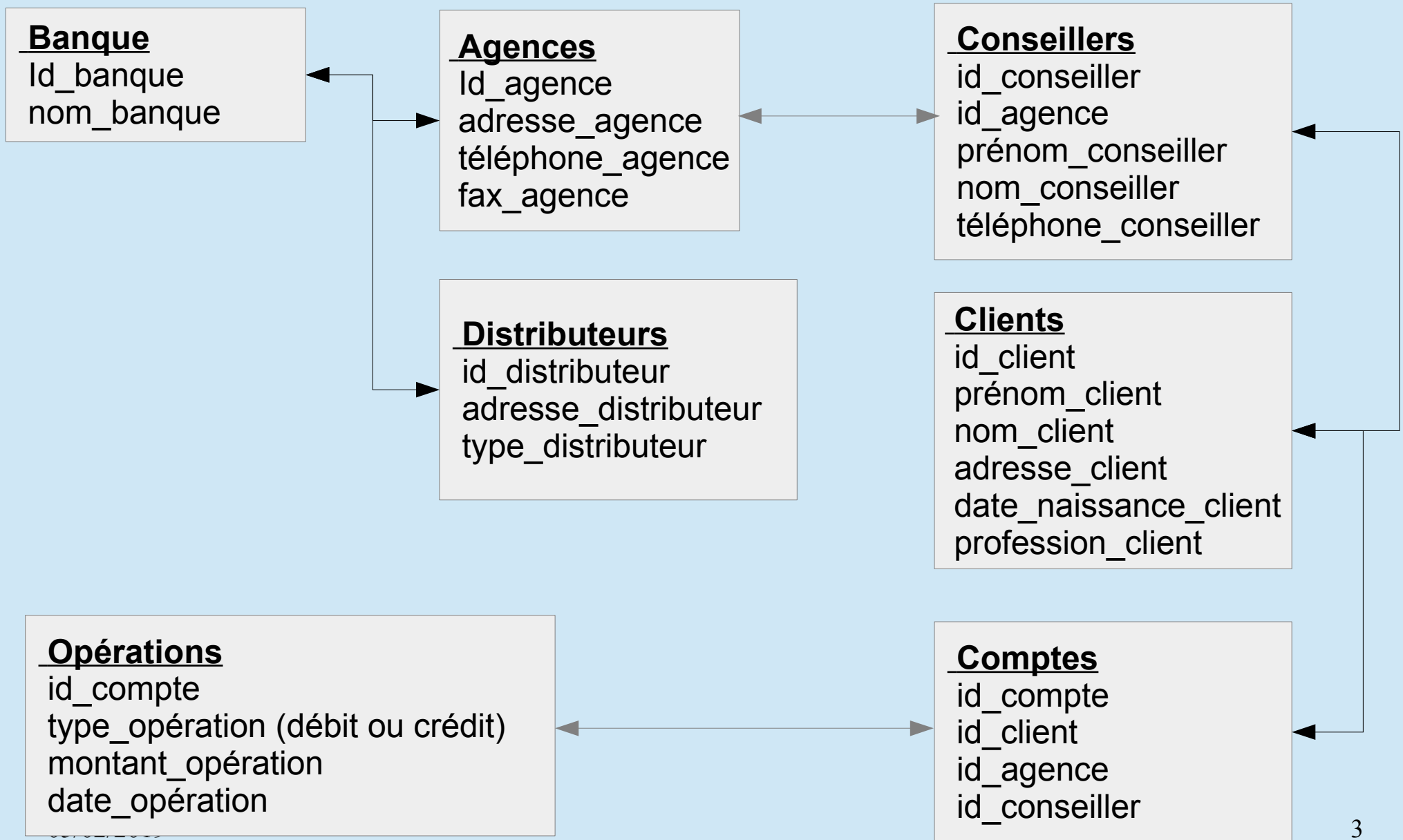


# **Introduction aux bases de données**

# **Exemple classique : Gestion de banque**

# Exemple classique : Gestion de banque



# Exemple classique : Gestion de banque

On a besoin de modéliser l'information de l'activité bancaire dans une base de données avec plusieurs tables comme :

- Table des agences
- Table des distributeurs
- Table des conseillers
- Table des clients
- Table des comptes bancaires
- Tables des opérations
- ...

# Exemple classique : Gestion de banque les besoins

- **Besoins de description** : Pouvoir décrire les données de l'application (clients, comptes, opérations..)
- **Besoins de création/modification des données** :
  - Pouvoir créer la base de données initiale avec les données représentant la banque
  - Pouvoir créer au fur et à mesure les données des clients et les nouvelles opérations bancaires de chaque jour.
  - Pouvoir modifier (et éventuellement supprimer) toute donnée déjà rentrée

# Exemple classique : Gestion de banque les besoins

**Besoins d'interrogation** : Pouvoir répondre à toute demande d'information portant sur les données contenues dans la base. Par exemple:

- C'est quoi le montant de la dernière opération du compte numéro 1098712333653 ?
- C'est qui le conseiller du client « John Doe » ?
- C'est quoi la somme totale des opérations du mois de janvier 2018 du compte 1098712333653 ?
- C'est quoi le nombre de clients de l'agence de Créteil ?
- C'est quoi le nombre d'agences du département 94 ?

# Exemple classique : Gestion de banque les besoins

**Besoins d'exactitude / cohérence:** Pouvoir exprimer toutes les règles imposées sur les données de façon à éviter toute erreur qui peut être détectée. Par exemple:

- Il ne faut jamais donner le même numéro de compte à 2 clients différents
- La date d'une opération bancaire doit correspondre à un vrai jour (pas 29 février 2019, 31 mars 2017,...)
- Chaque client doit avoir absolument un conseiller
- Une opération ne doit pas dépasser les plafond fixées pour le client

# Exemple classique : Gestion de banque les besoins

## Besoins de garanties :

- Il ne faut pas que les informations (par exemple, les opérations bancaires) soient perdues à cause d'un dysfonctionnement quelconque:
  - Erreur de programmation,
  - Panne système,
  - Panne de l'ordinateur,
  - Coupure de courant, ...
  - Donc il faut garantir la fiabilité de la BD
- Il ne faut pas qu'une action faite pour un client (par exemple, un virement) soit perdu du fait d'une autre action faite simultanément (par exemple un prélèvement).



# Exemple classique : Gestion de banque les besoins

## Besoins de confidentialité :

On a besoin que toute information doit pouvoir être protégée contre l'accès par des utilisateurs non autorisés soit pour

- la lecture des données (consultation)
- ou l'écriture des données (modification / suppression)

**Exemple** : Interdire par exemple aux clients de modifier le montant d'un virement après 24h.

# Exemple classique : Gestion de banque les besoins

## Besoin d'efficacité

Le temps de réponse du système doit être rapide et conforme aux besoins (par exemple pas plus de 3 secondes pour une consultation d'un compte)

Donc on a besoin de considérer pour les bases de données :

- mécanismes d'optimisation de temps de calcul et de l'espace de stockage
- répartition / duplication / réplication des données
- maître les données sur plusieurs sites (data centers)

# **Description des bases données**

# Description des bases données

## **BD= Base de Données**

C'est un ensemble de données utilisées par les applications des systèmes informatiques des entreprises.

## **Fonctionnalités d'une BD:**

- Une base données permet le stockage persistant / permanent des données .
- Contrairement à un système de fichier, la BD permet de stocker les données d'une manière typée et structurée.
- Dans une BD les données sont reliés entre elles.

# Description des bases données

## Fonctionnalités avancée d'une BD:

- Une BD est utilisé avec un langage très puissant SQL qui permet de :
  - Insérer de nouvelle données
  - Modifier ou supprimer des données existantes
  - Rechercher dans les données (même complexe)
- Une BD permet aussi de garantir la cohérence et la stabilité des données stockées
- Une base données permet aussi à plusieurs utilisateurs d'utiliser la base de données au même temps : on appelle ceci l'**Accès concurrent aux données**.

# Exemples réels

## Exemples de systèmes de BD :

Les bases de données relationnelles sont utilisées pour

- stocker des informations administratives (**la préfecture**)
- les systèmes de bibliothèque informatisés (**UPEC**),
- les systèmes de réservation de vol (**opodo**),
- les systèmes de réservation d'hôtel (**Ibis**),
- les systèmes informatisés d'inventaire (**rgis**)
- systèmes de gestion de contenu qui stockent des sites Web sous forme de collections de pages Web dans une base de données (**Wikipédia**)
- réseau sociaux (**Facebook**)
- commerce sur le web (**Amazon, ebay**)

# **Description des SGBD**

# Description des SGBD

**Système de Gestion de Bases de Données → SGBD**

Un système de gestion de base de données (**SGBD**) *relationnel* est un logiciel qui permet de manipuler le contenu des bases de données relationnelles.

**Exemple :**

- Oracle (Propriétaire)
- Microsoft SQL Server (Propriétaire)
- Teradata (Propriétaire)
- Microsoft Access (Version d'évaluation)
- MySQL (libre)
- SQLite (libre)
- PostgreSQL (libre)



# Description des SGBD

## Un système de gestion de base de données :

- Doit permettre un **usage** :
  - Efficace → performance de la lecture et l'écriture des données
  - Fiable → données accessible en toute circonstance même en cas de pannes de machine (duplication des serveurs, réplication des données)
  - Facile → l'existence d'un langage de haut-niveau pour la manipulation des données
  - Sûr → garantir la cohérences des données avant et après la manipulation des données
  - Concurrence ou Multi-utilisateurs → la base de données peut être utilisée par plusieurs utilisateurs au même temps

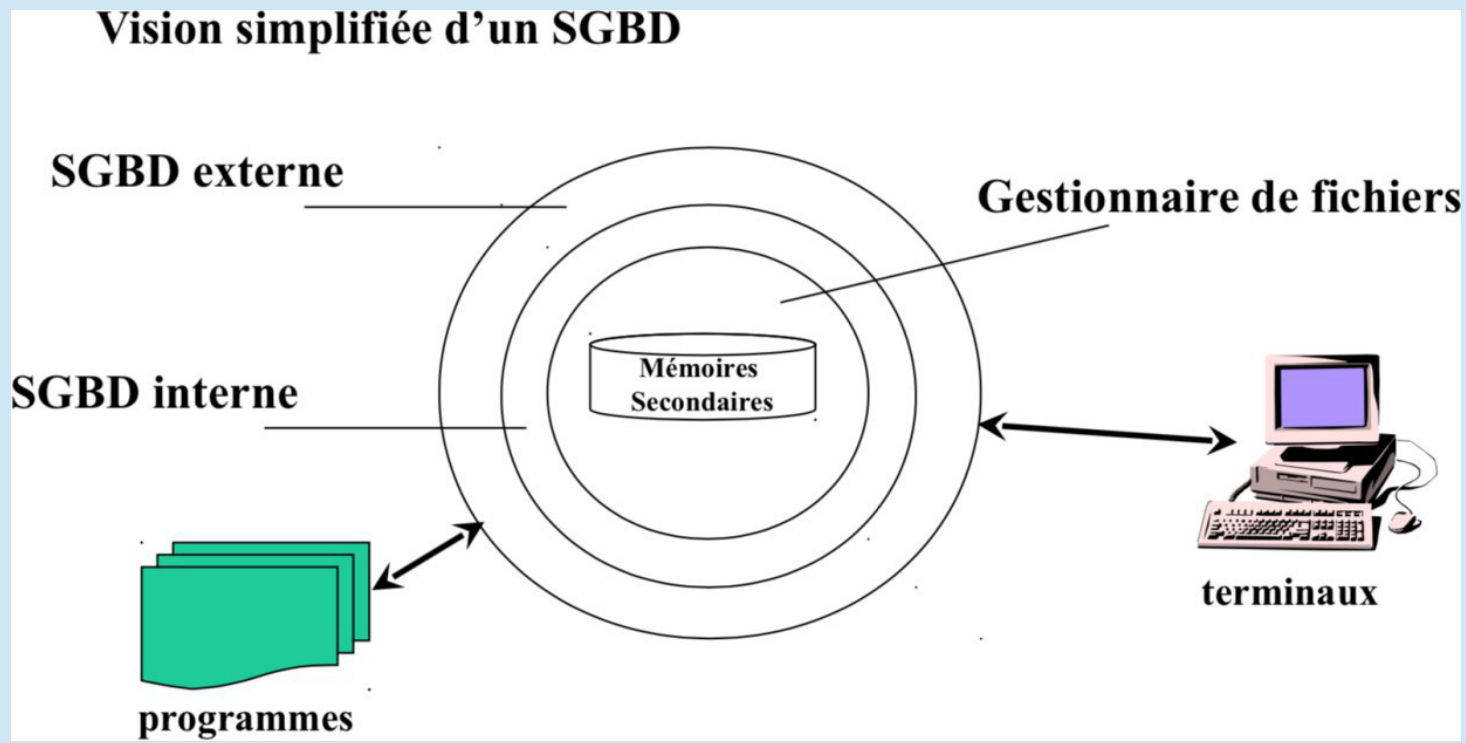
# Description des SGBD

**Un système de gestion de base de données :**

- Doit permettre un **stockage**:
  - Persistant → les données sont permanent dans la base de données
  - Massif → la base de données doit être capable de prendre en compte des tables de grandes tailles.

# Description des SGBD

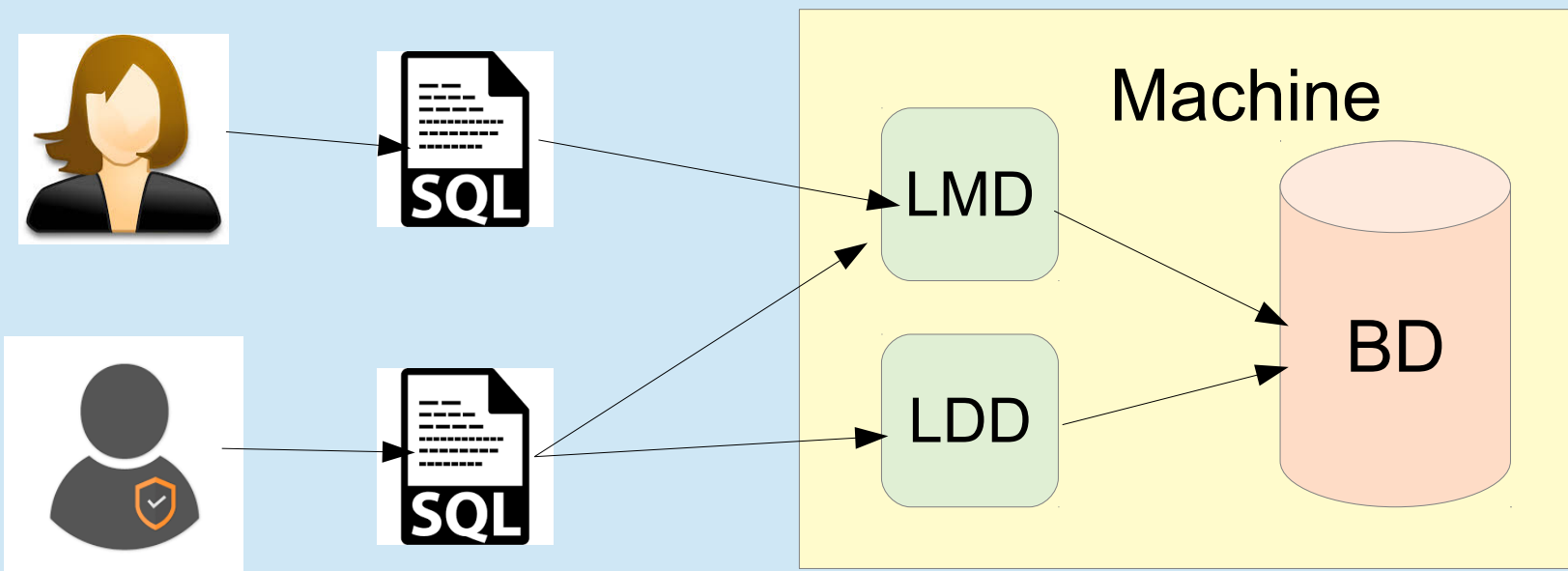
- **Mémoire secondaire** : c'est les disques dures
- **Gestionnaire de fichier** : c'est le système qui gère les fichiers qui contiennent la base de donnée
- **SGBD interne** : noyau de système de gestion de BD
- **SGBD externe** : couche qui permet la communication avec les utilisateurs



# Description des SGBD

Dans un SGBD on a deux types importants d'utilisateurs :

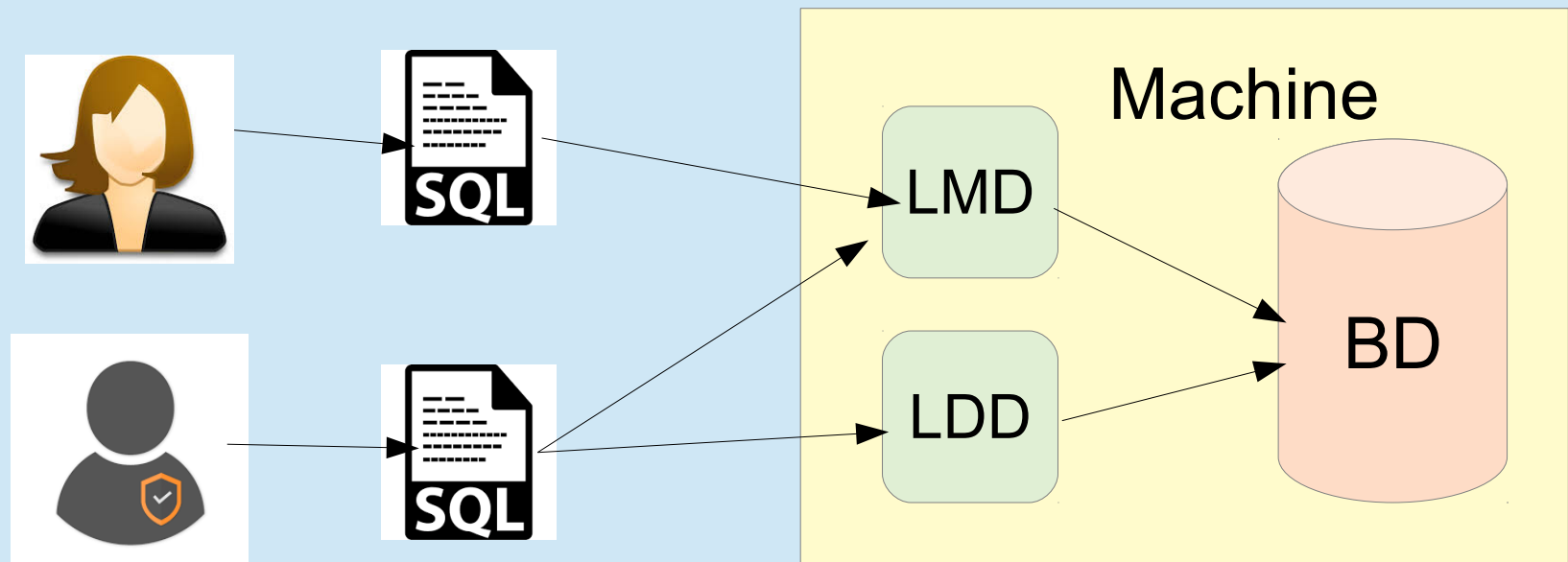
- Utilisateur conventionnel
- Administrateur



# Description des SGBD

**Utilisateur conventionnel:** c'est un utilisateur qui a le droit de manipuler les données en utilisant **LMD** ( **L**angage de **m**anipulation de **d**onnées).

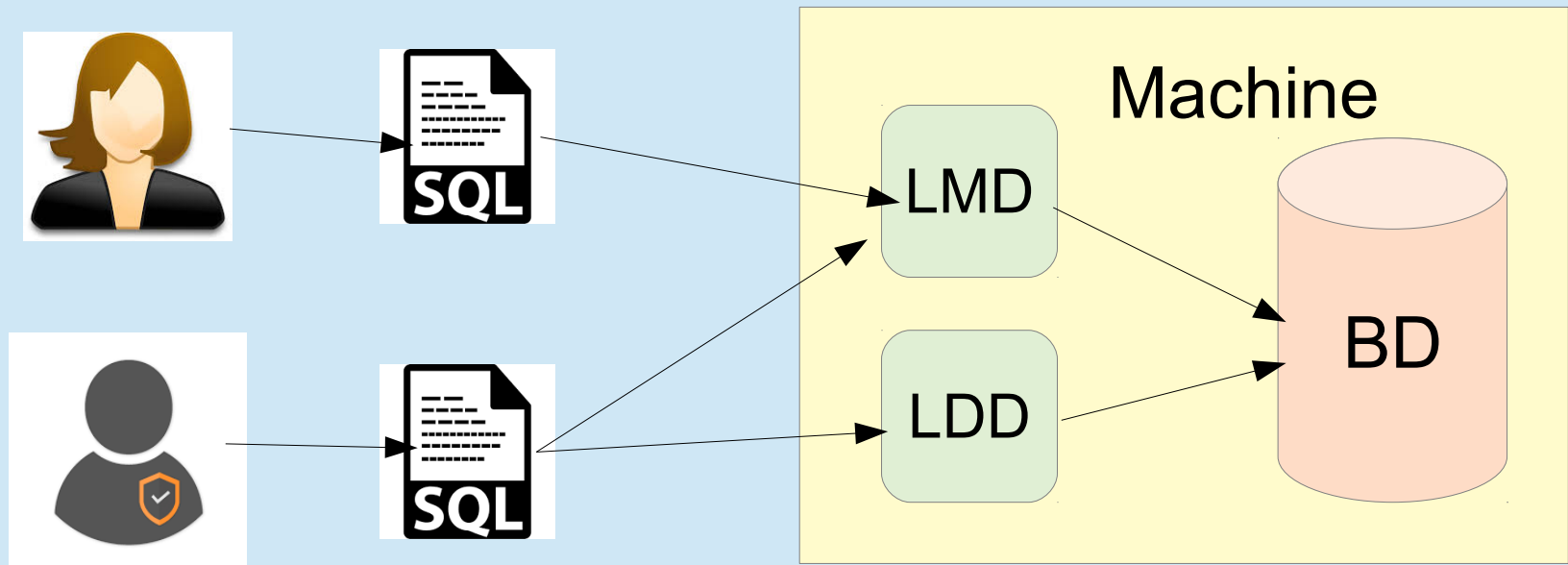
L'utilisateur écrit une requête qui exactement décrit ce qu'il veut faire comme opération, ensuite LMD traduit cette requête en opérations sur la machine qui contient la base de données.



# Description des SGBD

**Utilisateur administrateur:** c'est un utilisateur qui a le droit de manipuler les données et aussi de changer la structure des données en utilisant **LDD** ( **L**angage de **d**éfinition de **d**onnées).

L'administrateur écrit une requête qui décrit ce qu'il veut faire comme opération pour structurer la BD, ensuite LDD traduit cette requête en opérations sur la machine qui contient la BD :



# **Le modèle relationnel**

# Le modèle relationnel

## Les tableaux *relationnels* :

- Une **table** a un **nom** et contient :
  - Des lignes qu'on appelle **nuplets**
  - Des colonnes qu'on appelle **attributs**



# Exemple de BD relationnel

Nom de relation (table)

Attributs  
Champs  
Colonnes

Produit :

fabricant	modèle	type
A	1003	pc
A	1002	pc
A	1010	pc
A	2004	portable
A	2005	portable
A	2009	portable
B	1001	pc
B	1003	pc
B	3001	imprimante
B	3002	imprimante
B	3004	imprimante
B	1009	pc
C	1001	pc
C	2001	portable
C	1010	pc
C	2001	portable
C	2009	portable
C	3001	imprimante
D	1002	pc
D	1010	pc
D	2002	portable
D	2005	portable
E	2001	portable
E	2009	portable
E	3001	imprimante
E	1002	pc
E	1004	pc
F	2002	portable
F	3005	imprimante
F	3006	imprimante
F	3009	imprimante
F	2009	portable

nuplets  
Lignes

PC :

modèle	processeur	RAM	DD	DVD/Blu-ray	prix
1001	i3	2	400	8xDVD	400
1002	i3	4	450	16xDVD	500
1003	i5	4	500	8xBR	450
1004	i5	8	400	16xDVD	600
1005	i7	8	600	6xBR	650
1006	i7	16	800	8xDVD	700
1007	a4	4	500	16xDVD	350
1008	a4	8	800	6xBR	450
1009	a6	8	600	7xBR	500
1010	a8	16	800	12xBR	700

Portables :

modèle	processeur	RAM	DD	écran	prix
2001	a4	4	200	14	600
2002	a6	4	250	13	700
2003	a6	6	300	14	750
2004	a8	4	200	15	700
2005	a8	6	300	14	650
2006	i3	16	200	11	700
2007	i3	4	300	13	650
2008	i5	8	200	15	850
2009	i5	8	400	13	800

Imprimantes :

modèle	type	couleur	prix
3001	laser	oui	250
3002	jet d'encre	oui	150
3003	laser	non	100
3004	laser	non	90
3005	multifonction	non	350
3006	jet d'encre	oui	120
3007	laser	oui	230
3008	multifonction	oui	450

# Le modèle relationnel

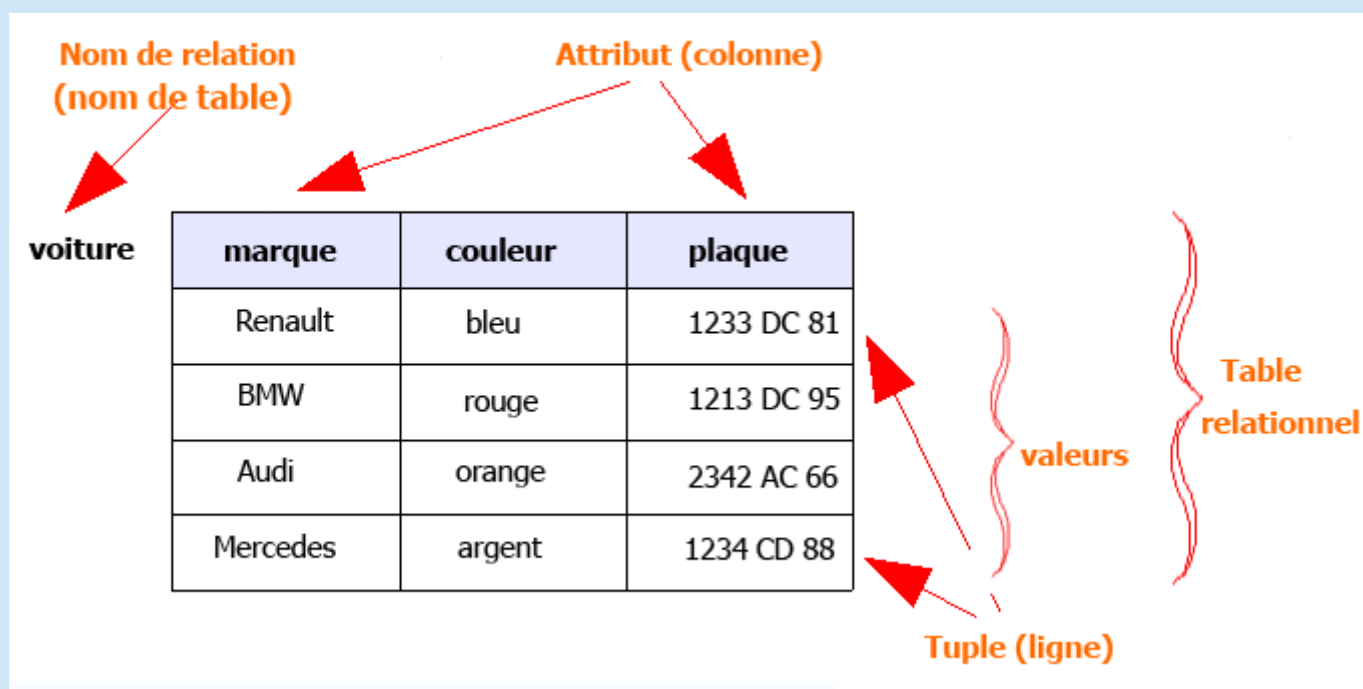
Aspects essentiels des **SGBD** *relationnels* :

- Toute **relation** est représentée par une **table**.
- Pour manipuler les données on utilise **calcul des relations**, qui est un ensemble réduit d'opérateurs et suffisant pour manipuler les tables et définir de nouvelles tables :
  - Projection → choix de certaine colonnes
  - Restriction → choix de certaine lignes
  - Jointure → croisement de plusieurs tables

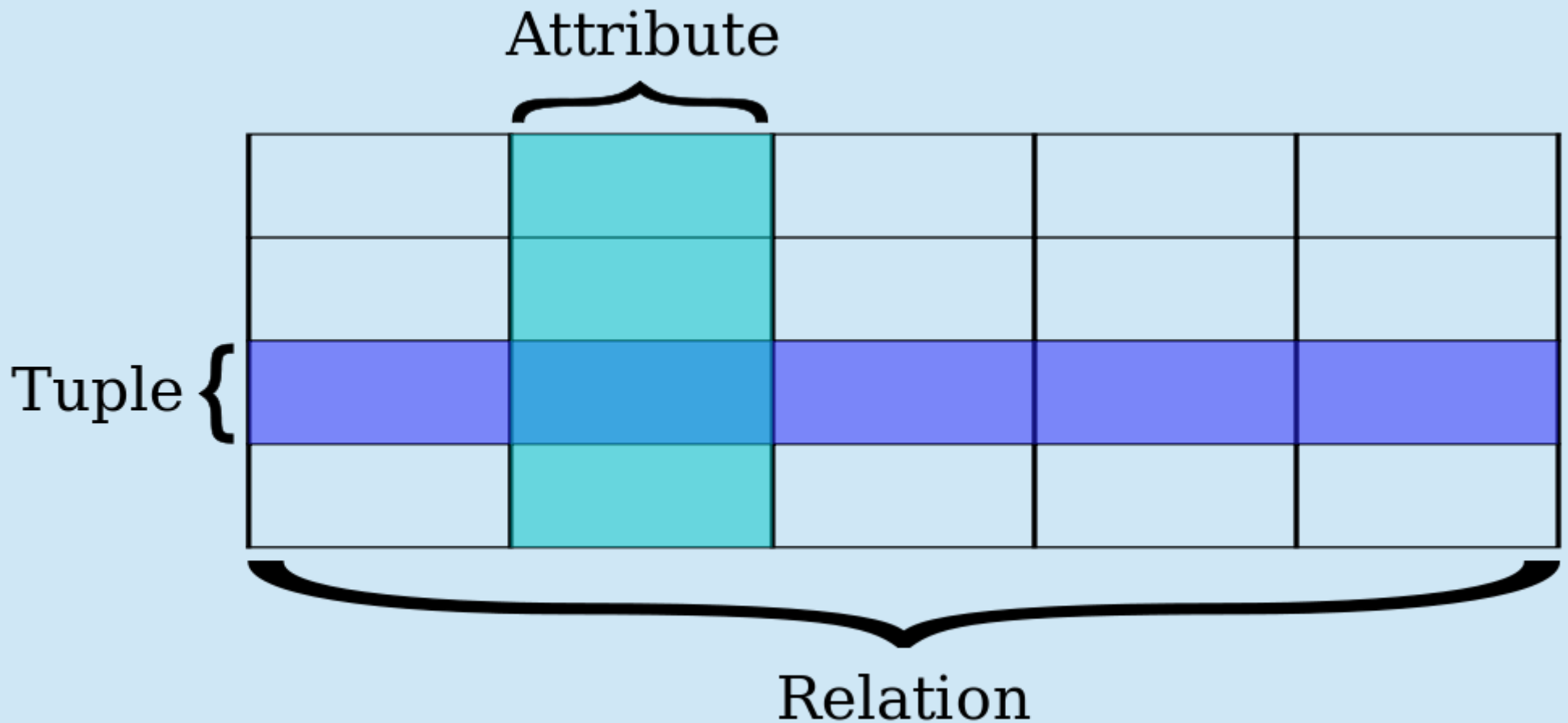
# **Algèbre relationnelle**

# Le modèle relationnel

- C'est un modèle de BD basé sur la manipulation des **tables** (tableaux), qu'on appelle des **relations**.
- Une **table** a un **nom** et contient :
  - Des lignes qu'on appelle **nuplets**
  - Des colonnes qu'on appelle **attributs**
  - L'ensemble des noms des attributs est appelé **schéma**



# Le modèle relationnel : à retenir



# Le modèle relationnel

- Toute **relation** est représentée par une **table**.
- Pour manipuler les données on utilise l'**algèbre relationnelle**, qui est un ensemble réduit d'opérateurs est suffisant pour manipuler les tables et définir de nouvelles tables.
- Opérateurs **ensemblistes** : opérations usuelles sur les ensembles :
  - Union
  - Intersection
  - Différence
- Opérateurs **relationnels**:
  - Projection → choix de certaine colonnes
  - Restriction → choix de certaine lignes
  - Jointure → croisement de plusieurs tables

# Algèbre relationnel : Opérateurs ensemblistes

**Union** : Soient A et B deux relations (tables) de même schéma (même attributs) alors **A U B** est la relation qui représente les lignes dans A ou B.

Relation étudiants

id	prénom	nom
303	Henri	Berger
304	Luigi	Zaffalon
305	Pierre	Breguet

Relation enseignant

id	prénom	nom
401	Assil	Guerid
402	Claire	Noirault

Relation étudiants **U** enseignant

id	prénom	nom
401	Assil	Guerid
402	Claire	Noirault
303	Henri	Berger
304	Luigi	Zaffalon
305	Pierre	Breguet

# Algèbre relationnel : Opérateurs ensemblistes

**Intersection:** Soient A et B deux relations (tables) de même schéma (même attributs) alors  $A \cap B$  est la relation qui représente les lignes qui sont au même temps dans A et B.

Relation étudiants\_Maths

id	prénom	nom
303	Henri	Berger
304	Luigi	Zaffalon
305	Pierre	Breguet

Relation étudiants\_Info

id	prénom	nom
401	Assil	Guerid
402	Claire	Noirault
304	Luigi	Zaffalon

Relation étudiants\_Maths  $\cap$  enseignant\_Info

id	prénom	nom
304	Luigi	Zaffalon



# Algèbre relationnel : Opérateurs ensemblistes

**Différence**: Soient A et B deux relations (tables) de même schéma (même attributs) alors **A - B** est la relation qui représente les lignes qui sont A mais pas dans B.

Relation L1\_Maths

id	prénom	nom
303	Henri	Berger
304	Luigi	Zaffalon
305	Pierre	Breguet

Relation L1\_Info

id	prénom	nom
401	Assil	Guerid
402	Claire	Noirault
304	Luigi	Zaffalon

Relation L1\_Maths - L1\_Info

id	prénom	nom
303	Henri	Berger
305	Pierre	Breguet

Relation L1\_Info - L1\_Maths

id	prénom	nom
401	Assil	Guerid
402	Claire	Noirault

# Algèbre relationnel : Opérateurs relationnels

**Restriction** appelé aussi **sélection**

Soit A une relations (table), et soit C une condition sur les attributs de A. La restriction  $\sigma_C(A)$  est la relation qui ne contient que les lignes de A qui vérifient la condition C.

Relation Clients

id	prénom	nom
401	Assil	Guerid
402	Claire	Noirault
303	Henri	Berger
304	Luigi	Zaffalon
305	Pierre	Breguet

Relation  $\sigma_{id=303}(Clients)$

id	prénom	nom
303	Henri	Berger

# Algèbre relationnel : Opérateurs relationnels

**Restriction** appelé aussi **sélection**

Soit A une relations (table), et soit C une condition sur les attributs de A. La restriction  $\sigma_C(A)$  est la relation qui ne contient que les lignes de A qui vérifient la condition C.

Relation Clients

id	prénom	nom
401	Assil	Guerid
402	Claire	Noirault
303	Henri	Berger
304	Luigi	Zaffalon
305	Pierre	Breguet

Relation  $\sigma_{id=401}(Clients)$

id	prénom	nom
401	Assil	Guerid

# Algèbre relationnel : Opérateurs relationnels

**Restriction** appelé aussi **sélection**

Soit A une relations (table), et soit C une condition sur les attributs de A. La restriction  $\sigma_C(A)$  est la relation qui ne contient que les lignes de A qui vérifient la condition C.

Relation Clients

id	prénom	nom
401	Assil	Guerid
402	Claire	Noirault
303	Henri	Berger
304	Luigi	Zaffalon
305	Pierre	Breguet

Relation  $\sigma_{id < 400}(Clients)$

id	prénom	nom
303	Henri	Berger
304	Luigi	Zaffalon
305	Pierre	Breguet

# Algèbre relationnel : Opérateurs relationnels

**Restriction** appelé aussi **sélection**

Soit A une relations (table), et soit C une condition sur les attributs de A. La restriction  $\sigma_C(A)$  est la relation qui ne contient que les lignes de A qui vérifient la condition C.

Relation Clients

id	prénom	nom
401	Assil	Guerid
402	Claire	Noirault
303	Henri	Berger
304	Luigi	Zaffalon
305	Pierre	Breguet

**Relation**  $\sigma_{\text{prénom}=\text{Henri}}(\text{Clients})$

id	prénom	nom
303	Henri	Berger

# Algèbre relationnel : Opérateurs relationnels

**Restriction** appelé aussi **sélection**

Soit A une relations (table), et soit C une condition sur les attributs de A. La restriction  $\sigma_C(A)$  est la relation qui ne contient que les lignes de A qui vérifient la condition C.

Relation Clients

id	prénom	nom
401	Assil	Guerid
402	Claire	Noirault
303	Henri	Berger
304	Luigi	Zaffalon
305	Pierre	Breguet

Relation

$\sigma_{nom - commence - par - 'B'}(Clients)$

id	prénom	nom
303	Henri	Berger
305	Pierre	Breguet

# Algèbre relationnel : Opérateurs relationnels

**Restriction** appelé aussi **sélection**

Soit A une relations (table), et soit C une condition sur les attributs de A. La restriction  $\sigma_C(A)$  est la relation qui ne contient que les lignes de A qui vérifient la condition C.

Relation Clients

id	prénom	nom
401	Assil	Guerid
402	Claire	Noirault
303	Henri	Berger
304	Luigi	Zaffalon
305	Pierre	Breguet

Relation $\sigma_{id>300-et-id<305}(Clients)$		
id	prénom	nom
303	Henri	Berger
305	Pierre	Breguet

# Algèbre relationnel : Opérateurs relationnels

**Projection** Soit A une relations (table), et soient  $X_1, X_2, \dots, X_n$  des attributs de A. La projection  $\pi_{X_1, X_2, \dots, X_n}(A)$  est la relation qui ne contient que les colonnes  $X_1, X_2, \dots, X_n$  de A.

Relation Clients

id	prénom	nom
401	Assil	Guerid
402	Claire	Noirault
303	Henri	Berger
304	Luigi	Zaffalon
305	Pierre	Breguet

Relation  $\pi_{nom, prénom}(Clients)$

nom	prénom
Guerid	Assil
Noirault	Claire
Berger	Henri
Zaffalon	Luigi
Breguet	Pierre



# Algèbre relationnel : Opérateurs relationnels

**Projection** Soit A une relations (table), et soient  $X_1, X_2, \dots, X_n$  des attributs de A. La projection  $\pi_{X_1, X_2, \dots, X_n}(A)$  est la relation qui ne contient que les colonnes  $X_1, X_2, \dots, X_n$  de A.

Relation Clients

id	prénom	nom
401	Assil	Guerid
402	Claire	Noirault
303	Henri	Berger
304	Luigi	Zaffalon
305	Pierre	Breguet

Relation  $\pi_{\text{prénom}}(\text{Clients})$

prénom
Assil
Claire
Henri
Luigi
Pierre

# Algèbre relationnel : Opérateurs relationnels

**Projection** Soit A une relations (table), et soient  $X_1, X_2, \dots, X_n$  des attributs de A. La projection  $\pi_{X_1, X_2, \dots, X_n}(A)$  est la relation qui ne contient que les colonnes  $X_1, X_2, \dots, X_n$  de A.

Relation Clients

id	prénom	nom
401	Assil	Guerid
402	Claire	Noirault
303	Henri	Berger
304	Luigi	Zaffalon
305	Pierre	Breguet

Relation  $\pi_{nom, id}(Clients)$

id	nom
401	Guerid
402	Noirault
303	Berger
304	Zaffalon
305	Breguet

# Algèbre relationnel : Opérateurs relationnels

**Produit cartésien** Soient A et B deux relations (tables).

A X B est la relation qui contient toutes les lignes possibles tel que chaque ligne de A X B est composée d'une ligne de A et d'une ligne de B.

Relation Livres

id_livre	titre	année	id_auteur
10233	Oracle	2005	301
10234	ADA	1999	302

Relation Auteurs

id_auteur	prénom	nom
301	Henri	Berger
302	Luigi	Zaffalon
303	Pierre	Breguet

Relation Livres X Auteurs

id_livre	titre	année	id_auteur	id_auteur	prénom	nom
10233	Oracle	2005	301	301	Henri	Noirault
10233	Oracle	2005	301	302	Luigi	Zaffalon
10233	Oracle	2005	301	303	Pierre	Breguet
10234	ADA	1999	302	301	Henri	Noirault
10234	ADA	1999	302	302	Luigi	Zaffalon
10234	ADA	1999	302	303	Pierre	Breguet

# Algèbre relationnel : Opérateurs relationnels

**Jointure** Soient A et B deux relations (tables).

$A \bowtie B$  est la relation qui contient toutes les lignes du produit cartésien  $A \times B$  tel que la valeur des colonnes communes sont égales.

Relation Livres

id_livre	titre	année	id_auteur
10233	Oracle	2005	301
10234	ADA	1999	302

Relation Auteurs

id_auteur	prénom	nom
301	Henri	Berger
302	Luigi	Zaffalon
303	Pierre	Breguet

Relation Livres  $\bowtie$  Auteurs

id_livre	titre	année	id_auteur	id_auteur	prénom	nom
10233	Oracle	2005	301	301	Henri	Noirault
10233	Oracle	2005	301	302	Luigi	Zaffalon
10233	Oracle	2005	301	303	Pierre	Breguet
10234	ADA	1999	302	301	Henri	Noirault
10234	ADA	1999	302	302	Luigi	Zaffalon
10234	ADA	1999	302	303	Pierre	Breguet

# Algèbre relationnel : Opérateurs relationnels

**Jointure** Soient A et B deux relations (tables).

$A \bowtie B$  est la relation qui contient toutes les lignes du produit cartésien  $A \times B$  tel que la valeur des colonnes communes sont égales.

Relation Livres

id_livre	titre	année	id_auteur
10233	Oracle	2005	301
10234	ADA	1999	302

Relation Auteurs

id_auteur	prénom	nom
301	Henri	Berger
302	Luigi	Zaffalon
303	Pierre	Breguet

Relation Livres  $\bowtie$  Auteurs

id_livre	titre	année	id_auteur	id_auteur	prénom	nom
10233	Oracle	2005	301	301	Henri	Noirault
10234	ADA	1999	302	302	Luigi	Zaffalon