



BDD

Base De Données
C'est la base, les données !

Par William Cloudless

Table des matières

Base de données	3
Définition	3
Application.....	4
Système d'informations (SI).....	4
Modèle Conceptuel de données (MCD)	4
Modèle Logique des Données (MLD) relationnel.....	6
Algèbre relationnelle	8
Opérateurs de base	8
Union	8
Différence	8
Produit Cartésien.....	9
Projection	9
Restriction.....	9
Opérateurs construits.....	10
Intersection.....	10
Quotient	10
Jointure.....	11
Opérateur de calcul	11
Compte	11
Somme.....	11
Autre.....	11

Base de données

Définition

Pour gérer une base de données, on utilise un SGBD (Système de Gestion de Base de Données)

Il permet :

- La persistance de l'information
- La structuration de l'information
- Permet la manipulation d'une grande masse de d'informations
- Permet plusieurs utilisateurs en simultanés
- Possède une interface entre l'utilisateur et la mémoire

Avec un SGBD, il faut retenir les points suivants :

1. On ne s'occupe pas de tous ce qui matériel
2. On doit pouvoir faire un système qui peut évoluer sans tout modifier pour le faire fonctionner
3. Un non-informaticien peut pouvoir l'utiliser sans problème avec l'utilisation des langages de haut niveau, déclaratifs
4. L'accès aux données doit être efficace avec l'utilisation des meilleurs algos de recherche
5. L'administration des données est centralisée
6. Il ne doit pas y avoir de la redondance dans le stockage des données
7. Il y a la possibilité d'accéder en simultanés aux données
8. Les données sont sécurisées avec la gestion de droits et autorisations
9. Il doit y avoir des cohérences dans les données avec des règles de gestion à respecter

Il existe 5 grands types de bases de données :

- **Les bases hiérarchiques et réseau**
Base navigationnelles (exemple : IMS/IBM, IDS II/Bull)
- **Les bases relationnelles (les plus utilisées)**
Base sous forme de tables – Algèbre relationnelle (SQL)
- **Les bases déductives**
Base sous forme de prédicts (DataLog)
- **Les bases objets**
Bases avec description de classes et héritage (exemple : O2, Gemstone, Realm)
- **Les bases noSQL**
données non structurées (exemple : MongoDB, Cassandra)

Dans ce document, on parlera uniquement des bases relationnelles.

Application

Système d'informations (SI)

Définition

Le SI (Système d'informations) représente l'ensemble du service que vous voulez mettre en place pour informatiser quelque chose. Il permet de collecter, stocker, traiter et distribuer de l'information.

Exemple : Système d'information de contrôle du trafic aérien. Celui-ci contient les avions, les vols, les employés, etc... On peut donc ajouter des avions, modifier des vols, supprimer des employés, afficher les vols par employés, etc...

Modélisation

Pour modéliser une SI, il faut bien penser aux tenants et aboutissants de ce que l'on veut réaliser. Il faut réfléchir de manière structurée ce que l'on veut mettre dans les entités et quelle sont les associations les plus adaptées. Certaines méthodes sont possibles, exemple : [Merise](#), [Axial](#).

Modèle Conceptuel de données (MCD)

Principe

Un MCD permet une description du SI à l'aide d'entités et d'associations. Celui-ci est contextuel, il n'est valable que dans une situation donnée. Le MCD ne peut se faire qu'après la modélisation du SI.

Définition des éléments du MCD

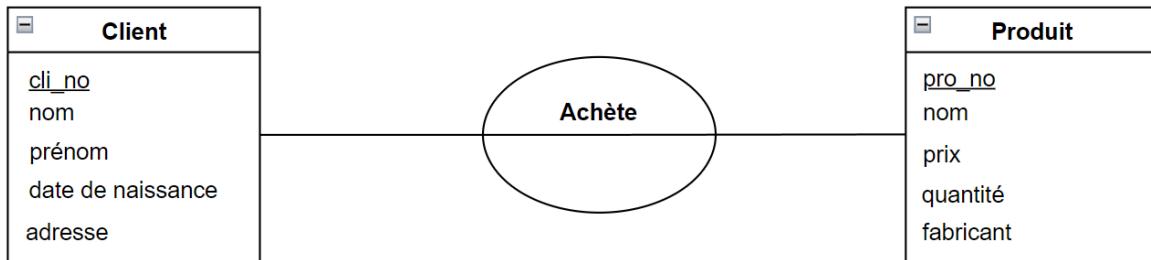
- **Propriété** : C'est la donnée que vous introduisez dans votre SI. Exemple : le nom d'un employé, la date de naissance d'un client, etc...
- **Dictionnaire de données** : Ensemble de toutes les propriétés du SI.
- **Entité** : Objet regroupant un ensemble de propriété de manière cohérente.
Exemple :

Client
nom
prénom
date de naissance
Adresse

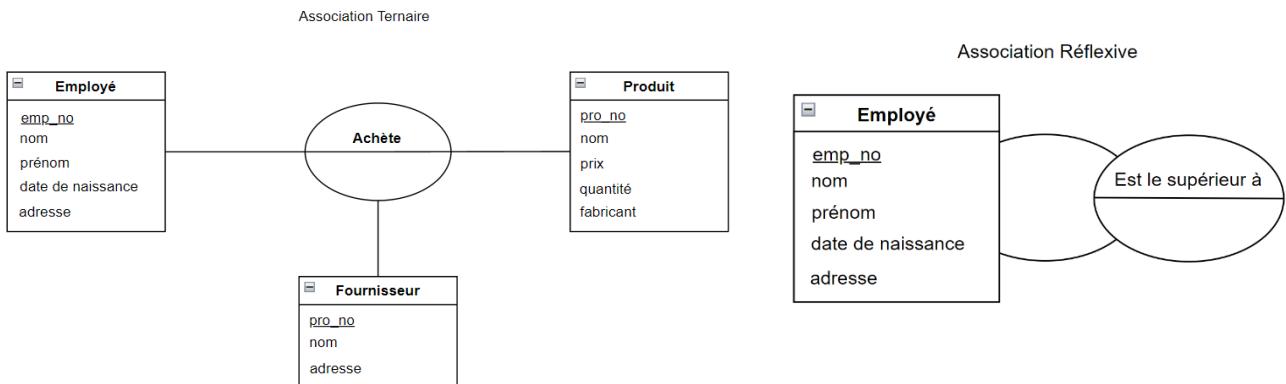
- **Entité forte** : Entité qui indiscutablement nécessaire dans l'existence de la SI.
- **Identifiant** : Propriété d'une entité qui permet d'identifier une instance de l'entité. Il est unique, deux instances ne peuvent pas avoir la même valeur d'identifiant.

Client
cli_no
nom
prénom
date de naissance
adresse

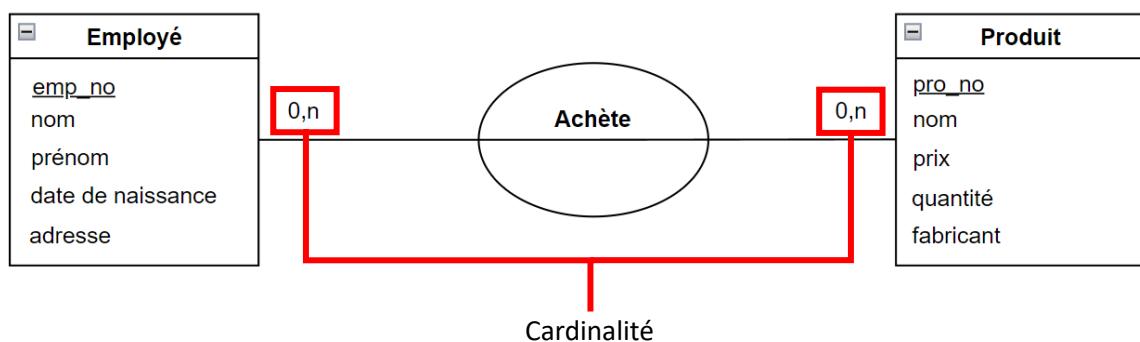
- **Association** : Lien entre différentes entités. Exemple (association binaire) :



Une association peut porter sur plusieurs entités. Exemple :



- **Cardinalité** : Quantification de la participation d'une entité à une association. Elle composé d'une borne minimale et d'une borne maximale. Exemple :



La borne minimale (chiffre à droite) désigne le nombre de fois minimum qu'une occurrence d'une entité participe aux occurrences de l'association. Il est généralement égal à 0 ou 1.

La borne maximal (chiffre à gauche) désigne le nombre de fois maximum qu'une occurrence d'une entité participe aux occurrences de l'association. Il est généralement égal à 1 ou n (n représente un naturel inconnu supérieur à 0).

- **Lien hiérarchique** : Association dont les cardinalités maximales sont 1 d'un coté et n de l'autre.
- **Lien maillé** : Association dont les cardinalités maximales sont n des deux côtés.

Règles pour modéliser

- Il ne doit pas avoir deux fois la même propriété dans un MCD
- Les attributs ne doivent pas avoir de dépendances avec d'autre attributs dans une entité.
- Une association maillée implique que la réunion des identifiants participants est unique.

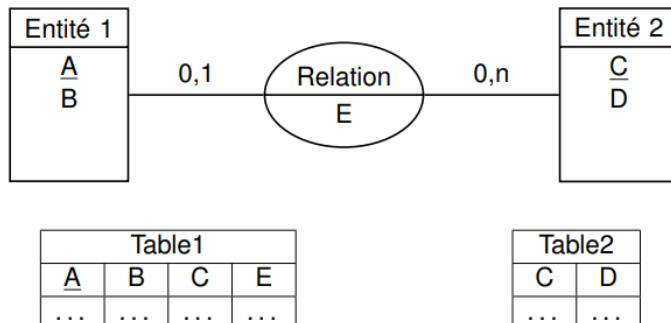
Modèle Logique des Données (MLD) relationnel

Dans la logique relationnelle, on transforme les entités en tables qui n'ont pas de lien physique entre elles puis les propriétés deviennent une colonne et l'identifiant d'une entité devient la clé primaire de la table correspondante. Cette logique est fidèle à l'algèbre relationnelle et correspond à la relation de la théorie des ensembles. On le manipule avec le SGBDR.

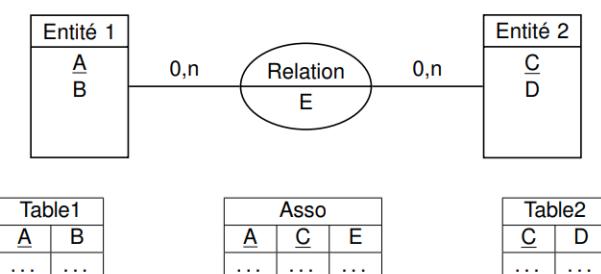
Pour chaque entité, on le met dans une table à qui on associe un nom. Dans cette table, on associe un type et un nom pour chaque colonne. Exemple :

Personne		
Pno	Nom	Prenom
1	lefebvre	christian
2	martin	franck
3	durand	franck

Une association avec un lien hiérarchique entraîne la migration d'une clé étrangère (0,n) vers la table de l'entité côté (0,1). Exemple :



Une association avec un lien maillé donne naissance à une nouvelle table contenant chacune des clés ainsi que les propriétés portées par l'association. Exemple :



Pour représenter une table, on applique le schéma suivant :

Personne(Pno,Nom,Prenom)

Pour représenter une base de données, on met à la suite toutes les tables de celle-ci. Exemple :

Fournisseurs(fno,nom,adresse,ville)

Produits(pno,libellé,prix,poids,couleur)

Commandes(cno,fno,pno,quantité)

Algèbre relationnelle

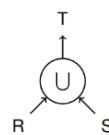
Introduit par Edgar F. Codd en 1970, il permet l'expression de requêtes permettant d'isoler des données. Il y a deux types d'opérateurs : Unaires (enlever des données) et binaires (regrouper les données de plusieurs tables). Il y aussi 3 familles d'opérateurs : les opérateurs de base, les opérateurs évolués et les opérateurs de calculs.

Opérateurs de base

Union

Cet opérateur permet d'ajouter les lignes deux tables dans une seule à la condition que les deux tables ont le même schéma.

On note $T = (R \cup S)$ ou $T = \text{union}(R, S)$ ou



. Exemple :

R	A	B	C
a	b	c	
d	a	f	
c	b	d	

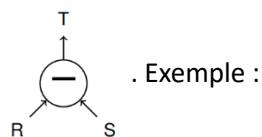
S	A	B	C
b	g	a	
d	a	f	

$R \cup S$	A	B	C
a	b	c	
d	a	f	
c	b	d	
b	g	a	

Différence

Cet opérateur permet de supprimer les lignes contenu dans une table par les lignes d'une autre table à la condition que les deux tables ont le même schéma.

On note $T = (R - S)$ ou $\text{minus}(R, S)$ ou



. Exemple :

R	A	B	C
a	b	c	
d	a	f	
c	b	d	

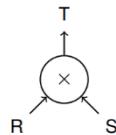
S	A	B	C
b	g	a	
d	a	f	

$R - S$	A	B	C
a	b	c	
c	b	d	

Produit Cartésien

Le produit cartésien de deux tables de schéma quelconque est une table qui contient toutes les combinaisons possibles de lignes de ces deux tables.

On note $T = (R * S)$ ou $T = \text{product}(R, S)$ ou



. Exemple :

R	A	B	C
a	b	c	
d	a	f	
c	b	d	

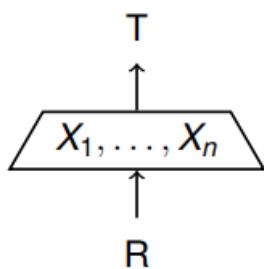
S	D	E	F
b	g	a	
d	a	f	

$R * S$	A	B	C	D	E	F
a	b	c	b	g	a	
a	b	c	d	a	f	
d	a	f	b	g	a	
d	a	f	d	a	f	
c	b	d	b	g	a	
c	b	d	d	a	f	

Projection

La projection permet récupérer les colonnes mis en paramètre.

On note $T = \pi_{X_1, \dots, X_n}(R)$ ou $T = \text{proj}_{X_1, \dots, X_n}(R)$



Restriction

La restriction permet de supprimer les lignes qui ne répondent pas à une condition indiquée en paramètre.

On note $T = \sigma_Q(R)$ ou $T = \text{select}_Q(R)$



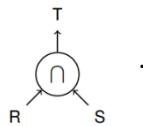
Opérateurs construits

Intersection

L'intersection permet de récupérer les lignes qui sont à la fois dans une table et dans une autre table de même schéma.

On note $T = (R \cap S)$ ou $T = \text{inter}(R,S)$ ou

Exemple :



R	A	B	C
a	b	c	
d	a	f	
c	b	d	

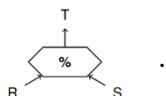
S	A	B	C
b	g	a	
d	a	f	

$R \cap S$	A	B	C
d	a	f	

Quotient

Le quotient d'une table par une sous-table est une table de toutes les combinaisons des lignes de la sous-table avec les lignes de la table originale qui sont valides.

On note $T = (R/S)$ ou $T = \text{div}(R,S)$ ou



Exemple :

R	A	B	C	D
a	b	c	d	
a	b	e	f	
b	c	e	f	
e	d	c	d	
e	d	e	f	
a	b	d	e	

S	C	D
c	d	
e	f	

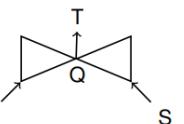
R/S	A	B
a	b	
e	d	

Jointure

La jointure consiste à faire un produit cartésien de 2 tables qui, par la suite, satisfait une qualification indiquée en paramètre.

Il y a différents types de jointure :

- **L'équi-jointure** : C'est une jointure avec une qualification d'égalité entre deux colonnes.
- **La jointure naturelle** : C'est une équi-jointure sur tous les attributs de même nom suivie de la projection qui permet de conserver un seul de ces attributs égaux de même nom.
- **L'auto-jointure** : C'est une jointure d'une table sur elle-même.



On note $T = (R \bowtie_Q S)$ ou $T = \text{join}_Q(R, S)$ ou $R \bowtie_Q S$.

Exemple :

R	A	B	C		S	D	E
1	2	3			3	1	
4	5	6			6	2	
7	8	9					

$R \bowtie_{B < D} S$	A	B	C	D	E
1	2	3		3	1
1	2	3		6	2
4	5	6		6	2

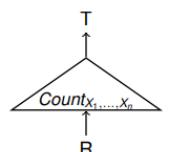
Opérateur de calcul

Compte

Cet opérateur permet de compter le nombre de ligne qui ont une même valeur d'attributs en commun.

Si aucun paramètre de regroupement n'est précisé alors l'opérateur renvoie uniquement le nombre de lignes de la table.

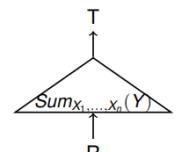
On note $T = \text{Compte}_{x_1, \dots, x_n}(R)$ ou $T = \text{Count}_{x_1, \dots, x_n}(R)$ ou



Somme

L'opérateur Somme additionne les valeurs d'une colonne, ou les valeurs de plusieurs colonnes regroupées par valeur.

On note $T = \text{Somme}_{x_1, \dots, x_n}(R, Y)$ ou $T = \text{Sum}_{x_1, \dots, x_n}(R, Y)$ ou



Autre

Il existe d'autre opérateur tel que moyenne, minimum et maximum qui fonctionne comme somme.