

N° ordre 40/2002

République Algérienne Démocratique et Populaire

Ministère de l'Enseignement supérieur et de la Recherche Scientifique

Université des Sciences et de la Technologie

Houari Boumedienne

U.S.T.H.B

INSTITUT D'INFORMATIQUE

Mémoire du Projet de Fin d'Etude
Pour l'obtention du diplôme

D'INGENIEUR D'ETAT EN INFORMATIQUE

Option: Software



Thème

Conception et réalisation d'un système d'analyse multidimensionnelle

Thème proposé et encadré par:
M^{me} F.Z. Bessai (CERIST)

Etudié par:
M^{elle} Zeddigha Ismahane
M^{elle} Boutouhami Sara

Soutenu le: 30 Octobre 2002
Devant le jury composé de:

M ^{me} Kadri	Présidente
M ^{elle} Mahdaoui	Examinatrice
M ^{elle} Kamel	Examinatrice

Organisme d'accueil : CERIST
Centre de Recherche sur l'Information Scientifique et Technique

PROMOTION: 2001/2002

Résumé

Les entrepôts de données ou Data Warehouse, est une collection de données orientées sujet, intégrées, non volatiles, historisées, et organisées pour le support d'un processus d'aide à la décision. Les bases de données multidimensionnelles sont une des applications des entrepôts de données destinées à l'aide à la décision. Par leur intermédiaire, une grande partie du travail de préparation des données est effectuée. Les données sont présentées sous forme de cubes multidimensionnels dont les cellules contiennent des mesures précalculées. Cette présentation est plus naturelle et plus lisible pour les analystes.

L'objectif de notre travail est la conception et la réalisation d'une base de données multidimensionnelle en dégagant les spécificités de telle base de données par rapport aux bases de données classiques et en étudiant les outils permettant de stocker, manipuler, interroger, et analyser ces bases.

Mots clés

Data Warehouse, Data Mart, OLAP, Analyse multidimensionnelle, cube, système d'aide à la décision.

Sommaire

INTRODUCTION GENERALE.....	1
CHAPITRE I : SYSTEME D'AIDE A LA DECISION.....	
I. INTRODUCTION.....	4
II. LES CONCEPTS DE BASE.....	4
II.1 Définition d'un système d'aide à la décision.....	4
II.1.1 L'infocentre.....	6
II.2 Architecture générale d'un système d'aide à la décision.....	7
II.2.1 Les éléments d'un système d'aide à la décision.....	7
II.2.2 Circuit de l'information dans le système d'aide à la décision	10
II.2.3 Modèle de données.	11
III. CONCLUSION.....	13
CHAPITRE II :DATA WAREHOUSE ET DATA MART.....	
I. INTRODUCTION.....	14
II. DATA WAREHOUSE.....	14
II.1 Définition	14
II.1.1 Données orientées sujet.....	15
II.1.2 Données intégrées.....	15
II.1.3 Données historisées.....	15
II.1.4 Données non volatiles.....	15
II.2 Caractéristiques d'un Data Warehouse.....	16
II.3 Conception d'un Data Warehouse.....	16
II.4 Implémentation d'un Data Warehouse.....	19
II.4.1 Architecture réelle.....	19
II.4.2 Architecture virtuelle.....	20
II.4.3 Architecture remote.....	20
II.5 Alimentation d'un Data Warehouse.....	20
II.5.1 Extraction.....	21
II.5.2 Nettoyage et transformation.....	22
II.5.3 Agrégation primaire.....	24
II.5.4 Les clés.....	25
II.5.5 Chargement des données.....	27

II.5.6	Rafraîchissement.....	27
III.	DATA MART.....	29
III.1	Définition.....	29
IV.	COMPARAISON ENTRE DATA MART ET DATA WAREHOUSE.....	29
V.	CONCLUSION.....	32

CHAPITRE III: BASE DE DONNEES MULTIDIMENSIONNELLE..

I.	INTRODUCTION.....	33
II.	ARCHITECTURE D'UNE BASE DE DONNEES MULTIDIMENSIONNELLE.....	33
II.1	Table de faits.....	34
II.2	Table dimensionnelle.....	35
II.3	Dimension temps.....	36
II.4	Les agrégats.....	38
II.5	Les clés.....	39
III.	CARACTERISTIQUES DE LA MODELISATION MULTIDIMENSIONNELLE.....	40
IV.	QUELQUES SITUATIONS PARTICULIERES.....	40
IV.1	Table de faits sans faits.....	40
IV.2	Dimension changeante.....	40
IV.3	Dimension changeante à évolution rapide.....	43
V.	MODELISATION DES DONNEES.....	45
V.1	Schéma en étoile.....	45
V.1.1	Avantages du schéma en étoile.....	46
V.1.2	Inconvénients du schéma en étoile.....	46
V.2	Schéma en flocon.....	47
V.2.1	Avantages du schéma en flocon.....	48
V.2.2	Inconvénients du schéma en flocon.....	48

V.3 Schéma en constellation.....	49
V.3.1 Dimension conforme.....	50
V.3.2 Fait conforme.....	50
VI. CONCLUSION.....	51

CHAPITRE IV: TECHNOLOGIE OLAP (On Line Analytical Processing)

I. INTRODUCTION.....	52
II. CUBE.....	52
III. LA TECHNOLOGIE OLAP.....	54
III.1 Définition.....	54
III.2 Les douze règles d'OLAP.....	55
III.3 ROLAP (Relational On Line Analytical Processing).....	57
III.3.1 Les caractéristiques de ROLAP	57
III.4 MOLAP (Multidimensional On Line Analytical Processing).....	58
III.4.1 Les caractéristiques de MOLAP.....	59
III.5 HOLAP (Hybride On Line Analytical Processing).....	59
IV. LES OPERATEURS DE NAVIGATION.....	61
V. SYSTEME OLAP VERSUS OLTP.....	62
VI. CONCLUSION	64

CHAPITRE V: CONCEPTION DU SYSTEME D'ANALYSE MULTIDIMENSIONNELLE (SAM)

I. INTRODUCTION.....	65
II. APPROCHE DE CONCEPTION DU PROTOTYPE SAM.....	65
II.1 Architecture générale du prototype SAM.....	67
II.2 Description du prototype SAM.....	68
II.2.1 Création et saisie d'un Data Mart.....	68
II.2.2 Analyse multidimensionnelle.....	68
III. CONCLUSION.....	69

**CHAPITRE VI: PRESENTATION ET EVALUATION DU
PROTOTYPE (SAM)**

I.	INTRODUCTION.....	70
II.	ENVIRONNEMENT DE DEVELOPPEMENT.....	70
III.	FONCTIONNEMENT GENERAL	71
IV.	REPRESENTATION DES INTERFACES.....	71
V.	EVALUATION DU LOGICIEL.....	80
VI.	CONCLUSION.....	93

	CONCLUSION GENERALE.....	94
--	---------------------------------	-----------

	BIBLIOGRAPHIE.....	96
--	---------------------------	-----------

ANNEXES

	ANNEXE 1 : GLOSSAIRE.....	101
--	----------------------------------	------------

	ANNEXE 2 : ESTIMATION DE LA TAILLE D'UN DATA MART.....	106
--	---	------------

	ANNEXE 3 : QUELQUES CHIFFRES ET QUELQUES PRODUITS.....	108
--	---	------------