

Introduction à la SGBD

Base de données & SQL



Dr Pape Abdoulaye BARRO

Enseignant – Chercheur

Spécialiste en Télémétrie et Systèmes Intelligents

Plan

- Généralités
- Le modèle conceptuel
- Le modèle relationnel
- Le langage SQL
- Etude et réalisation d'une base de données
- La sécurité des données

Plan

- **Généralités**
- Le modèle conceptuel
- Le modèle relationnel
- Le langage SQL
- Etude et réalisation d'une base de données
- La sécurité des données

Généralités

Objectif

L'**objectif** de ce chapitre est de:

- définir la **notion de base de données** ainsi que les principaux **concepts** qui s'y rattachent;
- présenter:
 - Une **méthodologie** permettant de les **concevoir**,
 - Quelques **applications** informatiques associées à leur **mise en œuvre**,
 - Les différents **métiers possible**.

Généralités

Définition

- C'est quoi une base de données ?
 - Il est très difficile de donner une définition exacte d'une base de données. Mais dès que des données sont stockées de manière organisée dans un ou plusieurs fichiers, on peut donc parler de base de données. Dès que l'on dispose d'un ensemble structuré, son exploitation (**ajout, suppression, mise à jour, etc.**) devient alors possible.

• ...

Une base de données est un ensemble de fichiers, informatiques ou non, structurés et organisés afin de stocker et de gérer de l'information.

Généralités

Définition

- Une **information** est un critère de choix parmi les éléments d'un ensemble de données qui permet de restreindre la taille de l'ensemble pour donner une réponse à une question.
 - Un **exemple d'ensemble de données** pourrait être la liste des élèves d'une classe. On pouvait se poser la question suivante: *qui est le meilleur élève de la classe ?*
 - On doit recherche parmi les meilleurs élèves (ceux qui ont obtenu une moyenne ≥ 12), Recherche : le meilleur d'entre eux (celui qui a la plus grande moyenne).
- L'**informatique** est le traitement automatique de l'information.

Généralités

Définition > Remarques

- Quelques remarques:

- Vous avez peut-être une idée intuitive des bases de données. Attention toutefois, le mot est souvent utilisé pour désigner n'importe quel ensemble de données. Il s'agit là d'un abus de langage qu'il convient d'éviter.
- Pour mériter le terme de base de données, un ensemble de données non indépendantes doit être interrogeable par le contenu.
 - Les données doivent pouvoir être retrouvées selon n'importe quel critère :
 - Il devrait être possible de trouver une liste d'articles qui coûtent moins de 500 francs.
 - Il doit également être possible de retrouver leur structure :
 - un produit a un nom, un prix, une quantité, ...

Généralités

Définition > Remarques

- Les bases de données se situent au cœur de l'activité des entreprises, de l'administrations, de la recherche et de bon nombre d'activités humaines désormais liées à l'informatique.
- Une base de données est une représentation partielle et très simplifiée du monde réel, que l'on a obtenu par un processus de **modélisation**.
- Pour les manipuler, on utilise généralement un logiciel spécialisé appelé **SGBD** (Système de Gestion de Bases de Données).
- Dans le domaine purement informatique, elles interviennent dorénavant à tous les niveaux.
 - Les développeurs d'applications s'appuient sur des bases de données externes pour gérer leurs données alors qu'auparavant elles étaient intégrées dans le programme.
- Les bases de données reposent sur des théories solides et sont à l'origine d'une des plus importantes disciplines de l'informatique : **l'ingénierie des systèmes d'information**.

Généralités

Définition > SGBD

C'est quoi un SGBD alors ?

- Un SGBD peut être perçu comme un ensemble de *logiciels systèmes* permettant aux utilisateurs d'**insérer**, de **modifier** et de **rechercher** efficacement des *données spécifiques* dans une grande masse d'informations (pouvant atteindre quelques milliards d'octets) partagée par de multiples utilisateurs.
 - Le SGBD rend transparent le partage, à savoir donne l'illusion à chaque utilisateur qu'il est seul à travailler avec les données.
- Les **SGBD** se distinguent clairement des *systèmes de fichiers*.
 - Les **SGBD** permettent la description des données (*définition des types par des noms, des formats, des caractéristiques et parfois des opérations*) de manière séparée de leur utilisation (*mise à jour et recherche*).
 - Ils permettent aussi de retrouver les caractéristiques d'un type de données à partir de son nom (*par exemple, comment est décrit un produit*).
 - Le **système de fichiers** est un composant de plus bas niveau ne prenant pas en compte la structure des données.
- La tendance aujourd'hui est d'intégrer le système de fichiers dans le SGBD.

Généralités

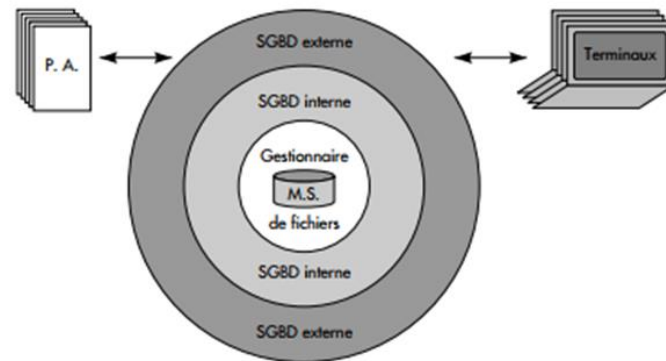
Définition > SGBD

- Un *SGBD* se compose en première approximation de trois couches emboîtées de fonctions (de la mémoire à utilisateur).
 - Le système de gestion de fichiers (ou gestionnaire de fichiers) fournit aux *couches supérieures* des mémoires secondaires idéales adressables par objets et capables de recherches par le contenu des objets (mécanismes d'indexation notamment).
 - le système d'accès aux données (ou SGBD interne) (*deuxième couche*) qui assure la gestion des données stockées dans les fichiers, l'assemblage de ces données en objets, le placement de ces objets dans les fichiers, la gestion des liens entre objets et des structures permettant d'accélérer les accès aux objets.
 - SGBD externe (*la troisième couche*) dont le rôle consiste à mettre en forme et à présenter des données aux programmes d'applications et aux utilisateurs interactifs. Il a pour rôle, d'assurer:
 - d'une part, l'analyse et l'interprétation des requêtes utilisateurs en primitives internes,
 - d'autre part, la transformation des *données extraites de la base* en *données échangeables avec le monde extérieur*.

Généralités

Définition > SGBD

- Au delà des SGBD, les systèmes d'informations intègrent plus souvent *des ateliers de génie logiciel* permettant de modéliser les données d'une base de données et de représenter les traitements associés à l'aide de graphiques et de langages de spécifications.
- *Ces outils d'aide à la conception*, bien que non intégrés dans le SGBD, permettent de spécifier les descriptions des données (supportés par les SGBD).



P.A. = Programmes d'Application
M.S. = Mémoires Secondaires

Composition d'un SGBD

Généralités

Historique des SGBD

Les toutes premières bases de données ont vu le jour vers les années 60 lorsqu'on utilisait les fichiers reliés par des pointeurs.

- C'est vers 1965, avec l'apparition des systèmes IDS.I de Honeywell et IMS.I de IBM (pour la mission APOLLO), que les SGBD modernes ont commencé à voir le jour. Cela a permis de constituer des chaînes d'articles entre fichiers et de parcourir ces chaînes.
- La première génération de SGBD est marquée par la séparation de la description des données et de la manipulation par les programmes d'application. Elle coïncide avec l'avènement des langages d'accès navigationnels.
 - Elle est basée sur les modèles réseau ou hiérarchique et a été dominée par les SGBD TOTAL, IDMS, IDS 2 et IMS 2. Elle traite encore aujourd'hui une partie importante du volume de données gérées par des SGBD.

Généralités

Historique des SGBD

- La **deuxième génération de SGBD** a fait son apparition vers les années 1970 dans les laboratoires (début du modèle relationnel) et vise à enrichir et à simplifier le SGBD externe afin de faciliter l'accès aux données pour les utilisateurs. Les données sont représentées sous forme de tables et les recherches et mises à jour sont effectuées à l'aide d'un langage non procédural standardisé appelé **SQL**(Structured **Q**uery **L**anguage).
 - Les systèmes de la deuxième génération sont commercialisés depuis 1980 et aujourd'hui, ils représentent l'essentiel du marché des bases de données.
 - Les principaux systèmes sont ORACLE, INGRES, SYBASE, INFORMIX, DB2 et SQL SERVER.

Généralités

Historique des SGBD

- La **troisième génération de SGBD** est apparue vers les années 1980 dans les laboratoires et est fortement adoptés dans le monde industriel. Elle supporte des modèles de données extensibles associant relationnel et objet, ainsi que des architectures mieux réparties, permettant une meilleure collaboration entre des utilisateurs concurrents.
 - Parmi les **systèmes objet-relationnels** nous pouvons citer Oracle 8, IBM DB2 Universal Database, Informix Universal Server et ObjectStore (plus novatrice).
 - Tous ces systèmes tentent de **répondre aux besoins** des nouvelles applications (multimédia, Web, CAO, bureautique, environnement, télécommunications, etc.).

Généralités

Historique des SGBD

- La **quatrième génération** quant à elle, est mise en place pour mieux supporter Internet (le Web), **les informations non structurées**, les objets multimédias, mais aussi faciliter l'extraction de connaissances (data mining) ainsi que **l'aide à la prise de décisions**. Il devient alors de plus en plus difficile de développer un nouvel SGBD. Cette nouvelle génération est juste une évolution des SGBD de 3e génération plutôt qu'à une nouvelle révolution.

Enfin, l'évolution des SGBD peut être perçue comme celle d'un arbre, des branches nouvelles naissant mais se faisant généralement absorber par le tronc, qui grossit toujours d'avantage.

Généralités

Les étapes de la conception

Le processus de conception d'une base de données peut être décomposé en plusieurs étapes allant de l'analyse du monde réel jusqu'à sa création et son utilisation.

- L'**analyse du monde réel** est l'étape de collecte des informations et des besoins des futurs utilisateurs. C'est dans cette étape que l'on détermine les objectifs du système d'information à mettre en place et identifie tous les éléments à prendre en compte.
 - Il permet d'identifier les objets et de modéliser les liens existant entre eux. Par exemple:
 - ❑ une personne achète un terrain à 5 000 000 FCFA.
 - Dans cette phrase, nous avons deux objets liés à savoir **personne** et **terrain** et le **prix** est un composant du lien.
 - Il permet également d'exprimer des règles de validité de données. Par exemple:
 - ❑ Les terrains dont nous disposons coutent au **minimum** 1 000 000 FCFA et au **maximum** 10 000 000 FCFA.

Généralités

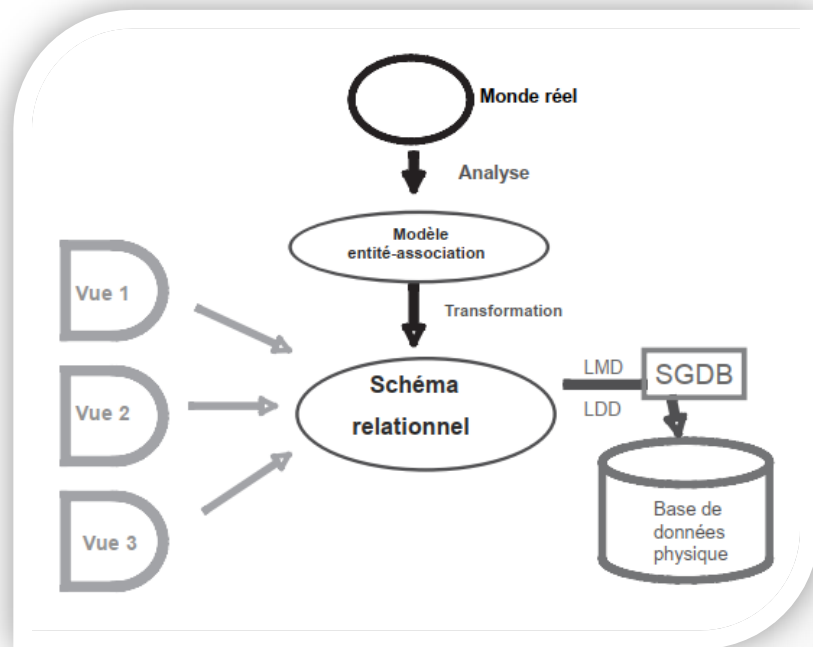
Les étapes de la conception

- L'exploitation des données ou des liens existant entre les entités est réalisée à l'aide du **modèle conceptuel** «*entité-association*» ou, plus couramment aujourd'hui, exprimé avec le langage **UML** (Unified Modeling Language).
- Le **passage au SGBD** est l'étape de transformation de la représentation précédente afin de la rendre acceptable par le SGBD choisi. Lors de cette étape, il est question de vérifier la qualité de la base de données en éliminant les redondances et autres. Le **modèle relationnel**, avec ses outils, permet d'arriver à cette fin.
 - On obtient alors un schéma des données qui fournira aux utilisateurs les informations nécessaires pour effectuer leurs **requêtes**, par exemple, la description des **noms de tables**, de **champs** et leurs **types**.
 - C'est dans cette phase que l'on définit les « **vues** » du système d'information qui sont adaptées à chaque **catégorie d'utilisateurs**.

Généralités

Les étapes de la conception

- Enfin, nous passons à l'étape de **création et d'utilisation de la base de données** en utilisant le schéma précédemment défini. Il s'agit de la création des tables qui constituent la base de données, et puis, d'insérer les valeurs dans les tables.
 - La création des tables ainsi que l'utilisation de la base de données nécessitent le *langage SQL*.



Étapes de conception d'une base de données

Généralités

Les différents métiers possibles

Dans le processus d'étude et de mise en place d'une base de données, divers acteurs peuvent y intervenir donnant ainsi naissance à d'intéressants métiers d'**ingénierie des systèmes d'information**. Parmi lesquels, nous avons:

- le métier de **Consultants/Analystes** qui prennent en charge la première étape de conception c'est-à-dire l'étape consistant à analyser les activités et les flux d'information du système (activités d'une entreprise par exemple) à modéliser. Cette phase nécessite beaucoup de dialogue et de psychologie pour bien parvenir à exprimer les besoins des futurs utilisateurs.
- Il n'est pas toujours nécessaire que l'acteur soit purement technique, il suffit juste qu'il soit capable de correctement exprimer les besoins afin de proposer un **modèle conceptuel de donnée** (le plus juste possible).

Généralités

Les différents métiers possibles

- le métier de *Concepteurs de base de données* qui s'occupent de la transformation du modèle conceptuel en un modèle logique pouvant être exploité par le SGBD. C'est le spécialiste qui doit mettre en place les tables, les vues, les schémas d'accès à la base en collaboration avec l'administrateur.
 - Parfois, il peut être *analyste et concepteur* en même temps, dans ce cas, le cahier des charges aura tendance à être trop technique. Il peut aussi être administrateur, ce qui ne pose aucun souci.
- Le métier d'*Administrateurs de base de données* qui a la responsabilité du fonctionnement général du SGBD.
 - Il crée les ressources(base, comptes) a la demande et donne les droits d'accès aux usagers.
 - Il vérifie la disponibilité des ressources, veille aux défaillance sécuritaire et est en étroite collaboration avec l'administrateur système et réseau de la structure.
- Le métier de *Programmeurs d'application* qui est un utilisateur standard du système d'information. Ils y ont accès grâce aux vues définies par le concepteur de la base. Pour mener à bien leur taches, les développeurs doivent avoir une très bonne connaissance de la SGBD mise en place.

Généralités

Résumé

- Il est très difficile de donner une **définition exacte** d'une base de données. Une base de données peut être considérée comme "*un ensemble de fichiers – informatiques ou non – structurés et organisés afin de stocker et de gérer de l'information*".
- Un **SGBD** peut être perçu comme **un ensemble de logiciels systèmes** permettant aux utilisateurs de pouvoir efficacement manipuler les données de la base.
- Un SGBD se compose, en première approximation, de trois couches à savoir la couche **gestionnaires de fichiers**, la couche **SGBD interne** et la couche **SGBD externe**.
- Les toutes **premières bases de données** ont vu le jour vers les **années 60** et leurs évolutions ont vu naître diverses générations qui ont suivies le même trajet, en prenant en compte, à chaque étape, les besoins ou exigences des utilisateurs.
- Pour **mettre en place une base de données** de manière efficace, il est nécessaire de passer par **quelques étapes** à savoir, **l'étude du monde réel**, la proposition d'un **modèle conceptuel**, la traduction du modèle conceptuel en **modèle relationnel** avant donc de pouvoir mettre en place la base de données.
- Ces différentes étapes impliquent des fois l'intervention de **plusieurs acteurs** faisant naître des **métiers connexes** comme le métier de **consultant/analyste**, le métier de **concepteur de base**, le métier d'**administrateur de base de données** et puis le métier de **programmeur d'application**.

Généralités

Discussions

- Q1: Quelles sont les différences majeures entre un fichier informatique et une base de données gérée par un SGBD ?
 - Il n'est pas nécessaire de connaître la méthode de stockage des informations sur le disque pour manipuler les données avec une base de données.
 - Un fichier informatique simple n'est pas conçu pour effectuer une recherche d'information par le contenu : pour retrouver les enregistrements, on est obligé de parcourir tout le fichier.
 - Les modifications de structure (ajout/suppression d'un champ ou modification de sa taille...) nécessitent de recréer un autre fichier et d'y recopier les données.
 - Une base de données contient en général plusieurs fichiers dont les enregistrements sont reliés entre eux.
 - Etc.

Généralités

Discussions

- Q2: Est-il possible de faire réaliser toutes les étapes de la conception d'une base de données par une même personne ? Si oui, quelles sont alors ses compétences minimales ?
 - Dans une petite structure, c'est souvent la même personne qui réalise l'ensemble du processus de construction d'une base de données. Ce n'est évidemment pas la bonne méthode, car la vision d'un système d'information élaboré par un administrateur de base de données est très orientée par le SGBD qu'il emploiera. On peut facilement faire le parallèle avec le développement de logiciels où un programmeur va avoir une approche déformée par les préoccupations liées au langage plutôt que d'adopter un point de vue sur la structure générale de l'application.
 - Au minimum, la personne devra disposer des compétences en conception de base de données et en administration du SGBD qui sera utilisé.

Généralités

Discussions

- Q3: On décide de **recopier** régulièrement **une base de données** complète sur chacun des **six sites de l'entreprise**. Quel est l'intérêt de cette solution ?
 - C'est une solution coûteuse en ressources, en particulier pour la synchronisation de toutes les mises à jour, mais qui peut remplir deux fonctions :
 - De toute évidence, l'idée est que les différents sites de l'entreprise accèdent à leur copie locale des données. Cela permet d'accélérer les accès et de répartir la charge sur les serveurs locaux de chaque site. Accessoirement, la circulation des requêtes d'interrogation et de mise à jour peut être limitée au réseau local, ce qui renforce la sécurité.
 - Il existe six copies des données sur des sites géographiquement séparés. En cas de sinistre, on peut repartir sans problèmes avec une des copies de la base.
 - Avant de mettre en place un tel dispositif, on doit se poser la question de la mise à jour des données. Est-ce que les modifications se font uniquement sur la base "maître", ce qui semble plus raisonnable, ou peut-on les effectuer sur toutes les bases et "consolider" ensuite ?

FIN

Feedback: pape.abdoutaye.barro@gmail.com



Dr Pape Abdoulaye BARRO

Enseignant – Chercheur

Spécialiste en Télémétrie et Systèmes Intelligents