



Université Djilali Liabès- Sidi Bel Abbès  
3ème Année Licence Mathématiques Appliquées

# Base de données avancée et système répartis

---

Mohammed Fethi KHALFI

2023-2024

Fethi.Khalfi@yahoo.fr

# Introduction en systèmes d'information et bases de données

---

# Présentation

---

- Contenu du cours :
  - Techniques pour l'analyse et la modélisation de données :
    - Modele Entité/Association ;
    - Modele relationnel ;
    - Algèbre relationnelle.
  - Langage de définition/manipulation des données : **SQL**.
- Objectifs :
  - Comprendre l'utilité et le fonctionnement des modèles conceptuels et relationnels.
  - Savoir utiliser un SGBD et connaître le langage SQL.

# Définition de SI

---

Système d'information: Un ensemble organisé de ressources:

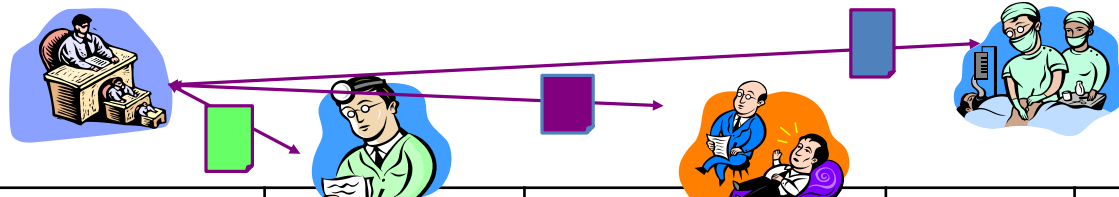
- matériels,
- logiciels,
- personnel,
- données
- et procédures.

qui permet :

- de regrouper,
- de classifier,
- de traiter,
- de diffuser de l'information sur un environnement donné.

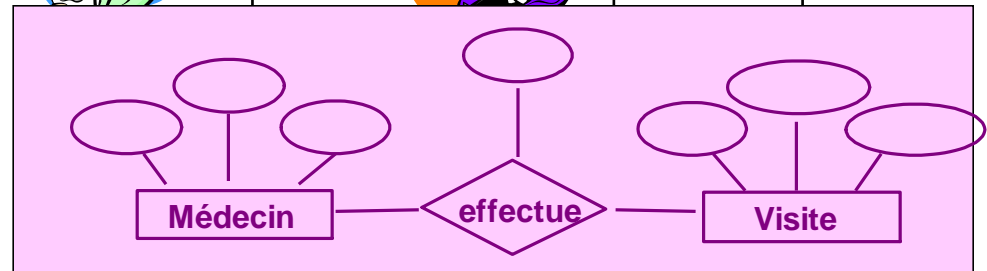
# Modélisation à plusieurs niveaux

Réel



Modèle  
conceptuel

- ◆ Indépendant du modèle de données
- ◆ Indépendant du SGBD



Modèle  
logique

- ◆ Dépendant du modèle de données
- ◆ Indépendant du SGBD

Relationnel

Objet

XML

Modèle  
Physique

- ◆ Dépendant du modèle de données
- ◆ Dépendant du SGBD



# Partons d'un exemple

## Votre recherche rapide

[train](#)
[vol](#)
[hotel](#)
[voiture](#)

### Composez votre voyage Alacarte® !

☒ Train seul
 ☐ Train & Hôtel

☐ Train & Voiture
 ☐ Train & Hôtel & Voiture

Au départ de  Départ (JJ/MM/AAAA) à partir de  12h ▼

A destination de  Retour (JJ/MM/AAAA) à partir de  15h ▼

Adulte ☐ 1e classe ☒ 2e classe

1 ▼

[Réserver votre billet](#) [Consulter les horaires](#)

12-25, Senior, Famille, Abonnés, S'Miles...

[France](#)
[Train](#)
[vol](#)
[hôtel](#)
[voiture](#)
[loisirs +](#)
[promos](#)

[Recherche](#)
[Résultats](#)

28/09/2006, 20:14

**Il n'y a pas d'autre trajet direct ou avec une correspondance, cliquez sur "Voir les trains suivants" pour afficher les trajets à 2 correspondances.**

**ALLERS disponibles pour le Samedi 30/09**

Trier par : | [Heure de départ](#) | [Heure d'arrivée](#)

| Gares de départ et d'arrivée | Horaires | Train(s) | Durée : 02h42                     |
|------------------------------|----------|----------|-----------------------------------|
| PARIS NORD                   | 18h18    | S051     | <a href="#">Choisir cet aller</a> |
| LONDON WATERLOO INT          | 18h58    |          |                                   |
| Gares de départ et d'arrivée | Horaires | Train(s) | Durée : 02h36                     |
| PARIS NORD                   | 18h19    | S055     | <a href="#">Choisir cet aller</a> |
| LONDON WATERLOO INT          | 20h54    |          |                                   |
| Gares de départ et d'arrivée | Horaires | Train(s) | Durée : 02h46                     |
| PARIS NORD                   | 20h43    | S051     | <a href="#">Choisir cet aller</a> |
| LONDON WATERLOO INT          | 22h28    |          |                                   |

### 3 Vos informations bancaires pour le paiement en ligne

Pour commander en toute confiance, les engagements de voyages-sncf.com sont :

- des pages sécurisées pour payer l'esprit tranquille
- l'envoi d'un email de confirmation récapitulant toutes les informations de votre commande
- un service client à votre disposition pour toutes questions à la suite de votre commande

Les symboles indiquent que votre transaction est sécurisée, vous pouvez remplir votre formulaire en toute confiance.

► Sélectionnez la carte de paiement de votre choix:

☐
☐
☐

► Saisissez votre numéro de carte bancaire:

Identifiant commerçant: 055204944722232

N° de carte :  Expire fin :  /

Code de sécurité :

[Voir les trains suivants](#)

[En savoir +](#)

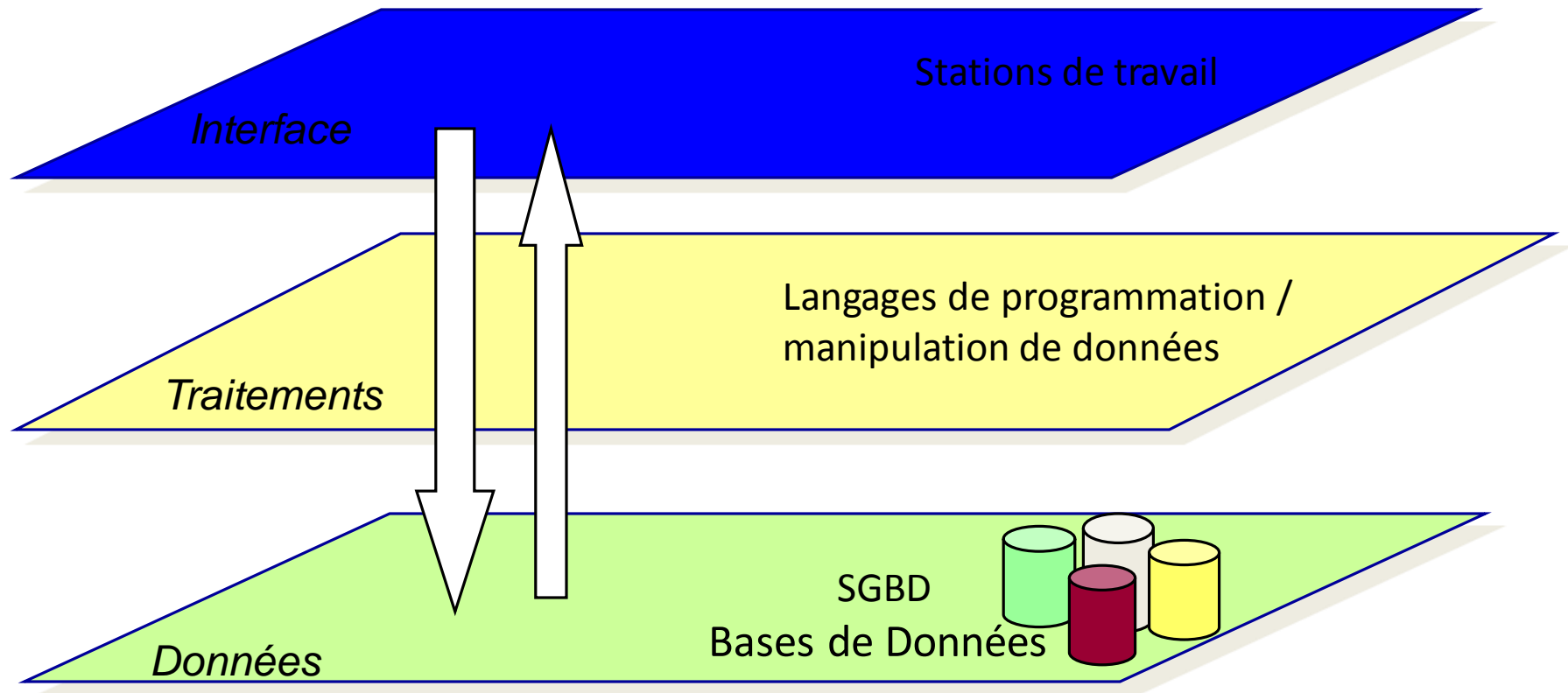
[En savoir +](#)

[En savoir +](#)

[En savoir +](#)

[Confirmer et payer votre commande](#)

# Les trois dimensions d'une application



# Exemples classiques d'applications BD

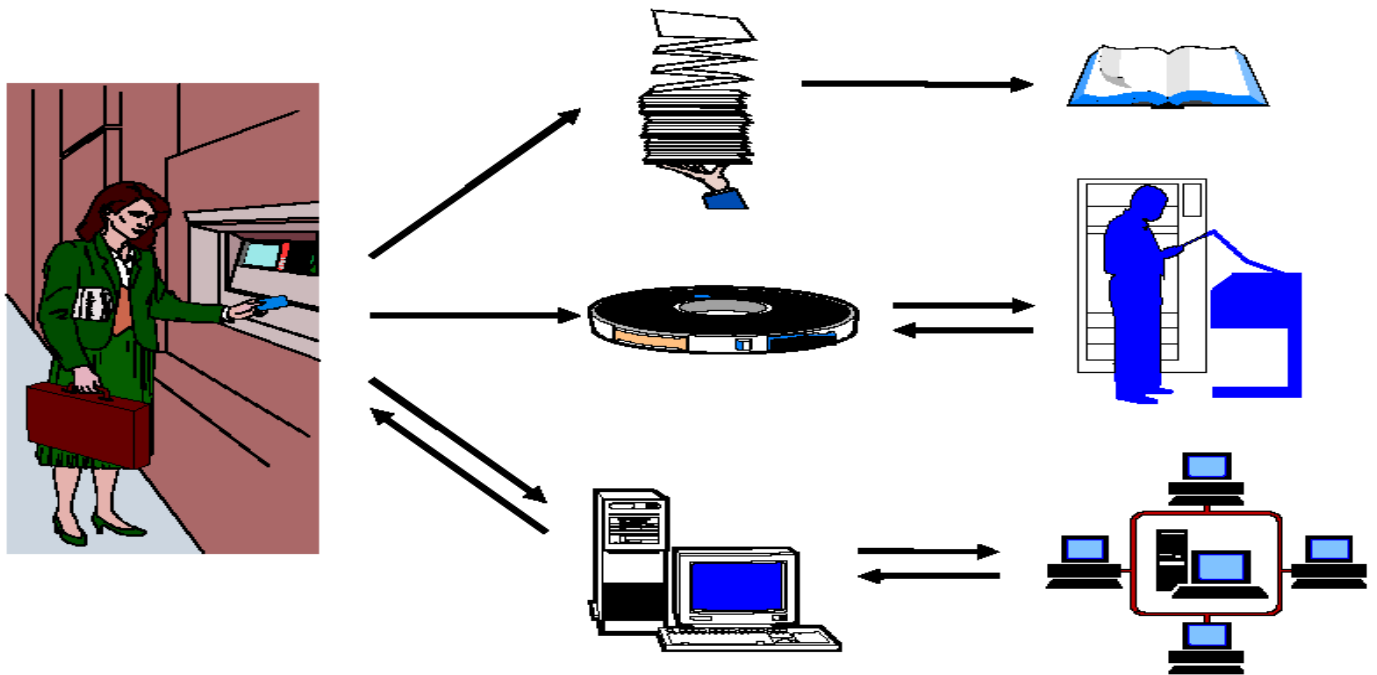
- Des relevés de banques clients, de cartes de crédit
- La consommation de téléphone
- Des inscriptions à des clubs, associations,
- Des horaires et disponibilités de transport
- Gestion des personnels, étudiants, cours, inscriptions, ... d'une université
- Système de réservation
- Gestion des commandes à Amazon.com



# Le stockage / l'accès

## Data and Files - Storage

◆ How is it stored?



# Des données ? Est ce important pour vous ?

- Assurer l'accès aux données
- Assurer la sécurité de ces données
  - Confidentialité
  - Authentification
  - Intégrité


# Un point dans le temps

- 1950-1960
  - Des fichiers séquentiels, du ‘batch’

Les enregistrements logiques sont accessibles les uns après les autres. Pour accéder à un enregistrement fixé, il faut utiliser un algorithme de recherche séquentielle (pour accéder à l'enregistrement de rang  $n$ , il faut parcourir les  $n-1$  précédents enregistrements).

Ci-dessous un fichier de caractères (un enregistrement logique = un caractère), ce fichier contient 13 enregistrements de taille fixe numérotés de 0 à 12) :

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |   |
|---|---|---|---|---|---|---|---|---|---|----|----|----|---|
| B | Y | B | B | L | K | C | N | Z | S | Y  | X  | S  | # |

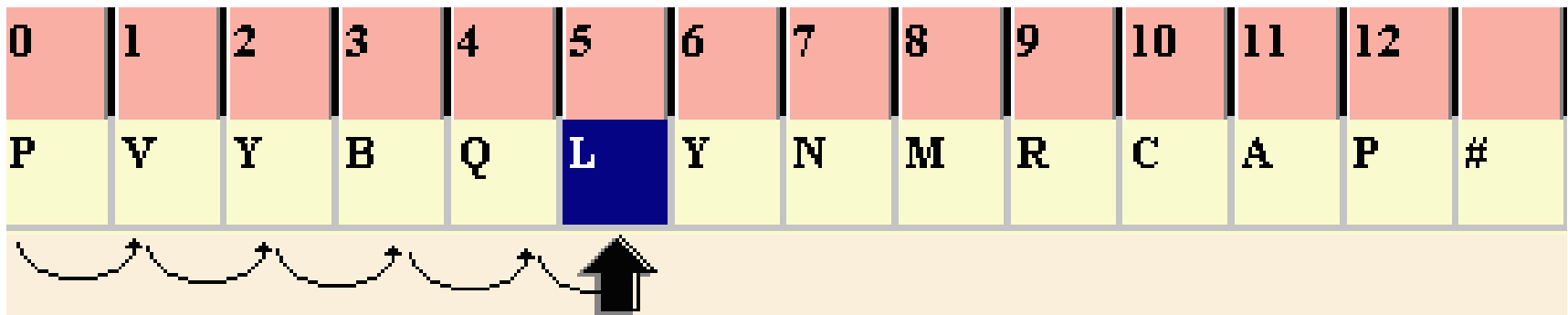


# Un point dans le temps

- 1950-1960
  - Des fichiers séquentiels, du ‘batch’

Pour lire le caractère 'L' de l'enregistrement que nous avons numéroté 5, il faut avoir lu les enregistrements P, V, Y, B et Q.

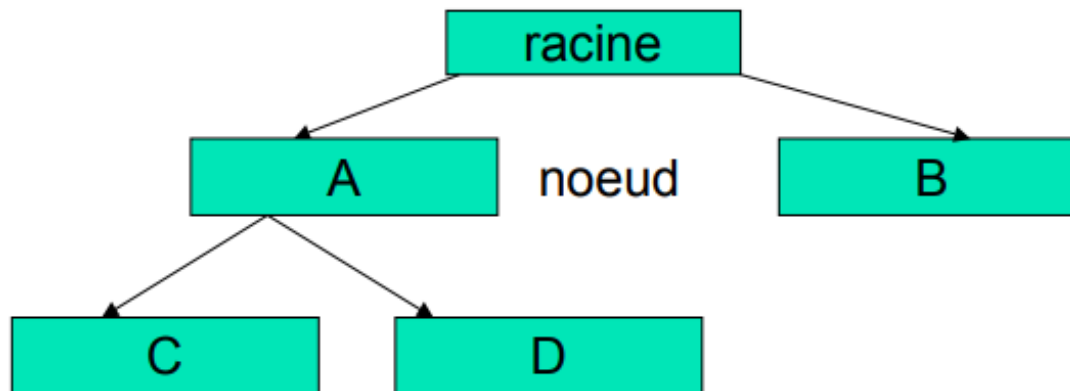
Dans ce cas de figure le pointeur est sur l'enregistrement contenant le caractère 'L' prêt à être lu :



# Un point dans le temps

- 1960 – 1970
  - Le début des bases de données hiérarchiques

Une BDH est une base de données dont le système de gestion lie les enregistrements dans une structure arborescente où chaque enregistrement n'a qu'un seul possesseur.

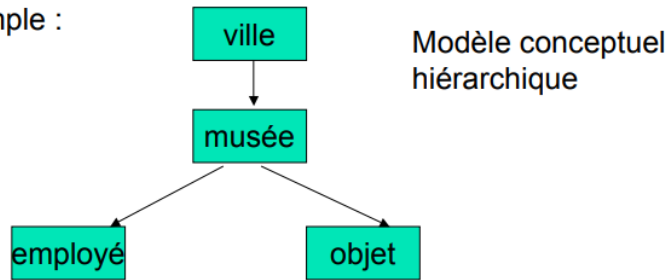


- Pour un nœud donné, un seul nœud père.
- Chaque nœud est une famille d'objets.

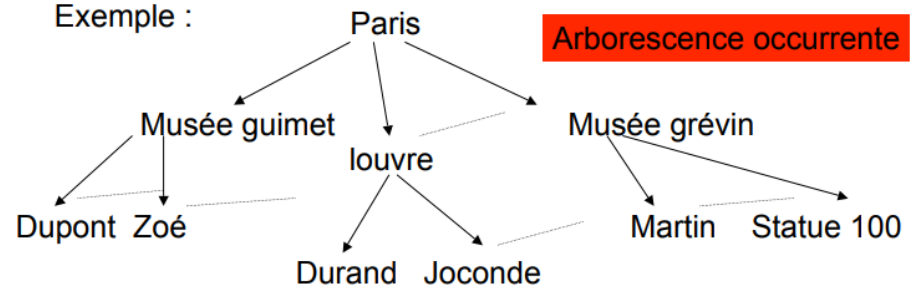
# Un point dans le temps

- 1960 – 1970
  - Le début des bases de données hiérarchiques

Exemple :



Exemple :



**Inconvénients** : la représentation arborescente ne décrit pas toujours la réalité.

**Exemple :**

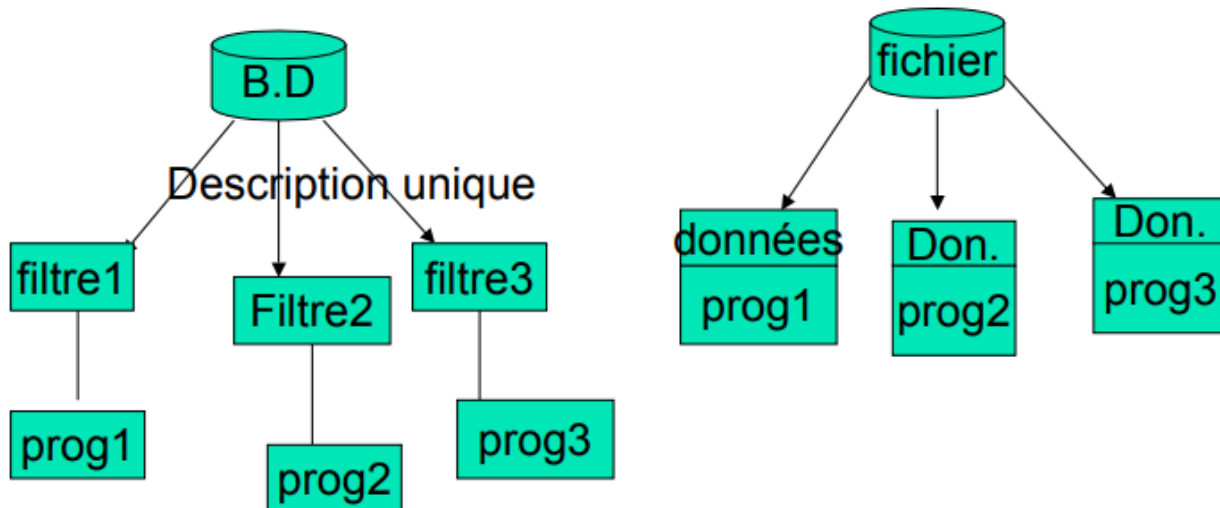
- **Modèle hiérarchique** : un module d'enseignement n'appartient qu'à un seul diplôme
- **Modèle « réel »** : un module d'enseignement peut être validé dans plusieurs « diplômes ».

# Un point dans le temps

- 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 multi tiers

# Les limites à l'utilisation des fichiers (1)

- Pour **des applications nouvelles**, l'utilisateur devra obligatoirement écrire de nouveaux programmes et il pourra être amené à créer de nouveaux fichiers qui contiendront peut-être des informations déjà présentes dans d'autres fichiers.
- Toute **modification** de la structure des enregistrements (ajout d'un champ par exemple) entraîne **la réécriture** de tous les programmes qui manipulent ces fichiers.





## Les limites à l'utilisation des fichiers (2)

- Les utilisateurs ne veulent plus de systèmes d'information constitués d'un ensemble de programmes inflexibles et de données inaccessibles à tout **non spécialiste**;
- ils souhaitent des systèmes d'informations globaux, cohérents, directement accessibles (sans qu'ils aient besoin soit d'écrire des programmes soit de demander à un programmeur de les écrire pour eux) et des réponses immédiates aux questions qu'ils posent.

# 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 SNCF
  - Annuaire électronique
  - Catalogue électronique d'une bibliothèque

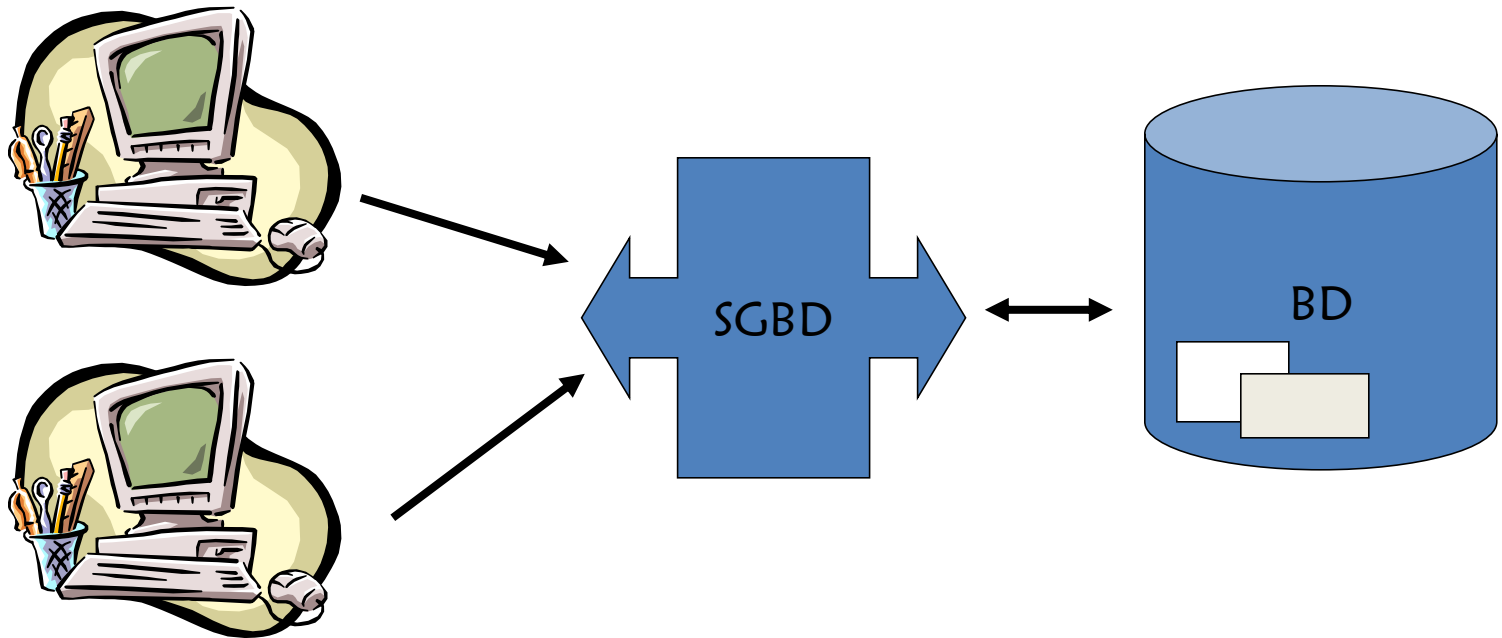
# Notions de bases

## 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 convenable (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**)

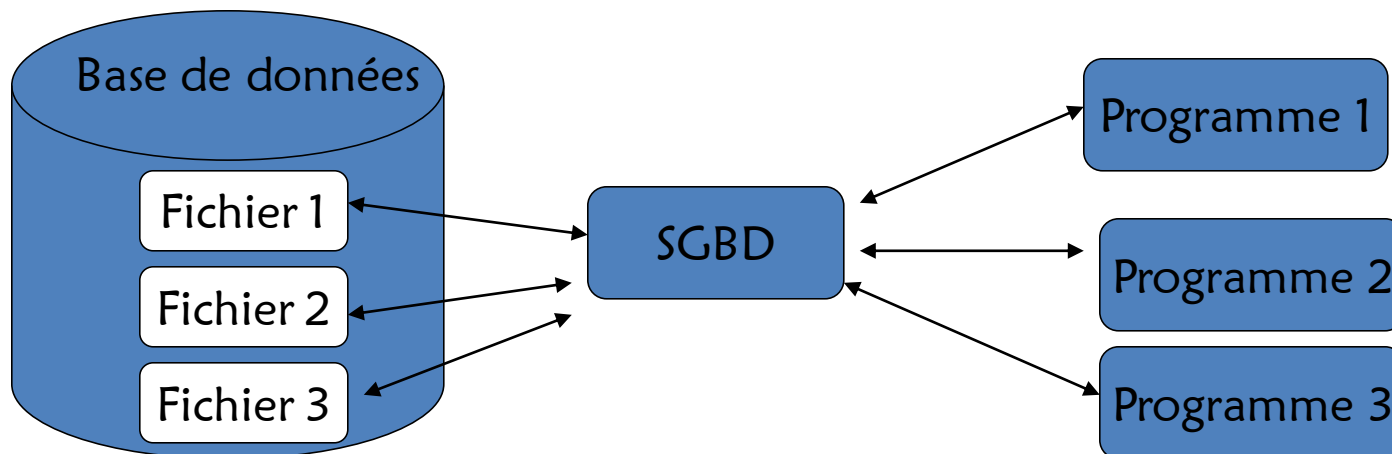


# SGBD (2)

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

Un SGBD facilite

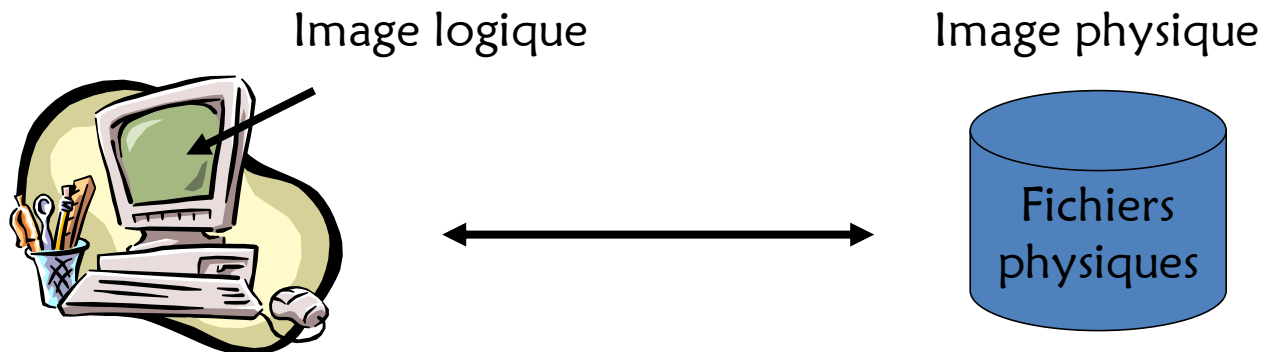
- la gestion de données, avec une représentation intuitive simple sous forme de table par exemple
- la manipulation de données. On peut insérer, modifier les données et les structures sans modifier les programmes qui manipulent la base de données



# 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 permet de décrire les données et les liens entre elles d'une façon logique. On parle alors d'image logique de la base de données, (ou aussi schéma logique ou conceptuelle). Ce schéma est décrit dans un modèle de données par exemple le modèles de tables, appelé **le modèle relationnel**.



## 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.
- 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 sûr, de façon complètement transparente pour l'utilisateur.

# Objectifs des SGBD (3)

- 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
  - -Il ne faut jamais donner la même place dans le même train à 2 clients

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.

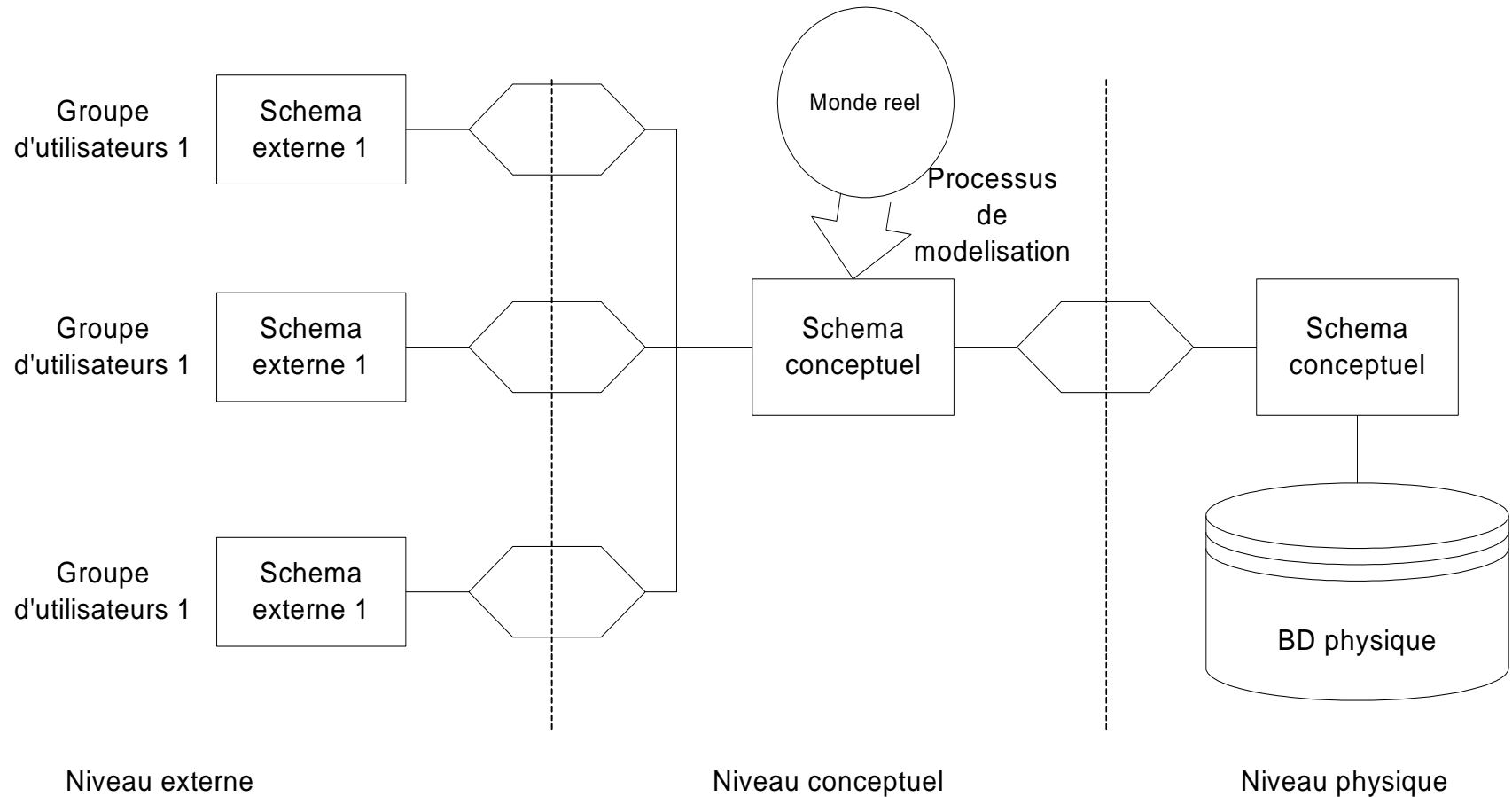
## Objectifs des SGBD (5)

- 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)

# Niveaux de représentation d'une BD



# 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

# 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.

Exemple : Il ne faut pas qu'une action faite pour un utilisateur (par exemple, l'enregistrement d'une réservation) soit perdue du fait d'une autre action faite simultanément pour un autre utilisateur (réservation de la même place).



## Contrôles réalisés par le SGBD

- Confidentialité : Toute information doit pouvoir être protégée contre l'accès par des utilisateurs non autorisés
- 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.

Exemple : Interdire aux clients de modifier les numéros des trains ou les horaires ou leur réservation.

# Contrôles réalisés par le SGBD

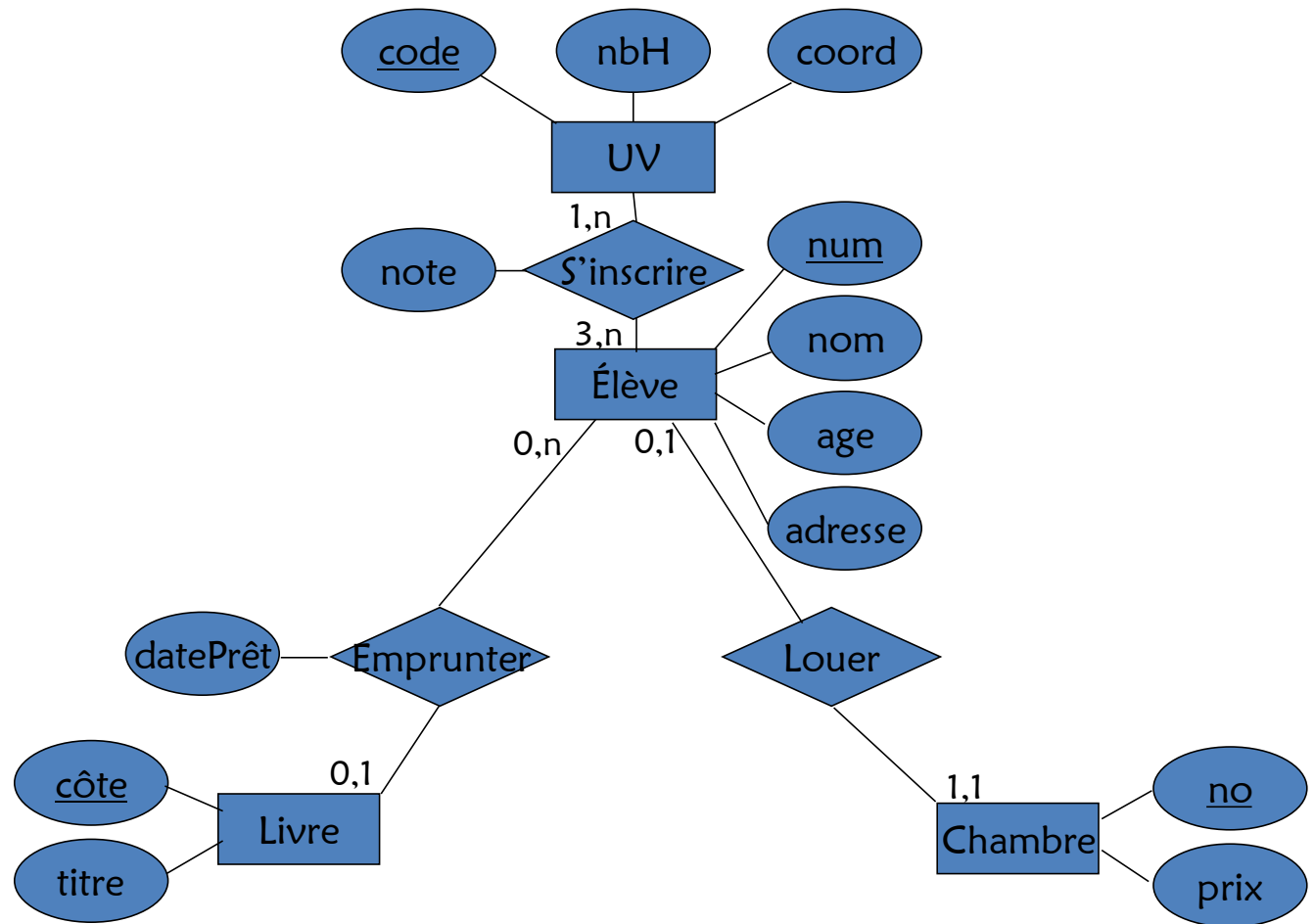
- Sécurité : une base de données est souvent vitale dans le fonctionnement d'une organisation, et il n'est pas tolérable qu'une panne puisse remettre en cause son fonctionnement de manière durable. Les SGBD fournissent des mécanismes pour assurer cette sécurité.
  - répartition / duplication des données
  - données sur plusieurs sites

# Modèles de SGBD

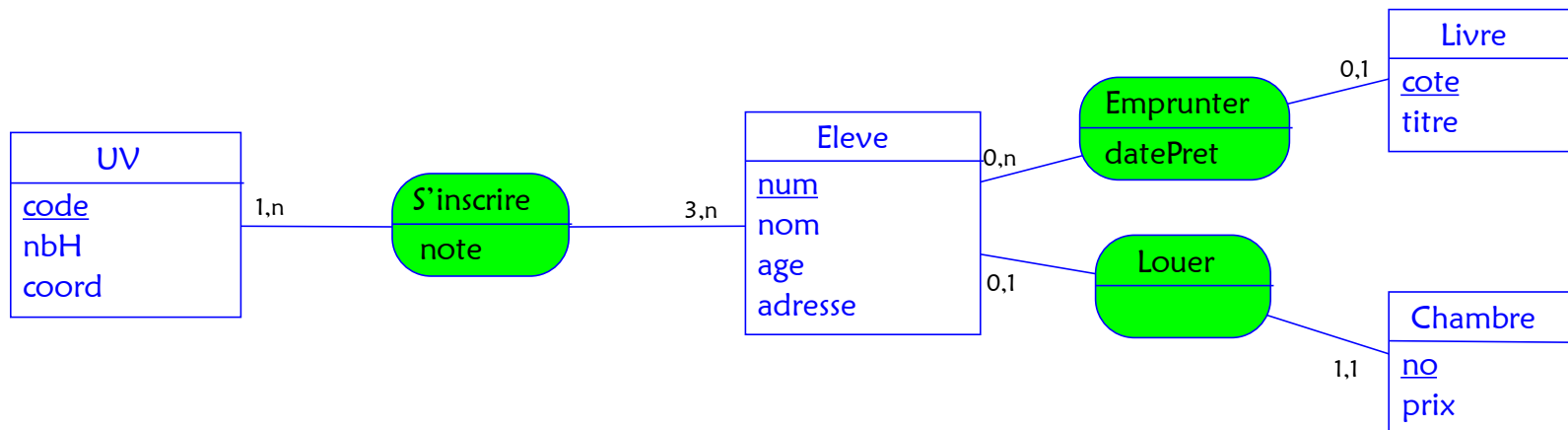
- Quelques modèles logiques :
  - Modèle hiérarchique
  - Modèle réseau
  - Modèle relationnel
  - Modèle objet
- Quelques SGBD (relationnels du marché)
  - Micro : ACCESS, Paradox, Dbase, PostgreSQL, MySQL, ...
  - Gros système : DB2, ORACLE, SYBASE, ...

# Modèle Entité/Association

## [Chen 76]



# Modèle Entité/Association - 2



# Critique du modèle Entité/Association

## ☹️ Inconvénients

- Uniquement un modèle de description de données
- Pas de langage de manipulation associé

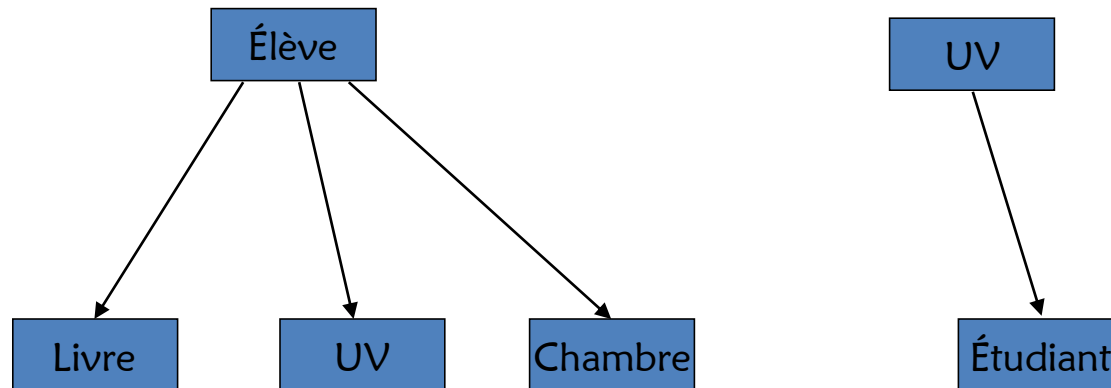
⇒ Pas un modèle  
d'implantation de BD

## • ☺️ Avantages

- Sémantique riche
  - Extension aux concepts objets (héritage, ...)
  - Aspect visuel
- ⇒ Modèle de  
conception de BD

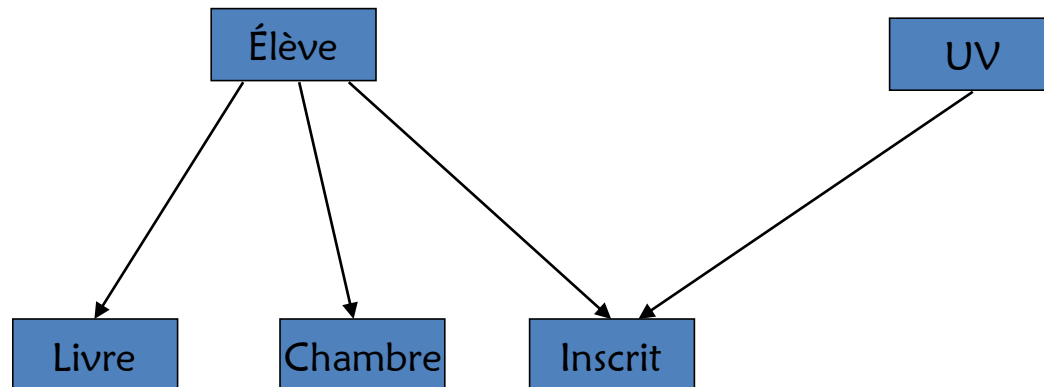
# Modèle hiérarchique

- Système IMS d'IBM conçu à la fin des années 60 pour le programme Appolo (NASA)
- Exemple



# Modèle réseau

- Modèle défini par le groupe DBTG du comité CODASYL en 1971 (revu en 1978)
- Exemple



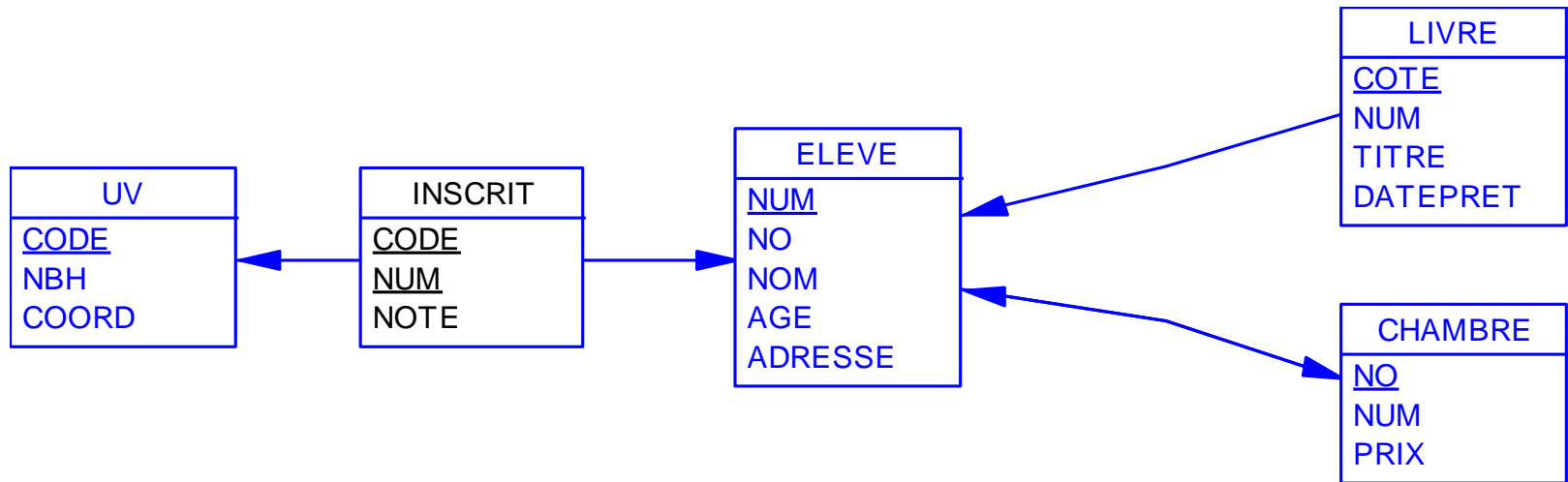


# Modèle relationnel

- [CODD 70] « A Relational Model for Large Shared Data Banks »
- Exemple
  - Élève(num, nom, adresse, age)
  - UV(code, nbH, coord)
  - Inscrit(numÉtudiant, codeUV, note)
  - Livre(côte, titre, numÉtudiant, datePrêt)
  - Chambre(no, prix, numÉtudiant)

# Modèle relationnel (2)

## « MLD » Ecole



# Modèle relationnel (3)

- Schéma BD
  - Structure de relation
- BD
  - Ensemble d'enregistrements reliés par des valeurs
- Langage de manipulation
  - Ensembliste
  - Déclaratif
  - Standard international [ SQL 86, 89, 92 ou SQL2]
- Prototypes de recherche
  - System/R chez IBM (1976), Ingres à Berkeley (1976)
- Systèmes commerciaux
  - SQL/DS et DB2 d'IBM (1982), Oracle (1983), Ingres (1983), Informix (1981), Sybase (1984), SqlServer (1998)
  - MySQL (1995)
  - Access

# Modèle relationnel (suite)

- Avantages
  - Indépendance logique/physique
  - Langage de manipulation simple
  - Basé sur une théorie mathématique solide
  - Standard

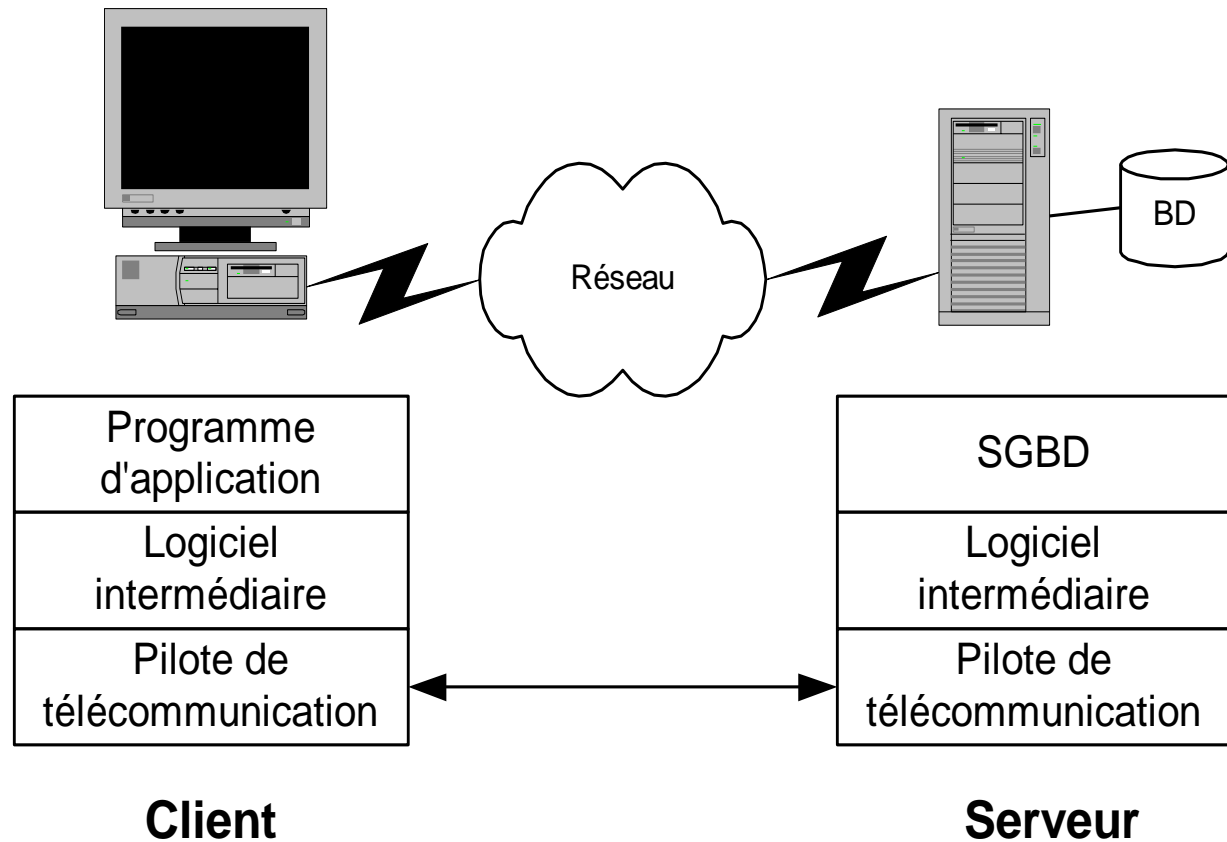
# L'architecture des SGBD

- Basée sur une architecture Client-Serveur
  - Données sur le serveur partagées entre N clients
  - Interfaces graphiques sur la station de travail personnelle
  - Communication par des protocoles standardisés
  - Clients et serveurs communiquant par des requêtes avec réponses

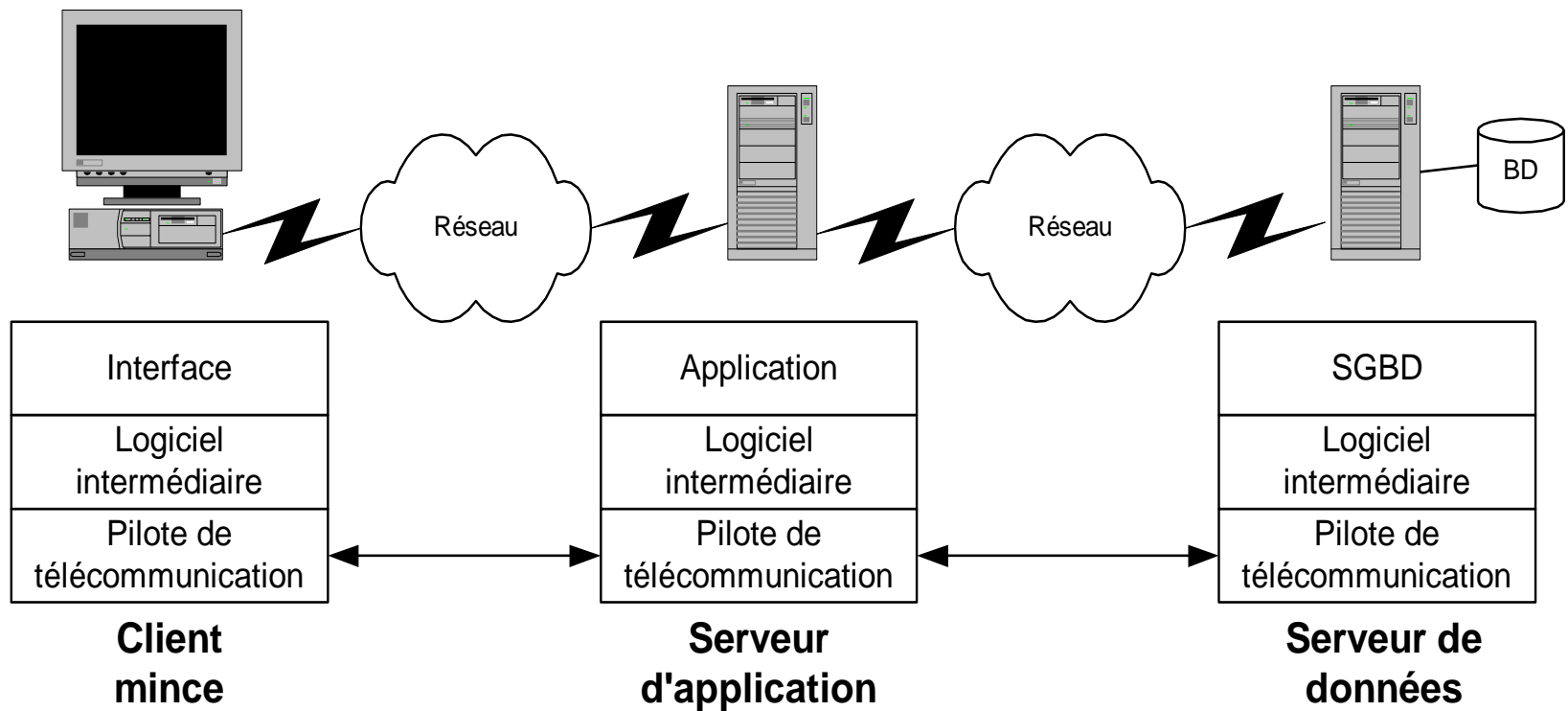
# Architecture

- *Architecture centralisée*
  - programme d'application et SGBD sur même machine (même site)
  - premiers systèmes
- *Architecture du type client-serveur (client-server architecture)*
  - programme d'application = *client*
    - interface (« GUI ») + traitement du domaine d 'application
  - SGBD = *serveur* (*de données « data server »*)
  - machines (sites) différentes

# Architecture client / serveur

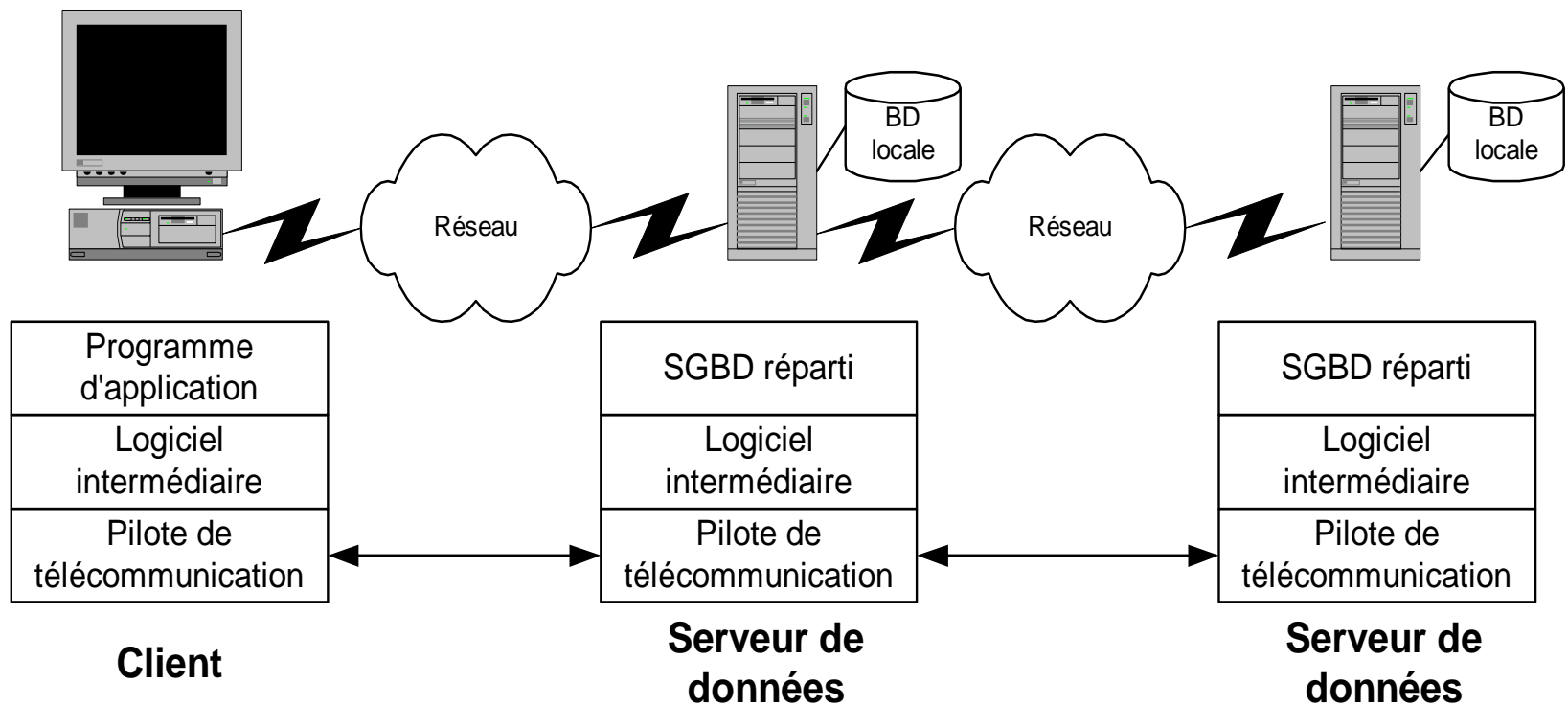


# Architecture 3 tiers

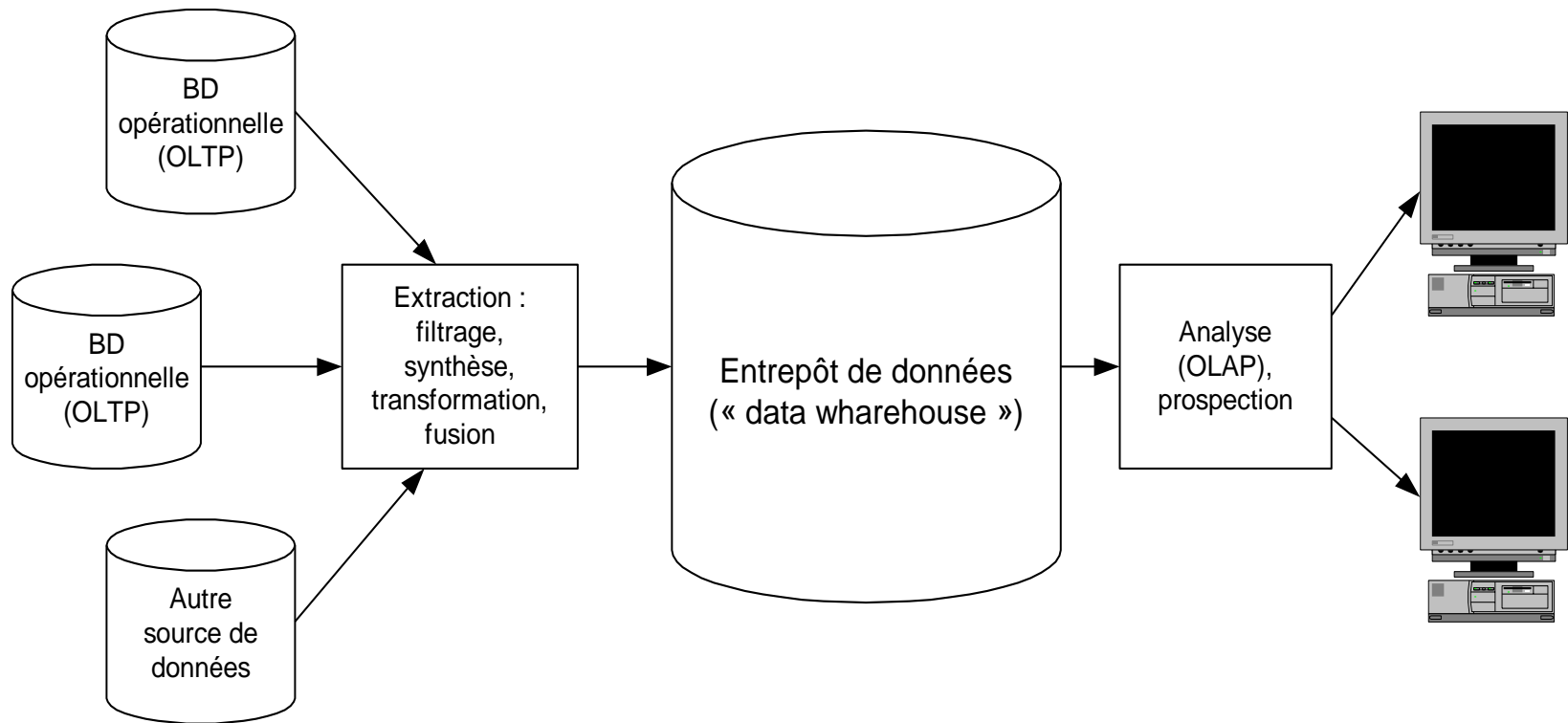




# Base de données distribuées



# Entrepôt de données



# Entrepôt de données

- *Base de données opérationnelle*
  - traitement des données quotidiennes et récentes
  - OLTP (“ *On Line Transaction Processing* ”).
- *Entrepôt de données (data wharehouse)*
  - grand volume de données historiques extraites de bases opérationnelles pour le support à la prise de décision
  - apprentissage machine , statistiques

# Démarche de construction d'une BD relationnelle

