

Chapitre 1

Concepts fondamentaux des bases de données

A decorative graphic element consisting of three horizontal lines of varying lengths and colors (teal, white, teal) extending from the right side of the slide.

Bases de données : Définitions

- **Entreprise** : Toute collectivité d'individus travaillant en coordination à la réalisation d'un objectif commun.
- **Système de gestion de bases de données** : logiciel qui assure la gestion des BDDs
- **Base de données** (BDD) : l'ensemble (cohérent, intégré, partagé) des informations nécessaires au fonctionnement d'une entreprise.
- **Base de données** : un ensemble structuré de données enregistrées sur des supports accessibles par l'ordinateur, représentant des informations du monde réel et pouvant être interrogées et mises à jour par une communauté d'utilisateurs n'ayant pas nécessairement les mêmes besoins de gestion.

Fichier et Base de données

- On utilise l'approche bases de données lorsque :
 - ✓ les données sont de natures diverses (exemples : étudiants, enseignants, cours, salles, etc.)
 - ✓ les données possèdent de nombreux liens entre elles (exemple : un étudiant suit un cours, un cours est assuré par un enseignant, un cours est enseigné dans une salle, etc.).
 - ✓ les données sont volumineuses
- on utilise l'approche par les fichiers lorsque :
 - ✓ les données à gérer sont homogènes (par exemple, Les abonnés d'une revue, le personnel d'une entreprise, les produits vendus par un magasin)
 - ✓ elles sont volumineuses

Autres définitions liées aux bases de données

- **Requête (Query):**
 - Recherche de données à l'aide d'une question posée dans un langage particulier
- **Transaction :**
 - Série d'opérations sur les données en bloquant éventuellement l'accès si besoin est
 - Effets non visibles tant que la transaction n'est pas validée
- **Administrateur de BD:**
 - Gestionnaire, responsable du bon fonctionnement de la BD

Omniprésence des bases de données

- Gestion des personnes : personnels d'une entreprise, étudiants, abonnés d'une bibliothèque, sécurité sociale, etc.
- Gestion des traitements : comptes bancaires, réservation des voyages, administration d'un réseau de machines
- Capitalisation d'information : archivage de documents, photothèque, données multimédia, ...
- Utilisées par des clients de plus en plus légers (téléphone, iPhone) et par des utilisateurs de plus en plus « novices »

cycle de vie d'une base de données

- On appelle **cycle de vie d'une base de données** la suite des phases conception, implantation et utilisation de la base.

Cycle de vie d'une base de données

conception de base de données

- On appelle **conception d'une base de données** la phase d'analyse qui aboutit à déterminer le futur contenu de la base.
- Cette phase se réalise en concertation entre concepteurs et utilisateurs futurs de la BDD
- Elle répond principalement à la question: **quelles informations placer dans la BDD pour assurer la bonne gestion de l'entreprise?**

Cycle de vie d'une base de données

implantation de la base de données

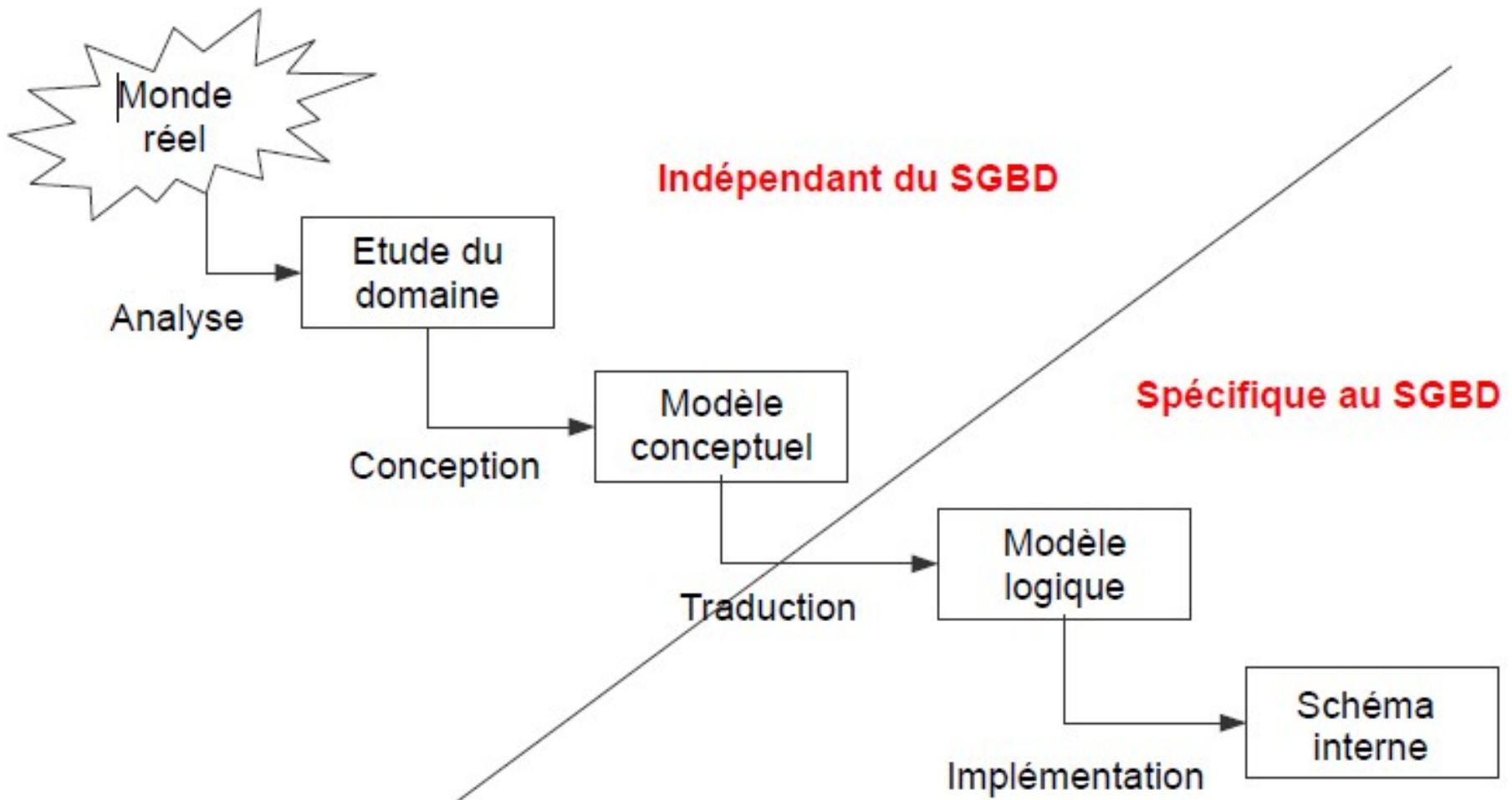
- On appelle **implantation de la base de données** la phase qui consiste à décrire la base de données dans le langage du SGBD choisi par l'entreprise et à construire cette première version.
- La description du contenu de la BDD : par le biais d'un **langage de description de données (LDD)**, correspondant au SGBD choisi.

Cycle de vie d'une base de données

utilisation de la base de données

- L'exploitation de la BDD se fait au moyen d'un langage appelé **langage de manipulation de la base de données (LMD)**,
- le LMD permet d'exprimer :
 - ✓ **des requêtes d'interrogation** (pour extraire des informations contenues dans la base)
 - ✓ **des requêtes de mise à jour** (pour ajouter de nouvelles informations, supprimer ou modifier des informations existantes).

Etapes de conception d'une BDD



Modèles et schémas d'une BDD

- **Modèle de données** : Ensemble des concepts qui permettent la description de données d'une base et les règles d'utilisation de ces concepts.
- **schéma d'une base de données**: l'expression de la description de la base de données d'une entreprise obtenue en employant un modèle de données.

Schémas d'une base de données :

Le schéma conceptuel

Le schéma logique

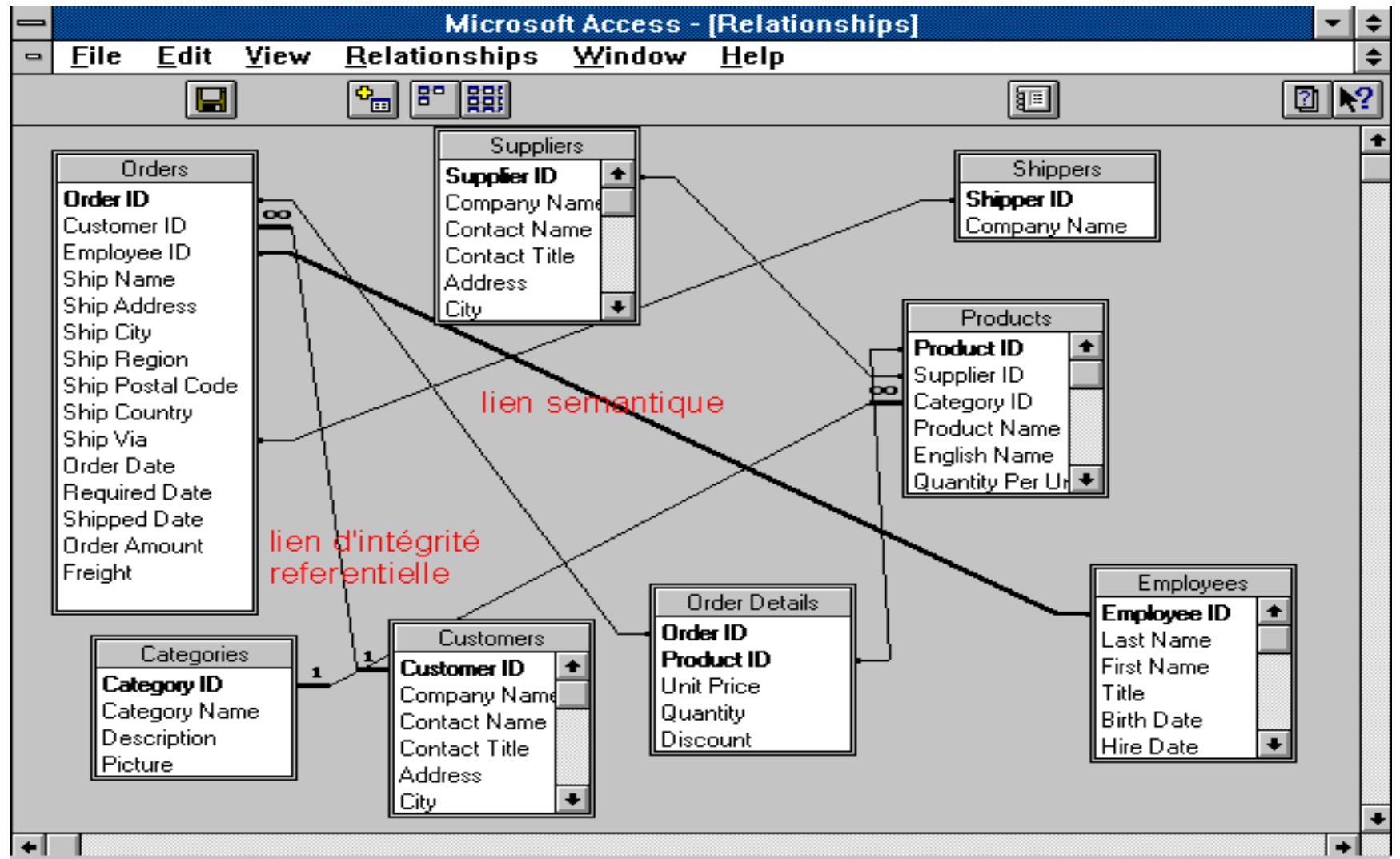
Le schéma interne

Les schémas externes

Le schéma conceptuel des besoins

- **Objectif:** quelles sont les informations nécessaires, et ce indépendamment de concepts qui font référence à l'informatique ?
- le modèle utilisé est dit «**conceptuel** »: il décrit la future base, indépendamment des choix techniques d'implantation
- Le modèle conceptuel comporte deux parties :
 - ❖ le **modèle statique**, concepts décrivant la structure des données
 - ❖ le **modèle dynamique**, permettant de décrire les traitements sur les données.

Schéma conceptuel de BDD: Exemple



Le schéma logique des données

- On appelle **modèle logique**, le modèle sur lequel est construit le SGBD actuel.
- plusieurs modèles logiques (**relationnel, CODASYL, hiérarchique, objet,...**).
- Le schéma obtenu en traduisant le modèle conceptuel des besoins en modèle logique est appelé le schéma logique de la base de données.

Le schéma interne de la base de données

- **Objectif** : définir les choix en matière de structuration des données sur les mémoires secondaires par les administrateurs système
 - quels types de fichiers ?
 - quels types d'organisation ?
 - quels index ?...
- **le schéma interne** de la base de données : décrit la manière dont les données vont être enregistrées dans les fichiers.
- Le schéma interne est spécifié dans un nouveau modèle : le modèle interne, où les concepts sont ceux de **fichier, organisation, index, chemin d'accès, clé,...**

Les schémas externes de la base de données :

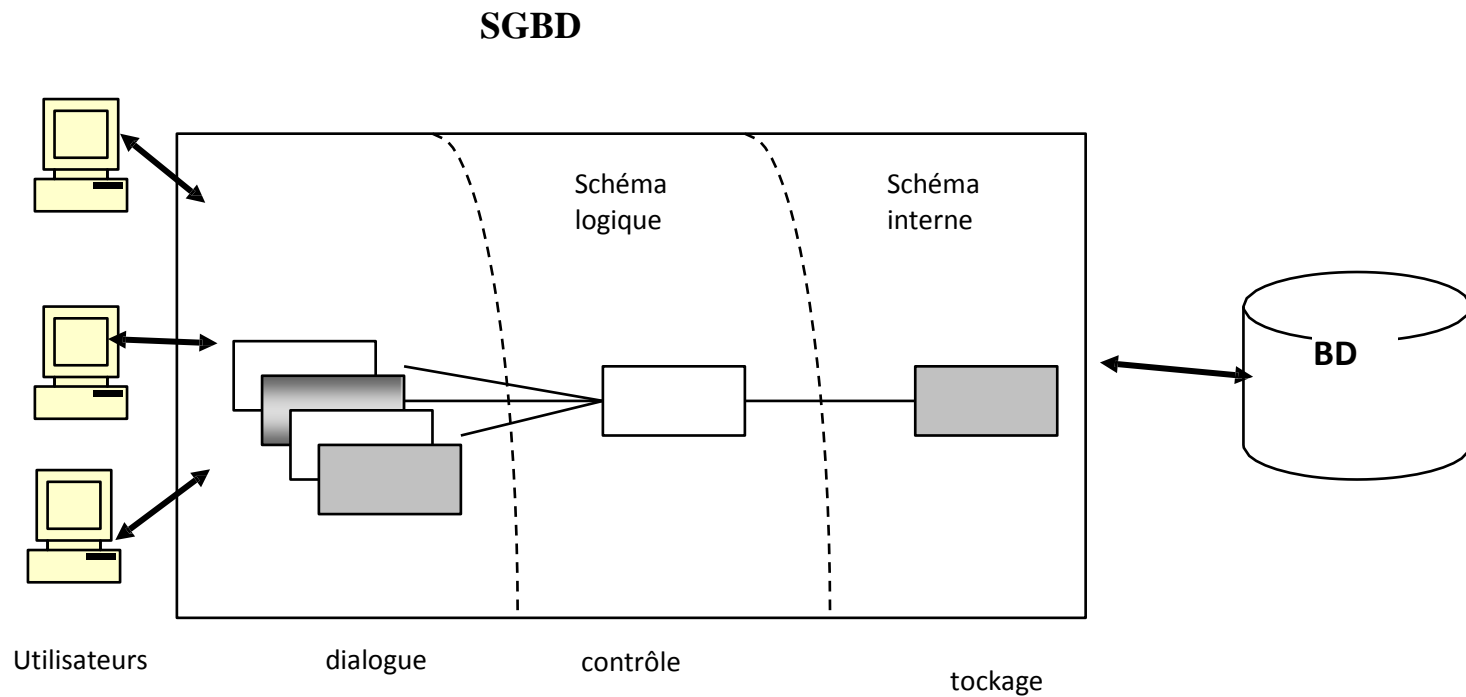
- élaborés au cours de la phase d'exploitation de la base de données, pour répondre aux besoins spécifiques des différents groupes d'utilisateurs.
- A chaque utilisateur est associé un schéma, dit son **schéma externe**, qui définit le sous-ensemble de la base de données auquel il a accès, structuré de façon à répondre à ses besoins spécifiques

Les schémas externes de la base de données :

Avantages de cette approche :

- **Simplicité** : chaque utilisateur n'a dans son schéma externe que ce qui l'intéresse.
- **Protection** : il n'est pas possible, que, par erreur ou par malveillance, un utilisateur accède aux données d'autres utilisateurs non décrites dans son schéma externe.

Les trois schémas de description d'une base de données.



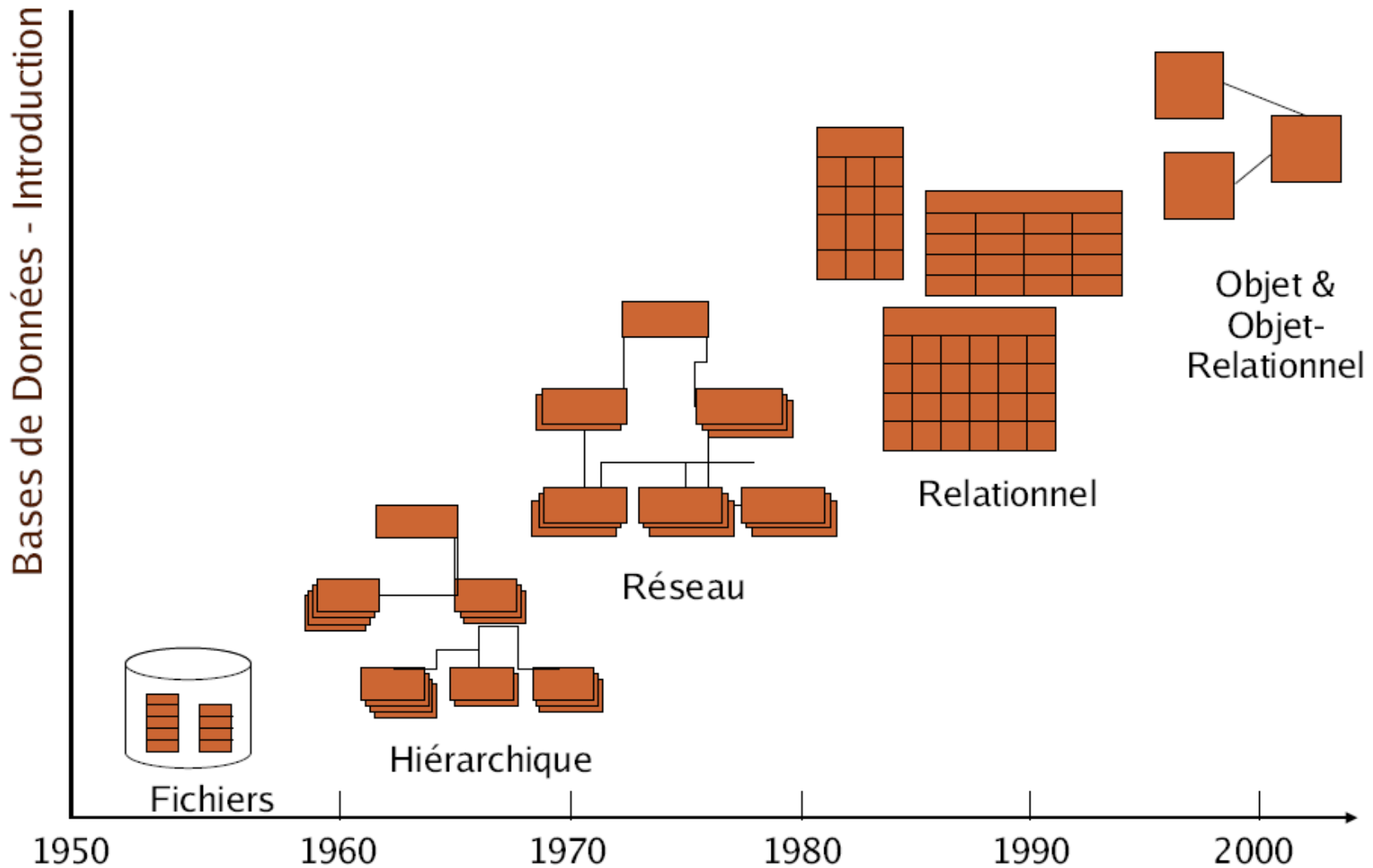
Intérêt des bases de données

- **Centralisation des données**
 - Non redondance (pas de duplications),
 - Intégrité,
 - Sécurité accrue
 - Langages / outils puissants de manipulation
 - Facilité de développement d'applications
- **Langages assertionnels de requêtes (SQL, QBE)**
- Décrit ce que l'on veut et non comment le faire
- Interrogation directe / réponse rapide
- Indépendance des données et des traitements

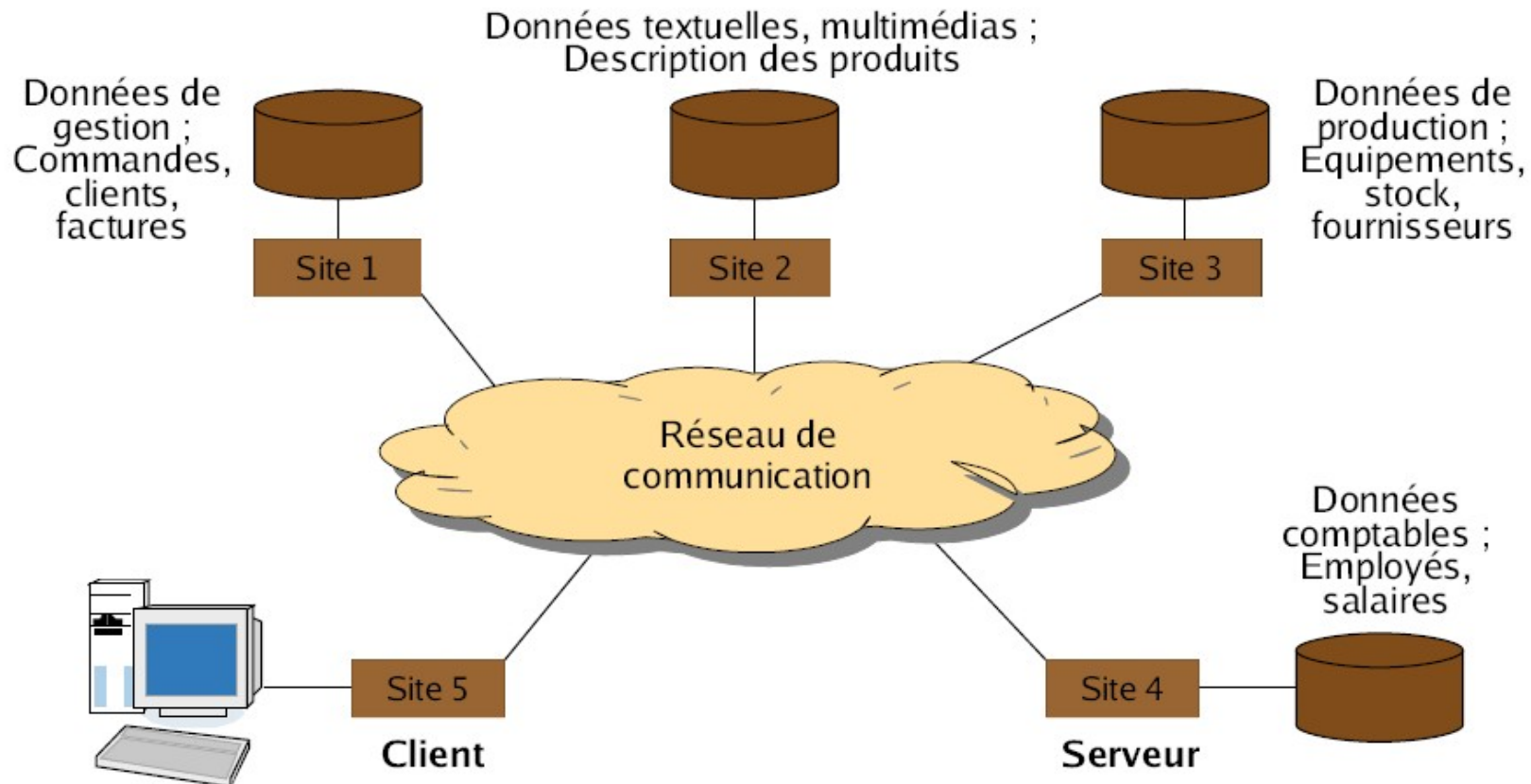
Catégories de BDD

- **BDD personnelle (10Ko- 100 Ko)**
 - OpenOffice (OOBase), Microsoft Access
- **BDD professionnelle (100 Ko – 100 Go)**
 - MySQL, Oracle
- **BDD professionnelle très grande(> 100 Go)**
 - Very Large Databases (VLDB)

Evolution des BDD:



Architecture distribuée des BDD



Partie 2 : Le SGBD

Objectifs et principes de
fonctionnement

Le système de gestion de bases de données (SGBD)

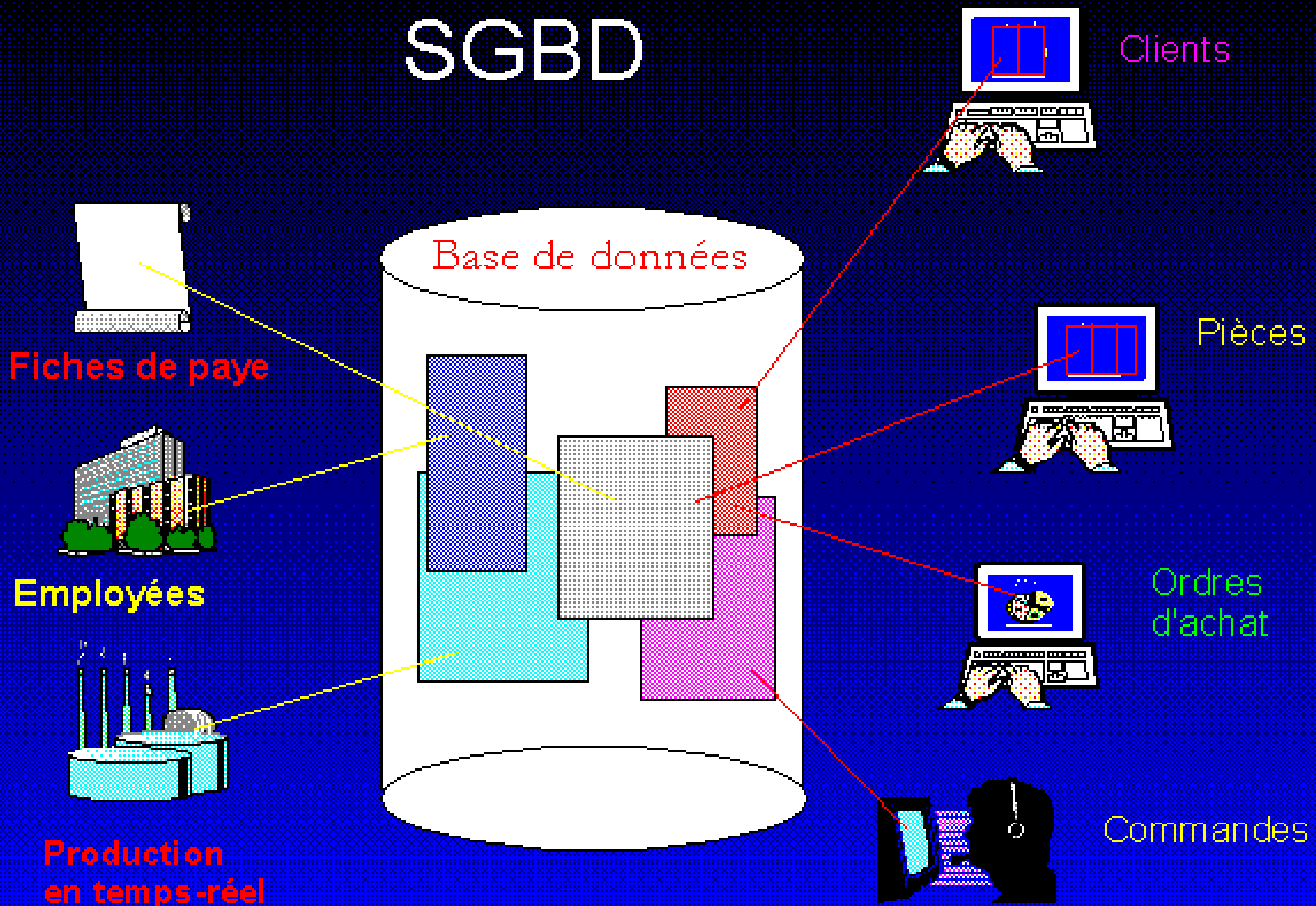
Définition

Un SGBD est un ensemble de programmes permettant la création et l'exploitation des bases de données. Le SGBD assure notamment l'interface entre la base de données physiques et les programmes d'application qui réalisent des accès, des traitements et des mises à jour.

SGBD

- La gestion et l'accès à une base de données sont assurés par un ensemble de programmes qui constituent le Système de Gestion de Base de Données (SGBD).
- Héberge généralement plusieurs bases de données, destinées à des logiciels ou des thématiques différents.

SGBD



Quelques SGBD

- *Borland : Interbase*
- *IBM : DB2 / UDB, Informix*
- *Microsoft : SQL Serveur, Access*
- *OpenOffice : OOBase*
- Oracle
- Sybase
- MySQL
- PostgreSQL

SGBD hiérarchiques & réseaux

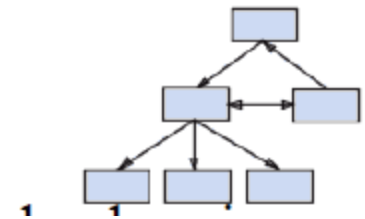
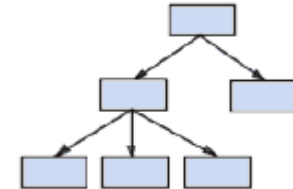
- **Structure en graphe : articles (noeuds) reliés par des pointeurs (arcs)**

- **SGBD Hiérarchiques**

- ◆ Structure arborescente descendante.

- **SGBD Réseaux**

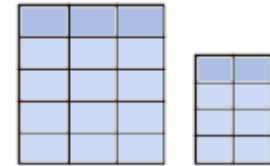
- ◆ Structure en graphe quelconque



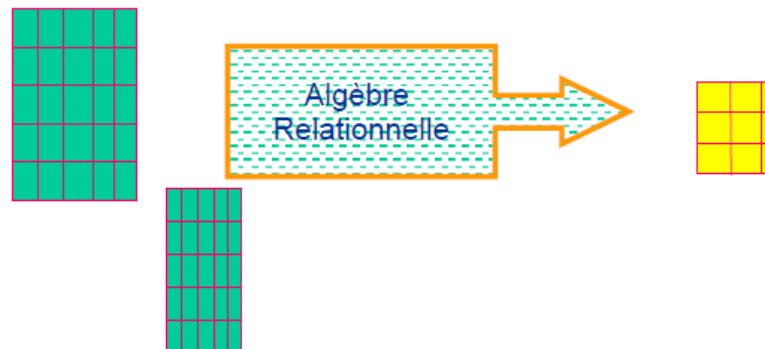
- **Utilisation de pointeurs pour déterminer le chemin aux données**
 - A indiquer dans les programmes
 - Non indépendant de la structure logique de la BD
 - Langage de manipulation complexe

SGBD Relationnels

- Basés sur la théorie mathématique des relations
- Données structurées en **tables (relations)**



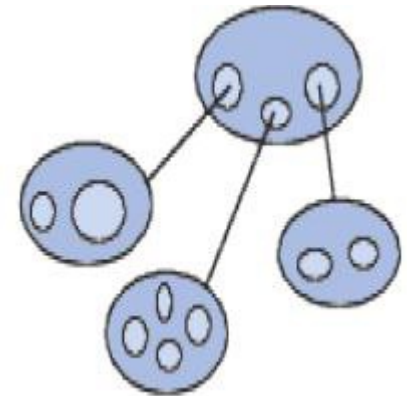
- Manipulations des données à l'aide de **l'algèbre relationnelle**
- La plupart des SGBD actuels sont relationnels



SGBD objet & objet-relationnels

- SGBD-OO : les données sont décrites par des classes (hiérarchie de classes et notion d'héritage)

- Permet de stocker des objets beaucoup plus complexes qu'avec le modèle relationnel

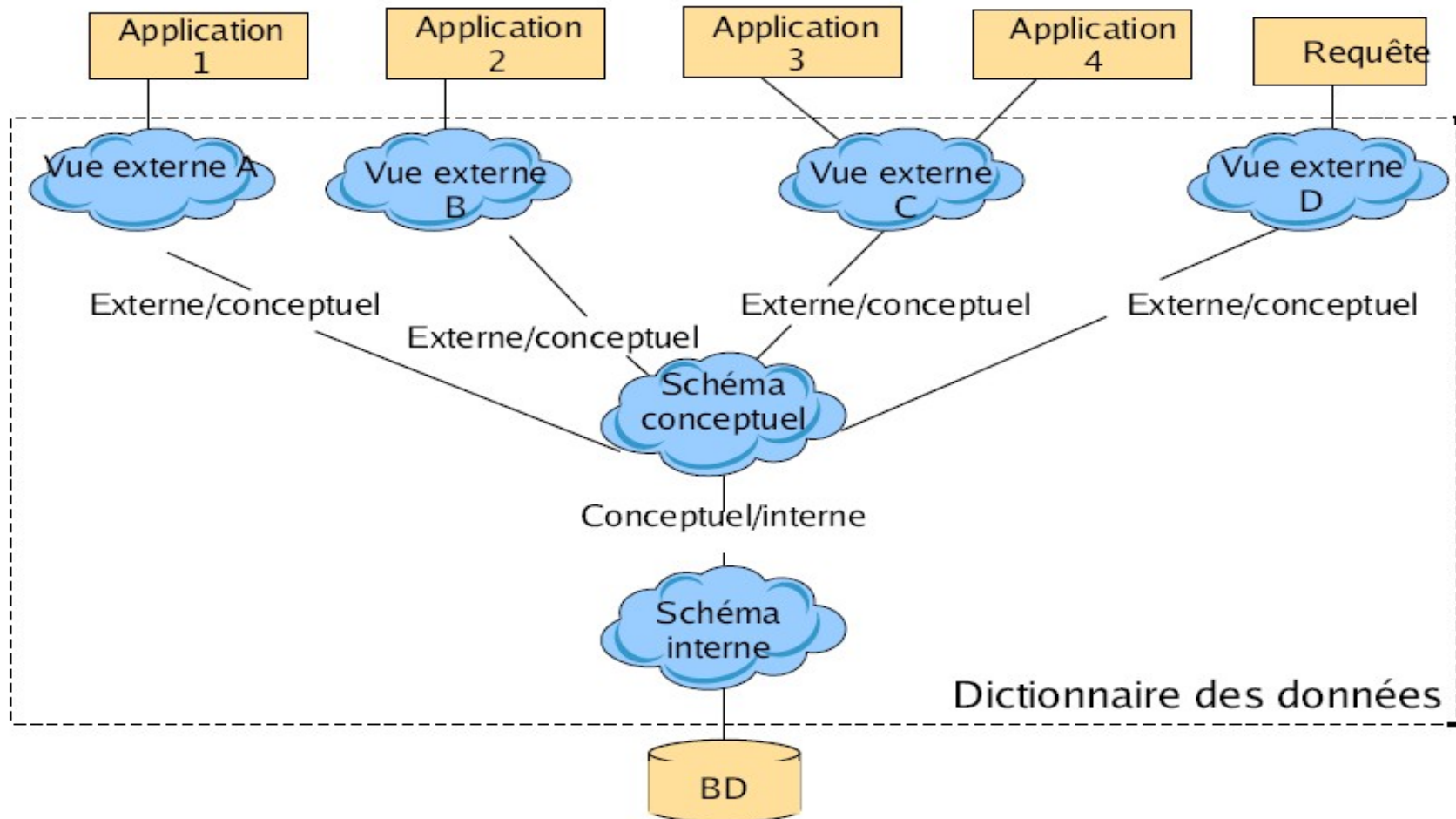


- Evolution vers les SGBD objet-relationnels (SGBD-OR) : norme SQL3

Architecture ANSI-SPARC

- Standard de l'architecture des BD proposé à partir de 1965, validé vers 1975
- Architecture à trois niveaux
 - Interne
 - Conceptuel
 - Externe
- Architecture fondamentale sur laquelle reposent tous les SGBD modernes

Architecture ANSI/SPARC



Architecture ANSI/SPARC

- **Schéma/vue externe**
 - Vue d'un utilisateur particulier
 - De la requête vers la BD
- **Schéma conceptuel**
 - Vision exhaustive des données du domaine métier
 - Description des contraintes sur les données
 - Description de l'organisation logique des données
- **Schéma interne**
 - Vue physique
 - Structures et méthodes de stockage des données

Objectifs et fonctionnalités de SGBD

(a) Indépendance physique

- Séparation nette entre les structures de données du monde réel perçues comme des entités et des liens entre ces entités et les structures de stockage sur mémoires secondaires perçues en termes d'articles, de fichiers, de chemins d'accès.
- **Avantage** : on peut changer le schéma physique sans remettre en cause le schéma conceptuel et les schémas externes (autrement dit, on peut modifier l'organisation physique des fichiers, rajouter ou supprimer des méthodes d'accès pour des fins d'optimisation des performances, etc..., sans réécrire les programmes d'application).

Objectifs et fonctionnalités de SGBD

(b) Indépendance logique

dénote de la capacité de changer le schéma conceptuel de la base de données sans remettre en cause les schémas externes ou les programmes d'application. Autrement dit, l'ajout ou le retrait de nouveaux concepts ne doit pas modifier les éléments qui n'y font pas explicitement référence

Objectifs et fonctionnalités de SGBD

(c) Manipulation des données par des non informaticiens

Les objectifs (a) et (b) s'ils sont atteints, doivent permettre à des utilisateurs non informaticiens d'exploiter la base en décrivant de façon déclarative les données qu'ils souhaitent extraire ou modifier sans décrire la manière de les retrouver ou les mettre à jour qui est propre au SGBD.

Objectifs et fonctionnalités de SGBD

(d) Non redondance des données

- La redondance d'information pose différents problèmes : (coût en temps, coût en volume et risque d'incohérence entre les différentes copies). Dans une approche base de données les données redondantes (utilisées par diverses applications) sont intégrées dans un seul fichier partagé par les diverses applications, afin d'éviter les mises à jour multiples.

Objectifs et fonctionnalités de SGBD

(e) Cohérence des données

- Les données sont soumises à des **contraintes d'intégrité**.
- **Contrainte d'intégrité** : propriété de l'application à modéliser devant être respectée.
 - **Les contraintes structurelles** (un salaire est compris entre 8000 et 30000, un employé a un supérieur hiérarchique et un seul)
 - **Les contraintes dynamiques** (un salaire ne peut diminuer). Les contraintes dynamiques doivent être codées dans les programmes d'application

Objectifs et fonctionnalités de SGBD

(f) Partageabilité des données

- partage des données entre les applications dans le temps
- Une application doit pouvoir accéder aux données comme si elle était **seule** à les utiliser.
- Le SGBD inclut un **mécanisme de contrôle de la concurrence** basé sur des techniques de verrouillage des données
- Le partage des données se fait également par la notion de **vue-utilisateur** qui permet de définir pour chaque groupe d'utilisateurs la portion de la base qui l'intéresse.

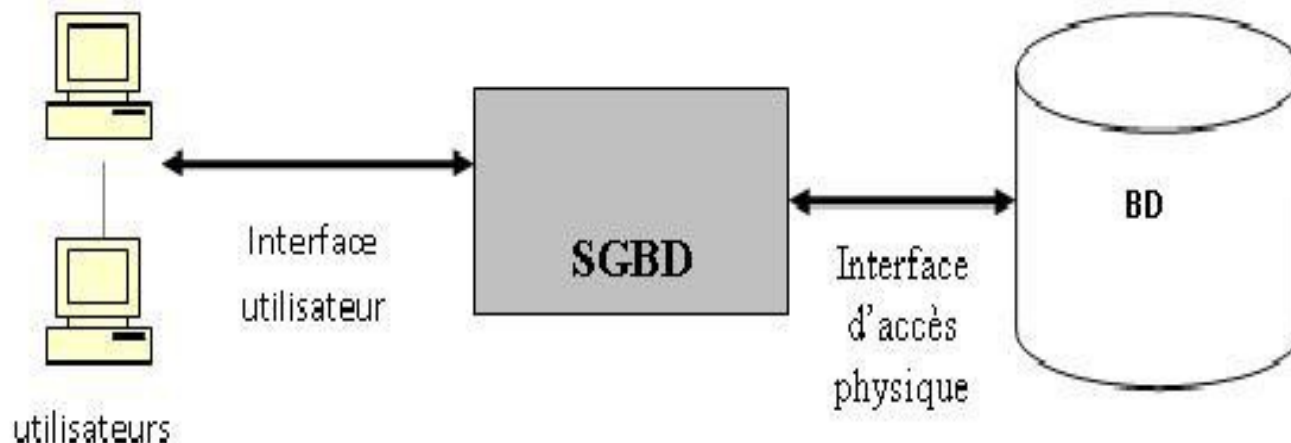
Objectifs et fonctionnalités de SGBD

(g) Sécurité et confidentialité des données

- **Protection** des données contre les accès non autorisés ou mal intentionnés.
- Le SGBD emploie des mécanismes adéquats pour **autoriser, contrôler ou supprimer les droits d'accès** de n'importe quel utilisateur à toute donnée de la base (un enseignant par exemple doit pouvoir accéder aux notes de ses élèves dans le cours qu'il leur dispense mais pas aux notes des autres cours ni aux notes des autres élèves).

Architecture d'un SGBD

- Au niveau d'abstraction le plus élevé, un SGBD peut être vu comme une boîte noire, assurant la gestion de la BD conformément aux requêtes de ses utilisateurs:



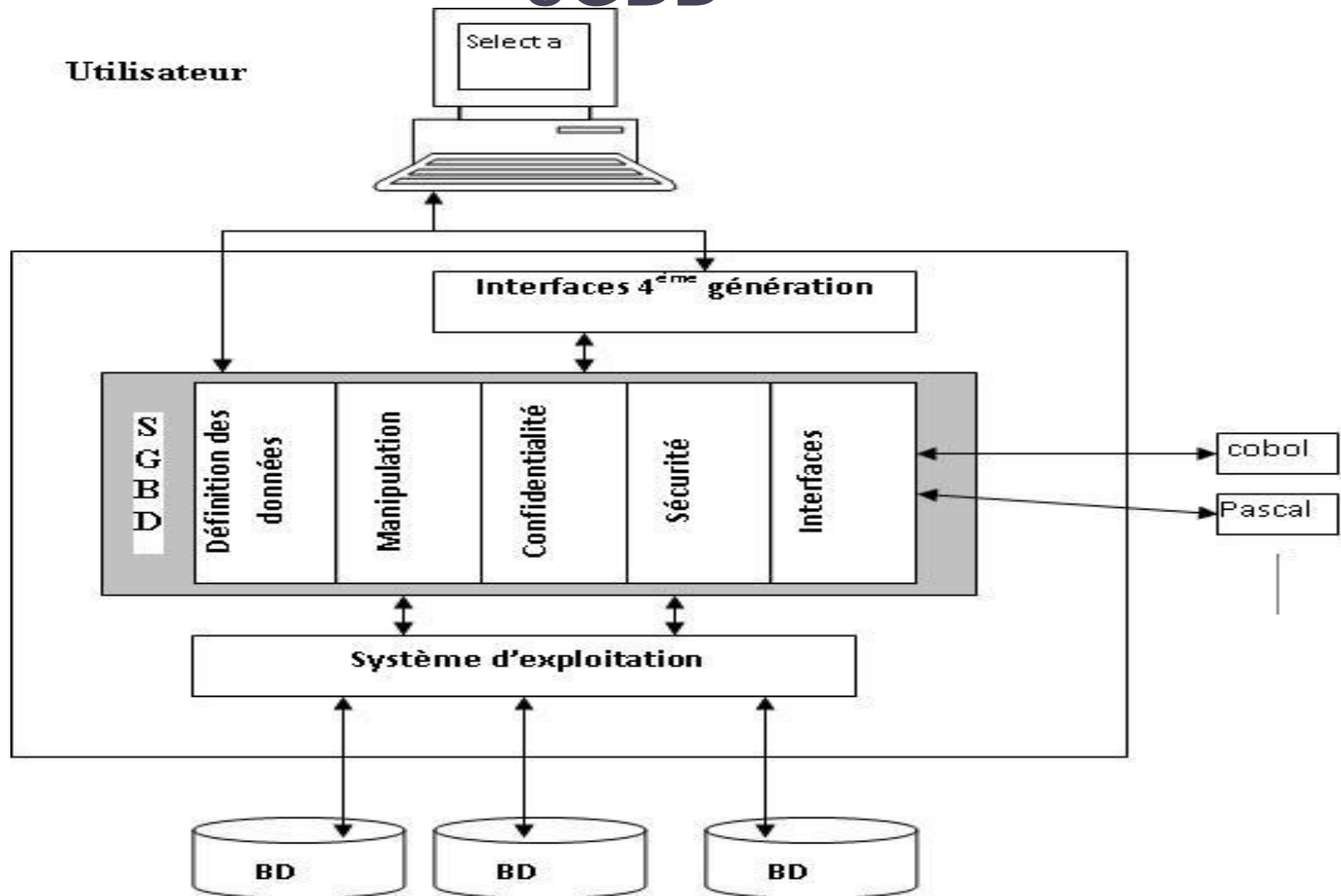
Interfaces de SGBD

- **L'interface utilisateur** permet aux utilisateurs d'exprimer des requêtes : soit pour définir le contenu de la base de données (avec le LDD), soit pour interroger la base de données (en extraire des informations), soit enfin pour apporter des modifications à ce qui a été déjà enregistré.
- **L'interface d'accès physique** permet au SGBD d'accéder aux données sur les supports de stockage.

Problèmes à gérer par le SGBD:

- Un SGBD gère des problèmes très différents, avec des objectifs particuliers :
- **Du côté de l'interface utilisateur** : compréhension, analyse et vérification des requêtes ; objectifs : convivialité de l'interface, puissance des langages de description et de manipulation.
- **Du côté interface d'accès physique** : optimisation du stockage des données (en termes d'espace occupé sur les supports) et de l'accès aux données (en temps) ; objectif : avoir les meilleures performances.

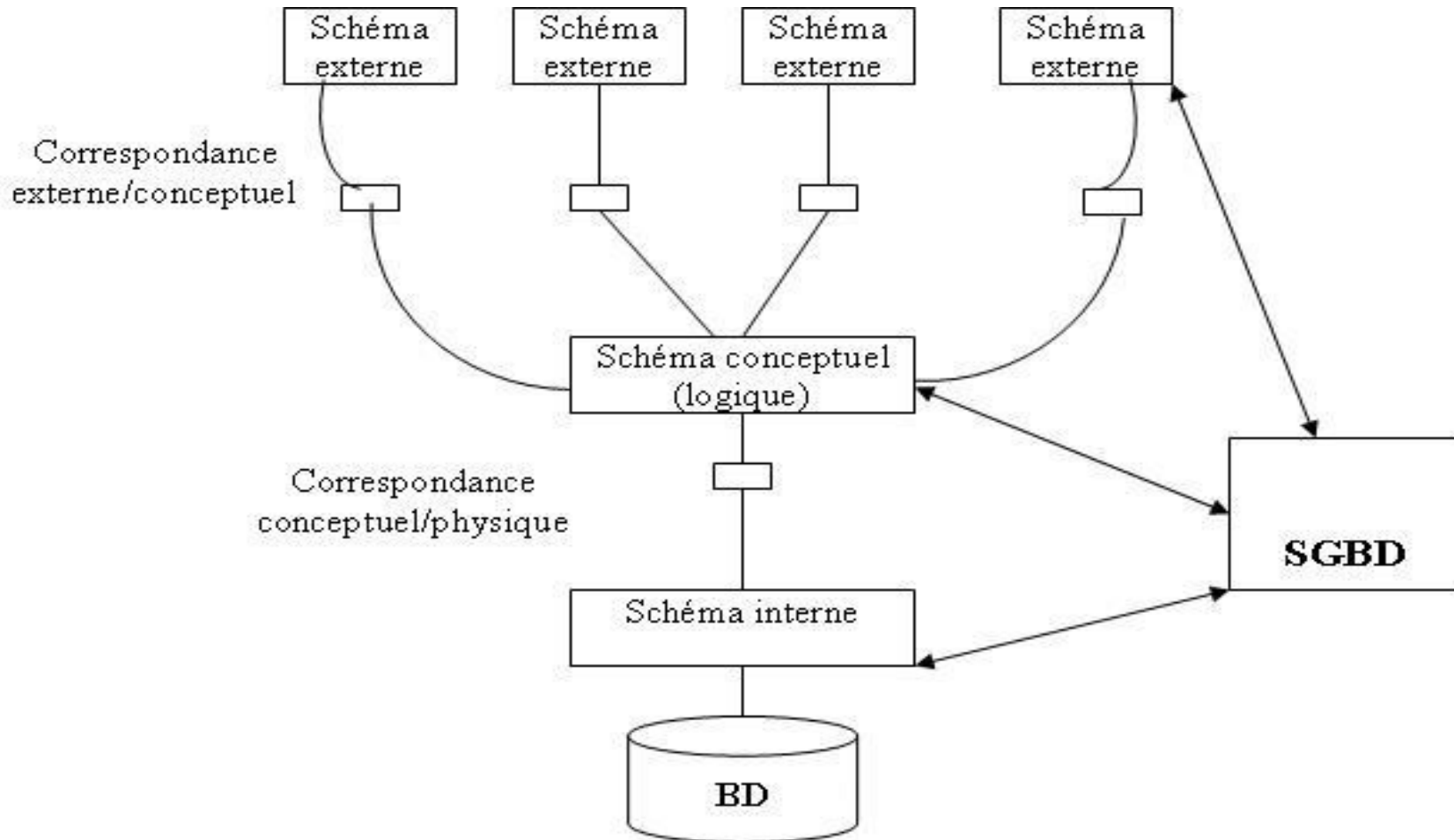
Architecture fonctionnelle d'un SGBD



Fonctions du SGBD

- **Niveau externe** : prend en charge le problème du **dialogue** avec les utilisateurs : l'analyse des demandes de l'utilisateur, le contrôle des droits d'accès, la présentation des résultats.
- **Niveau interne** : s'occupe du **stockage** des données dans les supports physiques et de la gestion des fichiers, des index et des clés.
- **Niveau logique** : assure les fonctions de **contrôle global** :
 - Optimisation globale des requêtes.
 - Gestion des conflits d'accès simultanés de la part de plusieurs utilisateurs.
 - Contrôle général de la cohérence des données
 - Coordination et suivi des processus en cours
 - Garantie du bon déroulement des activités même en cas de panne.

L'architecture à trois niveaux d'un SGBD.



Principe de fonctionnement d'un SGBD

- Une requête est exprimée par l'utilisateur dans un langage accepté par le système (LMD).
- La requête est d'abord validée dans la couche externe : elle est analysée du point de vue syntaxique (conformité à la grammaire du langage) puis d'un point de vue sémantique (les entités citées doivent être connues dans le schéma externe de l'utilisateur).
- La requête est traduite, pour son passage à la couche logique : les références aux entités du schéma externe sont remplacées par les références aux objets correspondants dans le schéma logique. On utilise pour cela la description des règles de correspondance entre schéma externe et schéma logique, établie, nécessairement, au moment de la définition du schéma externe.

Principe de fonctionnement d'un SGBD (suite)

- Des contrôles sur la confidentialité, la concurrence, etc. sont effectués. Si la requête est acceptée, elle est optimisée et découpée en sous-requêtes plus élémentaires qui sont transférées au niveau interne ; sinon elle peut être mise en attente ou refusée.
- Au niveau interne, chaque sous-requête reçue est traduite en une ou plusieurs requêtes physiques correspondantes (en fonction des informations contenues dans le schéma interne), puis le SGBD réalise l'accès physique aux données (pour extraction ou modification). Si c'est une requête d'interrogation, les résultats extraits sont passés à la couche logique, puis à la couche externe. Ici, elles sont réorganisées, selon le schéma externe de l'utilisateur, avant de les transmettre à l'utilisateur.