

Bases de Données

introduction

Table des matières

1. Introduction.....	2
2. Notion de Base de Données.....	2
2.1. Modèle de données.....	2
2.2. SGBD.....	2
3. Propriétés ACID.....	3
4. La création d'une base de données.....	3
4.1. Modèle Conceptuel des Données (MCD).....	4
4.1.1. propriété.....	4
4.1.2. Entités.....	4
4.1.3. Association.....	4
4.2. Modèle logique des données (MLD).....	6
4.3. Modèle Physique des Données (MPD).....	7

Une base de données ou banque de données (en anglais : database) est un outil permettant de stocker et de retrouver l'intégralité de données brutes ou d'informations en rapport avec un thème ou une activité. Dans la très grande majorité des cas, ces données sont très structurées, et la base est localisée dans un même lieu et sur un même support.



1. Introduction

Les bases de données (BDD) désignent tous les systèmes de stockage qui permettent de conserver des données à l'écart de l'application. Il peut s'agir de fichiers de texte, de fichiers au format XML, de systèmes de gestion de bases de données (SGBD), etc.

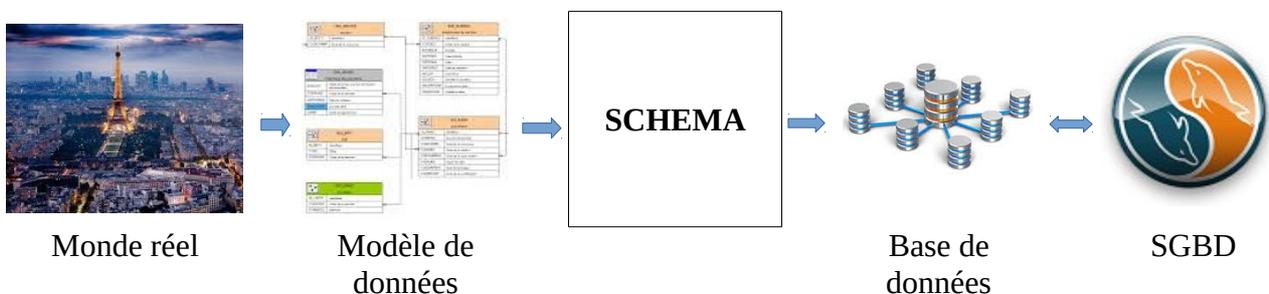
Une base de données relationnelle permet d'enregistrer les informations sans avoir aucun duplicata, ce qui réduit l'espace de stockage utilisé ainsi que les temps d'accès, tout en simplifiant la mise à jour. Les données y sont en relation les unes avec les autres et évite les doublons.

Le principe d'une BDD relationnelle est d'enregistrer les informations de manière hiérarchisée :

- Le SGBD contient des bases de données (exemples : "livre-d-or") : dans un tableur, ce sont les "fichiers" ;
- Chaque base de données est structurée en tables (exemples : "user", "message", "sujet") : dans un tableur, ce sont les "feuilles de calcul" ;
- Chaque table regroupe tous les champs d'une même entité (exemples : "id", "login") : dans un tableur, ce sont les "colonnes" ;
- Un champ est une valeur (exemples : "8480", "BrYs", "30724", "Yogui") : dans un tableur, ce sont les "cellules".

2. Notion de Base de Données

Les bases de données sont basées sur une approche de « STRUCTURATION » du monde réel qui donne une représentation abstraite (le SCHEMA) résultant de l'application d'un MODELE de DONNEES (data model).



2.1. Modèle de données

Le modèle de données comporte 3 grandes parties :

- Structures des données associées qui permettent de structurer l'univers réel.
- Opérateurs de manipulation des structures des données
- Règles d'intégrité liées aux contraintes induites par ces structures.

2.2. SGBD

Le schéma conceptuel est indépendant de l'implémentation dans un moteur de Base de Données (SGBD).

Il permet la création d'un réservoir de données (data pool).

Pour gérer ce contenu (définir, contrôler et manipuler les données), on utilise un SGBD.

3. Propriétés ACID

Dans le domaine des bases de données, une opération sur les données est appelée une transaction. Par exemple, un transfert de fonds d'un compte de banque à un autre, même s'il implique plusieurs actions comme le débit d'un compte et le crédit d'un autre, est une seule transaction.

Les propriétés ACID¹ sont un ensemble de propriétés qui garantissent qu'une transaction informatique est exécutée de façon fiable.

Jim Gray a défini les propriétés qui garantissent des transactions fiables à la fin des années 1970 et a développé des technologies pour les mettre en œuvre automatiquement.

1. Atomicité : La propriété d'atomicité assure qu'une transaction se fait au complet ou pas du tout : si une partie d'une transaction ne peut être faite, il faut effacer toute trace de la transaction et remettre les données dans l'état où elles étaient avant la transaction. L'atomicité doit être respectée dans toutes situations, comme une panne d'électricité, une défaillance de l'ordinateur, ou une panne d'un disque magnétique.
2. Cohérence : La propriété de cohérence assure que chaque transaction amènera le système d'un état valide à un autre état valide. Tout changement à la base de données doit être valide selon toutes les règles définies, incluant mais non limitées aux contraintes d'intégrité, aux rollbacks² en cascade, aux déclencheurs de base de données, et à toutes combinaisons d'événements.
3. Isolation : Toute transaction doit s'exécuter comme si elle était la seule sur le système. Aucune dépendance possible entre les transactions. La propriété d'isolation assure que l'exécution simultanée de transactions produit le même état que celui qui serait obtenu par l'exécution en série des transactions. Chaque transaction doit s'exécuter en isolation totale : si T1 et T2 s'exécutent simultanément, alors chacune doit demeurer indépendante de l'autre.
4. Durabilité : La propriété de durabilité assure que lorsqu'une transaction a été confirmée, elle demeure enregistrée même à la suite d'une panne d'électricité, d'une panne de l'ordinateur ou d'un autre problème. Par exemple, dans une base de données relationnelle, lorsqu'un groupe d'énoncés SQL ont été exécutés, les résultats doivent être enregistrés de façon permanente, même dans le cas d'une panne immédiatement après l'exécution des énoncés.

4. La création d'une base de données

La construction d'une base de données commence par une phase d'analyse. Cela se fait au moyen de méthodes d'analyse comme Merise³ ou UML⁴. Cette analyse conceptuelle permet de se représenter le fonctionnement du système d'informations pour lequel on cherche à construire une base de données.

L'ANSI SPARC⁵ propose différentes approches de schéma de données qui ont été reprises par ces méthodes :

1 atomicité, cohérence, isolation et durabilité

2 méthode permettant d'annuler l'ensemble des requêtes que l'on vient de réaliser (le fait inverse du commit)

3 http://fr.wikipedia.org/wiki/Merise_%28informatique%29

4 http://fr.wikipedia.org/wiki/UML_%28informatique%29

5 http://fr.wikipedia.org/wiki/Architecture_Ansi/Sparc

- « schéma conceptuel »,
- « schéma logique »
- « schéma physique »

4.1. Modèle Conceptuel des Données (MCD)

Le Modèle Conceptuel des Données (ou Modèle entité-association), permet de représenter la structure du système d'information, du point de vue des données, et définit également les dépendances ou relations entre ces différentes données.

Les concepts de base du modèle conceptuel de données sont : l'entité, l'association, la propriété et les cardinalités.

4.1.1. propriété

La propriété peut être définie comme une donnée élémentaire ou atomique. Appelées également attributs ou colonnes (dans le modèle relationnel), les propriétés servent à décrire les entités et les associations. Elles prennent des valeurs appelées occurrences de la propriété.

Exemples : Pour les propriétés NomEmployé et dateEmbauche, on peut avoir les occurrences suivantes

Propriété	Occurrence
NomEmployé	Dupont
	Pierre
	Joseph
DateEmbauche	10-20-2002
	25-05-2006

4.1.2. Entités

Une entité peut être définie comme un objet concret ou abstrait du monde réel perçu :

- pourvu d'une existence propre
- conforme aux besoins de gestion de l'entreprise

Exemple : Exploitant, parcelle.

Les entités possèdent un ensemble de propriétés. Par exemple, si l'on considère l'entité "Exploitant", les propriétés peuvent être : NoExploitant, nomExploitant , prenomExploitant, dateNaissanceExploitant, etc.

Une occurrence de l'entité est composée d'une occurrence de chacune de ses propriétés. Par exemple une occurrence de l'entité exploitant peut être: 002 Pierre Robenson 20-04-1977 .

Dans une entité, une propriété joue un rôle particulier. Il s'agit de l'identifiant (encore appelée la clef). Ses valeurs sont discriminantes. Deux occurrences distinctes de l'entité ne peuvent avoir même valeur pour la propriété identifiante. Elle est généralement soulignée.

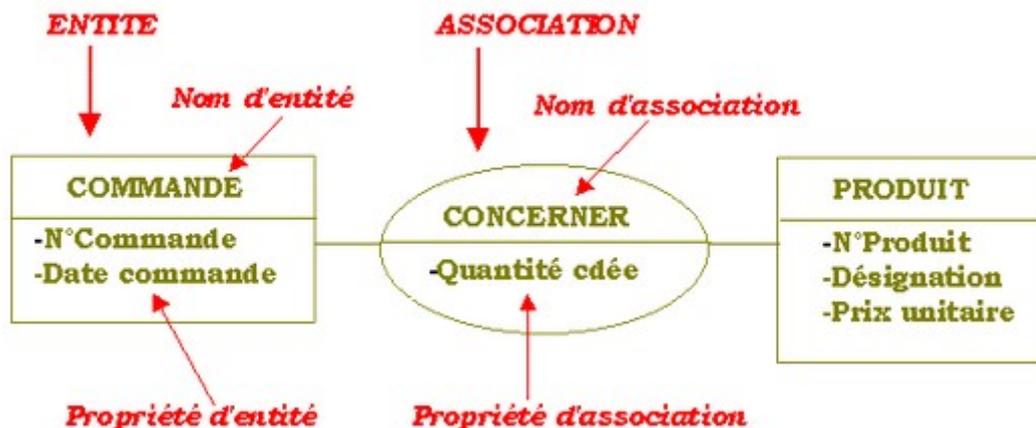
4.1.3. Association

L'association est un lien sémantique entre une ou plusieurs entités : l'association peut être réflexive,

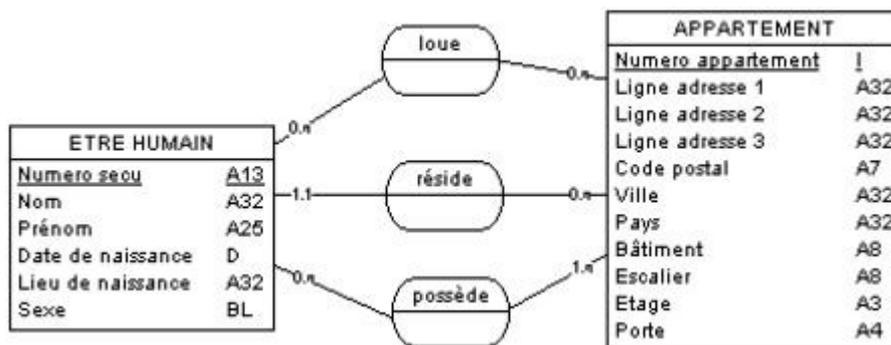
de préférence binaire, parfois ternaire, voire de dimension supérieure. Elle peut également être porteuse d'une ou plusieurs propriétés.

Les associations se représentent dans une ellipse (ou un rectangle aux extrémités rondes), reliée par des traits (pattes) aux entités qu'elles lient logiquement.

Exemple:



La **dimension** d'une association est le nombre d'entités concernées par celle-ci (nombre de pattes de l'association).



Exemple : l'association loue (entre "Être Humain" et "appartement") est de dimension 2 (binaire).

On distingue trois **types de liaisons** entre deux entités X et Y participant à l'association:

- Liaison de type 1 à 1 : à toute occurrence de X correspondant une et une seule occurrence de Y et réciproquement. Par exemple: un homme est marié à une femme et réciproquement
- Liaison de type 1 à plusieurs (1 à n) : à toute occurrence de X correspondant une et plusieurs occurrences de Y et à toute occurrence de Y une seule de X. Exemple : Un "être humain" réside dans un appartement, mais dans un appartement, on peut avoir plusieurs "être humain".
- Liaison de type plusieurs à plusieurs (n à m) : à toute occurrence de X correspondant une et plusieurs occurrences de Y et réciproquement. Exemple : Un "être humain" peut posséder plusieurs appartements et un appartement peut être possédé par plusieurs "être humain".

La notion de **cardinalité** minimum/maximum est liée aux type de liaison inter-entités.

- La cardinalité minimum est le nombre minimum d'occurrences d'une entité X dans l'association considérée. Par exemple, dans la relation "être Humain" réside dans un appartement. Du côté "être Humain" la cardinalité minimum est 1 et celle du côté

appartement est 0.

- La cardinalité maximum est le nombre maximum d'occurrences d'une entité X dans l'association considérée .Par exemple, dans la relation "être Humain" réside dans un appartement. Du côté "être Humain" la cardinalité maximum est 1 et celle du côté appartement est n.

Par conséquent, les valeurs de cardinalités sont en générales 0, 1, n.

Valeur	Définition
0	Une occurrence de l'entité peut exister sans participer à l'association
1	Une occurrence de l'entité participe nécessairement au moins une fois à une occurrence d'association

Valeur	Définition
1	Une occurrence de l'entité participe au plus une fois
N	Une occurrence de l'entité peut participer plusieurs fois

Combinaisons possibles :

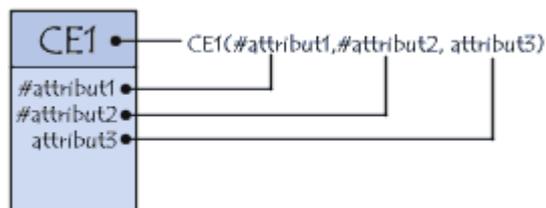
0,1	Une occurrence participe au moins 0 fois et au plus 1 fois à l'association
1,1	Une occurrence participe exactement 1 fois à l'association
0,N	Une occurrence peut ne pas participer ou participer plusieurs fois
1,N	Une occurrence participe au moins 1 fois, voire plusieurs

4.2. Modèle logique des données (MLD)

Le modèle logique des données consiste à décrire la structure de données utilisée sans faire référence à un langage de programmation. Il s'agit donc de préciser le type de données utilisées lors des traitements.

Ainsi, le modèle logique est dépendant du type de base de données utilisé.

Chaque classe d'entité du modèle conceptuel devient une table dans le modèle logique. Les identifiants de la classe d'entité sont appelé clés de la table, tandis que les attributs standards deviennent des attributs de la table, c'est-à-dire des colonnes.



Le passage du modèle conceptuel au modèle logique au niveau des classes de relation se fait selon les cardinalités des classes d'entité participant à la relation :

- si une des classes d'entités possède une cardinalité faible, la table aura comme attributs, les attributs de la classe ayant une cardinalité faible, puis le (ou les) attribut(s) de relation et enfin les attributs de la seconde classe précédé du nom de la classe
- si les deux classes d'entités possèdent une cardinalité forte, la table aura comme attributs, les attributs des deux classes de relation précédés des noms des classes respectives, puis le (ou les) attribut(s) de relation

4.3. Modèle Physique des Données (MPD)

Le Modèle Physique des Données peut être composé soit de tableaux décrivant le schéma, soit du code SQL de création du schéma : le code SQL est une déduction directe du MPD en fonction du SGBD choisi et de sa version.

Il s'agit simplement de décrire le type (chaîne de caractères, numérique...) de chacun des champs des tables.

Exemple :

Table : user

Nom	Type	Clef
id	numérique	primaire
login	chaîne	unique
password	chaîne	
website	chaîne	

Table : message

Nom	Type	Clef
id	numérique	primaire
user_id	numérique	étrangère
title	chaîne	
body	texte	