

Bases de Données 1

#0 - Intro

Matthieu Nicolas
Polytech S5 - II

Slides réalisées à partir de celles de Claude Godart et Malika Smaïl

Se présenter

- matthieu.nicolas@univ-lorraine.fr
- Doctorant, équipe COAST, LORIA
- Algorithmes pour l'édition collaborative pair-à-pair

Quelques règles

- Présence obligatoire
- 5min de retard tolérées
 - mais le cours commence à l'heure
- Si possible, chacun sur sa machine
 - encouragés à réfléchir en binôme

Organisation

- **Volume horaire : 18h**
- **Enseignement intégré : CM et TD fusionnés**
- **2 notes** pour l'évaluation
 - 1 DS de 2h
 - 1 TD/TP noté de 2h

Objectifs

- Découvrir les notions de **BD** et **SGBD**
- Apprendre à **concevoir** une BD
- Apprendre à **utiliser** une BD

Syllabus

- *Conception d'une BD*
 - Diagramme de classes
 - Schéma relationnel
 - Normalisation
- *Utilisation d'une BD*
 - Algèbre relationnelle
 - Déploiement d'une BD + intro SQL

Liens avec autres modules

- *Base de Données 2*
 - Poursuite apprentissage SQL
 - Interagir avec une BD depuis un langage de programmation
- *Projet transversal II*
 - Conception et développement d'une appli interagissant avec une BD

Avant de commencer

- Qui êtes-vous ?
 - Formations précédentes ?
 - Déjà vu la notion de BD ?
 - Déjà utilisé une BD ?
- Pour la suite
 - Déjà fait du PHP/MySQL ?
 - Déjà utilisé Git ?

Plan

- **Problématiques et besoins** liés au stockage de données
- Notions de **Base de Données (BD)** et de **Système de Gestion de BD (SGBD)**
- **Fonctionnalités** d'un SGBD

Problématiques et besoins liés au stockage de données

Base de Données 1
#0 - Intro

Stockage de données

- Chaque application dispose de ses données
- Besoin de stocker ces données
 - Pour une utilisation ultérieure
 - Pour les partager avec les autres composants de l'application ou d'autres applications
- Historiquement, utilise des fichiers

Limites du Système de Gestion de Fichiers - 1

- Mauvaises performances pour un gros volume de données
 - Lecture et écriture lente
- Comment faire lorsque des données sont réparties entre plusieurs fichiers ?

Limites du Système de Gestion de Fichiers - 2

- Comment faire lorsque des données sont dupliquées dans plusieurs fichiers ?
- Demande un effort de programmation important pour exploiter les données
 - Pas de standard
 - Pas d'outils dédiés

Redondance de données

- Répétition **non-nécessaire** de données
- Possède plusieurs inconvénients
 - **Coût** en volume et en argent **dupliqué**
 - Risque d'**incohérence** entre les différentes copies
- On souhaite donc **supprimer** la redondance des données

Notions de Base de Données (BD) et de Système de Gestion de BD (SGBD)

Base de Données 1
#0 - Intro

Base de Données (BD)

- **Collection importante de données**
 - Cohérentes
 - Structurées indépendamment d'une application
 - De redondance minimale
 - Accessibles par plusieurs utilisateurs à la fois
- Apparaît au début des années 60

Exemples d'utilisation de BDs

- **Twitter** enregistre tweets, likes, retweets et réponses dans sa BD
- **Steam** conserve librairies, succès, et messages dans sa BD
- **Netflix** stocke historiques de visionnage et likes dans sa BD

Systeme de Gestion de BD (SGBD)

- Un ensemble de **programmes** permettant à des utilisateurs de **créer** et **d'utiliser** des BDs

Objectifs d'un SGBD - 1

- Indépendance données-programme
 - **Indépendance physique** : un changement de l'organisation physique des données n'entraîne pas de changement dans les programmes
 - **Indépendant logique** : un changement de l'organisation logique des données (ex. nouvelle rubrique) n'entraîne pas de changement dans les programmes non concernés

Objectifs d'un SGBD - 2

- **Manipulation aisée des données** : par des utilisateurs non informaticiens (interrogation et mise à jour)
- **Administration aisée des données** : un SGBD doit fournir des outils pour décrire les données et autoriser leur évolution (tâche de l'administrateur)

Objectifs d'un SGBD - 3

- **Partage des données** : utilisation simultanée par différentes applications
- **Efficacité des accès aux données** : garantie d'un bon *débit* (nombre de transactions exécutées par seconde) et d'un bon *temps de réponse* (temps d'attente moyen par transaction)

Fonctionnalités d'un SGBD

Base de Données 1

#0 - Intro

Fonctionnalités d'un SGBD

- **Définition des données**
- **Manipulation des données**
- **Intégrité des données**
- Partage des données
- Confidentialité des données
- Sécurité de fonctionnement
- Support de persistance

Définition de données

- Un **Langage de Définition de Données** (LDD) permet de décrire :
 - des **objets** (*personnes, véhicules*)
 - des **attributs** d'objets (nom, *n° d'immat*)
 - des **liens** entre objets (*personne possède véhicule*)
 - des **contraintes** sur objets, attributs et liens (*un véhicule n'a qu'un seul propriétaire*)
- **Schéma** = description d'une BD à l'aide du LDD

Manipulation de données

- Un **Langage de Manipulation de Données** (LMD) permet de *créer, rechercher, supprimer et modifier des données*
- Offre plusieurs types d'interfaces
- **Interfaces orientées utilisateur final** : langages de requêtes déclaratifs comme SQL avec mise en oeuvre graphique, interface de type formulaire...
- **Interfaces orientées programmeurs d'applications** : interface avec des langages de programmation classiques (PHP, Java...)

Intégrité des données - 1

- Un schéma BD se compose d'une **description des données**, de leurs **relations**, ainsi que d'un ensemble de **contraintes d'intégrité** (CI)
- Une CI est une **propriété de l'application à modéliser**
- Les données stockées dans une BD doivent vérifier ces CI

Intégrité des données - 2

- 2 types de contraintes :
 - Contraintes **structurelles** : *un employé a un et un seul chef*
 - Contraintes **dynamiques** : *un salaire ne peut pas diminuer*

Partage de données

- Plusieurs utilisateurs peuvent accéder à la même information en même temps
- Mécanismes de gestion de conflits d'accès
 - **autorisation** d'accès multiples en consultation (lecture de la BD)
 - **verrouillage** en cas d'accès en modification (écriture dans la BD)

Confidentialité des données

- Cache certaines informations à certains utilisateurs
- Mise en oeuvre :
 - comptes utilisateurs, **mots de passe**, privilèges d'accès
 - utilisation de **sous-schémas**

Sécurité de fonctionnement

- 1

- Mécanisme de **transaction**
 - Séquence d'opérations faisant passer une BD **d'un état cohérent à un nouvel état cohérent**
 - Une transaction est **exécutée** complètement **avec succès**, ou est **annulée**

Sécurité de fonctionnement

- 2

- Remettre rapidement une BD dans un état opérationnel après un incident hardware ou software
- **Journalisation** des opérations réalisées sur la BD
(*dans le journal ou log*)
- **Ré-exécution** automatique en cas d'incident

Support de persistance

- Durée de vie des données > à celle des programmes qui les ont générées
 - **Stockage** des données **sur disques** externes (*mémoire secondaire*)
- Pour garantir un bon **débit** et **temps de réponse**
 - Gestion de **cache** pour un accès efficace (*index, hachage*)
 - Techniques d'**optimisation de requêtes** pour minimiser les transferts

Résumé

- Besoins de stocker une **large collection de données**
 - Système de **fichiers non-adapté**
 - Utilisation d'un **SGBD recommandé**
- Un SGBD nous permet de :
 - **Définir le schéma**
 - **D'interroger et de manipuler** efficacement
 - Tout en garantissant sa **sécurité et robustesse**

Types de SGBD

- Il existe différents types de SGBD : *relationnels, orienté documents, orientés graphes, orienté séries chronologiques...*
- Il n'y a **pas de type “ultime”** de SGBD...
- ... ils sont justes conçus pour des **applications différentes**
- Le **choix** du type de SGBD **doit dépendre** :
 - Schéma des données
 - Utilisation des données

SGBD relationnels - 1

- Se focalise sur les SGBD relationnels dans le cadre de ce module
- Modèle proposé par E.F. Codd en 1970 : *A Relational Model of Data for Large Shared Databanks*
- Le type le plus utilisé encore de nos jours
- Fondé sur la notion de **relation** : *existe un lien entre les valeurs des attributs d'un objet*

SGBD relationnels - 2

- Existe de nombreux SGBD relationnels
 - Oracle
 - MySQL
 - PostgreSQL
 - MariaDB
 - ...

Mini-bibliographie

- *Fundamentals of database systems* - R. Elmasri et S. B. Navathe, The Benjamin/Cummings Pub., 2000
- *Bases de données et systèmes d'information* - N. Boudjlida, Dunod, 1999
- *Bases de données : objet et relationnel* - G. Gardarin, Eyrolles, 1999
- *Bases de données : les systèmes et leurs langages* - G. Gardarin, Eyrolles, 1984