

Bases de données

Informatique BUT 2^e année et L2

Stéphane DEVISMES

Professeur à l'Université de Picardie Jules Verne

Anaïs DURAND

Maîtresse de conférences à l'IUT de l'Université Clermont Auvergne

Pascal LAFOURCADE

Professeur des universités à l'IUT de l'Université Clermont Auvergne

DUNOD

Direction et conception graphique de la couverture :
Nicolas Wiel - Pierre André Gualino (graphiste)

© Dunod, 2025

11 rue Paul Bert, 92240 Malakoff
www.dunod.com

ISBN 978-2-10-089160-3

Avant-propos

Cet ouvrage est la suite du premier volume qui s'adressait à la première année de Bachelor Universitaire de Technologie (BUT) et qui permettait d'appréhender le fonctionnement d'un Système de Gestion de Base de Données (SGBD) et de faire une introduction au langage *Structured Query Language* (SQL). Le but de ce second volume est d'aller plus loin dans les concepts fondamentaux des bases de données.

Le premier chapitre de cet ouvrage aborde donc les grands principes d'optimisation des bases de données. Le second chapitre traite de la concurrence d'accès dans les bases de données. Le langage *Procedural Language/PostgreSQL* (PL/pgSQL), qui permet d'interroger une base donnée via un langage simple, est présenté dans le troisième chapitre. Dans le chapitre suivant, les mécanismes de mise à jour automatique d'une base de données sont abordés. Ensuite dans les deux chapitres suivants, les notions de forme normale sont présentées ainsi que les algorithmes sous-jacents. Le chapitre 7 s'attarde sur le concept de dénormalisation. Le chapitre suivant propose une petite introduction aux langages NoSQL. Enfin, le dernier chapitre présente MongoDB un SGBD orienté documents.

Cet ouvrage couvre le programme de deuxième année de BUT Informatique. Il s'adresse donc en priorité aux étudiants et enseignants de cette formation. Il s'adresse également aux étudiants et enseignants de licence et d'école d'ingénieur en informatique, ainsi qu'à toute personne souhaitant découvrir et travailler avec des bases de données.

Les auteurs expriment leur gratitude à Matthieu Daniel, Laurent Desrosiers et Anne Le Duc pour leurs commentaires et suggestions de modifications constructifs, à la suite de leurs relectures assidues.

À Amiens et Clermont-Ferrand, 2 octobre 2025.

Stéphane DEVISMES, Anaïs DURAND et Pascal LAFOURCADE

Table des matières

VI

1	Optimisation des requêtes	3
1.1	Est-ce que la requête est bien écrite ?	3
1.2	Index	8
1.2.1	Quelles colonnes indexer ?	10
1.2.2	Quels types de colonnes utiliser ?	11
1.2.3	Gains et pertes	11
1.3	Cache	12
1.4	Plan d'exécution	12
1.4.1	Contenu du plan d'exécution	13
1.4.2	Statistiques	15
1.5	TP1 : Optimisation	16
2	Concurrence d'accès	19
2.1	Transaction et modèle ACID	19
2.2	Niveaux d'isolation	20
2.3	Lecture cohérente et ROLLBACK	22
2.4	Gestion des accès concurrents	24
2.5	Interblocage	28
2.6	TP2 : Concurrence d'accès	30
3	PL/pgSQL	33
3.1	Bloc PL/pgSQL	33
3.2	Déclaration de variables	35
3.3	Traitements	36
3.3.1	SELECT ... INTO	36
3.3.2	Conditions	37
3.3.3	Boucles WHILE	38
3.3.4	Boucles LOOP	39
3.3.5	Boucles FOR ... IN	39
3.4	Fonctions	40
3.4.1	Paramètres	40
3.4.2	Fonction retournant un type "simple"	41
3.4.3	Fonction retournant une table	42
3.4.4	Utilisation d'une fonction	43
3.5	Gestion des erreurs	44
3.5.1	Traiter les exceptions	45
3.5.2	Exceptions PostgreSQL	46
3.5.3	Signaler une erreur	47
3.6	Curseurs	49
3.6.1	Parcours « manuel » avec curseur	49
3.6.2	Parcours avec boucle FOR sur curseur	49
3.6.3	Modifications avec curseur	51
3.7	TP3 : PL/pgSQL	52

4	Déclencheurs	55
4.1	Création d'un trigger	55
4.2	Fonction trigger	56
4.3	Gestion des triggers	57
4.4	Trigger pour garantir des cardinalités	58
4.5	TP4 : Déclencheurs	62
5	Formes normales	63
5.1	Dépendances fonctionnelles	63
5.1.1	Hypothèses et notations	64
5.1.2	Propriétés des dépendances fonctionnelles	65
5.1.3	Dépendances fonctionnelles singletons et élémentaires	67
5.2	Dépendances et clés	68
5.2.1	Clôture	69
5.3	Formes normales	70
5.3.1	Première Forme Normale (1NF)	70
5.3.2	Deuxième Forme Normale (2NF)	71
5.3.3	Troisième Forme Normale (3NF)	71
5.3.4	Forme Normale de Boyce-Codd (BCNF)	73
5.3.5	Autres formes normales	74
5.4	Pour aller plus loin	76
5.4.1	Preuve des propriétés dérivées des axiomes d'Armstrong	76
5.4.2	Algorithme de clôture d'un ensemble d'attributs	77
5.4.3	Preuve de l'algorithme de clôture	79
5.4.4	Table des clôtures	82
5.4.5	Équivalences entre définitions des formes normales	85
6	Algorithmes de normalisation	89
6.1	Décomposition de relations	89
6.1.1	Rappel sur la jointure naturelle	89
6.1.2	Décomposition	90
6.1.3	Cinquième Forme Normale (5NF).	91
6.1.4	Théorème de décomposition	91
6.1.5	Décomposition : redondances et cohérence	92
6.2	Base minimale	93
6.2.1	Algorithme de détermination d'une base minimale	93
6.3	Normalisation BCNF : l'algorithme récursif	95
6.4	Normalisation 3NF : l'algorithme de synthèse	99
6.5	Exemple récapitulatif	101
6.6	Pour aller plus loin	103
6.6.1	Preuves des propriétés et théorème de décomposition	103
6.6.2	Preuve de l'algorithme de détermination d'une base minimale	105
6.6.3	Preuve de l'algorithme récursif	106
6.6.4	Preuve de l'algorithme de synthèse	108

7	Dénormalisation	111
7.1	Avantages et inconvénients de la dénormalisation	111
7.2	Techniques classiques de dénormalisation	113
7.2.1	Attributs calculés	113
7.2.2	Liens directs entre tables	114
7.2.3	Fusion de tables	115
7.2.4	Découpage de tables	116
7.3	Mécanismes utiles en cas de dénormalisation	118
7.3.1	Vues matérialisées et rafraîchissement	118
7.3.2	Tables et triggers	119
8	Introduction au NoSQL	121
8.1	Petit historique des bases de données	121
8.1.1	Cartes perforées	121
8.1.2	Modèle hiérarchique et en réseau	123
8.1.3	Modèle relationnel	124
8.1.4	Big Data et NoSQL	125
8.2	Panorama des modèles NoSQL	126
8.2.1	SGBD orientés clé-valeur	126
8.2.2	SGBD orientés documents	127
8.2.3	SGBD orientés colonnes	127
8.2.4	SGBD orientés graphes	127
8.2.5	Autres paradigmes	128
8.3	Relationnel vs. NoSQL	128
8.3.1	Structure des données	129
8.3.2	Requêtage	129
8.3.3	Évolutivité	129
8.3.4	Cohérence des données	131
8.3.5	Performances	132
9	MongoDB	133
9.1	Documents et collections	133
9.1.1	Format <i>JavaScript Object Notation</i> (JSON)	133
9.1.2	Types de données	134
9.1.3	Base de données et collection	135
9.1.4	Exemples de documents	136
9.2	Architecture MongoDB	137
9.3	Consulter des documents	138
9.3.1	Limiter le nombre de résultats.	139
9.3.2	Trier le résultat	140
9.3.3	Champs imbriqués	140
9.3.4	Tableaux	141
9.3.5	Opérateurs de comparaison	142
9.3.6	Opérateurs logiques	144
9.3.7	Recherche de motifs	145
9.4	Gestion des documents	146
9.4.1	Insertion de documents	146

9.4.2	Suppression de documents	148
9.4.3	Mise à jour de documents	148
9.5	Pipelines d'agrégation	151
9.5.1	Fonctionnement global d'un pipeline d'agrégation	151
9.5.2	Étapes du pipeline d'agrégation	151
9.5.3	Calcul lors d'une étape du pipeline	159
9.6	TP5 : MongoDB	163
Correction des exercices		167
Correction des TP		179
Liste des figures, tables, exercices, TP et abréviations		203
	Liste des figures	203
	Liste des exercices	203
	Liste des TP	204
	Liste des abréviations	204
Index		207

Comment utiliser cet ouvrage ?

Chaque chapitre de cet ouvrage contient des exemples illustrant les différentes notions abordées, mais également des exercices. Le lecteur est invité à résoudre ces exercices au fur et à mesure de sa lecture. Les corrections sont disponibles à partir de la page 167. Certains chapitres sont accompagnés d'un sujet de Travaux Pratiques (TP), où le lecteur est invité à pratiquer les concepts du chapitre sur son ordinateur.

Pour la partie SQL et PL/pgSQL (TP1 à TP4), il est conseillé d'installer un serveur PostgreSQL en local sur son ordinateur, ainsi qu'un client textuel (comme `psql`) ou graphique (comme `pgadmin`). PostgreSQL étant un logiciel libre, cette installation est gratuite. Des explications pour installer PostgreSQL sur sa machine sont disponibles sur la documentation officielle : <https://www.postgresql.org>

Similairement, pour la partie MongoDB (TP5), il est conseillé d'installer un serveur MongoDB en local sur ordinateur, ainsi qu'un client textuel (comme `mongosh`) ou graphique (comme `compass`). Même si MongoDB offre des solutions hébergées sur le cloud (MongoDB Atlas) payantes, MongoDB est libre et l'installation en auto-hébergement est gratuite. Plus d'explications sur cette installation sont disponibles sur la documentation officielle : <https://www.mongodb.com/docs/manual>

Pour réaliser les TPs, des fichiers sont à récupérer à l'adresse suivante :



<http://dunod.link/r01fctx>

La correction des TPs est disponible à partir de la page 179.

Syntaxe des commandes. Lorsque la syntaxe des langages SQL, PL/pgSQL et MongoDB est présentée dans les différents chapitres, certains caractères spéciaux sont utilisés :

- [a] signifie que a est une option, non obligatoire.
- a|b signifie qu'il est possible d'utiliser soit a, soit b (mais pas les deux en même temps!).

