

# Bases de données

## Perspectives

BD001  
v201a

2020-08-31

Département d'informatique  
Faculté des sciences



Christina.Khnaisser@USherbrooke.ca  
<http://info.USherbrooke.ca/ckhnaisser>  
Luc.Lavoie@USherbrooke.ca  
<http://info.USherbrooke.ca/llavoie>

## PLAN

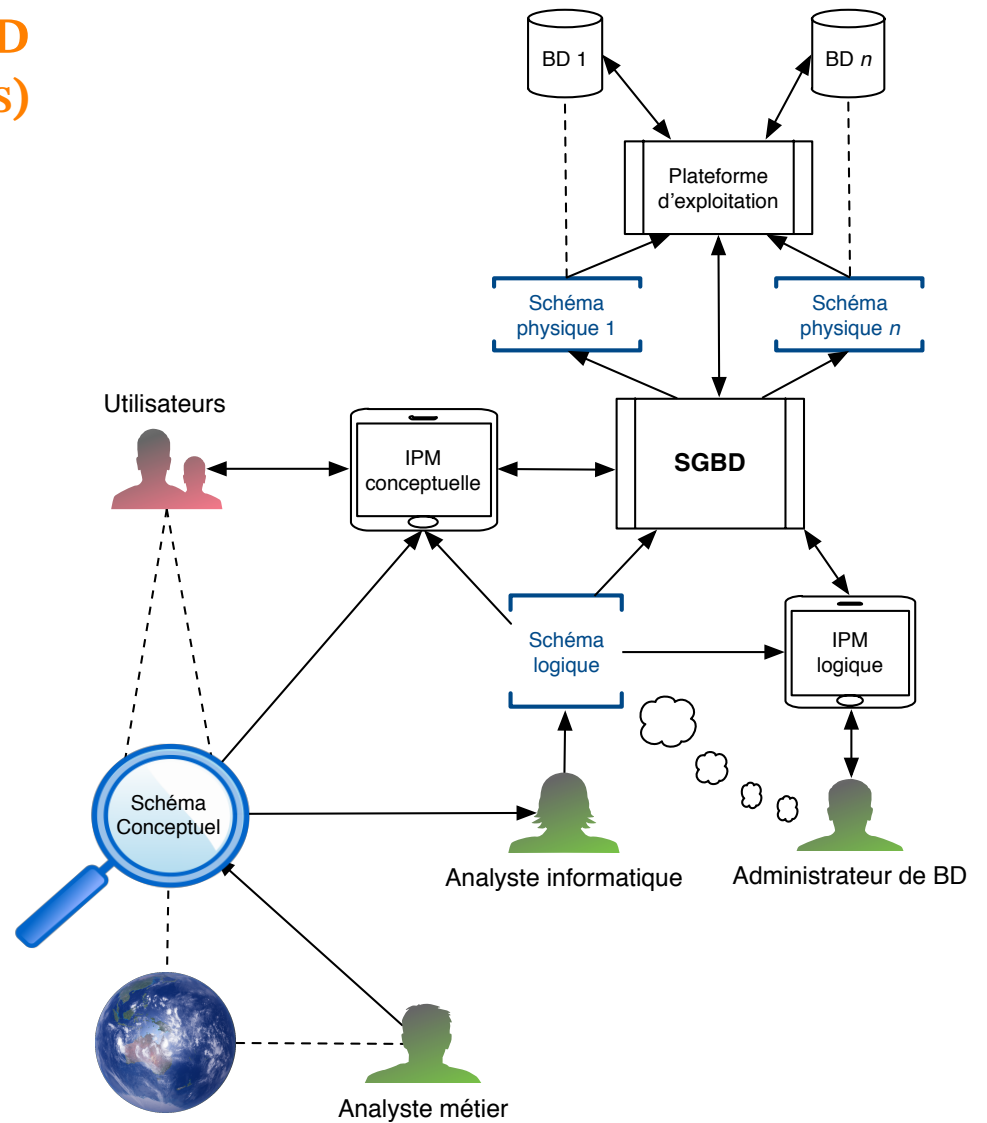
- Tour d'horizon
- Historique
- SGBD, un deuxième coup d'oeil
- Vocabulaire
- Références



## TOUR D'HORIZON

- Aperçu d'un SGBD
- Besoins et propriétés
- Fonctions et services
- Rôles et utilisateurs

# TOUR D'HORIZON — APERÇU D'UN SGBD (SYSTÈME DE GESTION DE BASES DE DONNÉES)



## TOUR D'HORIZON SGBDR

- Un SGBDR est une *machine* permettant de créer, modifier, détruire, gérer et interroger concurremment plusieurs bases de données **relationnelles**.
- Une machine est aussi appelée un système ou un service.

## TOUR D'HORIZON

### BASE DE DONNÉES RELATIONNELLE

- Une base de données relationnelle est un ensemble de *variables* respectant un schéma **relationnel**.
- Toute variable d'une base de données relationnelle est de « type relation », c'est-à-dire que la valeur référée par la variable est une relation.

## TOUR D'HORIZON

### SCHÉMA RELATIONNEL

- Un schéma relationnel est un ensemble de déclarations de **relations** soumises un ensemble de contraintes.
- Une déclaration de relation est la définition d'un « type relation » sur la base d'un entête de relation ET la déclaration d'une variable de ce type.

## TOUR D'HORIZON CONTRAINTE

- Une contrainte est une expression logique (un prédicat) dont les variables sont des variables de relation.
- Une contrainte est aussi appelée, selon le contexte, invariant.



## TOUR D'HORIZON

---

### Besoins à satisfaire

---

- Acquisition
- Stockage (rémanent)
- Traitement
- Accès
  - concurrent
  - contrôlé
- Traçabilité
- Disponibilité

L'importance de ces besoins induit un système complexe de (très) grande envergure

### Propriétés recherchées

---

- Exactitude
- Cohérence
- Intégrité
- Fiabilité
- Robustesse
- Performance
- Évolutivité dynamique
- ...

Le niveau de qualité recherché induit une réalisation soignée selon des méthodes rigoureuses mobilisant des experts de premier rang

## TOUR D'HORIZON

---

### Fonctions requises

---

- Définition
- Construction
- Modification
- Interrogation
- Contrôle
- Évolution
- ...

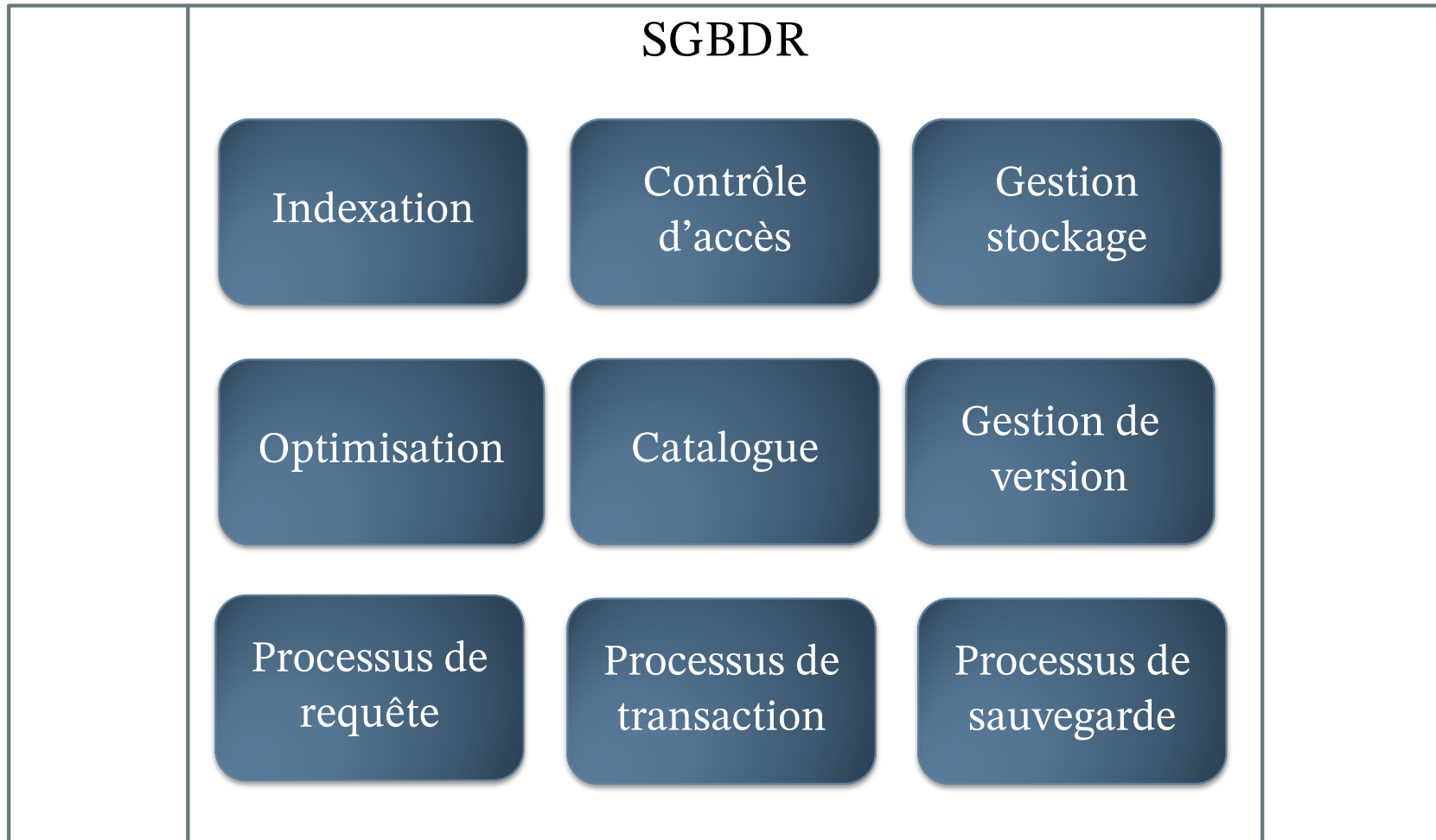
Le nombre, l'ampleur et la complexité des fonctions et des services induisent des interfaces variées et complexes

### Services offerts

---

- Autodescription (catalogue)
  - référence unique, cohérente, intègre
- Encapsulation et modularisation
  - abstraction
  - indépendance
    - entre logiciel et données
    - entre gestion et application
- Particularisation et spécialisation
  - spécificité (vues)
  - sécurité (règles)
- Traitement transactionnel
  - partage et concurrence
- ...

## TOUR D'HORIZON



## TOUR D'HORIZON

---

### Rôles diversifiés

---

- Exploitation
  - fourniture
  - saisie
  - validation
  - analyse
  - gestion
- Conception
  - base de données
  - application
  - système de gestion (SGBD)
  - outils

### Utilisateurs diversifiés

---

- Utilisateurs
  - occasionnels
  - experts
  - autonomes
- Préposés à l'exploitation
  - gestionnaire de BD,  
*data base administrator (DBA)*
  - technicien de BD,  
*database operator*
- Informaticiens
  - concepteur logiciel
  - analyste logiciel (de BD)
  - analyste d'affaires
- Gestionnaires
  - directeur informatique
  - directeur des TI  
*chief information officer (CIO)*

# HISTORIQUE

- Montée en puissance
- Théorie relationnelle
- Crise 1985-2005

## HISTORIQUE — LA MONTÉE EN PUISSANCE

### 1960

- Modèles primitifs
  - fichiers
  - index
  - hiérarchies

### 1970

- Modèles à 2 niveaux
  - Modèle réseau
    - CODASYL
  - **Début** du modèle relationnel
    - Alpha, R, SEQUEL
    - QBE, Ingres...

### 1980

- Modèles à 3 niveaux
- **Généralisation** relationnelle
  - Optimisation locale
  - Normalisation de SQL
    - SPARC, ANSI

### 1990

- **Domination** relationnelle
  - Optimisation globale
  - Normalisation de SQL
    - ISO
  - Répartition et distribution
- Apparition des BDOO

## HISTORIQUE — LA THÉORIE RELATIONNELLE


### 1969-1976

- Recherche
  - Travaux de Codd
  - Travaux de Date
  - ...
- Prototypes
  - Systèmes R (IBM)
  - INGRES
  - SEQUEL (IBM)
  - QBE (IBM)
  - SQL
  - Oracle
  - ...

### 1976-2015

- Évolution
  - Normalisation
  - OLAP (Codd)
  - Travaux d'Ullman
  - Travaux de Date
  - Typage (Cardelli)
  - Temporalité (Snodgrass et Lorentzos)
  - Co-relationalité (noSQL)
  - Néo-relationalité (newSQL)

## HISTORIQUE — LA CRISE 1985-2006

- Grandeur et décadence du modèle relationnel
    - ANSI SQL-89
    - ANSI SQL-92
    - ANSI SQL-99
  - Parallèlement, une approche objet qui ne s'impose pas... *et pour cause!*
    - CORBA
    - ODMG
  - La souplesse et ses limites
    - XML
  - La poursuite d'une normalisation de moins en moins suivie
    - ISO 9075:1999
    - ISO 9075:2003
    - ISO 9075:2006
    - ISO 9075:2008
    - ISO 9075:2011
    - ISO 9075:2016
- 



## HISTORIQUE — LES DÉFIS

- Intégrer le flou (2000...)
  - noSQL
- Intégrer le temps
  - BCDM
  - DDLM (ITL)
  - ...
- Intégrer l'intégration
  - MOLAP, ROLAP, HOLAP, DOLAP
  - Dimensionnalité (étoiles et flocons)
  - Entrepôts (ETL et médiateurs)

---

Allons-nous vers une renaissance du modèle relationnel avec le newSQL?

...  
*D*  
*Vertica*  
*Hanna*  
*VoltDB*  
...

# SQL, NOSQL, NEWSQL

Le point de vue  
de  
Bharat Upadrasta  
et  
Austin Chungath

LL :

- (1) **Oui**
- (2) **Grand**
- (3) **Grandissante.**

Caractéristiques	SQL	noSQL	newSQL
Propriétés ACID (intégrité des transactions)	Oui	Non	Oui
OLAP/OLTP	Oui	Non	Oui
Analyse des données	Oui	Non	Oui
Schéma défini (connu à l'avance)	Oui	Non	Peut-être <sup>(1)</sup>
Flexibilité du format des données	Non	Oui	Peut-être
Calcul distribué	Oui	Oui	Oui
Passage à l'échelle ( <i>Scale</i> )	Oui	Oui	Oui
Performance : l'augmentation des données	Rapide	Rapide	Très rapide
Utilisation ressources	Grand	Moyen	Minimal <sup>(2)</sup>
Popularité /Support de la communauté	Grand	Exp. rapide	Exp. lente <sup>(3)</sup>

source :

<http://www.informationweek.com/big-data/big-data-analytics/nosql-newsql-or-rdbms-how-to-choose/a/d-id/1297861>

## HISTORIQUE — EN GUISE DE CONCLUSION

- On assiste à un retour triomphal (?) du relationnel « pur » avec le « newSQL » en raison de :
  - sa capacité de maintenir les propriétés ACID (parce que, finalement, une réponse exacte... c'est souvent utile, voire nécessaire),
  - sa capacité de modéliser la temporalité (grâce à la logique des intervalles),
  - ses performances insurpassées (grâce aux représentations verticales et maillées),
  - aux capacités de plus en plus grandes des mémoires primaires (RAM).

## SGBD — UN DEUXIÈME COUP D'OEIL

- Modèles
- Langages
- Architectures
- Plateformes et produits
- Avantages et contre-indications
- Conclusion

## SGBD — MODÈLES

---

### Modèles de schémas

---

- **b**ischématique
  - externe (logique)
  - interne (physique)
- **t**rischématique
  - externe (conceptuel)
  - logique
  - interne (physique)
- **q**uadrischématique
  - connaissance
  - information (conceptuel)
  - données (logique)
  - représentation (physique)

### Modèles d'application

---

- monolithique
  - 1 niveau
- client-serveur
  - 2 niveaux
- multi-niveaux
  - $\geq 3$  niveaux

### Portée

---

- externe (conceptuel)
  - LRD
    - requêtes
- logiques
  - LDD (définition)
    - définition des schémas
    - définitions des types
  - LMD (manipulation)
    - modification des données
    - interrogation des données
  - LCD (contrôle)
    - contrôle d'accès aux données
    - contrôle des transactions
- interne (physique)
  - LSD
    - stockage et représentation

### Style

---

- Procédural, axiomatique ou fonctionnel
- Ensembliste ou algébrique
- Textuel ou graphique

*À l'origine, SQL se voulait  
axiomatique, ensembliste et  
textuel...*

## SGBD — ARCHITECTURE (INTERFACES)

### ○ Externe

- utilisateur → (formule des) requêtes
- programmeur → (élabore des) logiciels
- administrateur → (gère) schémas et données

### ○ Interne (système d'exploitation)

- gestionnaire de fichiers (mémoire rémanente)
- méthodes d'accès
- gestionnaire de tampons systèmes
- pilotes
- ...

## SGBD — ARCHITECTURE (COMPOSANTS)

### ○ SGBD

- Compilateur de LDD
- Analyseur de LMD
- Analyseur de LCD
- Logithèque (types et opérateurs)
- Gestionnaire de catalogue
- Gestionnaire de BD
  - ... (détail sur la diapositive suivante)...



## SGBD — ARCHITECTURE (SOUS-COMPOSANTS)

### ○ SGBD

- Gestionnaire de BD
  - Gestionnaire d'accès (sécurité)
  - Processeur de commandes
  - Vérificateur d'intégrité
  - Optimisateur de commandes
  - Gestionnaire de transactions
  - Planificateur
  - Gestionnaire de restauration
  - Gestionnaire de tampons
  - ...

## SGBD — PLATEFORMES ET PRODUITS

- Super-calculateur
  - Serveur
  - Équipement spécialisé (*appliance*)
  - Poste autonome
  - Assistant mobile
- Adabase
  - IDMS
  - Oracle
  - DB2
  - MS-SQL
  - INGRES
  - Sybase
  - Informix
  - PostgreSQL
  - MySQL
  - SQL lite
  - Berkeley DB
  - FileMaker
  - Access
  - ...

## SGBD

### AVANTAGES DES SGBD COMPARATIVEMENT AUX SOLUTIONS AD HOC

Avantages SGBD	Contre-indication
Normalisation	Problèmes hors norme. (on cherche encore lesquels 😊)
Évolutivité	Usage unique (par un seul client, dans un environnement figé!)
Multi-accès	Accès unique en tout temps (sur et depuis un seul ordinateur)
Économie d'échelle	Coût de développement initial faible et garanti (qui peut <b>garantir</b> cela?)
Mise à disposition rapide	Temps de développement initial court et garanti (qui peut <b>garantir</b> cela?)
Très haute fiabilité	Spécification formelle et programmation automatisée du système
Auto-documentation	Documentation complète pré-existante
Traçabilité	Il faut choisir entre performance et traçabilité
...	Application de type temps réel critique

# SGBD – EN GUISE DE CONCLUSION

2015

Figure 1. Magic Quadrant for Operational Database Management Systems



2017

Figure 1. Magic Quadrant for Operational Database Management Systems



# SGBD – EN GUISE DE CONCLUSION

2018



2019

Figure 1. Magic Quadrant for Operational Database Management Systems



## VOCABULAIRE — CONCEPTS DE BASE (RAPPELS – VOIR BD000)

- informatique
- information
  
- donnée
- représentation
  
- valeur
- domaine
- type

## VOCABULAIRE

### ○ schéma

- schéma d'une base de données
- modèle logique de données

### ○ base de données

- une instance d'un schéma
- une instance d'un modèle logique de données
- un contenant regroupant les données représentant un état de la réalité telle que modélisée grâce au modèle conceptuel à l'origine du schéma!

## VOCABULAIRE — SCHÉMAS

- schéma conceptuel (aussi nommé schéma externe)
  - décrit les concepts d'intérêt du monde réel selon la vision de l'utilisateur ou d'un expert du métier (p. ex. entité- association, orienté-objet, ontologie).
- schéma logique
  - décrit l'organisation et les contraintes applicables aux données (p. ex. relationnel, réseau, hiérarchique).
- schéma physique (aussi nommé schéma interne)
  - décrit la structure de stockage et les méthodes d'accès.

Pour définition complète :

[Deputy Chief Information Officer 2015]

<http://dodcio.defense.gov/TodayinCIO/DoDArchitectureFramework.aspx>.



## VOCABULAIRE — MODÈLES

### ○ conceptuel

- EA (entité-association)  
ER (entity-relationship)
- EAE (EA étendu)  
ERE (extended ER)
- « ontologique »  
(ex. notation OWL)
- ...

### ○ logique

- hiérarchique
- réseau
- relationnel
- objet
- ...

### ○ physique

- adressage dispersé (hashing),  
arbre de recherche  
(B-Tree)
- horizontal,  
vertical,  
mixte,  
KV
- ...

## RÉFÉRENCES

- Bases de données (Databases) :
  - [Date 2004], chapitres 1 et 2;
  - [Elmasri and Navathe 2004], chapitres 1 et 2;
  - [Elmasri and Navathe 2016], chapitres 1 et 2;
  - [Ullman and Widom 2008], chapitre 1;
  - [https://fr.wikipedia.org/wiki/Base\\_de\\_données](https://fr.wikipedia.org/wiki/Base_de_données)
  - <https://en.wikipedia.org/wiki/Database>
- Mégadonnées (Big Data) :
  - [https://fr.wikipedia.org/wiki/Big\\_data](https://fr.wikipedia.org/wiki/Big_data)
  - [https://en.wikipedia.org/wiki/Big\\_data](https://en.wikipedia.org/wiki/Big_data)

## LES COLLES DU PROF

- Qu'est-ce qu'un SGBDR ?
- Qui sont les principaux utilisateurs d'un SGBD ?
- Quelles sont les fonctionnalités d'un SGBDR ?
- Quel est le modèle prédominant dans les bases de données ?
- Quel modèle doit-on choisir pour développer un logiciel de gestion de prescriptions de médicament ?
- Quel modèle doit-on choisir pour développer pour un blogue sur les technologies ?
- Quel modèle doit-on choisir pour développer pour un site web d'achat en ligne ?

