'utilisateur.

Objectifs des SGBD (3)

- Administration centralisée des données : Des visions différentes des données (entre autres) se résolvent plus facilement si les données sont administrées de façon centralisée.
- Cohérence des données. Les données sont soumises à un certain nombre de contrainte d'intégrité qui définissent un état cohérent de la base. Elles doivent pouvoir être exprimées simplement et vérifiées automatiquement à chaque insertion, modification ou suppression de données, par exemple :
 - l'âge d'une personne supérieur à zéro
 - Salaire supérieur à zéro
 - Etc

Dés que l'on essaie de saisir une valeur qui ne respecte pas cette contrainte, le SGBD le refuse.

Objectifs des SGBD (4)

- Non redondance des données : Afin d'éviter les problèmes lors des mises à jour, chaque donnée ne doit être présente qu'une seule fois dans la base.
- Partageabilité des données : Il s'agit de permettre à plusieurs utilisateurs d'accéder aux mêmes données au même moment. Si ce problème est simple à résoudre quand il s'agit uniquement d'interrogations et quand on est dans un contexte mono-utilisateur, cela n'est plus le cas quand il s'agit de modifications dans un contexte multi-utilisateurs. Il s'agit alors de pouvoir :
 - Permettre à deux (ou plus) utilisateurs de modifier la même donnée « en même temps »;
 - Assurer un résultat d'interrogation cohérent pour un utilisateur consultant une table pendant qu'un autre la modifie.

Objectifs des SGBD (5)

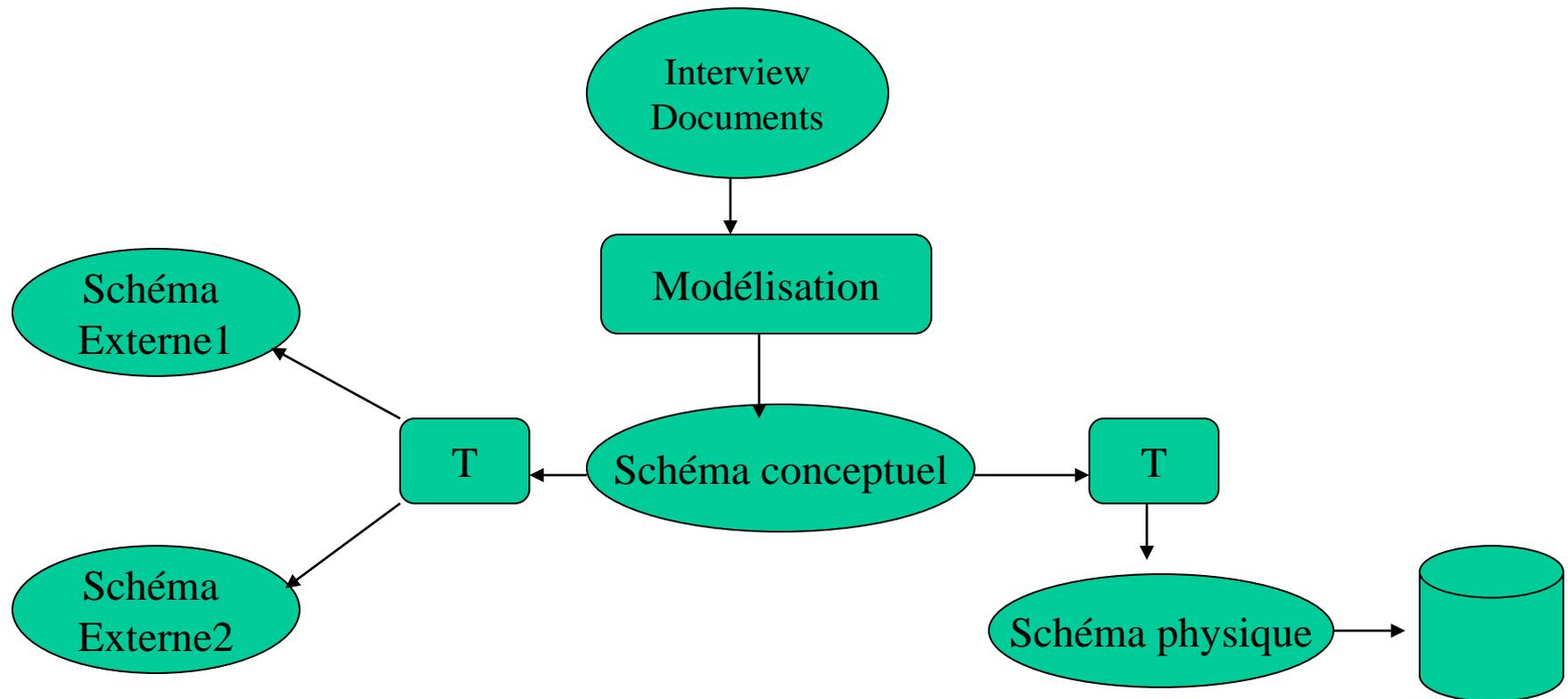
- Sécurité des données. Les données doivent pouvoir être protégées contre les accès non autorisés. Pour cela, il faut pouvoir associer à chaque utilisateur des droits d'accès aux données.
- Résistance aux pannes : Que se passe-t-il si une panne survient au milieu d'une modification, si certains fichiers contenant les données deviennent illisibles? Les pannes, bien qu'étant assez rares, se produisent quand même de temps en temps. Il faut pouvoir, lorsque l'une d'elles arrive, récupérer une base dans un état « sain ». Ainsi, après une panne intervenant au milieu d'une modification deux solutions sont possibles : soit récupérer les données dans l'état dans lequel elles étaient avant la modification, soit terminer l'opération interrompue.

Trois Fonctions d'un SGBD

- Description des données : codification structuration, grâce à un Langage de Description de Données (LDD)
- Manipulation et restitution des données (insertion, mise à jour, interrogation)
 - Mise en œuvre à l'aide d'un Langage de Manipulation de Données (LMD)
 - S.Q.L. (Structures Query Langage) : Langage standard
- Contrôle (partage, intégrité, confidentialité, sécurité)

Définition et description des données

3 niveaux de description



Définition et description des données niveau logique (conceptuel)

- Permet la description
 - Des objets : exemple OUVRAGES, ETUDIANTS
 - Des propriétés des objets (attributs) : exemple Titre de OUVRAGES
 - Des liens entre les objets : un OUVRAGE peut être emprunté par un ETUDIANT
 - Des contraintes : le nombre d'exemplaires d'un OUVRAGE est supérieur à zéro
- Cette description est faite selon un modèle de données.
- Un modèle de données est un ensemble de concepts permettant de décrire la structure d'une base de données. La plupart des modèles de données incluent des opérations permettant de mettre à jour et questionner la base. Le modèle de données le plus utilisé est le modèle relationnel
- Cette description va donner lieu à un schéma de base de données. Un schéma de base de données se compose d'une description des données et de leurs relations ainsi que d'un ensemble de contraintes d'intégrité.

Définition et description des données niveau physique

- Description informatique des données et de leur organisation : en terme de fichiers, d'index, de méthodes d'accès, ...
- Passage du modèle logique au modèle physique tend à être assisté par le SGBD : transparent et/ou semi-automatique
- Objectifs : optimiser les performances

Définition et description des données niveau externe

- Description des données vues par un utilisateur (ou un groupe d'utilisateurs)
 - Objectifs : simplification, confidentialité
 - Exemple : OUVRAGES édité par des éditeurs français

Manipulation et restitution des données

- Afin de réaliser les opérations suivantes
 - Insertion : saisir des données
 - Supprimer
 - Modifier
 - Interroger : rechercher des données via des requêtes

La manipulation des données est mise en œuvre à l'aide d'un Langage de manipulation de Données (LMD). SQL (Structured Query Language) est le langage standard de manipulation de BD

Contrôles réalisés par le SGBD

- Partage de données : accès à la même information par plusieurs utilisateurs en même temps. Le SGBD inclut un mécanisme de contrôle de la concurrence basé sur des techniques de verrouillage des données (pour éviter par exemple qu'on puisse lire une information qu'on est en train de mettre à jour)
- Intégrité des données grâce à la définition de contraintes sur les données. Le SGBD veille à ce que toutes les contraintes soient vérifiées à chaque insertion, suppression, ou modification d'une donnée.
- Confidentialité : plusieurs utilisateurs peuvent utiliser en même temps une base de données, se pose le problème de la confidentialité des données. Des droits doivent être gérés sur les données, droits de lecture, mise à jour, création; ... qui permettent d'affiner.
- Sécurité : une base de données est souvent vitale dans le fonctionnement d'une organisation, et il n'est pas tolérable qu'une panne puisse remettre en cause son fonctionnement de manière durable. Les SGBD fournissent des mécanismes pour assurer cette sécurité.

Utilisateurs des SGBD

Les différents rôles que doivent jouer un individu ou un groupe d'individus pour concevoir, créer, mettre en œuvre et exploiter une base de données.

- ***L'administrateur de la base de données*** : Il est chargé de décrire les entités de la base de données et indiquer les liaisons existant entre ces entités, ceci au moyen du DDL offert par le SGBD.
- ***Le programmeur d'application*** : Il est chargé d'élaborer les programmes pour exploiter la base de données en fonction de la description qui a été faite par l'administrateur d'application. Le programmeur d'application utilise le LMD offert par le SGBD ainsi que d'autres sous-programmes conservés généralement dans une librairie (i.e. bibliothèque de sous-programmes).

Utilisateurs des SGBD

- ***L'utilisateur*** : Il s'agit de caractériser ici la personne qui se sert simplement de la base de données et qu'on appelle couramment *l'utilisateur final (End User en anglais)*.

- ***Exemple*** :

Dans une agence de réservation de billets d'avion, la personne qui tape sur son terminal quelques commandes pour effectuer une réservation est une utilisatrice au même titre qu'un chef d'entreprise qui lui aussi demande de temps en temps à une base de données de son entreprise un certain nombre d'informations reflétant l'état de son entreprise (produits non vendus, commandes en attente, etc.).

Modèles de SGBD

- Quelques modèles logiques :
 - Modèle hiérarchique
 - Modèle réseau
 - Modèle relationnel
 - Modèle objet
- Quelques SGBD (relationnels du marché)
 - Micro : ACCESS, Paradox, Dbase, PostgreSQL, MySQL, ...
 - Gros système : DB2, ORACLE, SYBASE, ...

Architecture des SGBD