

**République Algérienne Démocratique et Populaire**  
**Ministère de l'Enseignement Supérieur et de la Recherche Scientifique**  
**Université Mouloud Mammeri Tizi-Ouzou**  
**Faculté de Génie Electrique et d'Informatique**  
**Département d'informatique**

# Mémoire



**De fin d'étude**  
**En vue de l'obtention du Diplôme d'Ingénieur**  
**d'Etat en Informatique**  
**Options : Systèmes Parallèles et Distribués**  
**Systèmes d'Information**

## Thème

# construction et maintenance des entrepôts de données

**Dirigé par :**

- M<sup>r</sup> OUAMRANE M.
- M<sup>me</sup> Dellal B.

**Présenté par :**

- M<sup>elle</sup> Kebaili Malika
- M<sup>r</sup> Bouaine Tarik

**Année universitaire 2007-2008**

Introduction générale.....	2
I- Concepts et définitions.....	2
I-1- Système d'information décisionnel .....	2
I-2- Source de données hétérogènes.....	2
I-3- Datawarehouse.....	2
I-4- Datamart.....	2
I-5- Data warehousing.....	2
I-6- ETL (Extract Transform Load).....	3
I-7- OLAP.....	3
I-8- EIS.....	3
I-9- Datamining.....	3
I-10- Reporting.....	3
I-11- Portail.....	3
I-12- Indicateur.....	3
<b>Chapitre 1: OLAP (On-Line-Analytical-Processing)</b>	
OLAP .....	4
I-1- Les fonctionnalités d'olap.....	4
I-1-1- vue multidimensionnelle .....	4
I-1- 2- transparence .....	4
I-1-3 - accessibilité .....	4
I-1- 4- performance continue dans les restitutions.....	4
I-1-5- architecture Client / Serveur .....	4
I-1-6- Dimensions Génériques .....	4
I-1-7- gestion dynamique des matrices creuses .....	5
I-1-8-Support Multi-Utilisateurs .....	5
I-1-9- opérations entre dimensions .....	5
I-1-10- manipulation intuitive des données.....	5
I-1-11- flexibilité des restitutions .....	5
I-1-12- nombre de dimensions et de niveaux de hiérarchie illimités .....	5
I-2- Niveau conceptuel.....	6
I-2-1- Hypercube OLAP .....	6
I-2-2- Concepts du modèle multidimensionnel.....	7
I-2-2-1- Mesure ou fait.....	7
I-2-2-2-Dimension .....	7
I-2-2-2-1- <b>Dimension dégénérée</b> .....	8
I-2-2-2-2- La dimension Temps .....	8
I-2-2-3- Hiérarchie .....	8
I-2-2-4- Attribut.....	8
I-2-2-5- Agrégat.....	8
I-2-2-6- Granularité des dimensions.....	8
I-2-2-7- Niveau et hiérarchie.....	8
I-3- Niveau logique.....	9
I-3-1- Modèle en étoile .....	9
I-3-1-1- Table de faits.....	9
I-3-1-2- Table de dimension .....	10
I-3-2- Modèle en flocon .....	10
I-3-3- Modèle en constellation .....	11
I-4- Niveau physique.....	12

I-4-1- Le ROLAP.....	12
I-4-2- MOLAP (Multidimensionnal OLAP).....	12
I-4-3- HOLAP (Hybride OLAP) .....	13
I-4-5- OOLAP (Object-OLAP).....	13
I-4-5- Opération sur les cubes.....	13
I-4-5-1-Rotation (Pivot).....	13
I-4-5-2-Permutation (Switch).....	14
I-4-5-3-Décomposition (split).....	15
I-4-5-4-Imbrication/désimbrication (nest/unnest).....	15
I-4-5-5-Concaténation (push/pull).....	15
I-4-5-6-Projection (slice) .....	16
I-4-5-7-Projection agrégative.....	16
I-4-5-8-Sélection (Dice).....	17
I-4-5-9-Roll up/Drill- down .....	17
I-4-5-10-Composition d'opérations.....	18
I-4-5-11-Jointure .....	18
I-4-5-12-Union .....	19
I-5-Maintenance de modèle multidimensionnel.....	19
Conclusion.....	20

## Chapitre 2 : Datawarehouse (Entrepôt de données)

1Introduction.....	21
2- Datawarehouse.....	21
2-1- Définition .....	21
a)-Orientées sujet.....	21
b)-Intégrées.....	21
c)-Non volatiles (rien ne se perd).....	22
d)-Historisées.....	22
e)-Organisées.....	22
2-2- Les quatre problématiques .....	23
2-2-1- Lien implicite à la stratégie de l'entreprise.....	23
2-2-2- Logique d'amélioration continue .....	23
2-2-3- Les évolutions technologiques .....	24
2-2-3- Les évolutions technologiques .....	24
2-2-4- Niveau de maturité différent selon les entreprises .....	24
2-3- Les architectures.....	24
2-3-1- L'architecture réelle.....	24
2-3-2- L'architecture virtuelle.....	24
2-3-3- L'architecture remote.....	24
2-4- Les modèles de Data Warehouse.....	24
2-4-1- L'info-centre.....	25
2-4-2- Le datamart.....	26
2-4-3- Le datawarehouse centralisé.....	27
2-4-5- Les datamarts d'entreprise.....	27
2-5- Données sources.....	28
2-6- Données d'un Datawarehouse.....	28
2-6-1- Les données détaillées.....	28
2-6-2- Les données agrégées.....	29

2-6-3- Les méta données.....	29
2-6-4- Les données historisées.....	29
2-5- Limites et risques.....	29
2-5-1- Mauvais dimensionnement.....	29
2-5-2- Mauvaise agrégation.....	29
2-5-3- Mauvaise granularité.....	29
2-5-4- Mauvaises périodes d'historisation.....	30
2-5-5- Problématique de la mise à jour.....	30
2-6- Fonction de Restitution.....	30
2-6-1- Les requêteurs et outil de reporting.....	31
2-6-1-1-Outil de reporting.....	31
2-6-1-2- Requêteur.....	31
2-6-2- EIS, OLAP.....	31
2-6-2-1- EIS.....	31
2-6-2-2- OLAP.....	32
2-6-3-Datamining.....	32
2-6-3-1- Les méthodes de déduction.....	32
2-6-3-2- Les méthodes d'induction.....	32
3- Optimisation et maintenance du Datawarehouse.....	33
3-1- Optimisation de performance du datawarehouse.....	33
3-1-1- Structures redondantes.....	33
3-1-1-1- Index.....	33
3-1-1-1-1- types d'index.....	33
3-1-1-1-1-1- Index en B-arbre.....	33
3-1-1-1-1-2- Index de hachage.....	33
3-1-1-1-1-3- Index bitmap.....	34
3-1-1-1-1-4- Index de jointure.....	34
3-1-1-1-1-5- Index de jointure en étoile.....	34
3-1-1-1-1-6- Index bitmap de jointure.....	34
3-1-1-1-1-7- Index de jointure de dimensions.....	34
3-1-1-1-2- Problème de sélection d'index.....	35
3-1-1-2- Vues matérialisées.....	36
3-1-1-2-1- Problème de sélection de vues.....	37
3-1-1-2-1-1- Le PSV statique.....	37
3-1-1-2-1-2- Le PSV dynamique.....	37
3-1-1-3- Répartition de l'espace entre les vues matérialisées et les index dans les entrepôt de données.....	37
3-1-2- Structures non redondantes.....	38
3-1-2-1- La fragmentation des données.....	38
3-1-2-1-1- Les types de fragmentation.....	39
3-1-2-1-1-1- La fragmentation horizontale dérivée.....	39
3-1-2-1-1-2- La fragmentation dans les entrepôts de données.....	39
3-1-2-2- Traitement parallèle.....	39
3-2- Maintenance du datawarehouse.....	40
3-2-1- Maintenance de vues matérialisées.....	40
3-2-1-1- Maintenance des données.....	40
3-2-1-1-1- Maintenance des entrepôt de données homogènes.....	40
3-2-1-1-2- Maintenance des entrepôts de données hétérogènes.....	41
3-2-1-2- Maintenance des structures.....	41

**Chapitre 3: ETL (extraction, transformation et chargement)**

1- Introduction.....	42
2- Définition de ETL.....	42
3- Cinq phases d'alimentation.....	43
3-1- Extraction des données.....	43
3-2- Fusion des données.....	45
3-3- Nettoyage des données.....	45
3-4- Transformation des données.....	46
3-5- Chargement et filtrage des données.....	46
4- Zones de stockage intermédiaire.....	46
5- Classement des ETL.....	47
5-1- Les serveur ETL.....	47
5-2- Fonctions ETL intégrées dans le SGBD.....	47
5-3- Outils générateurs de code.....	48
5-4- Outil ETL développé maison.....	49
6-1- Le générateur de code de l'ETL.....	49
6-2- L'interface graphique de l'ETL.....	49
7- Temps réel /Batch.....	50
7-1 Temps réel – mode synchrone.....	50
7-2 Mode batch – mode asynchrone.....	50
8- ETL et gestion des métadonnées.....	50
8- Évolution de l'ETL.....	51

**Chapitre4 : Standards pour l'échange de métadonnées**

1-Introduction.....	52
2-La MDA (Model driven Architecture).....	52
2-2- Les différents modèles du MDA.....	53
2-2-1- CIM (Computation Independent Model).....	53
2-2-2- Le PIM (Platform Independent Model).....	53
2-2-3- Le PSM (Platform Specific Model).....	53
2-2-4- Le PDM (Platform Description Model).....	53
2-3- La transformation des modèles du MDA.....	54
2-3-1- De PIM vers PIM.....	54
2-3-2- De PIM vers PSM.....	54
2-3-3- De PSM vers PSM.....	54
2-3-4- De PSM vers PIM.....	54
3- Architecture de modélisation à quatre niveaux.....	55
4- Quelques modèles de l'approche MDA.....	57
4-1- MOF: (Meta Object Facility).....	57
4-2-UML (Unified Modeling Language).....	57
4-3-XMI:(XML Metadata Interchange).....	58
4-4- Common Warehouse Metamodel (CWM).....	58
5- Les implantations du CWM.....	59
6- Les profils UML.....	59
7- Patron de conception (Design pattern).....	60
8-Mondrian.....	61
8-1-Couches d'un système Mondrian.....	61
8-1-1- La couche de présentation.....	61
8-1-2- La couche de calcul.....	61

8-1-3- La couche d'agrégation.....	61
8-1-4- La couche de stockage.....	61
8-2-Schéma Mondrian.....	62

## Chapitre 5: Analyse et conception

1-Introduction.....	64
2-Solution proposée.....	64
-Conception.....	65
Table dimensionnelle.....	65
Table fait.....	65
3-Les diagrammes de cas d'utilisation.....	69
3-1-Description des cas d'utilisation.....	69
3-2-Détermination des cas d'utilisation.....	69
3-3- Diagramme de cas d'utilisation « créer les tables métadonnées.....	69
3-3-1-Acteur.....	69
3-3-2-L'objectif.....	69
3-3-3- Scénarios.....	70
3-4-Diagramme de cas d'utilisation globale.....	71
3-4-1-Description.....	71
3-4-2- Acteur.....	71
3-4-3- L'objectif.....	71
3-4-4- Scénarios.....	72
3-5- Diagrammes de cas d'utilisation détaillés.....	73
3-5-1-Diagramme de cas d'utilisation « contrôle attribut».....	73
3-5-1-1-Description.....	73
3-5-1-2-Objectif.....	74
3-5-1-3-Scénarios.....	74
3-5-2-Diagramme de cas d'utilisation « contrôle hiérarchie ».....	75
3-5-2-1-Objectif.....	75
3-5-2-2-Scénarios.....	75
3-5-3-Diagramme de cas d'utilisation « contrôle niveau ».....	76
3-5-3-1-Objectif.....	77
3-5-3-2-Scénarios.....	77
3-5-4-Diagramme de cas d'utilisation « contrôle dimension ».....	78
3-5-4-1-Description.....	78
3-5-4-2-les Scénarios.....	78
3-5-5-Diagramme de cas d'utilisation « contrôle fait ».....	79
3-5-5-1-Objectif.....	79
3-5-5-2-Scénario.....	80
4-Les diagrammes de séquence.....	80
4-1-Description de diagramme de séquence.....	80
4-2-Diagramme de séquence pour le cas « ajouterAttributs » du diagramme « contrôle Attribut».....	81
4-3-Diagramme de séquence pour le cas « contrôle colonne » de diagramme «contrôle Attribut».....	82
4-4-Diagramme de séquence pour le diagramme de cas d'utilisation « Contrôle dimension ».....	83
4-5-Diagramme de séquence pour le diagramme de cas d'utilisation « contrôle fait ».....	84
4-6-Diagramme de séquence pour le diagramme de cas d'utilisation « contrôle hiérarchie ».....	85
4-7-Diagramme de séquence pour le diagramme de cas d'utilisation « contrôle niveau ».....	86

4-8-Diagramme de séquence pour le diagramme de cas d'utilisation	
« créer les tables Métadonnées ».....	87
5-Les Diagrammes de classes .....	88
5-1- Diagrammes de classe globale.....	89
5-2-Diagramme de classe pour la couche présentation.....	90
5-2-Diagramme de classe montrant L'héritage et l'implémentation de l'interface	
ActionListener par les classes de package couchePresentation .....	91
6-Description du diagramme de package .....	92
6-1-Les packages de l'application.....	92
6-1-1-Une couche présentation .....	92
6-1-2-Une couche service.....	92
6-1-3-Une couche application .....	92
6-2-Diagramme de classe pour le package coucheApplication .....	93
6-3- Le diagramme de classe pour le package CoucheService .....	94
6-4-Relation entre le package couche application et le package couche service .....	95
7-Conclusion.....	95
Réalisation.....	96
Conclusion générale.....	100
Bibliographie	