



Université Saad Dahleb de Blida

Faculté de science  
Département Informatique



En vue d'obtenir le diplôme de Master  
Domaine : informatique

Option  
Ingénierie de Logiciel



*Thème*

CONCEPTION ET MISE EN ŒUVRE D'UNE  
SOLUTION D'AIDE A LA DECISION  
DOMAINE COMMERCE EXTERIEUR  
ALGERIEN

Promotrice :  
M<sup>me</sup> Cherfa Imane

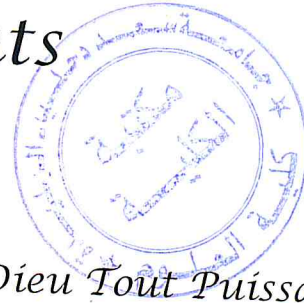
Présenté par :

Encadreur :  
M<sup>r</sup> Benmessahel Mouloud

M<sup>r</sup> Benacer Abdalah.  
M<sup>r</sup> Merdas Hichem.

MA-004-140-1

# Remerciements



*Nous tenons à remercier tout d'abord Dieu Tout Puissant,  
Tout Clément.*

*Nous remercions madame « cherfa Imane » notre  
promotrice de l'université de Blida.*

*Nous remercions monsieur BENMESSAHEL Mouloud et  
madame BOUCEDER Naziha nos encadreur dans  
l'entreprise ALGEX*

*Nous remercions spécialement nos très chers amis  
DAWED Younesse et CHEF Riad qui nous ont soutenus.*

*Nos remerciements s'adressent aussi aux enseignants du  
département Informatique de l'Université de Blida et les  
membres de jury qui nous ont fait l'honneur de juger ce  
travail.*

# SOMMAIRE

## Introduction générale

I.1 Contexte général.....	1
I.2 Problématique .....	3
I.3 Objectifs .....	4
I.4 Organisation du mémoire.....	4

## Partie I :

### Etat de l'art

#### Chapitre I : Le commerce Extérieur

1. Un peu d'histoire sur le commerce.....	5
2. Le commerce .....	6
3. Le commerce international .....	6
4. L'évolution de commerce international.....	7
5. Différents types de commerce international .....	8
6. La politique du commerce extérieur de l'Algérie.....	9
Evolution du commerce extérieur de l'Algérie par régions économiques .....	10
6.1. Importations dans le période: Années 2000 à 2010.....	10
6.2. Exportation dans le période : Années 2000 à 2010.....	10
Conclusion .....	10

#### Chapitre II : Les systèmes décisionnels

Introduction.....	11
1. Informatique Décisionnelle « Business Intelligence » .....	11

2. Historique des systèmes décisionnels.....	11
2.1. Les EIS (Executive Information System) .....	12
2.2. Infocentre.....	12
2.3. Entrepôt de données.....	13
2.4. Les bases de données multidimensionnelles.....	14
3. Le système décisionnel au sein du système d'information.....	17
Conclusion.....	18

### Chapitre III : Les entrepôts de données

Introduction.....	19
1. Définition d'un entrepôt de données.....	19
2. Objectifs de l'entrepôt de données.....	19
3. Le processus ETL .....	20
3.1. L'extraction des données.....	20
3.2. La transformation et le contrôles des données.....	20
3.3. Le chargement et transfert des données.....	20
4. Approche par cycle de vie.....	21
4.1. Définition des besoins.....	21
4.2. Modélisation dimensionnelle des données.....	22
4.3. Mise en œuvre de l'architecture.....	22
4.4. Le déploiement.....	22
5. L'Architecture de l'entrepôt de données.....	22

5.1. Système sources.....	23
5.2. Zone de préparation des données.....	23
5.3. Serveur de présentation d'un entrepôt de données.....	24
5.3.1. Le Data Mart.....	24
5.3.2. OLAP « ON LINE ANALYTICAL PROCESSING ».....	25
5.3.2.1. OLAP Multidimensionnel (MOLAP) .....	25
5.3.2.2. OLAP Relationnel (ROLAP) .....	26
5.3.2.3. OLAP hybride (HOLAP) .....	26
5.4. Portail de restitution.....	26
5.4.1. Data mining.....	26
5.4.2. Le tableau de bord.....	27
5.4.3. EIS .....	28
5.4.4. Requêteur.....	28
5.4.5. KPI (Key Performance Indicators) .....	28
5.4.6. Data Webhouse.....	29
6. La modélisation multidimensionnelle.....	29
7. Vocabulaire de la modélisation multidimensionnelle.....	29
7.1. Définition de dimension.....	29
7.2. Définition de fait (mesure) .....	30
7.3. Table de faits.....	30
7.4. Table de dimension.....	31
7.5. Relier les faits aux dimensions.....	32
7.5.1. Schéma en étoile.....	32

7.5.2. Flocon de neige.....	34
7.5.3. Le schéma en constellation.....	35
8. Approches de développement d'un entrepôt de données.....	36
8.1. Ascendante Bottum-up.....	36
8.2. Descendante Top-down.....	37
Conclusion.....	37

## Partie II :

### Conception et implémentation

#### Chapitre I : Etude d'existant

1. L'historique de l'organisme d'accueil -ALGEX- .....	38
2. Organisme d'accueil.....	38
2.1 Mission et objectifs de l'entreprise.....	39
2.2 Organisation de l'entreprise.....	39
3. Système existant.....	43
3.1 Analyse des données d'ALGEX.....	45
4. Analyse du système existant.....	46

#### CHAPITRE II : Conception

1. Définition des besoins.....	48
2. Conception de l'entrepôt de données.....	52
2.1. Modélisation dimensionnelle des données.....	52
2.2. Construction de l'entrepôt de données .....	63

2.3 Conception de la zone d'alimentations.....	66
2.3.1 Etude et planification.....	66
2.3.2 Processus de chargement.....	66
2.4 Restitution.....	67

### CHAPITRE III : déploiement

1. Introduction .....	69
2. Architecture du système.....	69
3. Configuration du système.....	70
3.1 Système de gestion de base de données.....	70
3.2 ETL Pentaho Data Integration «PDI» .....	70
3.3 Architecture d'une application Web JEE:.....	75
3.4 Hibernate .....	75
3.5 Serveur d'application TOMCAT.....	76
4. Fonctionnalités du système.....	77
4.1 Interface administration.....	78
4.2 Interface utilisateur.....	79
4.3 Les rapports.....	77
5. Sécurité du système.....	80
Conclusion générale.....	81



- Figure 1 : Une base de données relationnelle
- Figure 2 : Un cube de données
- Figure 3 : La prise de décision
- Figure 4 : entrepôt de données
- Figure 5 : Adaptation de l'approche de cycle de vie de Ralph Kimball
- Figure 6 : l'architecture détaillée de l'entrepôt de données
- Figure 7 : Tableau de bord
- Figure 8 : Exemple de table de faits
- Figure 9 : Exemple de table de dimension
- Figure 10 : Schéma en étoile
- Figure 11 : Exemple de modélisation en flocon neige
- Figure 12 : Exemple de modélisation en constellation
- Figure 13 : Organigramme d'ALGEX
- Figure 14 : Diagramme d'activité modélisant le traitement et l'édition de l'information
- Figure 15 : Schéma relationnel de la base de données « BaseALGEX »
- Figure 16 : Adaptation de l'approche de cycle de vie dimensionnel
- Figure 17 : Diagramme cas d'utilisation
- Figure 18 : Table de fait balance par pays
- Figure 19 : Balance group pays par période
- Figure 20 : taux importation et exportation par util\_f et par pays
- Figure 21: taux importation et exportation par période et par group pays et par util\_f
- Figure 22 : taux importation et exportation par période et par groupe pays
- Figure 23 : valeur d'exportation de l'entreprise
- Figure 24 : Diagramme d'activité du processus d'alimentation



- Figure 25 : interface représentant la balance par période et pays
- Figure 26 : Architecture physique du système
- Figure 27 : tache chargement du notre entrepot de donneees
- Figure 28 : etapes de la tache chargement de la table fait balance par groupe pays
- Figure 29 : etapes de la tache chargement de la table fait taux par groupe pays
- Figure 30 : etapes de la tache chargement de la table fait entreprise
- Figure 31 : etapes de la tache chargement de la table fait taux par groupe pays et utilisation final
- Figure 32 : etapes de la tache chargement de la table fait taux par pays et utilisation final
- Figure 33 : Architecture en couches d'une l'application JEE
- Figure 34 : Structure de l'application java EE
- Figure 35 : exemple de code java
- Figure 36 : console administrateur
- Figure 37 : console utilisateur
- Figure 38 : rapport statistique de balance par groupe pays
- Figure 39 : rapport statistique de balance par pays

Tableau 1 : Exportations de marchandises

Tableau 2 : Evolution de l'importation de l'Algérie dans le période: Années 2000 à 2010

Tableau 3 : Evolution de l'exportation de l'Algérie dans le période: Années 2000 à 2010

Tableau 4 : OLTP versus OLAP

Tableau 5 : Liste des tables de notre entrepôt de données

Tableau 6 : Description des attributs

## Résumé

Afin de prendre les bonnes décisions, les entreprises ont besoin d'accéder de manière efficace à des informations précises. A l'ère du tout-technologique, la source de données consultables parait infinie. Comment naviguer sur cet océan de données importantes, souvent contradictoires et trouver la bonne information qui aidera les analystes dans leurs prises de décisions ? Le Data warehouse ou le Data mart est le cœur de pratiquement toutes les solutions de Business Intelligence. Ils permettent de faire toutes les opérations analytiques et donner aux décideurs des moyens chiffrés pour évaluer les faits voulus.

Notre projet consiste à développer un outil d'aide à la décision pour la Service Statistique d'ALGEX. L'outil devra exploiter les données provenant la base de données et fournir aux décideurs les moyens de les visualiser et d'en extraire de la connaissance afin d'améliorer la gestion et les performances de l'entreprise.

## Abstract

To make good decisions, businesses need efficient access to accurate information. In the era of all-technological source of information and data seems endless. How to navigate this sea of important data, often contradictory and find the right information to assist analysts in making their decisions? The Data Warehouse or Data Mart is the heart of virtually every business intelligence solutions. They can do all the analytical operations and provide decision makers with quantitative means to assess the facts desired.

Our project is to develop a tool for decision support for Statistics of ALGEX Service.

The tool will use the data from the database and provide decision makers with the means to view and extract the knowledge to improve the management and performance of the company.

## ملخص

لاتخاذ القرارات السليمة الشركات بحاجة للوصول إلى معلومات دقيقة تنسم بالكفاءة, في عصر التكنولوجيا المعلومات لا نهاية لها و للإبحار في هذا الكم الهائل من المعلومات و الحصول على المعلومات الهامة التي تساعد المحللين في اتخاذ قراراتهم , مستودع البيانات هو قلب كل حلول في مجال ذكاء الأعمال تقريبا . فإنها يمكن أن تفعل كل العمليات التحليلية و تزود صانعي القرار بالحقائق المرجوة .

مشروعنا يتمثل في تطوير أداة لدعم القرار من أجل مصلحة الإحصاء في الشركة ألكاس

تقوم الأداة باستخراج المعلومات اللازمة لاتخاذ القرارات و ذلك لتحسين أداء و إدارة الشركة ألكاس

## INTRODUCTION GENERALE

### 1. Contexte générale :

Dans un environnement économique très pressant, tous les utilisateurs de l'entreprise ayant à prendre des décisions doivent pouvoir accéder en temps réel, aux données de l'entreprise, doivent pouvoir les traiter et en extraire l'information pertinente afin de prendre les bonnes décisions. Ils se posent alors des questions du type : « Quelle est l'évolution des chiffres d'affaires par type de magasin et par période », ou encore : « Quelle sont les résultats des ventes par gamme de produit et par région pour l'année dernière ? »...

Le système d'information opérationnel ne peut satisfaire ces besoins pour au moins deux raisons principales :

- Les bases de données opérationnelles sont trop complexes pour pouvoir être appréhendées facilement par tout utilisateur.
- Le système opérationnel ne peut être interrompu afin de pouvoir répondre à des questions qui nécessitent des calculs importants.

Une nouvelle tendance a commencé à apparaître consistant à développer un système d'information orienté vers la décision et est séparé du système d'information opérationnel. Et pour cela, il faut donc garder un historique et restructurer les données de production.

Les entrepôts de données sont la réalisation de ces nouveaux systèmes d'information.

Cette apparition a été rendue possible grâce progrès technologique à couts constants.

Parmi ce progrès, on cite :

- L'augmentation importante des capacités de stockage
- L'introduction des techniques du parallélisme dans l'informatique de gestion, technique qui permettent des accès rapides à de grandes bases de données.

L'entrepôt de données n'est pas une usine à produire l'information, mais plutôt un moyen de la mettre à la disposition des utilisateurs de manière efficace et organisée.

La mise en œuvre d'un entrepôt de données est un processus complexe. Il permet la mise en place d'un outil décisionnel s'appuyant sur les informations pertinentes pour l'entreprise, centré sur le métier utilisateur.

L'objectif à atteindre est de recomposer les données disponibles pour en donner :

- Une vision intégrée et transversale aux différentes fonctions de l'entreprise
- Une vision métier à travers différents axes d'analyse
- Une vision agrégée ou détaillée suivant le besoin des utilisateurs

Et cet objectif découle l'avantage de :

- Rendre l'utilisateur autonome (par rapport au système opérationnel) dans l'exploitation des informations. Il est, dès lors, en mesure d'analyser les procédures de gestion et l'état du système, de prendre des décisions et de mettre en œuvre des actions stratégiques.
- Consolider et organiser les données historique d'activités afin de les analyser et d'aider ainsi les gestionnaire à prendre des décisions concernant l'organisation.

## 2. Problématique

ALGEX Algérie est venue pour apporter un support efficace aux exportations, à travers ses différents produits, avec un ensemble de réseaux et de base de données d'informations spécialisées, l'accès à l'information est disponible et gratuit.

Parmi les principaux rôles d'ALGEX est d'aider à faire connaître les produits et services sur les marchés extérieurs, à travers les différents rapports d'analyses réalisés par les statisticiens, ainsi que la constitution des fichiers contenant les données sur les importations et les exportations, réalisées par la sous-direction « Communication et Relation Publique », se trouvant confronté à des problèmes de gestion d'une grande masse de données. Elle dispose de multitudes systèmes opérationnelles au niveau de toutes les sous directions. Ces données couvrent la majeure partie de ses activités de base. Au niveau de la structure opérationnelle de la sous-direction d' « Analyse du Commerce Extérieur », plusieurs traitements sont effectués chaque mois, et les données produites sont transférées de cette structure vers d'autres structures pour élaborer chaque année, en concertation avec d'autres informations enregistrées lors des visites des exportateurs à ALGEX, aux chambres de commerce, et celle collectés par les cadres d'ALGEX

lors de leur missions et visite aux foires et salon nationaux et internationaux, une base de données des exportateurs algériens qui est utilisée pour la création d'un annuaire des entreprises exportatrices, l'alimentation du répertoire du site, ainsi que différents documents d'analyses.

Mais avec le volume important de données, les employés d'ALGEX se trouvent dans l'incapacité d'analyser, ou d'acheminer de façon rapide et correcte par ces moyens disponibles, le volume de ces informations régulièrement.

Les principales difficultés auxquelles elle devrait faire face sont :

- Grande masse de données est stockée dans plusieurs tables, ce qui rend l'utilisation et l'analyse très difficile ;
- L'information traverse un long chemin pour qu'elle soit utilisable ou pour arriver à l'analyse, ce long cheminement implique la diminution de la fiabilité de l'information.
- Difficulté de garder l'historique même sur une période de quelques mois, ce qui rend impossible le suivi avec les entreprises qui exporte régulièrement.
- La redondance des données due à de multiples écritures.
- Temps d'accès potentiellement important si les requêtes sont complexes.

En plus d'une exploitation réduite des données, le processus de collecte et de présentation de l'information aux décideurs se fait d'une manière archaïque.

Nous nous sommes intéressés dans ce travail, à la conception et à la réalisation d'un système décisionnel qui résorberait les anomalies citées ci haut.

### **3. Objectifs :**

Notre projet consiste à développer un outil d'aide à la décision pour le service « Statistiques » d'ALGEX. L'outil devra exploiter les données provenant de la base de données, comme il devra fournir aux décideurs les moyens de les visualiser, et d'en extraire de la connaissance afin d'améliorer la gestion et les performances de l'entreprise.

#### 4. Organisation de mémoire

Dans ce présent mémoire nous allons présenter le travail qui nous a été confié, qui a pour thème « Conception et mise en œuvre d'une solution d'aide à la décision » dans le domaine du Commerce Extérieur Algérien. Pour cela, nous l'avons organisé en deux parties.

La première partie a été consacrée à l'état de l'art sur les systèmes décisionnels, et sur le commerce en Algérie. Elle comporte trois chapitres. Le premier chapitre aborde le domaine du Commerce International. Il expose les concepts, et les définitions du commerce international, ainsi que son évolution. Le deuxième chapitre présente le domaine des Systèmes Décisionnels, il introduit les différents outils d'aide à la décision, ainsi que leur évolution avec le temps, comme il discute l'intégration du système décisionnel au sein du système d'information. Le troisième chapitre est consacré aux entrepôts de données, il expose les concepts, les définitions et la terminologie des entrepôts de données, et propose la démarche suivie pour l'élaboration de l'entrepôt de données.

La deuxième présente notre solution, à savoir l'outil d'aide à décision, elle comprend également trois chapitres. Dans le premier chapitre de cette partie, nous présentons notre organisme d'accueil qui est l'entreprise ALGEX, nous abordons son organisation, ses missions, nous terminons ce chapitre par une analyse sur le fonctionnement du système existant. Le chapitre qui suit présente notre contribution, à savoir la conception de l'entrepôt de données. Le dernier chapitre sera consacré à la validation expérimentale de notre travail.

Nous terminons ce document par une conclusion générale dans laquelle seront présentés un bilan et les perspectives des travaux réalisés.

# Partie I

Etat de l'art



# **Chapitre I**

Commerce international

## Introduction

Dans ce chapitre, nous abordons quelques notions sur le commerce, le commerce international et l'évolution du commerce international. Nous allons voir en fin la politique du commerce extérieur en l'Algérie, ainsi que l'évolution du commerce extérieur en Algérie.[3]

### 1. Un peu d'histoire sur le commerce

Les habitants de l'antique cité de Carthage (l'emplacement de l'actuelle ville de Tunis, au nord de l'Afrique), qu'on appelait Carthaginois ou Phéniciens, étaient d'habiles navigateurs et des commerçants hors pair. Ils étaient connus des Égyptiens contemporains de Ramsès II sous le nom de Peuple de la mer. Ils furent à l'origine de notre alphabet phonétique.

Les cités phéniciennes comme Sidon ou Tyr créèrent des colonies tout autour de la mer Méditerranée, dans les anses rocheuses propices à l'amarrage des navires.

L'agriculture, l'arboriculture et la céramique furent les principales ressources de Carthage, mais le commerce compta aussi pour quelque chose dans sa prospérité. Les premiers échanges commerciaux se déroulaient comme suit.[3]

Les vendeurs carthaginois tirent leurs marchandises de leurs vaisseaux et les étalent sur le rivage. Ils réintègrent ensuite leurs bâtiments maritimes et y allument des feux, de manière à faire beaucoup de fumée. Les autochtones du pays, apercevant cette fumée, viennent au bord de la mer, étalent de l'or sur le sol en paiement de la marchandise offerte, puis s'éloignent.

Les Carthaginois sortent alors de leurs vaisseaux et examinent la quantité d'or apportée; si elle leur paraît suffisants, ils l'emportent sur leurs navires et quittent. En revanche, si la quantité d'or ne répond pas à leurs attentes, ils retournent dans leurs vaisseaux et attendent tranquillement de nouvelles offres. Les acheteurs reviennent alors et ajoutent de l'or jusqu'à ce que les Carthaginois soient satisfaits. Ni les uns ni les autres n'engagent d'échanges verbaux ou physiques violents. Les Carthaginois ne touchent jamais à l'or avant que la quantité offerte n'équivaille selon eux à la valeur de leurs marchandises, et les acheteurs n'emportent pas les marchandises avant que les Carthaginois aient pris l'or.

Ce fut là le début du commerce international. Depuis des millénaires, les échanges commerciaux dépassent les frontières nationales.[3]

### **2. Le Commerce**

Le commerce désigne l'activité d'achat et de revente de biens et de services, en particulier l'achat dans le but de revendre avec un profit ou un bénéfice. Le commerce a ses lois propres qui sont recueillies dans le code de commerce, et ses propres juridictions nationales ou internationales. Ayant d'autres finalités que de faire un profit, beaucoup d'activités économiques ou professionnelles comme l'éducation, la police, la culture, la médecine, la justice, l'architecture, la religion, etc. ne relèvent pas du domaine du commerce, mais des lois et des tribunaux civils ou administratif. Avec ses puissants relais d'opinion, l'organisation mondiale du commerce s'efforce d'étendre sa compétence et ses principes de profitabilité sur toutes les activités non commerciales, en leur contestant le droit d'avoir des législations et des valeurs désintéressées.[3]

### **3. Le commerce international**

Le commerce international est une source de prospérité pour les nations, bien qu'il semble davantage une cause seconde de cette prospérité qu'une cause première.

Le commerce dépend en amont de l'habileté productive, elle-même déterminée sur le long terme par les capacités d'organisation, le progrès technique et l'innovation, ainsi que la démographie.

Le commerce est - avec la guerre et la diplomatie- l'un des trois principaux modes par lesquels les états entrent en relation. Il est à la fois source de dialogue, de paix et de conflits.

L'histoire montre que des périodes de protectionnisme, ont coïncidé avec un fort développement économique endogène et un commerce international malgré tout florissant ; ces phases paradoxales illustrent simplement que ni l'idéologie du libre-échange intégral jamais réalisé- ni celle du protectionnisme, n'épuisent la réalité des rapports commerciaux, eux-mêmes largement déterminés par des logiques de souveraineté. Le commerce est également un attribut de la puissance. Avec l'investissement, il offre une capacité de projection et

d'influence au-delà des frontières. Il participe à l'établissement de la hiérarchie des puissances et contribue à leur renforcement.[3]

#### 4. L'évolution du commerce international

L'évolution du commerce mondial est donc étroitement liée à celle de la production des nations, de leur puissance, ainsi qu'à l'état de leurs relations. L'étude du commerce dans le long terme est donc indissociable de celle de l'environnement économique, démographique, institutionnel et géopolitique dans lequel il s'inscrit.

Ces deux dernières décennies, les échanges commerciaux internationaux ont grandement augmenté, plus particulièrement pour les pays développés, et pour les nouveaux pays industrialisés, favorisant la croissance de ces derniers. Les pays les moins avancés n'ont pas connu une telle hausse des échanges commerciaux internationaux. Le volume du commerce mondial a été multiplié par quinze à partir des années 1950 et 1960 et il a encore triplé entre la chute du Mur de Berlin en 1990 et 2010.

La figure suivante montre les exportations de marchandises entre l'année 1870 et l'année 1998, en pourcentage du PIB.

	1870	1913	1950	1973	1998
Europe Occidentale	8,8	14,1	8,7	18,7	35,8
Occident non-européen	3,3	4,7	3,8	6,3	12,7
Europe de l'est et ex URSS	1,6	2,5	2,1	6,2	13,2
Amérique latine	9,7	9	6	4,7	9,7
Asie	1,7	3,4	4,2	9,6	12,6
Afrique	5,8	20	15,1	18,4	14,8
Monde	4,6	7,9	5,5	10,5	17,2

Tableau N° 1 : Exportations de marchandises (1870 –1998, en % du PIB, prix constants 1990)[2]

## 5. Différents types de commerce international

Le commerce international comprend toutes les opérations sur le marché mondial. Il est l'organe regroupant les divers pays du monde engagés dans la production des biens destinés aux marchés étrangers.

Le commerce mondial comprend :

- Le commerce de concentration: qui consiste à assembler les petites productions locales ou régionales dans des comptoirs créés à cette fin, en quantités convenables pour être manipulés sur le marché mondial.
- Le commerce de distribution: consiste à se procurer les marchandises en très grandes quantités sur le marché mondial et à les emmagasiner pour les distribuer aux consommateurs sur le plan mondial.

Le commerce extérieur s'effectue entre les habitants de deux ou plusieurs pays. Il comprend les importations, achats à l'étranger et les exportations, ventes à l'étranger des biens produits à l'intérieur d'un pays.

Le commerce de transit n'est rien d'autre que la faculté accordé à un produit originaire du pays X et destiné à la consommation dans le pays Y, de traverser le pays Z sans acquitter les droits de douane.

On distingue le commerce général qui est l'ensemble de commerce d'importation, d'exportation et de transit, du commerce spécial qui ne comprend que le commerce d'importation et le commerce d'exportation. Il ne comprend pas non plus le commerce en relation avec des zones franches et des entrepôts douaniers.

## 6. La politique du commerce extérieur de l'Algérie

L'Algérie pratique des droits de douane à l'importation simplifiés en nombre et en structure et avec des taux réduits.

Le niveau de ces droits atteint 25 %, 40 % ou 50 % pour certaines marchandises alors que pour d'autres le niveau n'est que de 3,7 % ou 15 %.

Les exonérations de droits de douane sont rares. A noter que pour une détermination adéquate du prix à l'importation, est assis sur le prix CIF exprimé en dinars, les droits de douane, la redevance douanière (0,40 %), la taxe de dossier en douane (2 %) et le coût de transit et de manutention (0,17 %).

Au prix ainsi déterminé s'ajoute la TVA. Pour certains produits il faut ajouter la taxe spécifique additionnelle.

Le commerce d'importation ne fait l'objet d'aucune restriction quantitative. Aucune restriction ne touche l'exportation.

Pour stimuler l'exportation, la Caisse Algérienne de Garantie des Exportations (CAGEX) qui garantit le recouvrement des droits et protège contre les risques commerciaux, politiques, de non transfert et de catastrophes.

Quant au Fonds spécial de promotion des exportations, il finance la participation des entreprises à des foires et expositions internationales, la prospection de marchés extérieurs et les études d'amélioration des produits exportables.

Le crédit-bail international de biens immobiliers à usage professionnel est soumis à un régime douanier suspensif des droits douaniers.

A la fin de la location et en cas de levée de l'option de l'achat par le crédit-preneur, ce dernier procède au dédouanement pour la mise à la consommation du bien acquis.[2]

### 6.1. Evolution du commerce extérieur de l'Algérie par régions économiques :

#### Importations dans le période: Années 2000 à 2010

REGIONS ECONOMIQUES	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
U.E ( EX CEE )	5 256	5 903	6 732	7 954	10 097	11 255	11 729	14 427	20 985	20 772	20 704
O.C.D.E ( hors U.E )	2 194	2 125	2 485	2 242	3 071	3 506	3 738	5 363	7 245	6 435	6 519
Autres pays d'EUROPE	603	636	757	855	1 097	1 058	777	715	659	728	388
AMERIQUE DU SUD	142	269	385	567	1 166	1 249	1 281	1 672	2 179	1 866	2 380
ASIE ( sans les pays arabes )	599	579	943	1 206	1 952	2 506	3 055	4 318	6 916	7 574	8 280
OCEANIE	64	92	127	47	56	31	-	-	-	2	
pays Arabes ( sans UMA )	144	179	366	418	525	387	493	621	705	1 089	1 262
Pays du MAGHREB	52	72	127	120	169	217	235	284	395	478	544
Autres pays d'AFRIQUE	119	85	87	125	175	148	148	231	395	350	396
TOTAL	9 173	9 940	12 009	13 534	18 308	20 357	21 456	27 631	39 479	39 294	40 473

Tableau N° 2 : Evolution de l'importation de l'Algérie dans le période: Années 2000 à 2010[2]

**Exportation dans le période : Années 2000 à 2010**

REGIONS ECONOMIQUES	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
U.E ( EX CEE )	13 792	12 344	12 100	14 503	17 396	25 593	28 750	26 833	41 246	23 186	28 009
O.C.D.E ( hors U.E )	5 825	4 549	4 602	7 631	11 054	14 963	20 546	25 387	28 614	15 326	20 278
Autres pays d'EUROPE	181	87	130	123	91	15	7	7	10	7	10
AMERIQUE DU SUD	1 672	1 037	951	1 220	1 902	3 124	2 398	2 596	2 875	1 841	2 620
ASIE ( sans les pays arabes )	210	476	456	507	686	1 218	1 792	4 004	3 765	3 320	4 082
OCEANIE	0	23	38	0	-	-	-	55	-	-	-
pays Arabes ( sans UMA )	55	315	248	355	521	621	591	479	797	564	694
Pays du MAGHREB	254	275	250	260	407	418	515	760	1 626	857	1 281
Autres pays d'AFRIQUE	42	26	50	13	26	49	14	42	365	93	79
<b>TOTAL</b>	<b>22 031</b>	<b>19 132</b>	<b>18 825</b>	<b>24 612</b>	<b>32 083</b>	<b>46 001</b>	<b>54 613</b>	<b>60 163</b>	<b>79 298</b>	<b>45 194</b>	<b>57 053</b>

**Tableau N° 3 : Evolution de l'exportation de l'Algérie dans le période: Années 2000 à 2010[2]**

**Conclusion**

Aujourd'hui, ALGEX a à sa disposition une masse de données importante, pour cela elle se trouve dans l'incapacité d'analyser le volume de donnée de façon rapide, correcte et efficace.

Les principales difficultés auxquelles fait face l'entreprise ALGEX sont :

- Une grande masse de données est stocké dans plusieurs bases de données, ce qui rend difficile l'interrogation et l'analyse.
- Le retard dans l'établissement des rapports d'activité, cela peut prendre quelques mois. Ce qui fait que l'information n'est pas disponible au moment voulu.

Pour essayer de faire face à ses problèmes, nous avons essayé de faire recours à l'informatique décisionnelle, qui dans nos jours, a montré ses preuves pour ce genre de problèmes.

Les outils d'aide à la décision ont connu l'évolution suivante :

## 2.1 Les EIS (Executive Information System)

L'étape suivante est née en même temps que le PC, l'ordinateur personnel et sa démocratisation chez les cadres dirigeants. Leur donnant l'impression de contrôler l'activité de leur entreprise, les EIS (Executive Information System) proposaient un tableau de bord permanent des indicateurs clefs de l'entreprise.

Un des grands groupes français du bâtiment avait ainsi installé sur le bureau de son PDG un PC équipé d'un EIS qui lui servait uniquement à visualiser quelques indicateurs tels que le cours de l'action, le nombre d'employés, le montant des contrats en cours, et celui des nouveaux contrats signés. Cette facilité apparente dans la mise à disposition des informations a été le cauchemar des informaticiens qui avaient bien du mal à expliquer à leur patron que, même si l'affichage d'un graphique sur son poste lui paraissait simple, il supposait une mécanique complexe en amont de remontée et d'agrégation des informations.

C'est néanmoins par l'EIS que l'informatique décisionnelle est sortie de la tour d'ivoire des services informatiques pour prendre place sur le bureau des utilisateurs.

## 2.2 Infocentre

La première solution trouvée était une solution de bon sens. Puisque les analystes avaient besoin des mêmes données que les opérationnels, et en même temps, mais que les requêtes des premiers bloquaient les requêtes des seconds, il suffisait donc de donner à chacun sa copie des données.

On a donc dupliqué les bases de données de production, c'est-à-dire que chaque jour, chaque semaine, ou chaque mois, une copie des informations de production était réalisée, à l'identique, sur un autre ordinateur, spécialement pour les analystes. Les équipes opérationnelles pouvaient continuer à utiliser leurs applications, sans être nullement perturbés par les requêtes analytiques, et les analystes pouvaient prendre le risque de lancer des requêtes complexes, analysant par exemple le chiffre d'affaires suivant plusieurs dimensions (clients, produits, fournisseurs), sur plusieurs années, mois par mois, sans prendre le risque de bloquer le système opérationnel.

L'infocentre est un concept commercial IBM de la fin des années 1970. Les informaticiens et les utilisateurs spécialisés dans la rédaction de requêtes pouvaient accéder rapidement à des données demandées par les dirigeants. [Gar, 2000]



Avant l'infocentre, la réponse à une requête supposait le développement d'un programme informatique qui générait des délais d'attente importants et l'incompréhension des décideurs qui avaient le sentiment que leurs demandes étaient pourtant plutôt simples.

Dans l'infocentre les requêtes étaient réalisées sur des répliqués des bases de production à l'aide de langage de requêtes graphiques tel le QBE, query by example.

### 2.3 Entrepôt de données

Après plusieurs années d'utilisation des infocentres, c'est-à-dire la duplication de données de production, les services informatiques ont imaginé une évolution intelligente de ce mode de stockage. Ils ont en effet constaté que les informations traitées dans les applications opérationnelles étaient très différentes de celles interrogées dans les applications décisionnelles. Le nombre de tables, de fichiers interrogés dans une même requête est bien plus important dans l'aide à la décision, le nombre d'indicateurs calculés également. En revanche, les applications décisionnelles se contentent presque toujours de lire les données. Elles n'ont jamais à écrire de nouvelles informations dans les bases de données.

Autre constat, les questions posées par un décideur impliquent fréquemment des informations stockées dans plusieurs applications ou base de données. Très souvent, lorsqu'il s'agit de calculer la rentabilité des clients, le système décisionnel doit exporter les données sur la gestion commerciale (factures, commandes), les données comptables (délais de règlement, impayés), mais également de la gestion de production (coût des produits fabriqués). Le fait de dupliquer ces bases de données dans un infocentre ne simplifie en rien ces extractions.

On a donc pensé à mettre en place, en sortie des bases de production, un entrepôt de données. Cet entrepôt est dédié uniquement au stockage des données décisionnelles, permet de réconcilier les différentes sources initiales de données, et les applications de production.

Le tableau suivant donne une comparaison entre l'Infocentre et l'entrepôt de données.

Infocentre	Entrepôt de données
Collection de données	Collection de données
Orientées sujet	Orientées sujet
Intégrée	Intégrée
Volatiles	Non volatiles
Actuelles	Historisées
Organisées pour le support d'un processus de décision ponctuelle	Organisées pour le support d'un processus d'aide à la décision
Outils	Architecture

**Tab4** : Comparaison entre Infocentre et entrepôt de données

## 2.4. Les bases de données multidimensionnelles

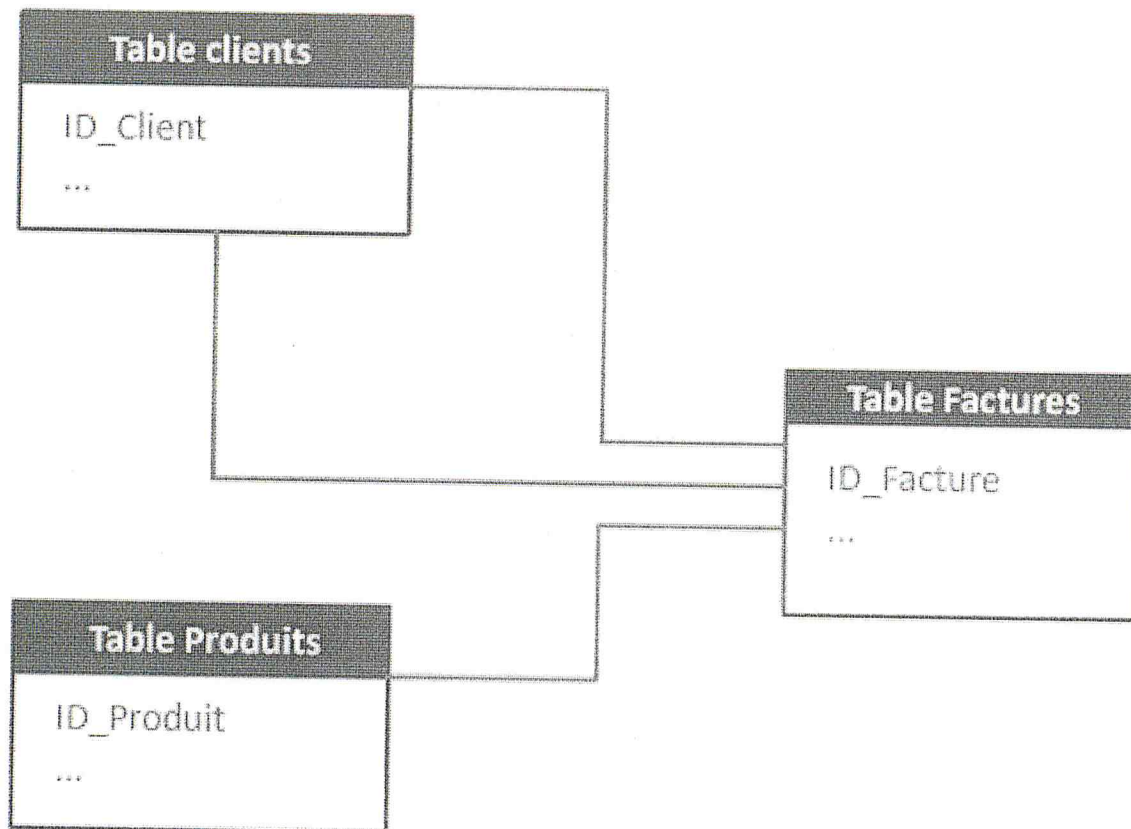
Quand un manager parle de ses résultats, il leur associe toujours une variable. Il parle d'une progression d'un mois sur l'autre, d'une gamme de produit plus performante que l'autre, d'un bénéfice plus important dans un pays que dans l'autre, d'un commercial plus efficace que son collègue.

Tout simplement parce que d'un point de vue métier, l'information de base n'a aucune importance si elle n'est pas affectée d'une dimension. On cherche le chiffre d'affaires par trimestre, la rentabilité par client, la marge par commercial et par région.

Chacun de ces attributs est appelé une dimension. Il s'agit d'un axe suivant lequel les informations seront analysées.

Un système décisionnel est donc constitué d'indicateurs, analysés suivant des dimensions.

Or les bases de données relationnelles sont constituées d'un ensemble de tables à deux dimensions (voir Figure 1). Comme une suite de feuilles Excel, composées de lignes et de colonnes, qui seraient conservées dans un même classeur. Mais l'intérêt des bases de données relationnelles est dans le relationnel, c'est-à-dire dans la capacité à mettre en relation les feuilles les unes avec les autres. Ainsi une base de données clients comportera plusieurs tables : une table facture, une table client. Dans la première sera stocké le détail des factures, dans la seconde seront stockées les coordonnées des clients. Inutile donc de répéter les coordonnées dans chaque facture, un lien permettra d'aller chercher pour chaque ligne de la table facture les coordonnées correspondantes dans la table client.



**Figure 1 :** Une base de données relationnelle [Kimball, 2003]

Lorsque l'on lance des requêtes décisionnelles comportant de nombreuses dimensions d'analyse, comme par exemple, « le calcul de la marge par client, et par gamme de produits, mois par mois depuis un an », le système a besoin d'accéder à de nombreuses tables. Une telle requête peut être très longue à exécuter, voir même parfois bloquer totalement le serveur.

On a, pour résoudre ce problème, inventé les bases de données multidimensionnelles. Ces bases se présentent sous forme d'un cube (voir figure 2). Chaque dimension du cube représente une dimension d'analyse. Il est difficile de représenter sur le papier plus de trois dimensions, mais un cube d'analyse peut parfaitement comporter dix ou vingt dimensions d'analyse simultanées. Ce nouveau modèle de bases multidimensionnelles a été inventé par le Docteur Edgar Codd, déjà considéré comme le Père des bases de données relationnelles (il a inventé ce modèle en 1969).

Ces bases de données relationnelles étaient nées du modèle OLTP (On Line Transactional Processing). Le Docteur Codd a poursuivi dans la même direction en définissant pour les bases multidimensionnelles le modèle OLAP (On Line Analytical Processing).

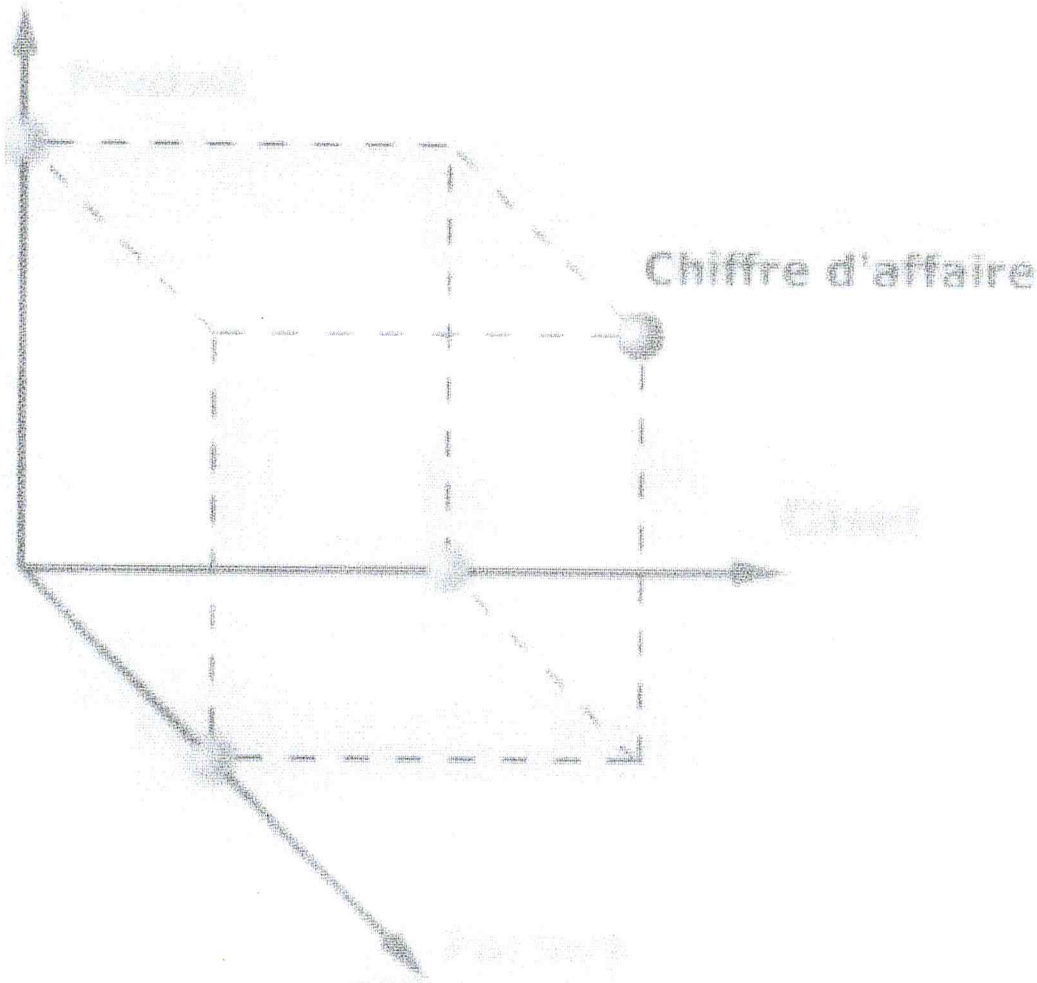


Figure 2 :Un cube de données

Chaque base de données multidimensionnelle pré-calculé et stocke toutes les informations, au croisement de chacune des dimensions. Le chiffre d'affaire réalisé par chaque client, sur chaque produit, dans chaque magasin, durant chaque mois, sera pré-calculé, et conservé dans le cube, au croisement de chacune des dimensions prévues. Ainsi, lors du lancement d'une requête décisionnelle complexe, le système aura simplement à lire les croisements de chaque colonne du cube, pour extraire des informations qui auront déjà été calculées.

Bien entendu, la phase d'alimentation, c'est-à-dire de pré-calcul et de stockage des informations agrégées, est plus longue. Mais nous l'avons vu en introduction, les requêtes décisionnelles sont rarement lancées en temps réel. Les données analysées sont fréquemment des données mensuelles, au mieux hebdomadaires, et la phase d'alimentation peut être programmée en automatique, de nuit ou le week-end.

La base de données multidimensionnelle OLAP est incontestablement le cœur de tout système décisionnel. La vision multidimensionnelle qu'elle apporte des données métiers en fait la pierre angulaire de la prise de décisions. Sans pour autant réduire l'importance des autres briques d'un système décisionnel, un véritable outil d'aide à la décision ne saurait être construit sans s'appuyer sur une base de données OLAP.

### 3. Le système décisionnel au sein du système d'information

Le développement de nouveaux marchés des entreprises passe par la maîtrise de la connaissance des divers environnements sur lesquels pourront croître ces nouveaux marchés, et par la connaissance d'informations légales, juridiques, fiscales, politiques, techniques et d'informations nécessaires aux marketings mix. Toutes ces sources d'informations issues du monde entier, alimentent par tous les canaux possibles et imaginables toutes les strates de l'entreprise qui doit alors les digérer pour en extraire les informations les plus importantes.

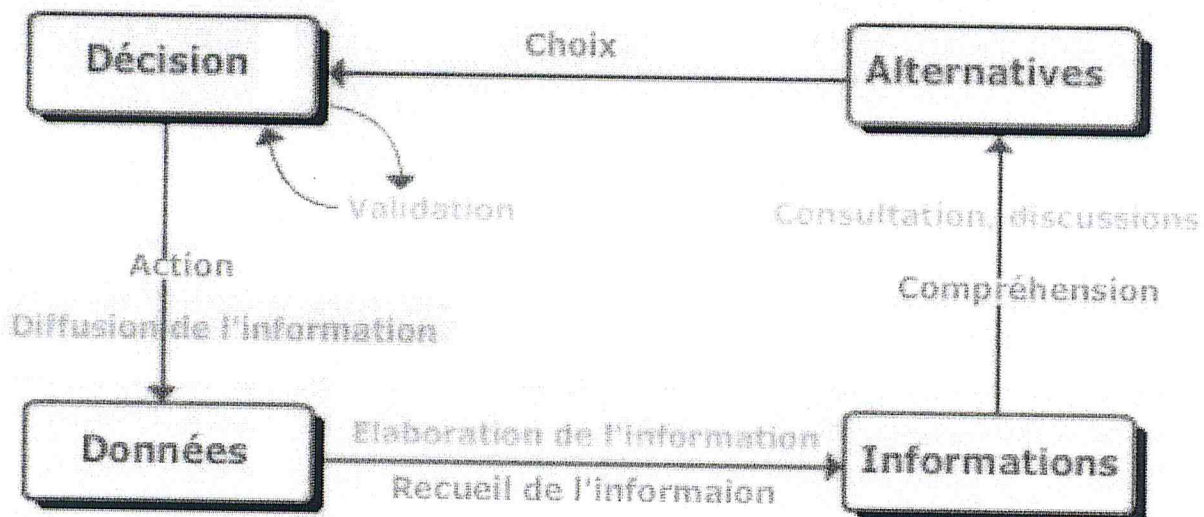


Figure 3: La prise de décision. [Goglin, 2001]

L'administration de l'entreprise va générer des informations légales et commerciales et ce par domaines (la paie, la comptabilité, la facturation, le suivi commercial, etc.). Ces informations vont alors pouvoir être analysées par domaines grâce à des indicateurs adaptés. Ces éléments, une fois synthétisés, vont permettre au décideur d'assumer pleinement sa tâche de prise de décision.

Dans la figure 3 on observe que la prise de décision dans une entreprise passe par un recueil d'information, ces informations sont nécessaire pour les analystes qui vont les consulter et les synthétiser afin de faciliter la tâche du décideur. Après la validation des décisions, de nouvelles données seront diffusées pour accomplir des objectifs précis.

Les bases sont alimentées par des services divers dans une entreprise qui utilisent une informatique nomade dans différents services tel que le service marketing, service commercial, service administratif et la direction.

Un système d'information n'est réellement complet que s'il est apte à intégrer les contraintes des systèmes d'information extérieurs. Autrement dit, un vrai système d'information n'est pas une vue orientée vers l'intérieur de l'entreprise, mais bien une ouverture vers l'extérieur et donc tourné vers le monde. [Goglin, 2001]

### **Conclusion**

Tous l'enjeu des systèmes d'information de demain consistera à concilier les finalités divergentes, pour ne pas dire contradictoire, qui sont la délégation sur le terrain de la prise de décision, avec la nécessité de cohérence, qui se traduit par une intégration croissante supposée mieux gérer la complexité, et garantir les objectifs financiers. Ces deux pressions ne s'excluent pas mais se complètent.

# **Chapitre III**

Les entrepôts de données

## Introduction

Dans ce troisième chapitre, nous posons les bases des études des entrepôts de données. Nous examinerons les principales composantes de l'entrepôt de données, y compris la modélisation multidimensionnelle.

### 1. Définition d'un entrepôt de données

Un entrepôt de données est une collection de données thématiques, intégrées, non volatiles et historiées organisées pour la prise de décision. [Bill Inmon, 2000].



Selon les caractéristiques citées par Inmon, on a les données orientées sujets dans le sens où les informations des différents métiers sont regroupées, et ne tient pas compte de l'organisation fonctionnelle des données. On a aussi des données intégrées, ou l'entrepôt de données est un projet d'entreprise, il concerne différentes sources de données, avant d'intégrer ces dernières, elles doivent être mise en forme et unifiées.

Ainsi, les données de l'entrepôt sont non volatiles, on parle ici de la traçabilité des informations et des décisions prises, ainsi, l'information une fois stockée dans l'entrepôt ne peut être supprimée.

L'historisation des données est nécessaire pour suivre dans le temps, l'évolution de différentes valeurs des indicateurs à analyser. On parle donc de la mise en place d'un référentiel temps, qui est associé aux données, afin de les identifier dans la durée de valeur précise.

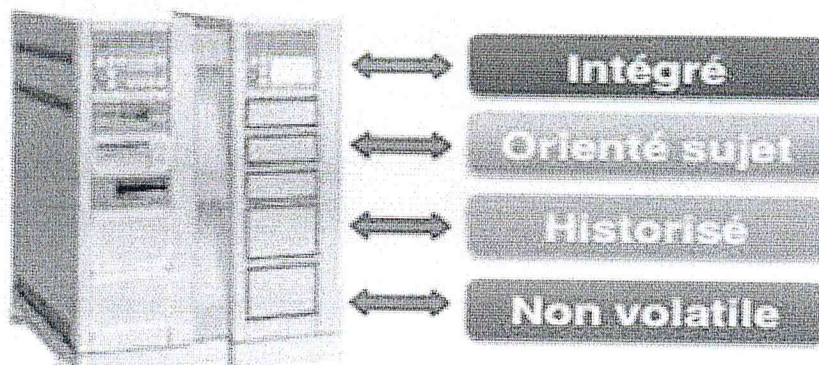


FIGURE 4 : Entrepôt de données [1]

### 2. Objectifs de l'entrepôt de données

L'atout principal d'une entreprise réside dans les informations qu'elle possède. Ces informations se présentent généralement sous deux formes : les systèmes opérationnels qui enregistrent les



données, et l'entrepôt de données. En bref, les systèmes opérationnels représentent l'emplacement de saisie des données, et l'entrepôt de données l'emplacement de la restitution. L'utilisation de l'entrepôt de données présente un certains nombres d'objectifs :

- Rend accessibles, et cohérentes les informations de l'entreprise ;
- Constitue une source d'information souple et adaptable ;
- Constitue la base décisionnelle de l'entreprise

### **3. Le processus ETL**

Le processus ETL est une opération de migration de données, (de différents systèmes à un entrepôt de données) qui consiste à la rendre facilement lisible et disponible. Ce processus représente une part majeure des traitements, et nécessite une attention régulière tout au long du cycle de vie du système, dans la mesure où il est garant de la qualité des données finale. C'est le processus ETL qui se décompose en trois phases : l'extraction, la réparation/transformation et le chargement.

#### **3.1. L'extraction des données**

L'extraction des données est la première des étapes des systèmes ETL. Le but de cette étape est premièrement de localiser dans le système opérationnel les données qu'il est nécessaire de prélever. Ensuite d'effectuer l'extraction des données depuis les systèmes sources.

#### **3.2. La transformation et le contrôles des données**

Le transformateur se charge de nettoyer les données afin de les rendre compatibles avec la granularité et le schéma de l'entrepôt de données. [Kimball1996]

Toutes les données ne sont pas utilisables telles qu'elles. Elles méritent d'être vérifiées, reformatées, nettoyées afin d'éliminer les valeurs aberrantes et les doublons, et consolidées.

#### **3.3. Le chargement et transfert des données**

Dans cette étape, les données doivent être insérées dans l'entrepôt. La gestion des données est une des fonctions primordiales de l'entreprise. Il est en effet important de garantir la qualité des données circulant au sein du système d'information, et plus particulièrement du système décisionnel.

#### 4. Approche par cycle de vie

Ce schéma représente la succession des tâches de haut niveau (macro tâches) nécessaires à la conception, au développement et au déploiement d'entrepôts de données efficaces. Il décrit le cheminement du projet dans son ensemble, chaque rectangle sert de poteau indicateur ou de borne.

Dans notre étude nous avons adoptés l'approche de cycle de vie de Ralph Kimball.

La figure suivante (fig.5) illustre en détail l'approche.

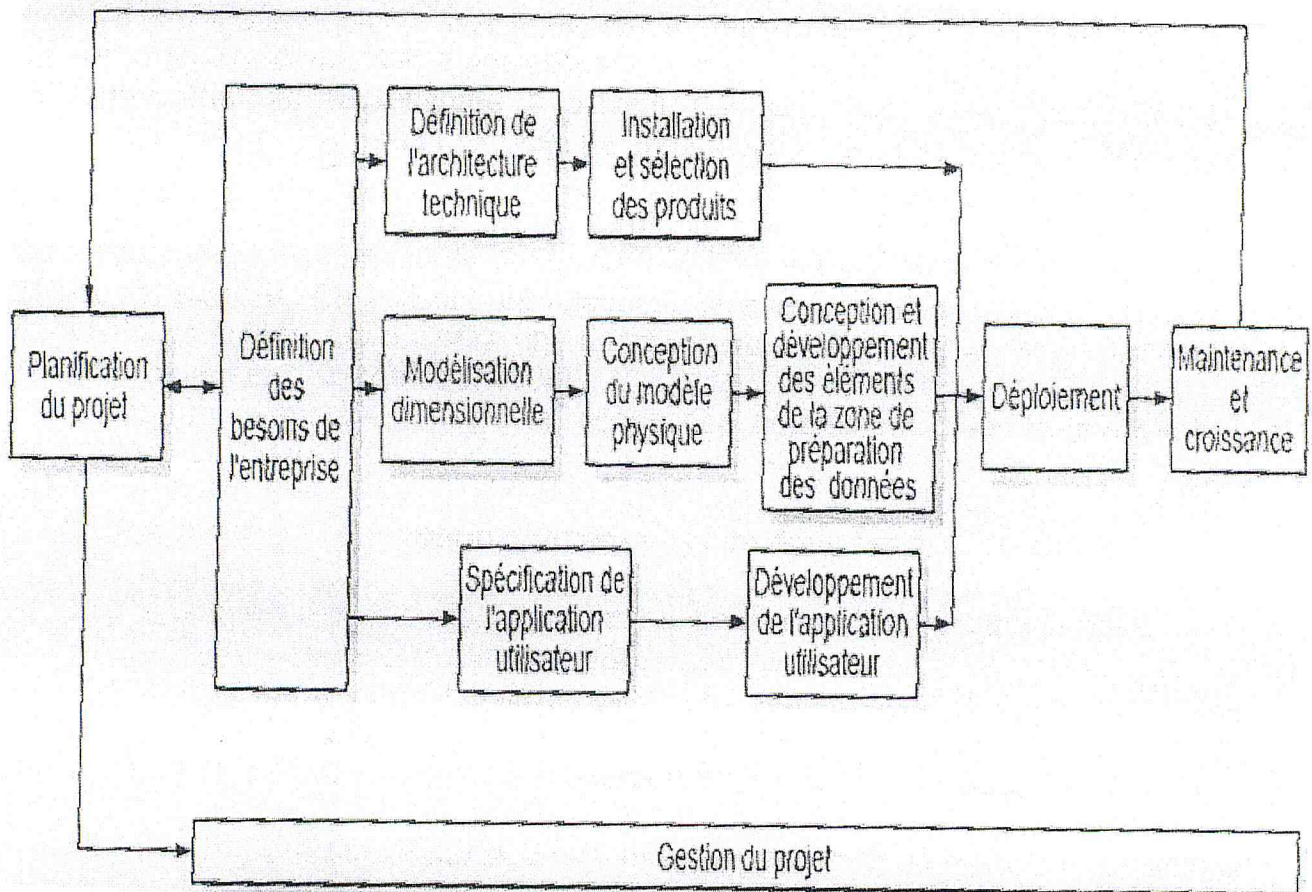


Figure 5 : Adaptation de l'approche de cycle de vie de Ralph Kimball [Kimball, 2003]

##### 4.1. Définition des besoins

Les chances de succès d'un entrepôt de données se trouvent considérablement accrues par la bonne compréhension des utilisateurs, et de leurs besoins. Sans elle, l'entrepôt de données deviendrait rapidement un exercice vain pour l'équipe des concepteurs. L'approche utilisée pour identifier les besoins analytiques des employés diffère de manière significative de la traditionnelle analyse des besoins pilotée par les données.

Ces besoins constituent le point de départ des trois trajectoires parallèles, qui sont la technologie, les données et les interfaces utilisateur.

### **4.2. Modélisation dimensionnelle des données**

La définition des besoins détermine les données requises pour répondre aux besoins d'analyse des utilisateurs. La conception de modèles de données destinés à assurer ces analyses nécessite une approche différente de celle utilisée lors de la conception de systèmes opérationnels.

### **4.3. Mise en œuvre de l'architecture**

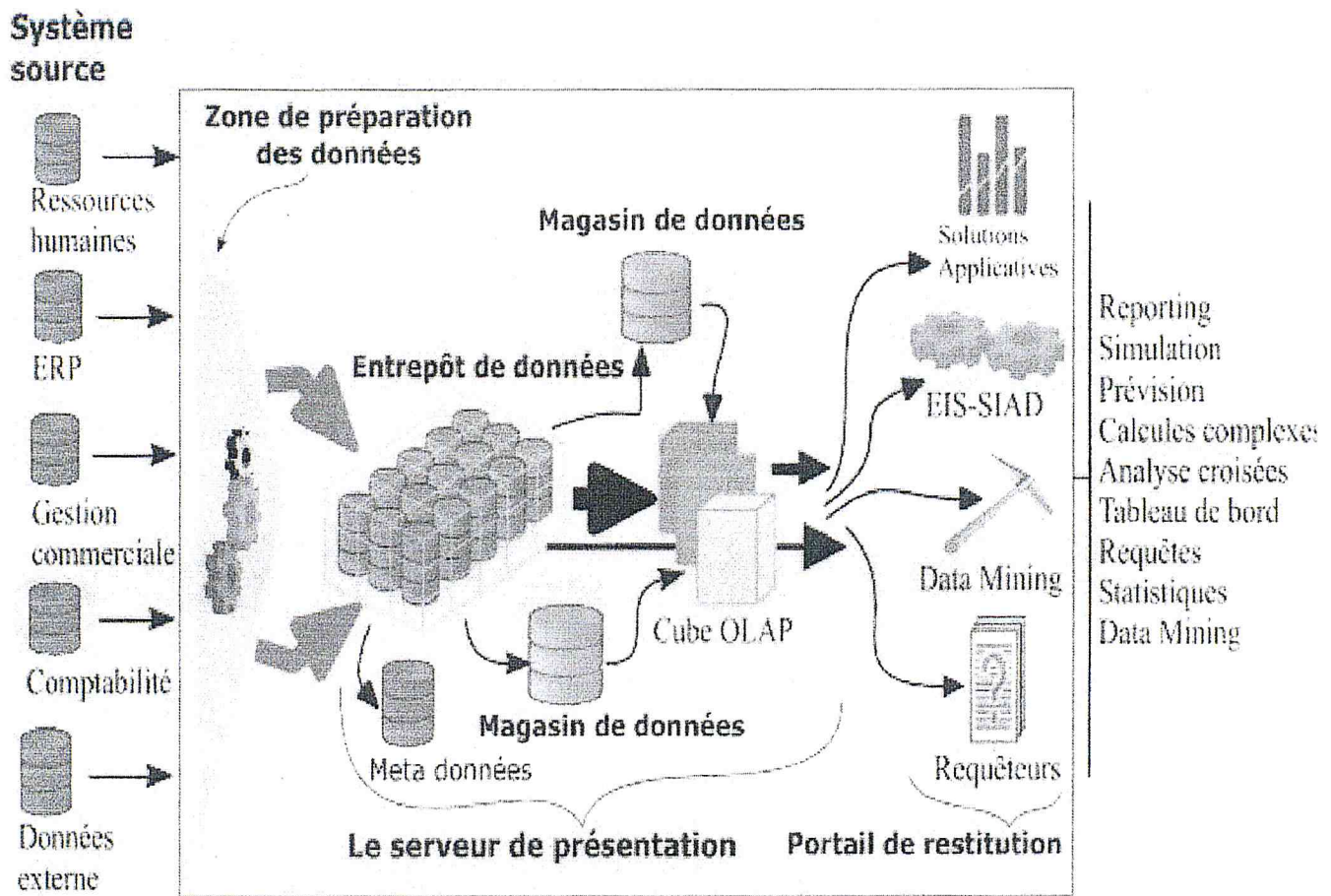
La mise en œuvre de l'architecture consiste à concevoir le modèle physique de notre entrepôt de données, ainsi que sa zone d'alimentation.

### **4.4. Le déploiement**

Le déploiement est le point de convergence de la technologie, des données et des applications utilisateurs. Une planification est indispensable pour gérer le déploiement qui comprend également la formation des utilisateurs, les processus de communication, le support utilisateur, la prise en compte des demandes d'évolution et de correction.

## **5. L'Architecture de l'entrepôt de données**

L'entrepôt de données peut être vu comme une grande base de données, regroupant une masse de données de diverses provenances que l'on appelle système source. La collecte des données à partir de ce système se fait à travers un processus appelé extraction, une fois les données collectées, elles doivent être stockées et doivent subir des traitements (transformations), dans une zone de préparation des données pour qu'elles soient prêtes à être peuplées, répliquées et récupérées vers les différents data marts, qui constituent le serveur de présentation de l'entrepôt de données. L'illustration des composants d'un entrepôt de données est présentée dans la figure (fig.6).



**Figure 6** : L'architecture détaillée de l'entrepôt de données [1]

### 5.1. Système source

C'est le système opérationnel d'enregistrement, dont la fonction consiste à capturer les transactions liées à l'activité de l'entreprise [Kimball, 2003]. Il s'agit souvent de ce que l'on appelle les applications de gestion.

### 5.2. Zone de préparation des données

Une fois le processus d'extraction accompli, les données doivent être accueillies dans un espace de stockage. Les SGBDR sont les plus utilisés pour cette tâche.

Comme déjà cité avant, les incohérences des données provoquées par le processus d'extraction de données, provenant de sources différentes, obligent la mise en marche d'un processus de transformation, traduit en plusieurs traitements. Mentionnant quelques-uns :

- Nettoyage des données.
- Transformation des données.
- Combinaison des données.
- Suppression des doublons.
- ...etc.

Ces traitements visent à lever toute sorte d'incohérence avant de procéder au chargement.

Les données stockées dans cette zone, sont de différentes sorte et éventuellement organisées de différentes manières.

L'organisation des données est un point fort pour une bonne performance de l'entrepôt de données, car elle ne permet de cibler que les données nécessaires à alimenter vers les Data Marts.

### **5.3. Serveur de présentation d'un entrepôt de données**

Cette zone contient tous les Data Marts constituant le serveur de présentation, ainsi que les serveurs OLAP. Chaque Data Mart est conçu à l'intérêt d'un seul sujet de l'entreprise, et en s'appuyant sur un modèle multidimensionnel. Ce dernier a fait son apparition pour mettre fin aux problèmes liés à la conception à la base entité/relation. Rappelons que, la modélisation relationnelle ne supporte pas l'analyse multiaxes. Dans ce qui suit, nous allons identifier le Data Mart, ainsi que les concepts de base de la modélisation multidimensionnelle, le niveau logique et le niveau physique.

#### **5.3.1. Le DataMart**

Le Data Mart est une base de données moins coûteuse que l'entrepôt de données, et plus légère puisque destinée à quelques utilisateurs d'un département. C'est une petite structure très ciblée, et pilotée par les besoins utilisateurs. Il a la même vocation que l'entrepôt de données (fournir une architecture décisionnelle), mais vise une problématique précise avec un nombre d'utilisateurs plus restreint. [Kimball, 2007]

Un entrepôt de données est souvent volumineux (plusieurs centaines de Go voire To) avec des performances inappropriées (temps de repense très long). Le Data Mail comporte moins de 50 GO, ce qui permet des performances acceptables.

Les Data Marts sont petits, non complexes et plus faciles à déployer que les entrepôts de données. Les Data Marts peuvent être un moyen de débiter un projet d'entrepôt de données. Les différents Data Marts indépendants peuvent être dynamiquement couplés pour se métamorphoser en entrepôt de données.

Les Data Marts ne se résument qu'à une seule information métier. Donc le Data Mart peut préparer la mise en place de l'entrepôt de données. Mais il faut penser grand, et adopter des technologies capables d'évoluer. [Kimball, 2003]

### **5.3.2. OLAP « ON LINE ANALYTICAL PROCESSING »**

Parmi les procédés existants du système décisionnel, il y a OLAP qui est une catégorie de logiciels axés sur l'exploration et l'analyse rapide des données, selon une approche multidimensionnelle à plusieurs niveaux d'agrégation.

L'outil OLAP repose sur la restructuration et le stockage des données, dans un format multidimensionnel issu de fichiers plats ou de bases relationnelles. Ce format multidimensionnel connu sous le nom d'hyper cube, organise les données le long de dimensions. Ainsi, les utilisateurs analysent les données suivant les axes propres à leur métier.

OLAP a pour principal but de fournir une assistance pour une analyse optimale des données sans se focaliser sur les moyens utilisés (abstraction), et d'apporter une visualisation multidimensionnelle des données (lacune des bases de données relationnelles). En plus, les types d'analyses permises par l'OLAP vont de la simple navigation (Roll-up, Drill-down, slice and dice...), aux analyses les plus complexes, telles que les suites temporelles et la modélisation complexe.

De nombreuses variétés d'outils OLAP sont aujourd'hui disponibles ils se répartissent dans des catégories en fonction de l'architecture employée pour stocker et traiter les données multidimensionnelles :

#### **5.3.2.1. OLAP Multidimensionnel (MOLAP)**

Les MOLAP sont des structures de données et des systèmes de gestion de bases de données multidimensionnelles (SGBDMD) spécialisés, qui organisent les données. Pour améliorer les performances des requêtes, ils agrègent et stockent d'une façon générale les données en fonction de leur usage prévisible.

### 5.3.2.2. OLAP Relationnel (ROLAP)

ROLAP est le type de technologie OLAP qui connaît actuellement le plus grand succès. Cette croissance répond aux demandes de la part des utilisateurs d'analyse de plus en plus grandes quantités de données, et à la constatation des utilisateurs qu'ils ne peuvent stocker dans des bases de données MOLAP, toutes les données dont ils ont besoin pour leurs analyses. Le ROLAP est compatible avec les produits de SGBDR, grâce à la présence d'une couche de métadonnées, ce qui évite la nécessité de créer une structure multidimensionnelle de données statiques. Ceci facilite la création de vues multidimensionnelles multiples à partir de la relation bidimensionnelle.

### 5.3.2.3. OLAP hybride (HOLAP)

Les outils d'OLAP hybride offrent des capacités d'analyse limitées, soit directement sur des produits de SGBDR, soit par l'entremise d'un serveur MOLAP intermédiaire. Les outils HOLAP livrent à une station de travail (ou à un serveur local) des données sélectionnées directement à partir du SGBD ou par l'intermédiaire d'un serveur MOLAP, sous la forme d'un cube de données, et la station de travail (ou le serveur local) les emmagasine, les analyse et les entretient localement.

## 5.4. Portail de restitution

Le portail de restitution regroupe tous les outils capables d'émettre des requêtes utilisateurs interrogeant l'entrepôt de données, et qui sont en mesure de produire une vue claire des rapports ou des graphismes résultats. Les utilisateurs de ces outils ont presque toujours une interface qui leur permet d'interagir avec l'entrepôt, un simple clic doit leur permettre d'obtenir le résultat attendu, en aucun cas ils doivent avoir un accès directe à l'entrepôt, seul l'administrateur des requêtes peut avoir ce privilège par souci de garder la cohérence des données.

### 5.4.1. Data Mining

Le terme datamining signifie littéralement l'exploitation des données. Comme dans toute exploitation, le plus recherché est de pouvoir extraire de la richesse. Ici, la richesse est la connaissance de l'entreprise. Fort du constat qui existe au sein des bases de données de chaque entreprise, une ressource de données cachées et surtout inexploitées. Le datamining permet de faire les apparaitre, et cela grâce à un certain nombre technique spécifiques.

Nous encadrons datamining l'ensemble des techniques qui permettent de transformer les données en connaissance.

Le datamining est un processus itératif qui met en œuvre un ensemble de techniques hétéroclites telles que des bases de données (entrepôt de donnée), de la statistique, de l'intelligence artificielle, de l'analyse des données, des interfaces de communication homme- machine.

Le résultat du datamining se présente sous différentes formats : texte plat, tableau, graphique,...

### 5.4.2. Le tableau de bord

Un tableau de bord n'est qu'une forme particulière d'état de rapports. On appelle tableau de bord, un rapport qui est:

- Synthétique : tout doit tenir sur une feuille A4, ou sur un écran d'ordinateur.
- Composite : présente plusieurs indicateurs, ce qui offre une représentation complète de l'activité de l'entreprise.
- Personnalisé : chaque acteur consulte un rapport qui concerne son activité.

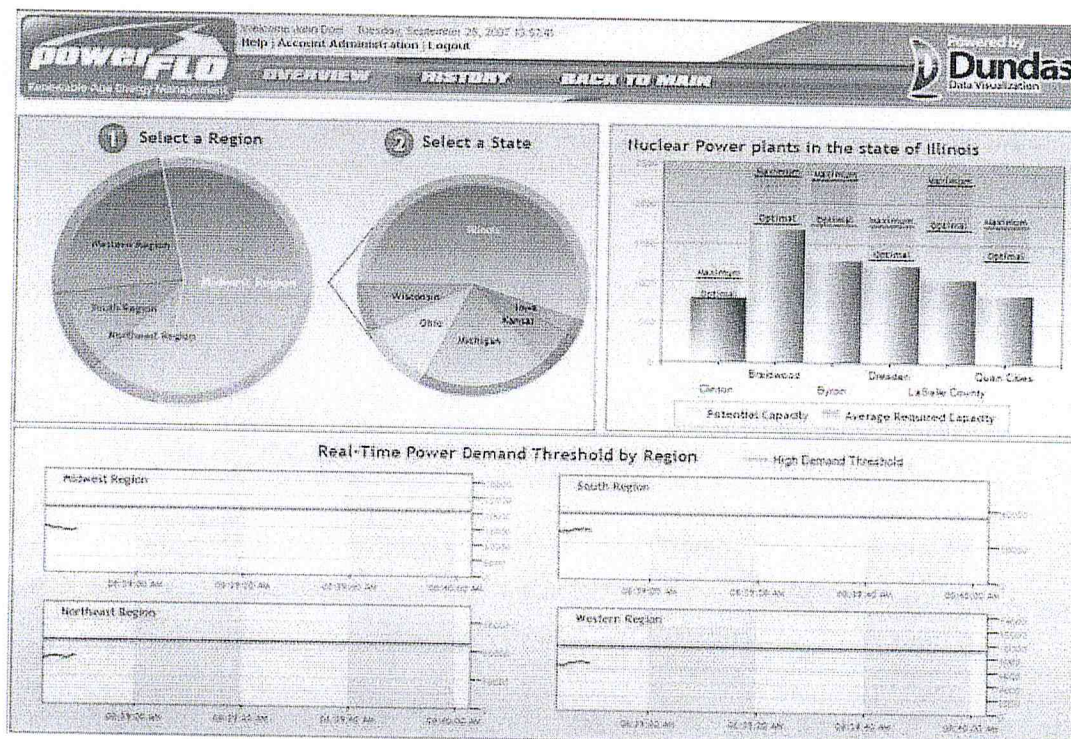


Figure 7 : Tableau de bord. [El Goll, 2008]



### 5.4.3. EIS (Executive Information System)

Comme déjà présenté dans le chapitre précédent, un EIS est un outil permettant d'organiser, d'analyser et de mettre en forme des indicateurs afin de constituer des tableaux de bord qui permettent différents niveaux de management d'accéder aux informations essentielles de l'organisation, de les analyser et de les présenter de façon élaborée. Ces outils sont dotés d'une interface graphique très conviviale et très esthétique. Ce type d'outils, facile à utiliser, ne permet de manipuler que des requêtes préalablement modélisées par le concepteur.

Destiné auparavant qu'à un nombre limité d'utilisateurs, les EIS conservent aujourd'hui tous les acteurs de l'entreprise, chacun à son niveau.

### 5.4.4. Le requêteur

Le requêteur va permettre à l'utilisateur de piocher lui-même dans les données de l'entreprise au libre accès, dans un langage propre de celui de son métier. Cette technologie est plus pratique pour répondre à des questions précises concernant des données primaires, non agrégées.

Le requêteur va offrir les mêmes fonctionnalités que la logique relationnelle:

Jointure de données, sélection, tri, projection. Ce qui est visé, est que les utilisateurs puissent poser des questions dans leurs propres termes et non dans le jargon informatique SQL (Standard Query Language)

De simples outils d'interrogation graphique, ils tendent désormais à s'orienter vers les métiers des utilisateurs pour permettre une utilisation plus intuitive. Les premières solutions nécessitaient la connaissance de la structure de la base de données et même parfois du langage SQL.

Aujourd'hui ces éléments sont transparents pour l'utilisateur final.

### 5.4.5. KPI (Key Performance Indicators)

Paramètre qui se veut le plus représentatif d'une activité de l'entreprise et qui permet d'évaluer la performance globale de cette dernière en fonction des objectifs à atteindre. Le but de l'informatique décisionnelle est de parvenir à mettre à jour en permanence ces indicateurs (par exemple les chiffres de vente de tel ou tel produit phare de l'entreprise) pour fournir aux managers de l'entreprise des tableaux de bord constamment à jour.

### 5.4.6. Data Webhouse

« Alors qu'un entrepôt de données classique est avant tout un outil d'aide à la décision pour la direction de l'entreprise, le data Webhouse a pour vocation d'offrir –via une interface web, bien entendu – de nouveaux services aux responsables de l'entreprise comptant à ses clients et partenaires : outils de pilotage pour le direction, Rapports d'analyse du comportement client pour les responsables marketing, relevé de compte des statistiques pour les partenaires commerciaux, service commercialisé pour le client final». [Kimball, 2000]

## 6. La modélisation multidimensionnelle

Le modèle entité-association est souvent utilisé pour la conception de la base de données. Ce modèle permet d'éliminer les redondances en introduisant un grand nombre de nouvelles entités. De ce fait, l'accès aux données devient plus compliqué et le schéma difficile à lire. L'utilisation de ce modèle pour la conception des entrepôts de données n'est donc pas appropriée. [Kimball, 1997]

La modélisation multidimensionnel souvent appelée modélisation OLAP se présente comme une alternative au modèle relationnel. Elle correspond mieux aux besoins du décideur tout en intégrant la modélisation par sujet. C'est une méthode de conception logique qui vise à présenter les données sous une forme standardisée intuitive, et qui permet des accès hautement performants. Elle aboutit à présenter les données non plus sous forme de tables mais de cube centré sur une activité. Un cube de dimension  $n$  ( $n > 3$ ) est aussi dit hyper cube.

Un cube est une construction multidimensionnelle formée de la conjonction de plusieurs dimensions. Chaque cellule est définie par une seule valeur de chaque dimension.

Cette modélisation a donné naissance aux concepts de fait et de dimension. [Kimball, 1996]

## 7. Vocabulaire de la modélisation multidimensionnelle

### 7.1. Définition de dimension

Lorsqu'on fait un schéma de BD pour un système d'information classique, on parle en termes de tables et de relations, une table étant une représentation d'une entité et une relation une technique pour lier ces entités. En Business Intelligence, on parle en termes de Dimension et de Faits. C'est une autre approche des données, on entend par dimensions les axes avec lesquels on veut faire l'analyse. Il peut y avoir une dimension client, une dimension produit, une dimension géographie (pour faire

des analyses par secteur géographique), etc.

Une dimension est tout ce qu'on utilisera pour faire nos analyses.

## 7.2. Définition de fait (mesure)

Le terme de fait est utilisé pour représenter une mesure économique. Par exemple, lors de la vente de produits sur un marché, on comptabilise les types de produits vendus, leur quantité et le montant de chaque vente au jour le jour et ce, pour chaque produit et pour chaque magasin. La mesure des quantités et des prix est réalisée à l'intersection de toutes les dimensions (produit, magasin, temps). Le nombre des dimensions détermine la finesse, la granularité de la table et indique la portée de l'indicateur. Un fait est tout ce qu'on voudra analyser.

## 7.3. Table de faits

Une table de faits est la table principale d'un modèle dimensionnel où les mesures de performance sont stockées, comme dans l'exemple de la figure 8.

Nous

nous efforçons de stocker les informations de mesure d'un processus d'entreprise dans un unique marché d'info. Les données de mesures étant très loin la portion la plus volumineuse d'un marché d'info, nous nous abstenons de les dupliquer à différents endroits de l'entreprise.

Table de faits des ventes journalières
Clé date (CE)
Clé Produit (CE)
Clé magasin (CE)
Quantité vendue
Montant des ventes

Figure 8 : Exemple de table de faits

Les faits les plus utiles sont les faits numériques, additifs, tels que les montants en euros.

L'additivité est cruciale parce que les applications d'entrepôts de données ne récupèrent presque jamais une seule ligne de table de faits. En fait, elle est en récupèrent des centaines, des milliers même des millions à la fois, et la chose la plus utile que l'on puisse faire avec un tel nombre de lignes est de les additionner. Dans la figure 8, quel que soit la tranche de données choisie par l'utilisateur, nous pouvons ajouter les quantités et les euros pour former un total valide.

Il existe des faits semi-additifs, ces faits ne peuvent être additionnés que pour certaines dimensions et les faits non-additifs ne peuvent pas du tout être additionnés. Dans le cas des faits non additifs,

nous en sommes obligés d'utiliser des comptages, ou des moyennes si nous souhaitons cumuler les lignes, sinon nous en sommes réduits à afficher les lignes de faits une par une, ce serait monotone dans le cas d'une table de faits d'un milliard de lignes.

Toutes les tables de fait ont deux clés étrangères ou plus, désignées par la notation **CE** sur la figure 8. Ces clés étrangères se connectent aux clés primaires des tables de dimension.

Par exemple, la clé produit de la table de fait correspond toujours à une clé de produit spécifique de la table de dimension produit. Lorsque toutes les clés de la table de fait renvoient effectivement aux clés primaires appropriées dans les tables de dimension correspondantes. Nous accédons à la table de faits par le biais des tables de dimension qui lui sont jointes.

La table de faits elle-même a généralement sa propre clé primaire faite d'un sous ensemble des clés étrangères. Cette clé est parfois appelée clé composite ou clé concaténée. Chaque table de fait d'un modèle dimensionnel a une clé composite et inversement, toute table qui a une clé composite est une table de faits. On peut dire aussi que dans un modèle dimensionnel, toute table qui exprime une relation de plusieurs-à-plusieurs doit être une table de faits. Toutes les autres tables sont des tables de dimension. [Kimball, 2003]

#### 7.4. Table de dimension

Les tables de dimensions sont les entités complémentaires à la conception de la table de faits. Elles contiennent, autant que possible, des attributs sous forme de descriptions textuelles permettant de qualifier ou d'expliquer l'activité. Des attributs de dimensions, nombreux, permettent de varier les possibilités d'analyse (par tranches ou en dés). Ces attributs rendent utilisables et intelligible les données de l'entrepôt de données. Ils établissent, en quelque sorte une interface homme/entrepôt de données. En général, les tables de dimensions tendent à être peu profondes mais elles sont larges (l'inverse de la table de faits), en d'autres termes elles ont peu de lignes(en général beaucoup moins qu'un million de lignes) mais beaucoup de colonnes.

Chaque dimension est définie par son unique clé primaire dénoté à la figure 8 par l'abréviation **CP**.

Table de dimension produit
Clé Produit (CE)
Description du produit
Numéro US
Description de la marque
Description de la catégorie
Description du rayon
Description du type d'emballage
Taille de l'emballage
Description matière grasses
Description type de régime
Poids
Unités de mesures de poids
Type de stockage
Type de durée étagère
largeur sur étagère
Hauteur sur étagère
Profondeur sur étagère
... et bien d'autre attributs.

Figure 9 : Exemple de table de dimension

Les attributs de dimensions ont la principale source des contraintes de requête, de regroupement et d'intitulés de colonne des états.

Les attributs de dimension jouent un rôle vital dans un entrepôt de données. Comme ils sont la source de pratiquement toutes les contraintes et de tous les intitulés intéressants dans les états, ils sont le moyen par lequel l'entrepôt de données est rendu utilisable et compréhensible. La puissance de l'entrepôt de données est directement proportionnelle à la qualité et à la profondeur des attributs de dimension. Les tables de dimension indiquent souvent des relations hiérarchiques dans une activité. [Kimball, 2003]

## 7.5. Relier les faits aux dimensions

Pour rendre les SGBD plus utiles pour les applications OLAP (On-line Analysis Process), les vendeurs ont ajouté de nouvelles fonctions. Pour fournir des temps d'accès rapide aux applications OLAP, les données sont organisées selon un schéma en étoile (star) ou en flocon (snowflake). [Moran, 2002]

### 7.5.1. Schéma en étoile

Un schéma en étoile consiste en une table de faits centrale et plusieurs tables de dimensions. Les mesures d'intérêt pour l'OLAP sont stockées dans la table des faits (ex. Ventes, stocke). Pour chaque dimension du modèle multidimensionnelle il existe une table (ex. région, produit, temps). Cette table stocke des informations relatives aux dimensions. La relation de faits contient les différentes mesures et une clé étrangère pour faire référence à chacune de leurs dimensions.

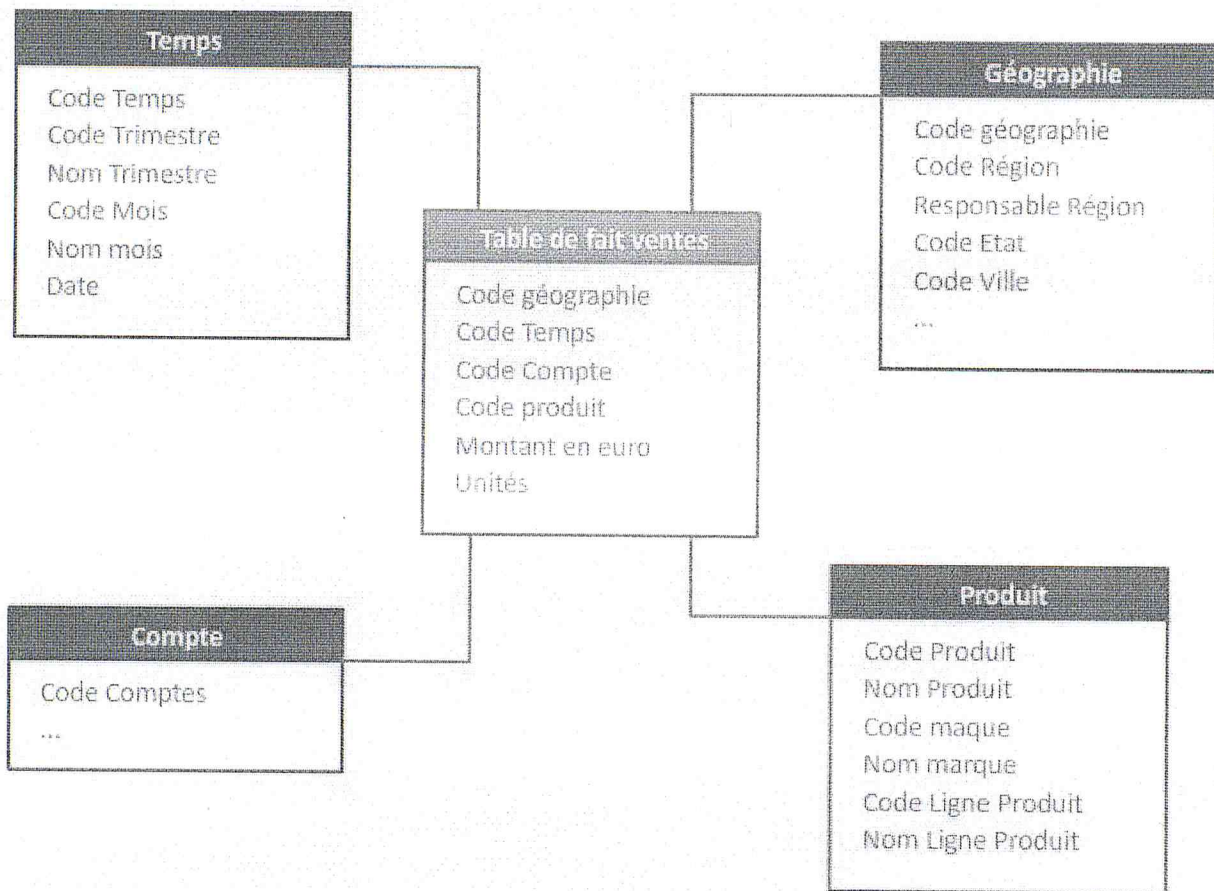


Figure 10 : Schéma en étoile. [Kimball, 2003]

### Avantages du schéma en étoile

Le schéma en étoile présente un certain nombre d'avantages qui sont :

- Facilité de navigation ;
- Performances élevées: nombre de jointures limité, gestion des données creuses ;
- Gestion des agrégats ;
- Fiabilité des résultats.

### Inconvénients du schéma en étoile

Ce schéma comporte un certain nombre d'inconvénients qui peuvent être résumés en :

- Toutes les dimensions ne concernent pas les mesures ;
- Redondances dans les dimensions ;
- Alimentation complexe.

### 7.5.2. Le schéma en flocon de neige

Il dérive du schéma précédent avec une relation centrale, et autour d'elle les différentes dimensions qui sont éclatées ou décomposées en sous hiérarchies. L'avantage du schéma en flocon de neige est de formaliser une hiérarchie au sein d'une dimension, ce qui peut faciliter l'analyse. Un autre avantage est représenté par la normalisation des dimensions, car nous réduisons leur taille.

Néanmoins dans [Kimball, 1996], l'auteur démontre que c'est une perte de temps de normaliser les relations des dimensions dans le but d'économiser l'espace disque. Par contre, cette normalisation rend plus complexe la lisibilité et la gestion dans ce type de schéma. En effet, ce type de schéma augmente le nombre de jointures à réaliser dans l'exécution d'une requête. [Trinidad, 2005]

Les hiérarchies pour le schéma en flocon de neige de l'exemple de la figure 11 sont :

Dimension **Temps**=**Jour** à **Mois** à **Année**

Dimension **Magasin**=**Commune** à **Département** à **Région** à **Pays**

La figure 11 montre le schéma en flocon de neige avec les dimensions **Temps** et **Magasin** éclatées en sous hiérarchies.

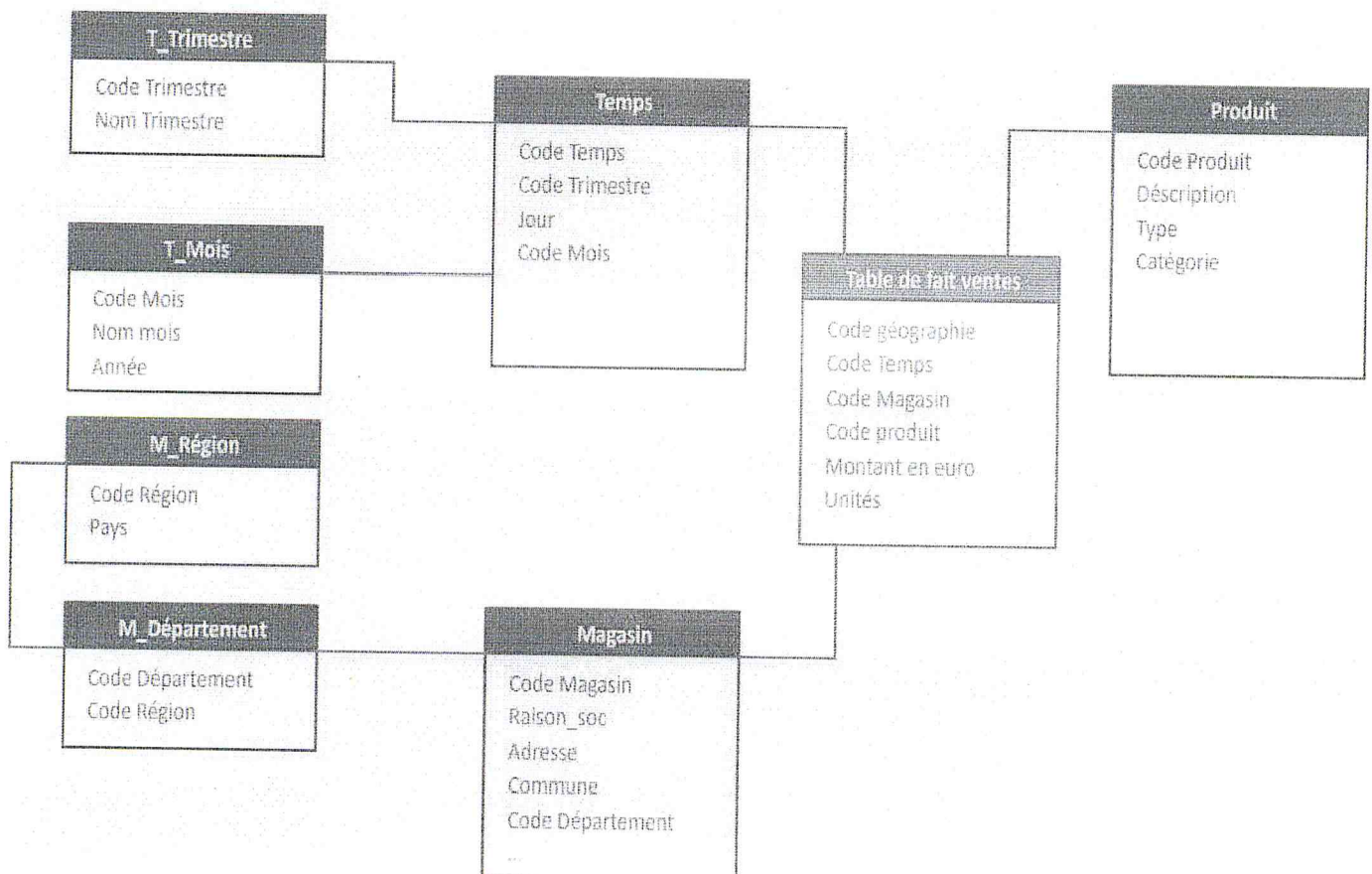


Figure11 : Exemple de modélisation en flocon neige. [Kimball, 2003]

Dans l'exemple ci-dessus, la dimension Temps a été éclatée en deux, Temps et T\_Mois. La deuxième dimension Magasin, a été décomposée en trois : Magasin, M\_Département et M\_Région.

#### **Avantages du schéma en flocon de neige:**

Ce modèle comporte un certain nombre d'avantages qui sont :

- Réduction du volume ;
- Gain de place de stockage ;
- Plus propre ;
- Permet des analyses par pallier (drill down) sur la dimension hiérarchisée.

#### **Inconvénients du schéma en flocon de neige**

De ce schéma découle un certain nombre d'inconvénients qui sont :

- Navigation difficile.
- Plus complexe pour l'utilisateur final.
- Nombreuses jointures, donc plus lent.

### **7.5.3. Le schéma en constellation**

Le schéma en constellation représente plusieurs relations de faits qui partagent des dimensions communes. Ces différentes relations de faits composent une famille qui partage les dimensions, mais où chaque relation de faits a ses propres dimensions. [Benitez, 2001]

La figure 12 montre le schéma en constellation qui est composé de deux relations de faits. La première s'appelle Ventes et enregistre les quantités de produits qui ont été vendus dans les différents magasins pendant un certain jour. La deuxième relation gère les différents produits achetés aux fournisseurs pendant un certain temps.

La relation de faits Ventes partage leurs dimensions Temps et Produits avec la table Achats. Néanmoins, la dimension Magasin appartient seulement à Ventes. Egalement, la dimension Fournisseur est liée seulement à la relation Achats. [Trinidad, 2005]



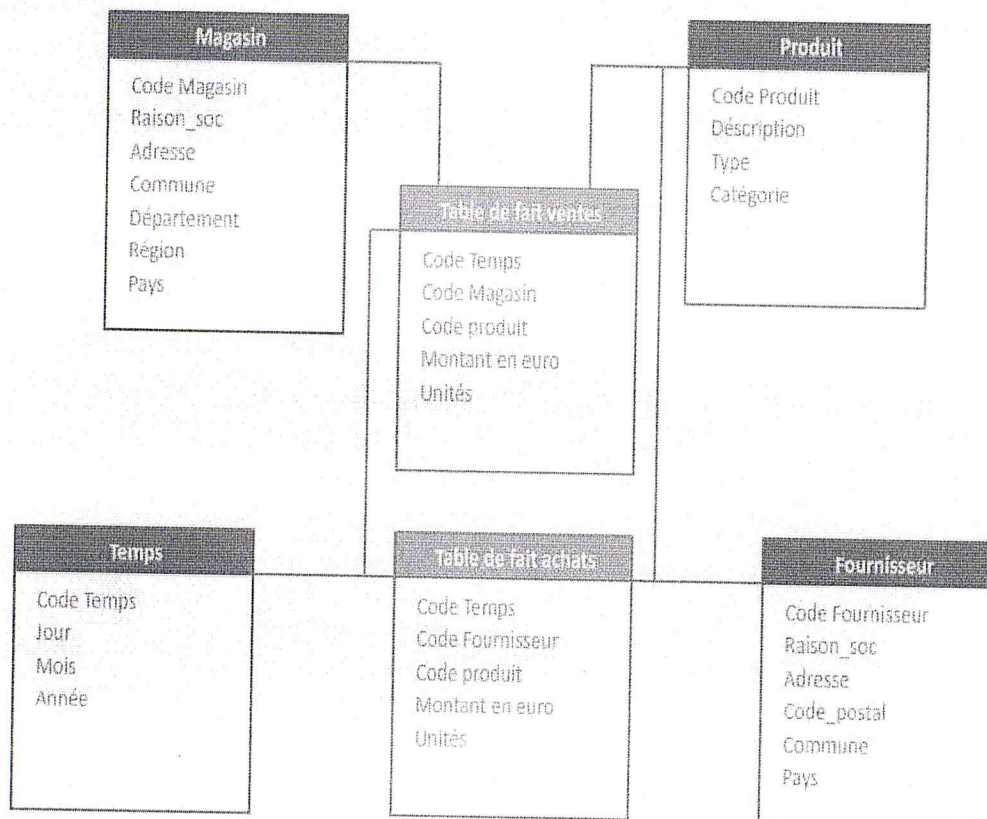


Figure12 : Exemple de modélisation en constellation. [Kimball, 2003]

## 8. Approches de développement d'un entrepôt de données

Il existe plusieurs approches pour mettre en place un DW. Par contre seulement trois approches sont communes. Il s'agit de l'approche "Top-Down" prônée par Inmon, l'approche "Bottom-up" de Kimball et de l'approche "Hybride" qui dérivent des deux premières approches.

### 8.1. Ascendante Bottom-up

Selon Kimball la première étape est la conception du modèle dimensionnel pour l'entrepôt des données, c'est-à-dire ayant une vue métier. Celui-ci placera les Data Marts au centre de l'architecture. Le reste sera composé d'un staging area temporaire. Dans cette approche, on dit que les Data Marts sont centraux car ils peuvent contenir à la fois des données atomiques et agrégées, et qu'ils offrent la possibilité de fournir une vue entreprise et une vue métier. Il est à noter que l'implantation des Data Marts se fait d'une façon incrémentale et intégrée. Pour finir, les utilisateurs ne peuvent effectuer des requêtes sur le staging area.

## 8.2. Descendante Top-down

Selon B.Inmon, c'est l'entrepôt qui est au centre de l'architecture. La première étape est donc de concevoir un modèle de l'entrepôt mais au niveau entreprise cette fois-ci et non plus métier. Cette architecture sera de type multi-tiers composée du staging area (zone temporaire de stockage de données sources et des données issues de l'ETL), de l'entrepôt de données et des Data Marts.

Le staging area sera permanent et les Data Marts seront interdépendants. C'est à dire qu'ils communiqueront entre eux.

La principale caractéristique de cette approche réside dans le fait que c'est l'entrepôt de données qui contient les données atomiques. Il utilise un modèle normalisé de toute l'entreprise. Il est donc orienté entreprise, à l'opposé des Data Marts qui eux orientés métier. Ces derniers utilisent des modèles dimensionnels orientés métier. Les données y sont stockées sous formes agrégées. Pour finir, les utilisateurs peuvent effectuer des requêtes sur l'entrepôt de données et sur les Data Marts.

### Conclusion

Dans ce chapitre nous avons abordé les différentes notions de base sur l'entrepôt de données, comme étant un outil de base dans le pilotage des entreprises. Nous avons décidé d'utiliser la technologie d'entrepôt de données, pour apporter une solution à la problématique déjà expliquée de l'entreprise ALGEX.

## **Partie II**

Conception et implémentation

# Chapitre I

Etude de l'existant

## Introduction

Dans ce chapitre, nous allons présenter l'entreprise ALGEX, qui est notre organisme d'accueil, ainsi que les différentes directions qu'elle comporte. Nous allons à la fin analyser son fonctionnement.

### 1. L'historique de l'organisme d'accueil -ALGEX-

Depuis quelque années, l'Algérie accord un très grand intérêt pour la promotion des exportations, sachant que l'économie algérienne est une économie dépendante c'est-à-dire. Que son revenu dépend de l'exportation des hydrocarbures qu'est une source non durable.

D'autre part, l'Algérie à adopte l'économie de marche et le libre échange ceci a rendu indispensable la création des instruments de promotion des exportations hors hydrocarbures tels que ALGEX, SAFEX, CAGEX.....etc. comme institution d'une importance particulière pour la promotion des exportations hors hydrocarbures.

### 2. Organisme d'accueil

L'agence nationale de promotion du commerce extérieur (ALGEX) à ces racines dans un établissement appelé « l'office algérien de promotion du commerce extérieur PROMEX », créé le 01/01/1996, par décret exécutif n°96-327.

Au début, avant l'ère du colonialisme l'Algérie était un pays à vocation agricole, mais après l'occupation française, elle s'est transformée en un pays d'extraction se basant sur des hydrocarbures. Avec la chute des prix des hydrocarbures en 1986, qui a affecté l'économie algérienne, l'état a décidé de prévoir d'autres moyens pour diversifier ces exportations et stabilisées le commerce extérieur.

Donc il fallait promouvoir des exportations hors hydrocarbure, c'est dans ce cadre que l'office a vu le jour en 1996.

L'Agence Nationale de Promotion du Commerce Extérieur, créée par décret exécutif n°04-174 du 12 Juin 2004, est un établissement public à caractère administratif sous la tutelle du Ministère du Commerce. Elle est chargée de la mise en œuvre de la politique nationale, en matière de promotion du commerce extérieur qui vise à encourager et à diffuser l'exportation des produits nationaux.

C'est un établissement créé pour cause de déséquilibre structurel et la faiblesse chronique des exportations hors hydrocarbures.

Parmi les missions et objectifs de l'office on distingue :

- La participation au développement du commerce extérieur.
- L'Observation et l'analyse des situations structurelles et conjoncturelles des marchés mondiaux.

## 2.1 Missions et objectifs de l'entreprise

Parmi les missions d'ALGEX on peut citer:

- La gestion des instruments de promotion des exportations hors hydrocarbures au bénéfice des entreprises exportatrices.
- L'analyse des marchés mondiaux et la réalisation d'études prospectives globales et sectorielles sur les marchés extérieurs.
- Concevoir un système d'information économique et commerciale articulé autour des chambres de commerce et industrie, les chambre d'agriculture et de la concurrence.
- Devenir un lieu recommandé pour les opérateurs import-export.

## 2.2 Organisation de l'entreprise

ALGEX Algérie repose sur sept (7) directions et vingt-quatre (24) sous-directions, la figure 1 montre la hiérarchie :

1- La Direction de l'Analyse des Marchés :

Est organisée en trois (03) sous directions ;

- Sous direction de l'analyse des marches de l'Afrique et pays Arabe.
- Sous direction de l'analyse des marches de l'Europe et Amérique du nord.
- Sous direction de l'analyse des marches de l'Amérique latine et de l'Asie.

2- La Direction de l'Analyse des Produits :

Est organisée en trois (03) sous directions ;



- Sous direction de l'analyse des produits agricoles.
- Sous direction de l'analyse des produits industriels.
- Sous direction de l'analyse des produits de l'artisanat du tourisme et des services.

3- La Direction des Services Spécialise :

Est organisée en trois (03) sous directions ;

- Sous direction de la logistique a l'export.
- Sous direction des politiques de financement et d'assurance des.
- Sous direction de la qualité emballage et la certification.

4- La Direction de la Formation, de la Coopération et de la Documentation :

Est organisée en quatre (04) sous directions ;

- Sous direction de la coopération.
- Sous direction de la formation.
- Sous direction de la documentation et de la recherche documentaire.
- Sous direction des publications et de la traduction.

5-La Direction de l'Information et de la Communication :

Est organisée en trois (03) sous directions ;

- Sous direction de l'information
- Sous direction d'informatique
- Sous direction de la communication et des relations publiques

6- La Direction de l'Administration et des Moyens :

Est organisée en trois (03) sous directions ;

- Sous direction de personnel.

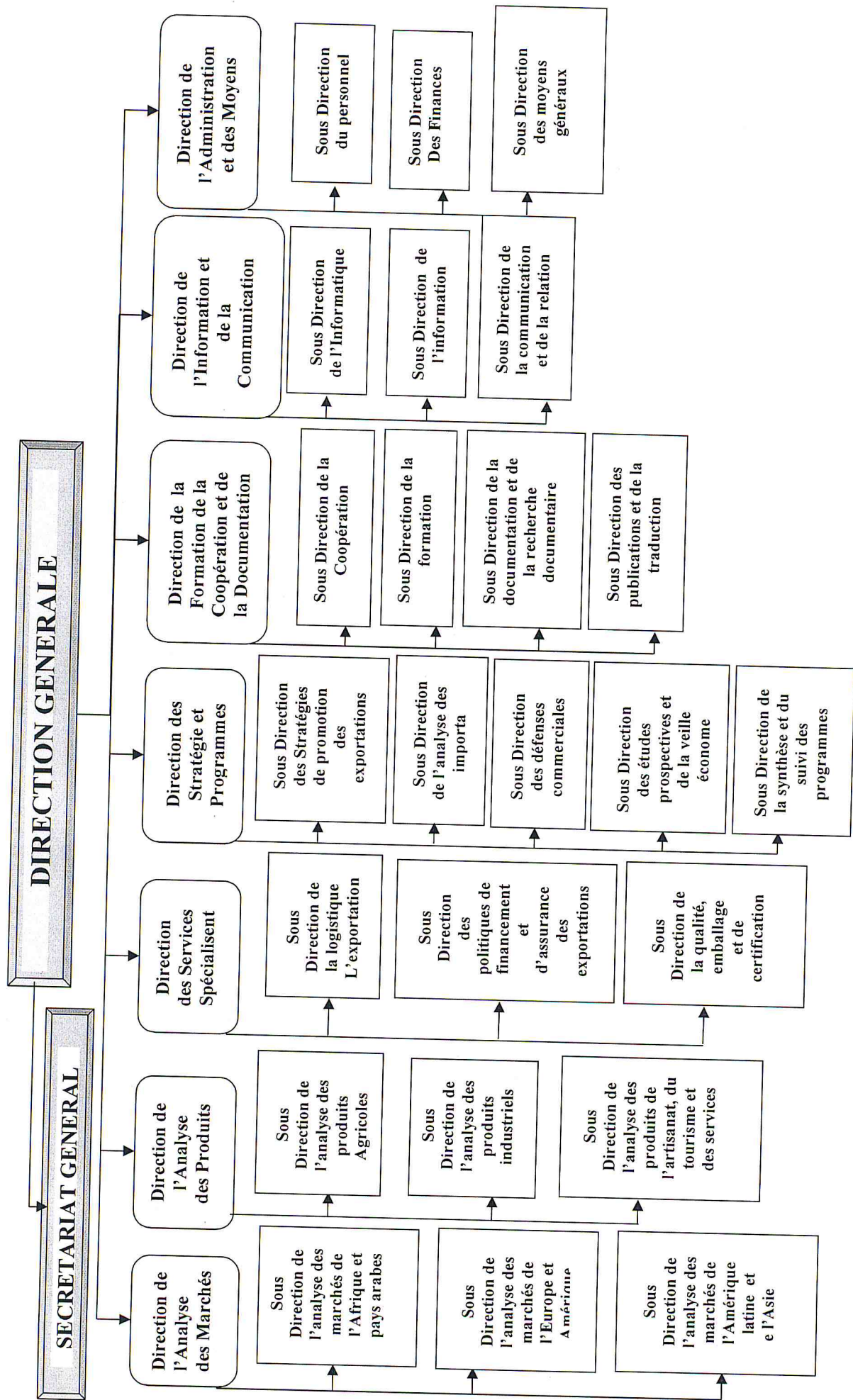


Figure 12 : Organigramme d'ALGEX



Ainsi notre projet se limite à la sous-direction de Communication et de Relations Publique.

## **Présentation et perspectives de la DIC**

Développer des projets de coopération bilatérale et multilatérale en matière d'information et de communication avec les organismes nationaux et étrangers similaires. Ainsi mettre en place des outils modernes d'analyse des banques de données de gestion et d'aide à la décision.

Les principales missions et responsabilités de la DIC

- Initier, proposer, et participer à la mise en place du système d'information relatif au commerce extérieur et d'en fixer les objectifs de l'organisation.
- Identifier, élaborer, et proposer les règles et procédures régissant le réseau national d'information relatif au commerce extérieur.
- Traiter l'information sur le commerce extérieur et assurer la diffusion.

### **3. Système existant**

Dans la suite de ce chapitre nous présenterons la situation actuelle issue d'une étude approfondie afin de réaliser une évaluation objective, en faisant ressortir des orientations pour la solution à concevoir d'une part, et d'autre part, identifier les anomalies tout en proposant des suggestions qui peuvent pallier aux insuffisances rencontrées.

Chaque mois, ALGEX récupère les données du commerce extérieur sur support amovible (flash disk) auprès du CNIS (Centre nationale d'informatique et de statistique Douane Algérienne) sous format DBF. Ces données sont consolidées manuellement dans une base de données ACCESS, par année. Cette dernière est analysée suivant plusieurs dimensions sur quatre indicateurs du commerce extérieur qui sont le tonnage, la quantité complémentaire, la valeur en dinars et la valeur en dollars. Pour chaque rapport ou tableau d'analyse, une requête est développée, exécutée et chargée dans une table agrégat dans la même base de consolidation ACCESS. Les données de cette table sont exportées vers un classeur Excel afin d'affiner le résultat souhaité par des formules et des Graphes. Une fois les rapports élaborés, ils sont diffusés à tous les acteurs du commerce extérieur. Le processus de circulation des données ainsi que la réalité de fonctionnalité du système est schématisé dans la figure 14

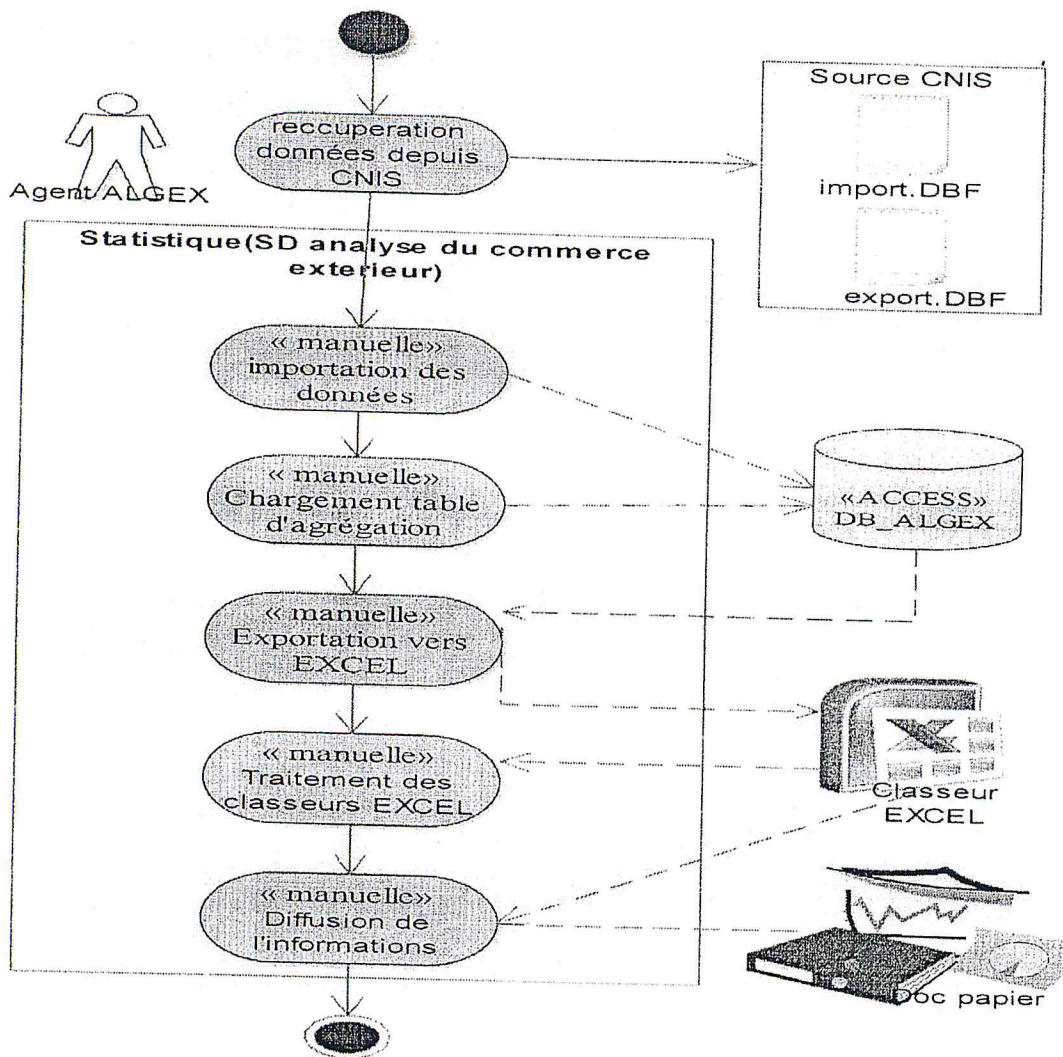


Figure 14 : Diagramme d'activité modélisant le traitement et l'édition de l'information

A partir de ce diagramme (Figure 14) on peut d'ores et déjà avoir une idée sur le nombre d'intervenants dans cette procédure qui reste une tâche très fastidieuse, surtout lors de la diffusion de l'information car dans cette étape il y'a les rapports sur les tendances ainsi que le répertoire des exportateurs qui est réalisé à partir des données traitées des statisticiens.

### 3.1 Analyse des données d'ALGEX

Les données du système d'information opérationnel actuel sont réparties sur des tables au sein d'un système de gestion de base de données « Microsoft Office Access » ainsi dans des fichiers dBase(.dbf) le nom qu'elle porte est « BaseALGEX », dans ce qui suit nous allons illustrer l'ensemble de la base « BaseALGEX » par des points ou on identifiera chaque table par une petite description en les rendant plus significatifs :

- **BALGUCUM** : cette table contient la quantité et valeur exportée ainsi qu'importée selon le groupe d'utilisation.
- **BALPAYCUM** : cette table contient la quantité et valeur exportée et importée selon le pays destination.
- **EXPPAYPROD** : cette table contient le produit exporté avec la quantité et valeur selon le pays destination.
- **BALPAYGUCUM** : cette table contient l'ensemble de données des deux tables « **BALGUCUM** » et « **BALPAYCUM** » car elle contient la valeur, quantité exporté et importé selon le pays et le groupe d'utilisation.
- **IMPCUM** : cette table présente l'importation d'un produit qui est présenté aussi par sa position tarifaire ainsi que la quantité en plusieurs unités et la valeur en Dinars Algérien et Dollars aussi le libellé du pays destination à une date données selon le groupe d'utilisation.
- **EXPCUM** : cette table contient les mêmes attributs que la table « **IMPCUM** » pour 1'exportation.
- **TZG27** : cette table contient le code du pays avec son nom ainsi que la zone géographique et le libellé de cette dernière.
- **TABPAYS** : cette table contient les codes et les libellés des pays.
- **TGU** : cette table contient le code de groupe d'utilisation ainsi que le libellé de chaque.
- **CHAPITRE** : cette table contient le code et le libellé du chapitre qui correspond à la branche d'activité.

La figure 15, relate le schéma de la base de données, qui va nous servir de source pour alimenter notre système décisionnel, les détails de nos deux sources de données et les tables qui les composent se situe dans Critique du système existant.

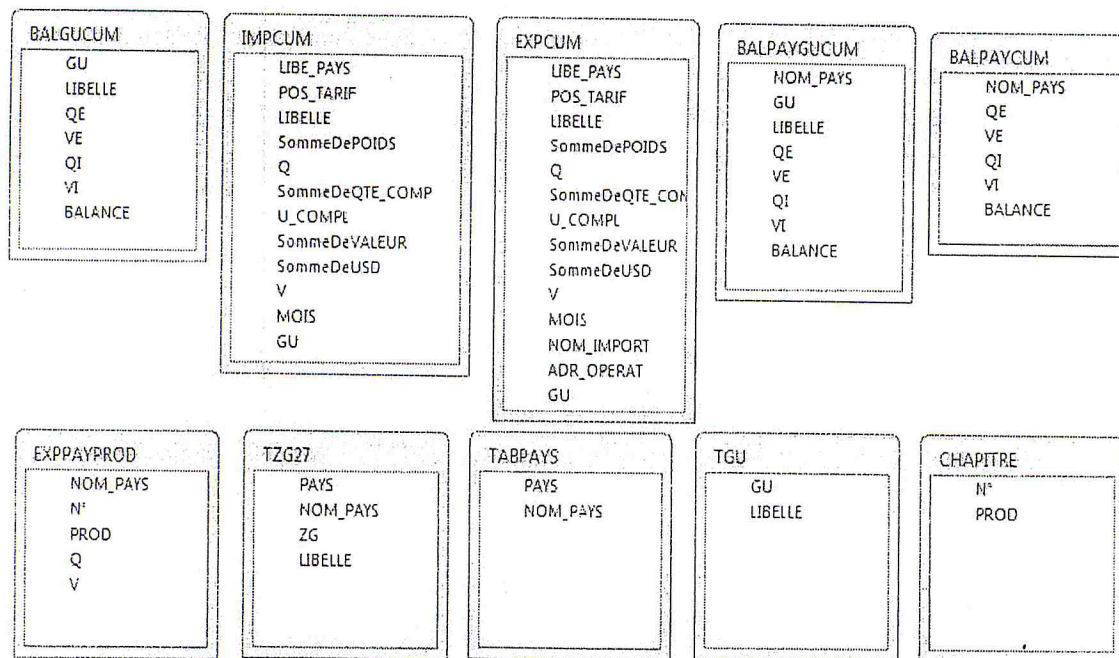


Figure 15: Schéma relationnel de la base de données « BaseALGEX »

#### 4. Analyse du système existant

L'analyse du système existant, nous a permis d'enregistrer un nombre de remarques, tels que :

- Lenteur du traitement et de transfert des données d'un service à l'autre pour cause d'inexistences d'informations et de cohésion dans le traitement des données ; manque de coordination entre direction.
- Les données existantes sont organisées dans une base de données Access puis basculé vers classeurs Excel, rendant la recherche d'une information très difficile et engendrant un grand risque d'erreur et de perte de données.
- L'absence de l'historique des mouvements même sur une période de quelque mois.
- Difficulté d'effectuer la consolidation des données récoltées régulièrement, pour avoir

une vue globale de l'entreprise à l'échelle nationale.

- Absence d'état de synthèse et de statistiques sur l'évolution des exportations que ce soit les produits exportés régulièrement ou non.
- Remplissage manuelle et itératif d'informations.
- Incohérences de données qui les rendent inutilisables et inexploitable due aux modifications d'informations causées des changements au niveau des entreprises.

### **Conclusion**

Nous avons présenté dans ce chapitre l'entreprise ALGEX, à travers son organisation, son fonctionnement dans le cadre de la promotion des exportations hors hydrocarbures. Nous avons terminé ce chapitre par une analyse du fonctionnement actuel de l'entreprise. Nous avons abouti à un certains nombres de remarques qu'ils vont nous aider à concevoir notre outil d'aide à la décision.

## **Chapitre II**

Conception

Dans notre étude nous avons adaptés l'approche de cycle de vie de Raplph Kimboll selon nos besoins La figure suivante (fig.15) illustre l'adaptation de l'approche.

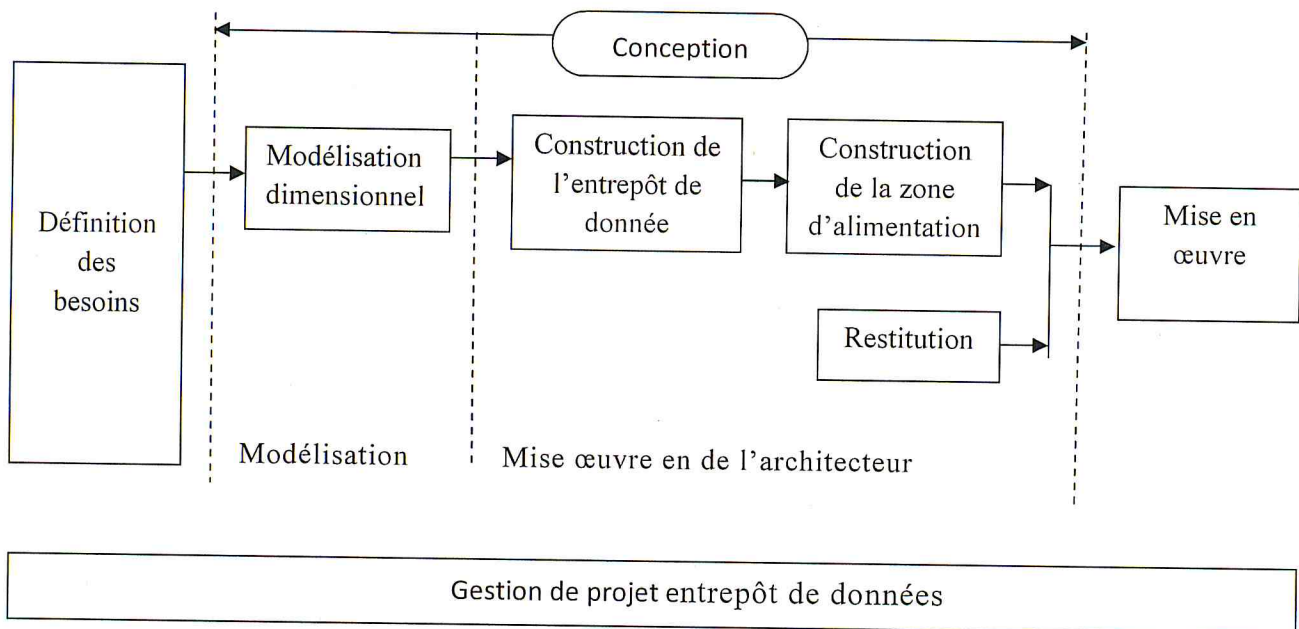


Figure 16 : Adaptation de l'approche de cycle de vie dimensionnel

## 1. Définition des besoins

Nous avons défini dans le premier chapitre les besoins de l'entreprise en matière de prise de décision et les raisons pour lesquelles les décideurs ont opté pour la réalisation d'un entrepôt de donnée. Après être familiarisé avec l'environnement opérationnel d'ALGEX, nous avons senti le besoin d'élaborer des questionnaires d'entretiens destinés aux décideurs de l'entreprise en vue de lever toutes ambigüités sur les objectifs à atteindre.

## Extrait de l'entretien

*Dans le cadre de notre projet de fin d'étude nous voulons concevoir et implémenter un entrepôt de donnée au profit du personnel de votre entreprise qui a les prérogatives de prendre des décisions en ce qui concerne la gestion d'annuaires des entreprises exportatrices. Après avoir consulté votre système opérationnel nous avons senti la difficulté d'avoir des informations historiées (datées).*

*L'objectif de cet entretien est de lever le voile sur certaines questions que nous allons vous poser durant cette heure précieuse que vous nous avez accordé.*

*Questions :*

- ✓ *Quels sont les objectifs de votre secteur d'activités ? qu'essayez-vous accomplir ?*
- ✓ *Quels sont les informations qui vous sont nécessaires pour prendre une décision concernant le suivi des entreprises ?*
- ✓ *Sur combien d'années vous avez gardé l'historique des exportations ?*
- ✓ *Quels sont les axes à prendre en compte lors d'une analyse ?*
- ✓ *Comment voulez-vous voir les informations obtenu à partir d'une requête d'analyse ?*

Exemple de questionnaire d'entretien avec un responsable.

Les informations obtenues de plusieurs entretiens nous ont permis de fixer les objectifs final en matière de besoins analytiques, besoins en information et les critères de réussite qu'on va présenter dans le compte rendu et le diagramme de cas d'utilisation.



Après de multiples entretiens programmés avec les employés de l'entreprise. Nous avons pu obtenir à plusieurs conclusions présentées point par point dans ce compte rendu

- **Attentes et objectifs :**
    - **Axes d'analyses :**
      - ✓ le temps, l'entreprise et leurs exportations.
    - **L'historique des exportations :**
      - ✓ L'historique est sauvegardé de façon anarchique.
      - ✓ L'historique est sauvegardé depuis l'année 2004, qui doit être récupéré de la du Commerce Extérieur afin de traiter ces données.
    - **Les vues des données issues du système décisionnel:**
      - ✓ Les rapports et les graphes seront les vues de multiples axes d'analyses.
    - **Critères de réussites :**
      - ✓ Un accès à l'historique pour au moins les 4 dernières années afin de pouvoir tendances.
- Solution proposée :**
- ✓ La solution propose doit répondre aux exigences des utilisateurs finals et doit aussi être exploiter pour une éventuelle synthétisation de toutes les données inclus dans l'analyse des exportations.

#### Compte rendu des entretiens avec les responsables

Ce compte rendu nous a permis de mettre deux principaux jalons pour notre conception. Le premier est de fournir une application, capable de satisfaire les utilisateurs finaux, à un niveau gestion des entreprises. Le deuxième est d'éviter une solution capable de satisfaire les utilisateurs finals, mais qui ne peut pas être exploitée dans son intégralité dans un entrepôt de données pour la gestion des entreprises exportatrices en Algérie.

Les besoins dans l'entrepôt de données seront présentés dans le cas d'utilisation illustré dans la figure 17.

• Diagramme de cas d'utilisation

Le diagramme de cas d'utilisation représente la structure des grandes fonctionnalités nécessaires aux utilisateurs du système. C'est le premier diagramme du modèle UML, celui où s'assure la relation entre l'utilisateur et les objets que le système met en œuvre. Il capture le comportement d'un système, d'un sous système, d'une classe ou d'un composant tel qu'un utilisateur extérieur le voit. Il scinde la fonctionnalité du système en unités cohérentes, les cas d'utilisation, ayant un sens pour les acteurs. Les cas d'utilisation permettent d'exprimer le besoin des utilisateurs d'un système, ils sont donc une vision orientée utilisateur de ce besoin.

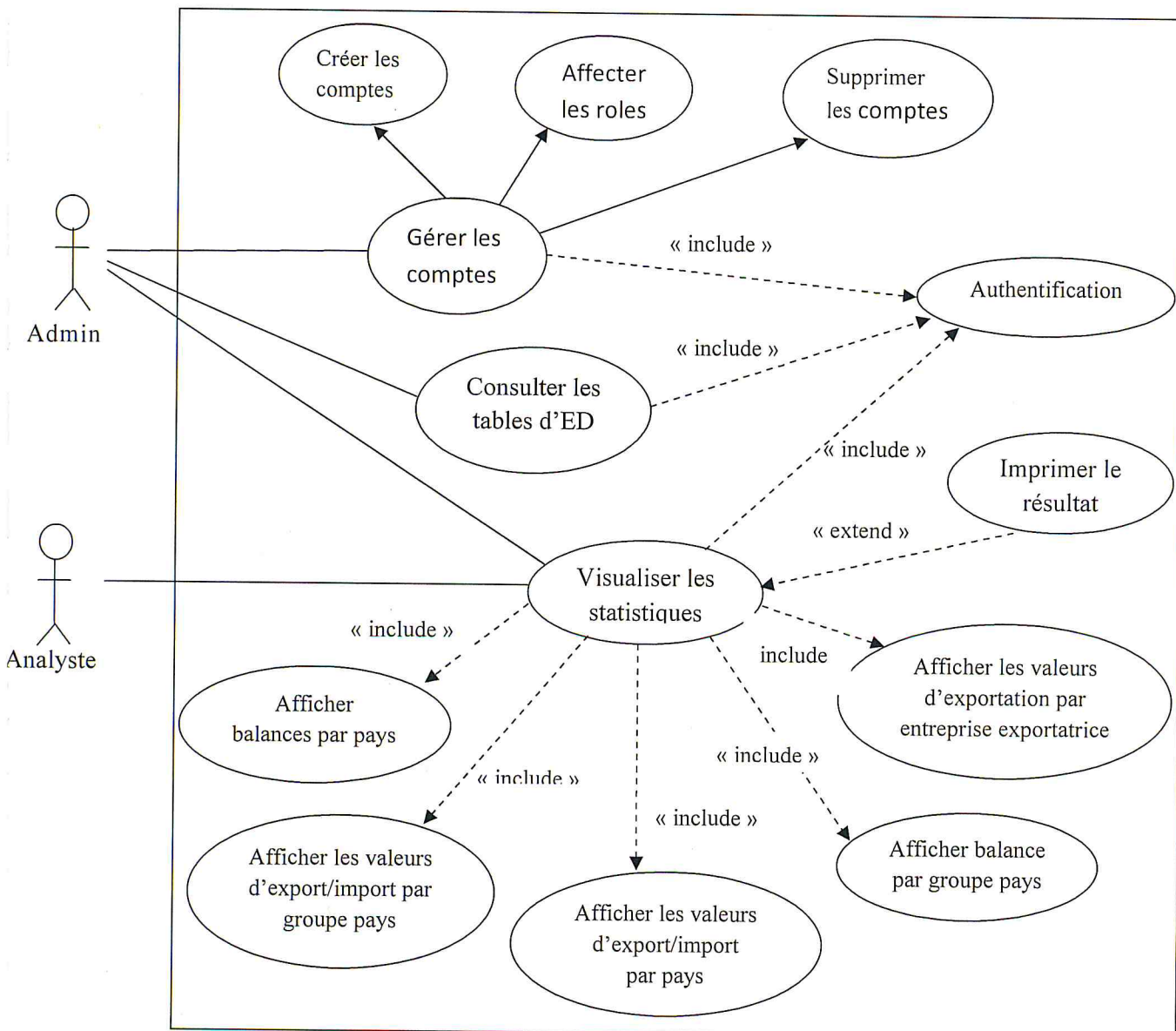


Figure 17: Diagramme cas d'utilisation

## 2. Conception de l'entrepôt de données

Une fois les besoins d'utilisations connus nous pouvons commencer à concevoir les volets de notre entrepôt de données. Pour cela, nous avons eu recours à la modélisation dimensionnelle qui est souvent associée aux entrepôts de données compte tenu de ses avantages.

Cependant, la modélisation multidimensionnelle d'un entrepôt de données va être faite en suivant l'approche Ascendante selon le cycle de vie dimensionnel.

### 2.1. Modélisation dimensionnelle des données

La modélisation multidimensionnelle consiste à considérer un sujet analyse comme un point dans un espace A plusieurs dimensions. Les données sont organisées de manière A mettre en évidence le sujet analyse et les différentes perspectives de l'analyse.

La conception logique d'un entrepôt de données passe par trois premières étapes :

- Choix du processus d'activité a modélisé.
- Choix des dimensions applicables a chaque table de faits.
- Choix des mesures que contiendra chaque enregistrement de la table de faits.

#### Modélisation dimensionnelle de l'activité :

##### **\*Balance par pays par période\***

##### **Le processus d'activité :**

Sur la base de notre étude, et grâce aux entretiens qui nous ont aides à mieux connaitre l'activité et les ressources de données disponibles. Le premier processus important choisi est **balance par pays et par période**.

## Les dimensions

D'après le processus d'activité de la table de faits, nous avons déterminé une série de dimensions principales suivantes : période, pays

Les tables de dimensions contiennent les descriptions textuelles de tous les attributs des dimensions.

### Dimension \*période \*

La seule dimension qui figure systématiquement dans tous les entrepôts de données, car en pratique tout entrepôt est une série temporelle.

### Dimension \*pays\*

L'Algérie fait des affaires avec de nombreux pays dans le monde, Cette dimension représente les pays que commerce avec l'Algérie

## Les faits mesures

Le fait modélise le sujet de l'analyse. Un fait est forme de mesures correspondant aux informations de l'activité analysée.

Les mesures d'un fait sont numériques et généralement valorisées de manière continue

Pour cette activité, les faits que nous avons enregistrés sont :

- Balance DA
- Balance \$

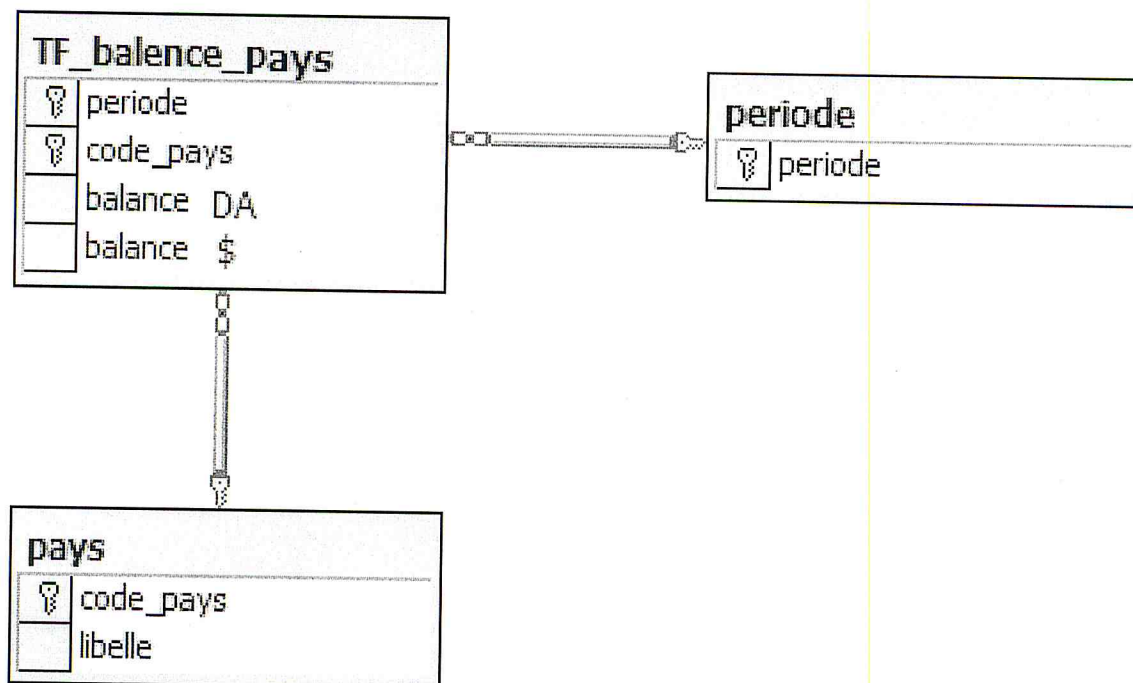


Figure 18 : Table de fait balance par pays

### Modélisation dimensionnelle de l'activité :

**\*Balance par group pays par période\***

#### Le processus d'activité :

Ce processus est important, parce que permet de calculer les balances par période entre l'Algérie et tous les groupes Pays exportateurs et importateurs

#### Les dimensions

D'après les Le processus d'activité de la table de faits, nous avons déterminé une série de dimensions principales suivantes : **période**, **group\_pays**

### Dimension \*période \*

La seule dimension qui figure systématiquement dans tous les entrepôts de données, car en pratique tout entrepôt est une série temporelle "temps".

### Dimension \*groupe\_pays\*

Cette dimension est nécessaire pour cette table de fait, qui représente les groupes pays que commerce avec l'Algérie

### Les faits mesures :

Pour cette activité, les faits que nous avons enregistrés sont :

- Balance\_G\_p DA
- Balance\_G\_p \$

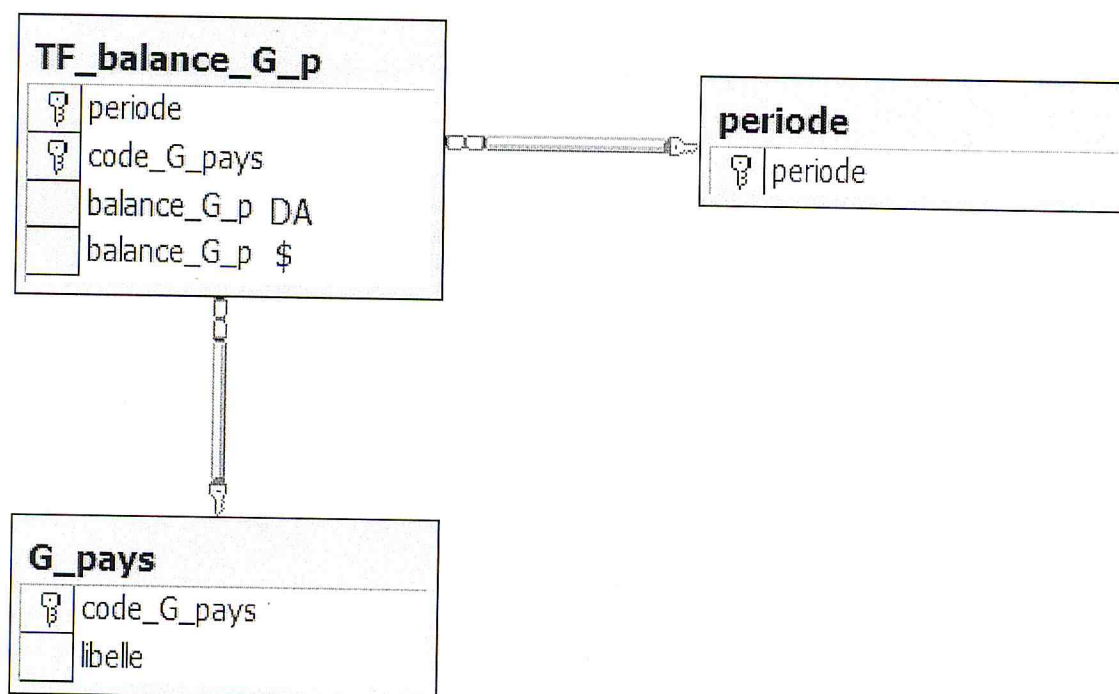


Figure 19: Balance group pays par période

## Modélisation dimensionnelle de l'activité :

### **\*Taux utilisation finale de produit par pays et par période \***

#### **Le processus d'activité :**

Ce processus est important, parce que permet de calculer les taux d'importation et les taux d'exportation par utilisation finale de produit et par période entre l'Algérie et tous les pays exportateur et importateur

#### **Les dimensions**

D'après les Le processus d'activité de la table de faits, nous avons déterminé une série de dimensions principales suivantes : **période, pays, utile finale**

#### **Dimension \*période \***

La seule dimension qui figure systématiquement dans tous les entrepôts de données, car en pratique tout entrepôt est une série temporelle "temps".

#### **Dimension \* pays\***

Cette dimension Est nécessaire pour cette table de faits, il représente les pays qui commercent avec l'Algérie

#### **Dimension \* utile finale\***

Cette dimension Est nécessaire pour cette table de faits, il représente l'utilisation finale de chaque produit

#### **Les faits mesures :**

Pour cette activité, les faits que nous avons enregistrés sont :

- Taux\_imp
- Taux\_exp

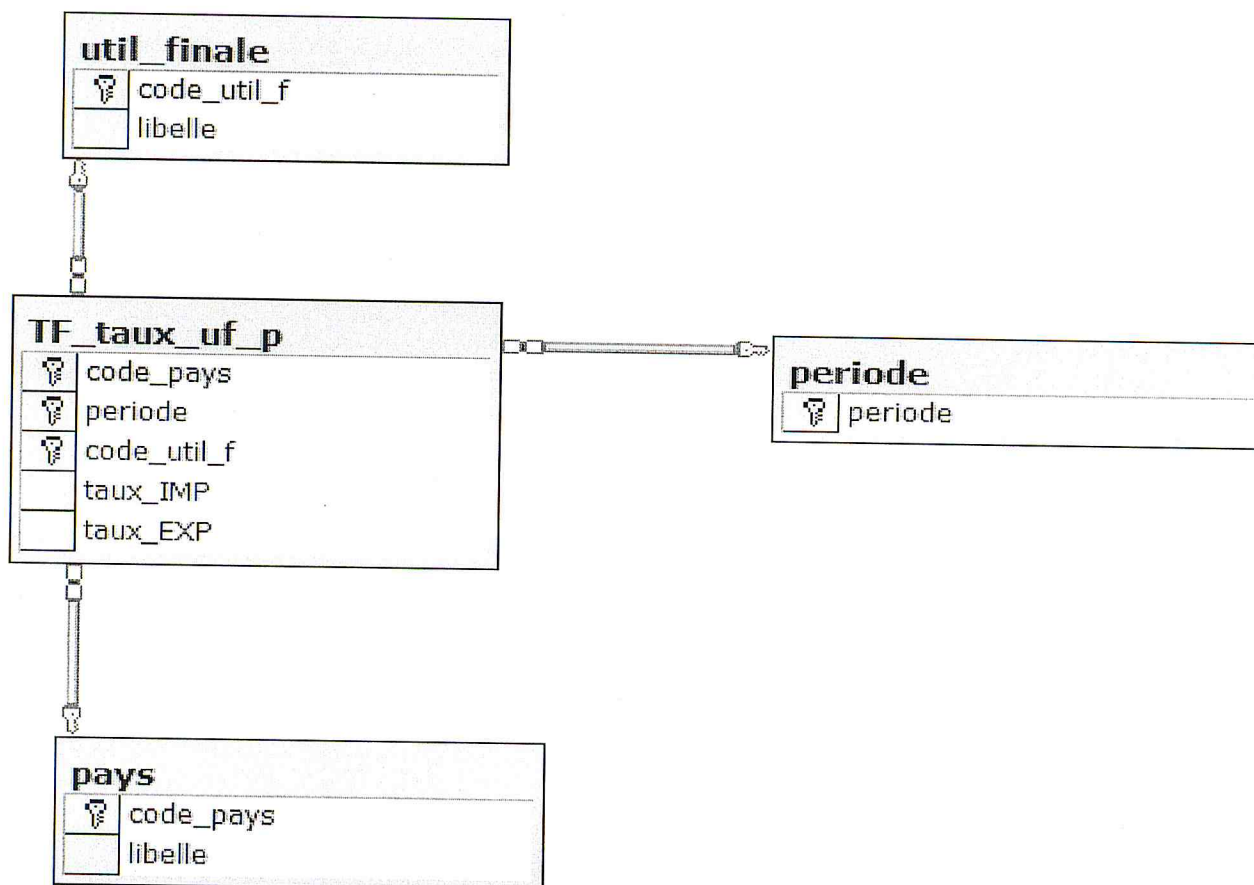


Figure 20 : Taux importation et exportation par util\_f et par pays

**Modélisation dimensionnelle de l'activité :**

**\*Taux utilisation finale de produit par group pays et par période \***

**Le processus d'activité :**

Ce processus est important, parce que permet de calculer les taux d'importation et les taux d'exportation par utilisation finale de produit et par période entre l'Algérie et tous les group pays exportateur et importateur



## Les dimensions

D'après le processus d'activité de la table de faits, nous avons déterminé une série de dimensions principales suivantes : **période, group pays, utile finale**

### Dimension \*période \*

La seule dimension qui figure systématiquement dans tous les entrepôts de données, car en pratique tout entrepôt est une série temporelle "temps".

### Dimension \*group pays\*

Cette dimension est nécessaire pour cette table de faits, il représente les group pays qui commercent avec l'Algérie

### Dimension \* utile finale\*

Cette dimension est nécessaire pour cette table de faits, il représente l'utilisation finale de chaque produit

## Les faits mesures :

Pour cette activité, les faits que nous avons enregistrés sont :

- Taux\_imp
- Taux\_exp

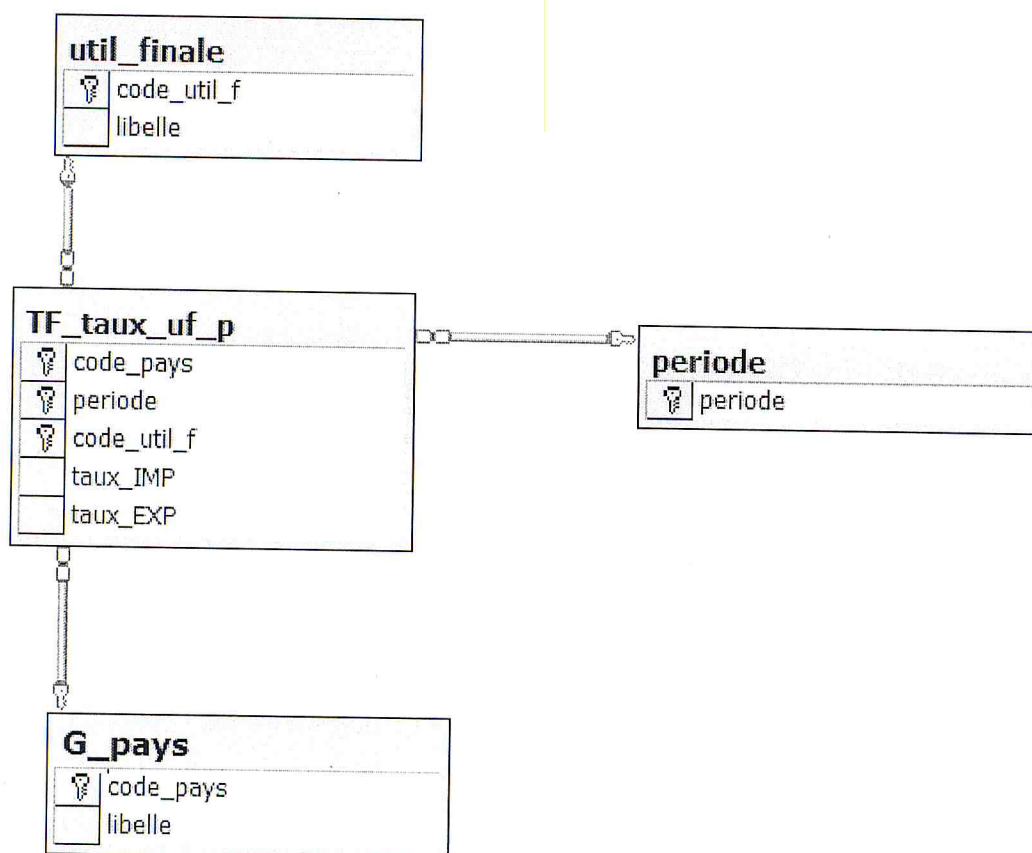


Figure 21 : Taux importation et exportation par période et par group pays et par util\_f

**Modélisation dimensionnelle de l'activité :**

**\*taux importation et exportation par group pays et par période \***

**Le processus d'activité :**

Ce processus est important, parce que permet de calculer la valeur d'exportation et importation pour chaque group pays par période

**Les dimensions**

D'après les processus d'activité de la table de faits, nous avons déterminé une série de dimensions principales suivantes : **période, group pays**

**Dimension \*période \***

La seule dimension qui figure systématiquement dans tous les entrepôts de données, car en pratique tout entrepôt est une série temporelle "temps".

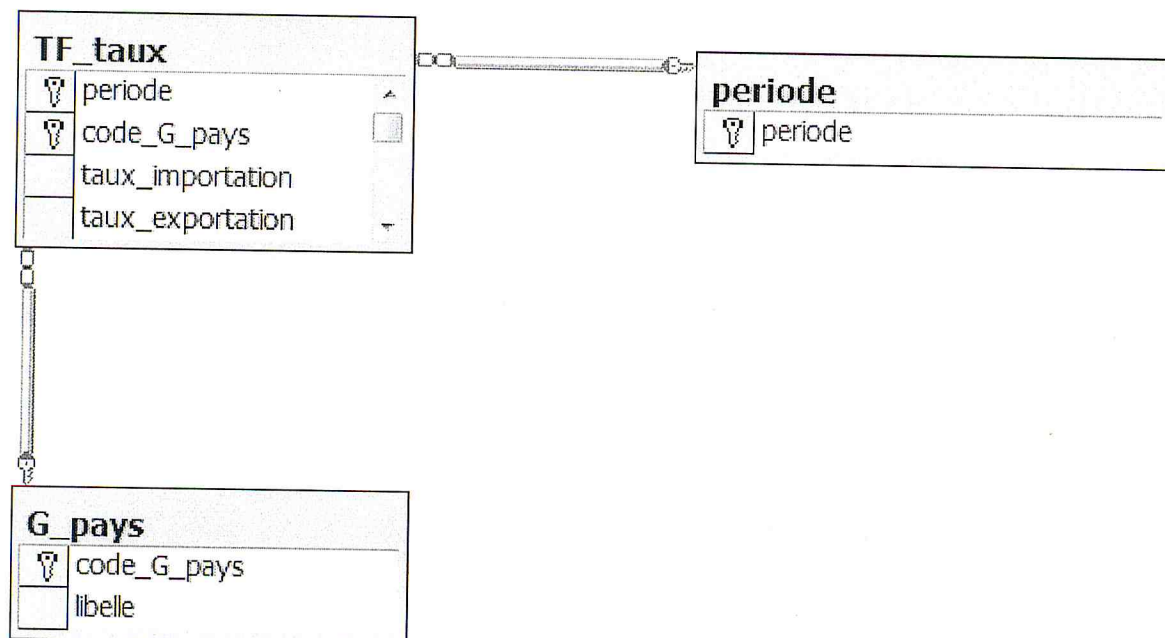
**Dimension \* group pays \***

Cette dimension est nécessaire pour cette table de fait, qui représente les groupes pays que commerce avec l'Algérie

**Les faits mesures :**

Pour cette activité, les faits que nous avons enregistrés sont :

- taux\_import
- taux\_export



**Figure 22 :** Taux importation et exportation par période et par groupe pays

## Modélisation dimensionnelle de l'activité :

### **\*Valeur d'exportation de l'entreprise \***

#### **Le processus d'activité :**

Ce processus est important, parce que permet de calculer la valeur d'exportation pour chaque entreprise par période

#### **Les dimensions**

D'après les processus d'activité de la table de faits, nous avons déterminé une série de dimensions principales suivantes : **période, entreprise**

#### **Dimension \*période \***

La seule dimension qui figure systématiquement dans tous les entrepôts de données, car en pratique tout entrepôt est une série temporelle "temps".

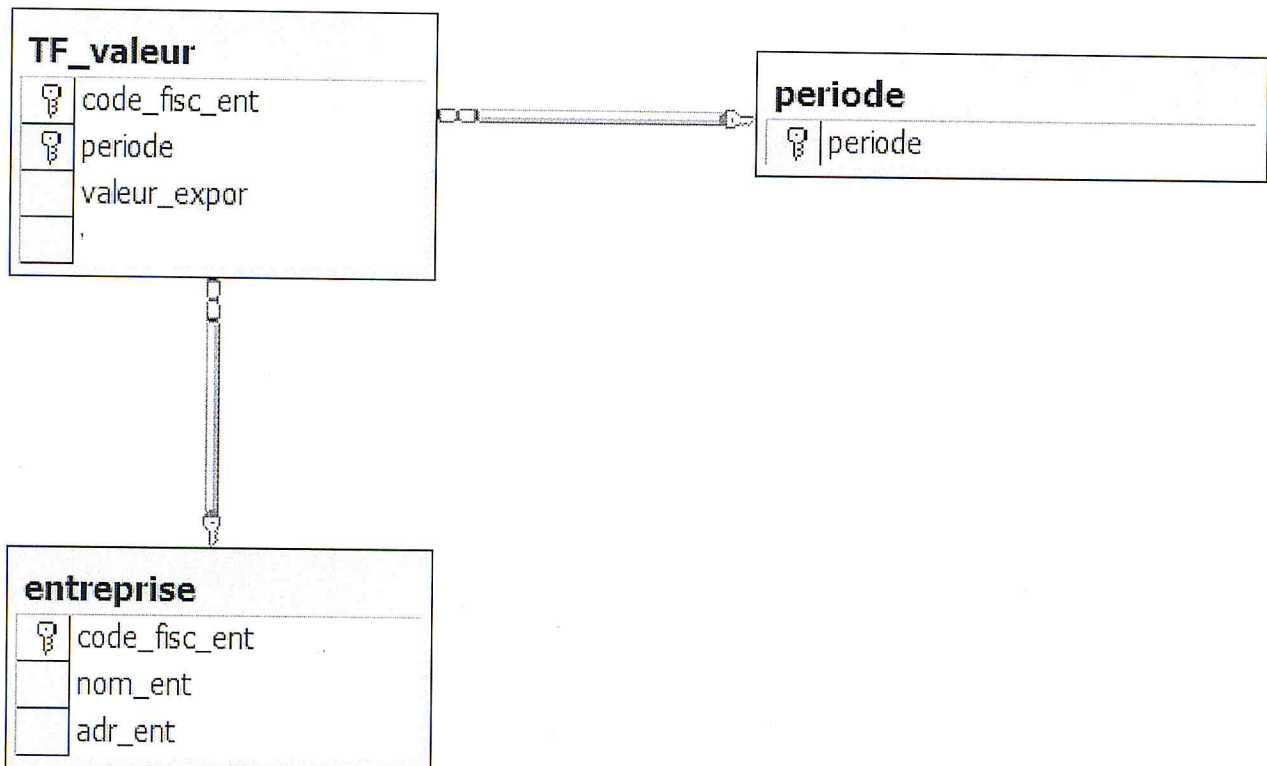
#### **Dimension \* entreprise \***

Cette dimension Est nécessaire pour cette table de faits, il représente l'entreprise exportatrice

**Les faits mesures :**

Pour cette activité, les faits que nous avons enregistrés sont :

- valeur\_export



**Figure 23 :** Valeur d'exportation de l'entreprise

## 2.2. Construction de l'entrepôt de données

Cette base répond à notre modèle de données, en d'autres termes c'est la transformation de la conception logique en une base de données physique.

Chaque entité dans notre modélisation, table de faits ou dimension, est transformée en une table de base de données relationnelle.

### Listes des tables qui constituent notre base décisionnelle

Table	Identifiant
Entreprise	code_fisca, nom_entreprise, adr_entreprise
G_pays	code_g_pays, libelle
Pays	code_pays, libelle
Periode	Periode
TF_balance_g_pays	code_g_pays, Periode, balance_g_pays
TF_balance_p	code_pays, Periode, balance_p
TF_taux	code_g_pays, Periode, taux_importation, taux_exportation
TF_taux_uf_g_p	code_g_pays, Periode, code_util_f, taux_importation_uf, taux_export, taux_import_D, taux_export_D
TF_taux_uf_p	code_pays, Periode, code_util_f, taux_importation, taux_export, taux_import_D, taux_export_D
TF_valeur	code_fisca, Periode, valeur_importation, valeur_expo_D
Util_final	code_util_f, libelle

Tab 5 : Liste des tables de notre entrepôt de données

## Description des attributs

La liste des attributs, leurs types de données et les tables à qui ils appartiennent sont résumés dans le tableau suivant

Description	Code	Table	Type de données
Code fiscal	Code_fisca	Entreprise	VARCHAR(50)
Nom de l'entreprise	Nom_entreprise	Entreprise	VARCHAR(50)
Adresse de l'entreprise	Adr_entreprise	Entreprise	VARCHAR(45)
Code groupe pays	Code_g_pays	GROUP_pays	VARCHAR(50)
Libelle groupe pays	Libelle	GROUP_pays	VARCHAR(50)
Code pays	Code_pays	Pays	VARCHAR(50)
Libelle pays	Libelle	Pays	VARCHAR(50)
Date	Periode	Periode	VARCHAR(7)
Mois	Mois	Periode	VARCHAR(2)
Code utilisation final	code_util_f	Util_final	VARCHAR(50)
Libelle utilisation final	Libelle	Util_final	VARCHAR(50)
Balance de groupe pays	Balance_g_pays	tf_balance_g_pay s	DECIMAL(20,0)
Balance de pays	Balance_pays	tf_balance_p	DECIMAL(20,0)
Taux d'importation	taux_importation	tf_taux	DECIMAL(20,0)
Taux d'exportation	taux_exportation	tf_taux	DECIMAL(20,0)

Taux d'importation	Taux_importation_uf	tf_taux_uf_g_p	DECIMAL(20,0)
Taux d'exportation	taux_export	tf_taux_uf_g_p	DECIMAL(20,0)
Taux d'importation	taux_import_D	tf_taux_uf_g_p	DECIMAL(20,0)
Taux d'exportation	taux_export_D	tf_taux_uf_g_p	DECIMAL(20,0)
Taux d'importation	taux_importation	tf_taux_uf_p	DECIMAL(20,0)
Taux d'exportation	taux_export	tf_taux_uf_p	DECIMAL(20,0)
Taux d'importation	taux_import_D	tf_taux_uf_p	DECIMAL(20,0)
Taux d'exportation	taux_export_D	tf_taux_uf_p	DECIMAL(20,0)
Valeur d'exportation	valeur_expo	tf_valeur	DECIMAL(20,0)

Tab 6 : Description des attributs



## 2.3 Conception de la zone d'alimentations

L'ETL est une étape des plus importantes dans un projet décisionnel. Cette étape a pour objectif d'assurer l'acheminement des données sources jusqu'à l'entrepôt de données, en passant par les différentes phases de transformation nécessaires.

La conception du processus d'alimentation nécessite les étapes suivantes

- Etude et planification.
- Conception des processus de chargement des tables de dimensions et des tables de faits.

### 2.3.1 Etude et planification

Notre source est une base de données Access que nous avons récupéré de l'entreprise ALGEX.

### 2.3.2 Processus de chargement

Le diagramme d'activités dans la figure (fig.22), décrit le processus général d'alimentation de l'entrepôt de données.

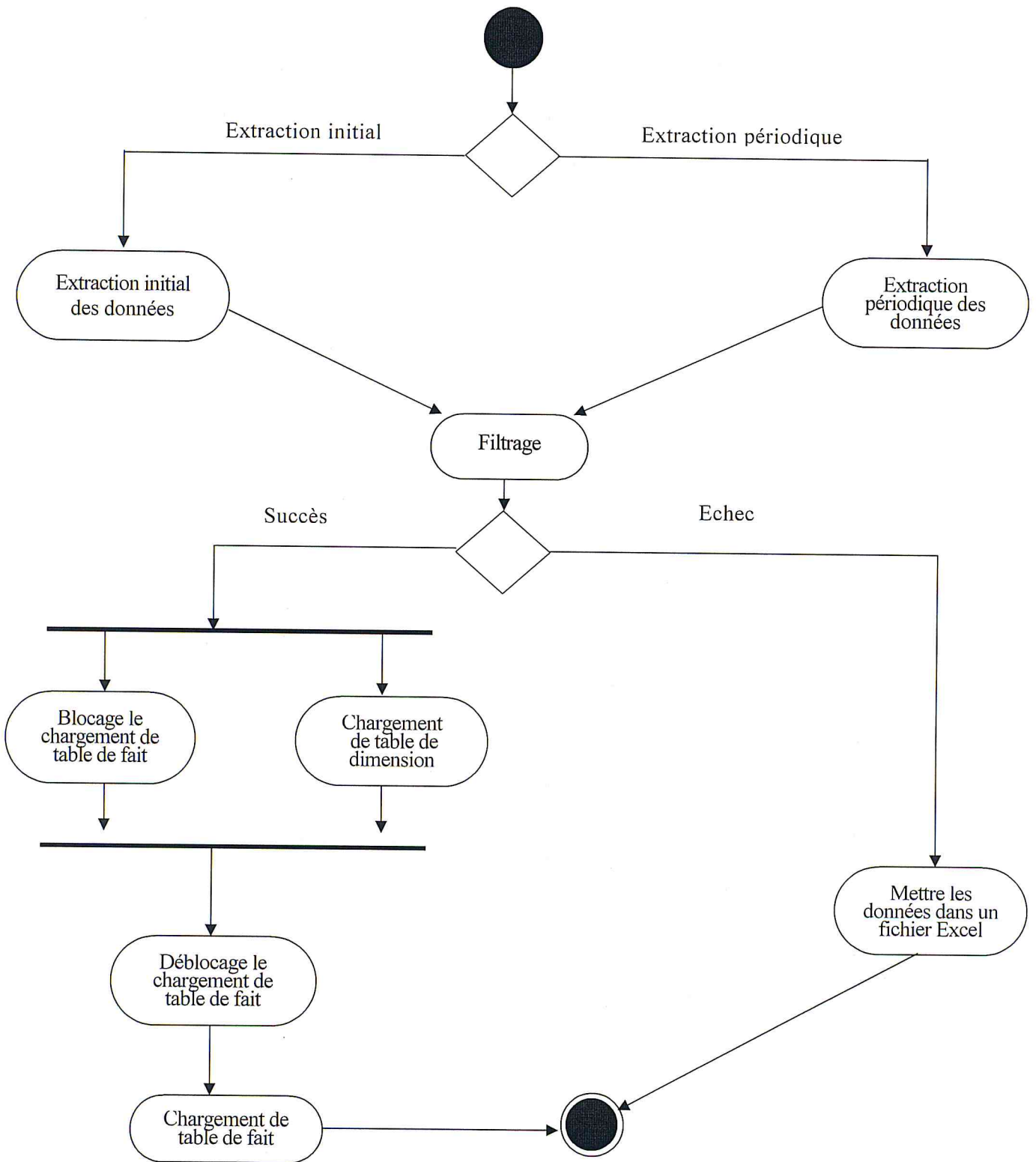


Figure 24: Diagramme d'activité du processus d'alimentation

## 2.4 Restitution

Après avoir présenté les aspects relatifs aux outils d'arrière plan de notre système décisionnel, ainsi que la conception et le développement de la base de données. Nous allons maintenant passer à la mise en œuvre d'une application qui permettra d'exploiter les données de notre entrepôt et de présenter l'information dans des formats utiles pour que les utilisateurs puissent accéder, comprendre et utiliser.

Cette partie est l'unique point d'entrée visible pour le consommateur de l'information. Pour cela, notre portail de restitution va lui permettre de manipuler les données.

Notre restitution va dépendre forcément des besoins des consommateurs de l'information, donc l'utilisation des informations de notre entrepôt de données va se baser sur :

### ➤ Le reporting :

Il s'agit de mettre en œuvre une interface utilisateur efficace, et les informations y sont pré structurées sous la forme de tableaux de bord ou d'états prédéfinis.

Code Pays	Pays	periode	Balance
025	AFRIQUE DU SUD	02/2011	-1228277225
025	AFRIQUE DU SUD	03/2011	-1264931816
025	AFRIQUE DU SUD	05/2011	-1223400787
025	AFRIQUE DU SUD	07/2011	-2541865723
025	AFRIQUE DU SUD	08/2011	-3832162375
025	AFRIQUE DU SUD	09/2011	-2504615539
025	AFRIQUE DU SUD	10/2011	-3873480127
031	ANGOLA	01/2011	830829
031	ANGOLA	04/2011	18712
031	ANGOLA	05/2011	1020352
031	ANGOLA	06/2011	683062
031	ANGOLA	10/2011	207634
043	CAMEROUN	01/2011	925882
043	CAMEROUN	02/2011	1473106
043	CAMEROUN	03/2011	1301794
043	CAMEROUN	05/2011	248570
046	CONGO	03/2011	3034124
046	CONGO	04/2011	1294246
046	CONGO	05/2011	733833
046	CONGO	10/2011	776169

Figure 25 : Interface représentant la balance par période et pays

# Chapitre III

Déploiement

## Introduction

Nous allons dans ce chapitre proposer une réalisation de l'entrepôt de données, ainsi que du tableau de bord. Nous allons présenter aussi l'ensemble des outils qui ont été utilisés, ainsi que l'architecture de notre application.

### 1. Architecture du système

Notre système décisionnel est destiné aux décideurs d'ALGEX pour avoir une meilleure vue sur les données de l'entreprise. L'architecture de notre solution est une architecture en trois-tiers :

**Architecture trois-tiers** est une architecture d'application dans laquelle on sépare la présentation, les traitements, et les bases de données. L'objectif ciblé est de permettre l'évolution de l'un de ces trois tiers de façon relativement indépendante des deux autres. L'implémentation physique de ces architectures est souvent soumise à plusieurs contraintes, car elles sont parfois mise en œuvre à travers des plates-formes différentes.

L'architecture que nous allons déployer est la suivante :

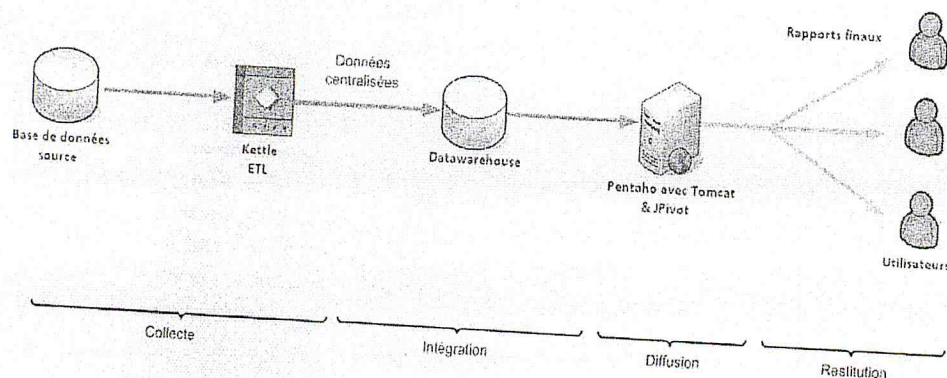


Figure 26 : Architecture physique du système

Les outils choisis sont :

1. Le SGDB MySQL 5.1 pour l'intégration de la base de données de notre entrepôt de données.
2. Pentaho Data Intégration (PDI) qui est un outil ETL intégré dans la solution pentaho.
3. serveur apache tomcat
4. Restitution : JSF Primefaces

### 3. Configuration du système

Notre système sera déployé comme suit :

#### 3.1 Système de gestion de base de données

Le SGBD choisit pour notre implémentation c'est le système de gestion de bases de données relationnel MySQL, est la base de données open source la plus populaire. Le SGBD MySQL 5.1 présente entre autre des améliorations importante qui ont été apportées en termes de performance qui offrent des niveaux supérieurs d'intégrité des données, de flexibilité dans les applications.

Le SGBD nous permet de définir tous les contraintes d'intégrité que peut comporter notre base : les clés primaires et secondaires, les indexes, les privilèges sur les données...

#### 3.2 ETL Pentaho Data Integration «PDI»

Le processus d'ETL est un processus primordial dans l'entrepôt de données, il permet la mise en place d'un lien entre le système source et le datamart, en faisant le transfert de données. Après le transfert, il entame l'épuration de données afin d'extraire les données utiles de charger dans la base de notre entrepôt de données.

Ce processus est réalisé grâce à Pentaho Data Intégration aussi connu sous le nom de KETTLE.

PENTAHO BI SUITE est un logiciel fourni par la compagnie PENTAHO, c'est une solution open source entièrement développée en java, qui permet de couvrir les principaux domaines d'un projet de Business Intelligence et ceci à travers différents logiciels de Pentaho ou intégrables dans l'offre de l'éditeur.

Pentaho Data Intégration est l'ETL de la suite décisionnelle libre Pentaho. Cet ETL est un «moteur de Transformation»: les données traitées et les traitements à effectuer sont parfaitement séparés.

Les traitements sont stockés dans un référentiel qui peut être soit au format XML (fichiers plats), soit dans une base de données (ce qui permet notamment le partage entre plusieurs designers).

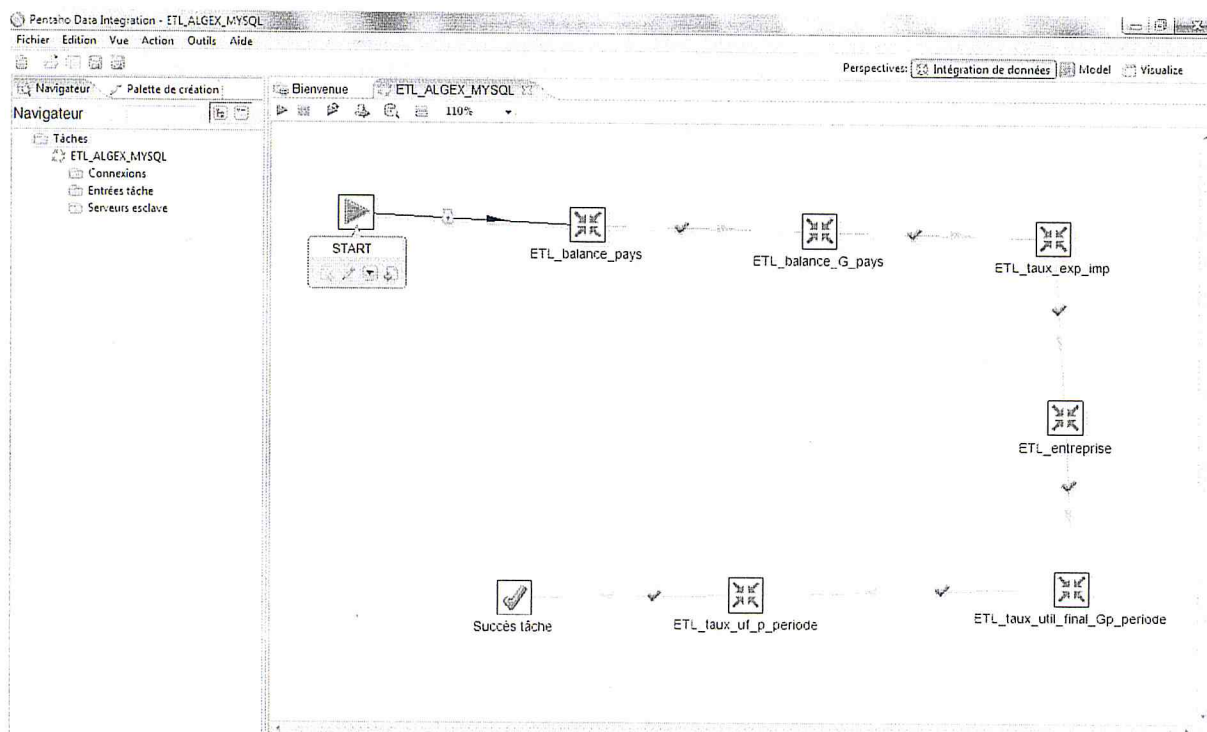
De nombreux types de SGBD sont supportés, ainsi que tous les types de fichiers plats (Excel, XML).

Pentaho Data Intégration dispose d'une interface graphique « Spoon » (basée sur SWT), depuis laquelle on peut créer deux types de traitements :

- Des transformations : celles-ci constituent les traitements de base d'intégration de données avec toutes les étapes (steps) nécessaires à l'extraction, la transformation, et le chargement des données.

- Des tâches (jobs) : celles-ci permettent le séquençement de plusieurs transformations avec des fonctionnalités plus orientées AIE (Enterprise Application Intégration): gestion des erreurs, envoi de mails de notification, transferts FTP, exécution de scripts Shell ou SQL.

Les figures suivantes représentent un exemple des étapes de transformations effectuées.



**Figure 27 :** Tache chargement du notre entrepot de données

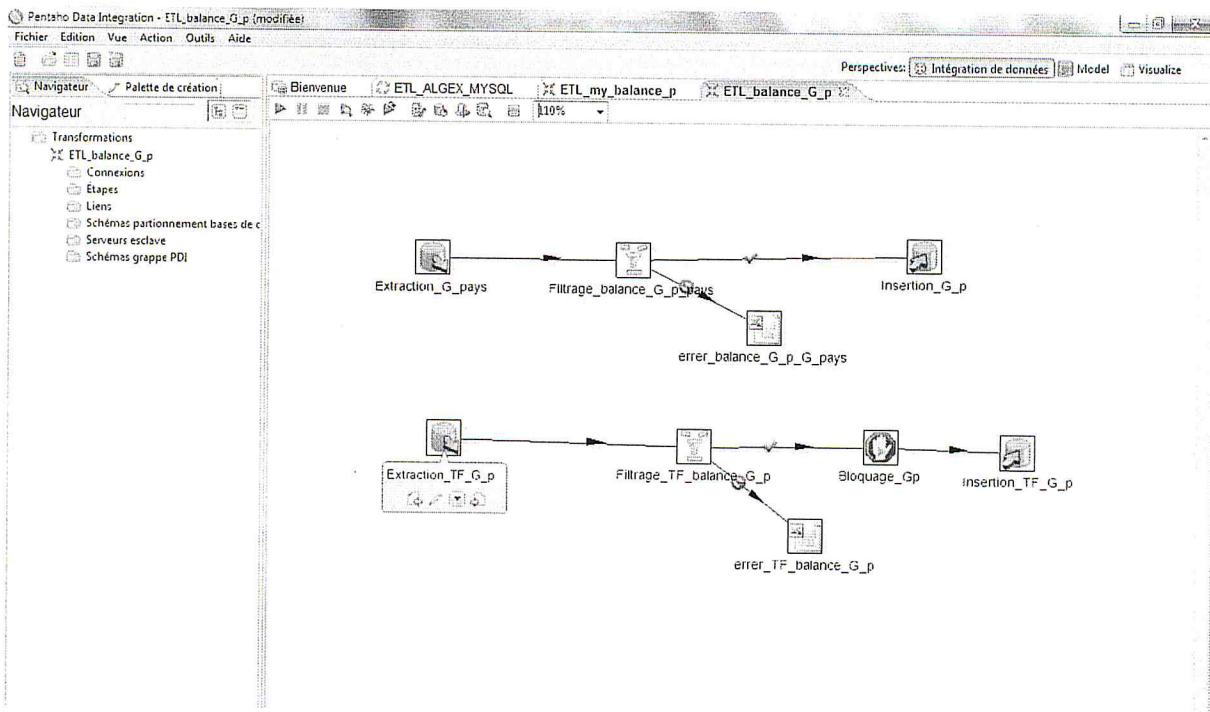


Figure 28 :Etapes de la tâche chargement de la table fait balance par groupe pays

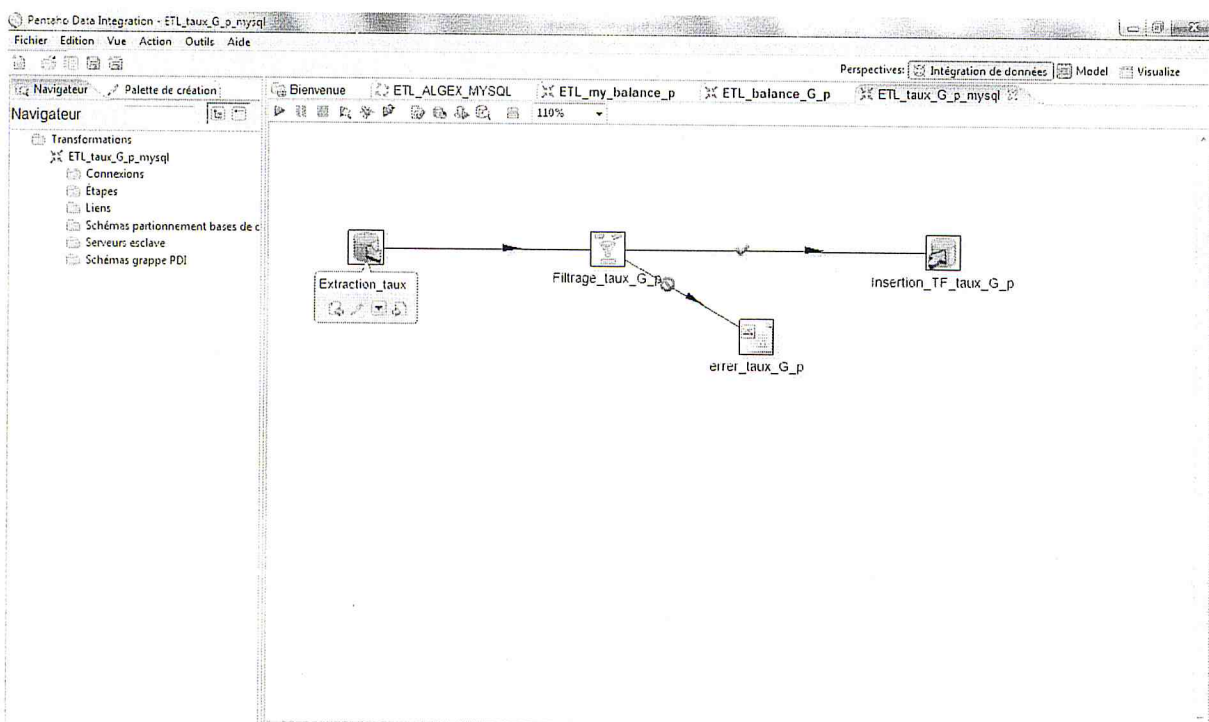


Figure 29 :Etapes de la tâche chargement de la table fait taux par groupe pays



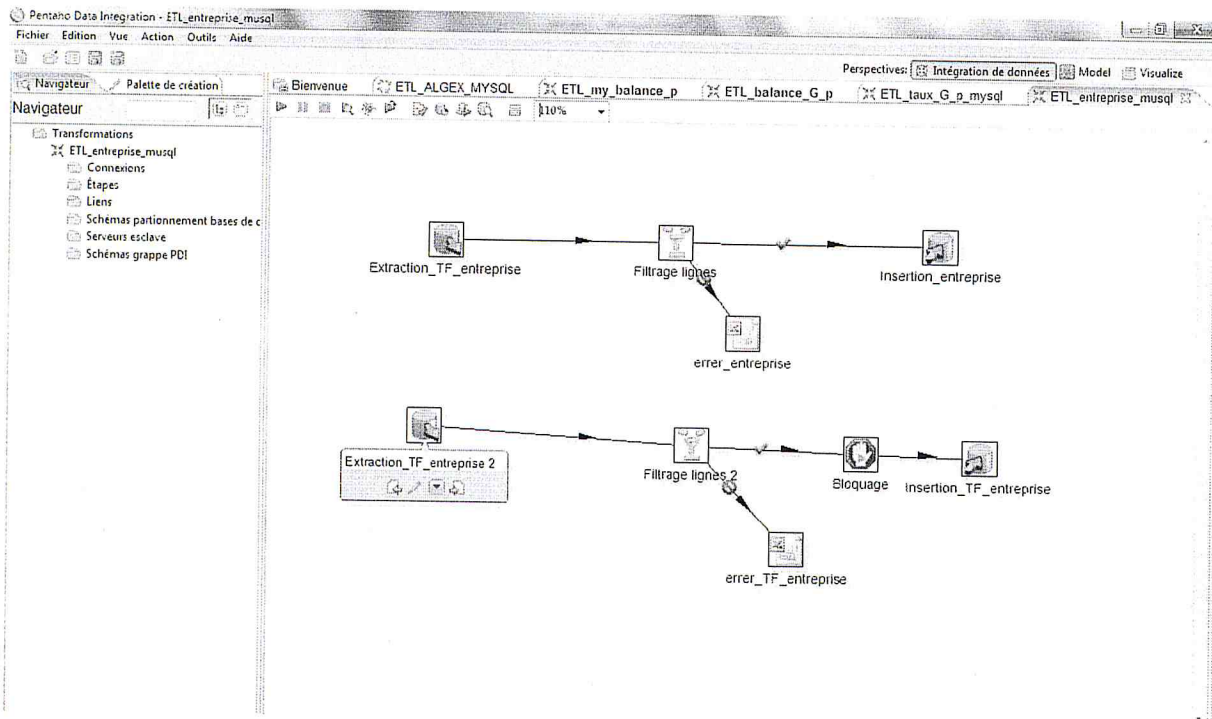


Figure 30 :Etapes de la tâche chargement de la table fait entreprise

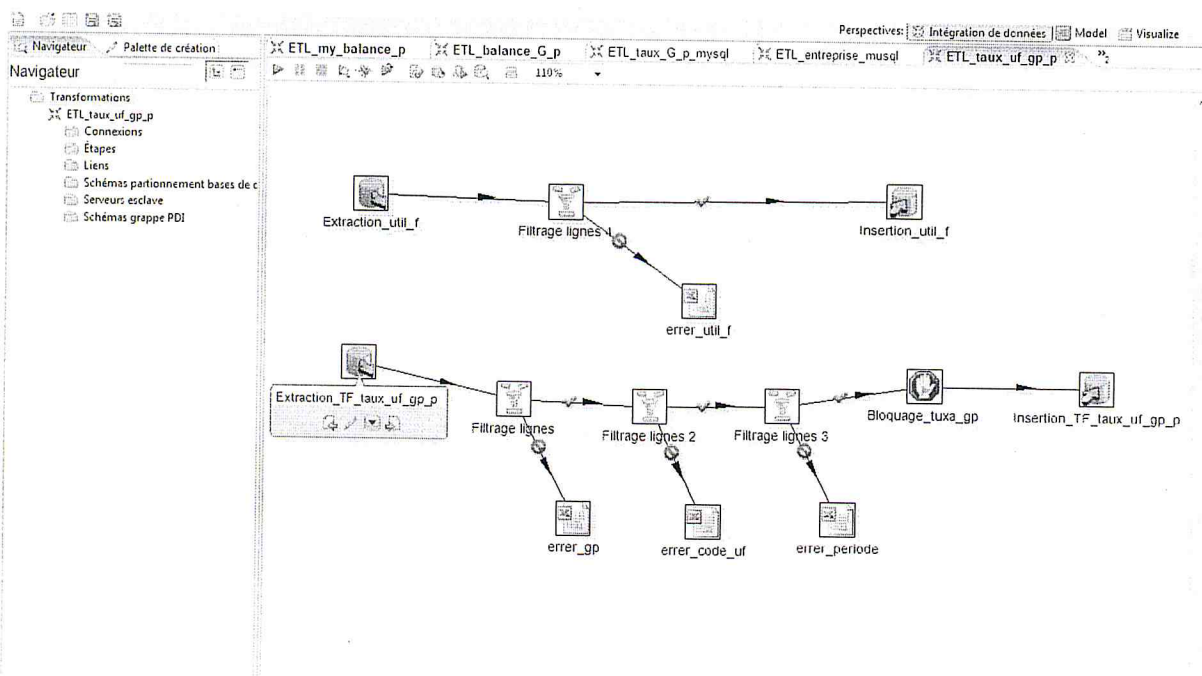


Figure 31:Etapes de la tâche chargement de la table fait taux par groupe pays et utilisation final

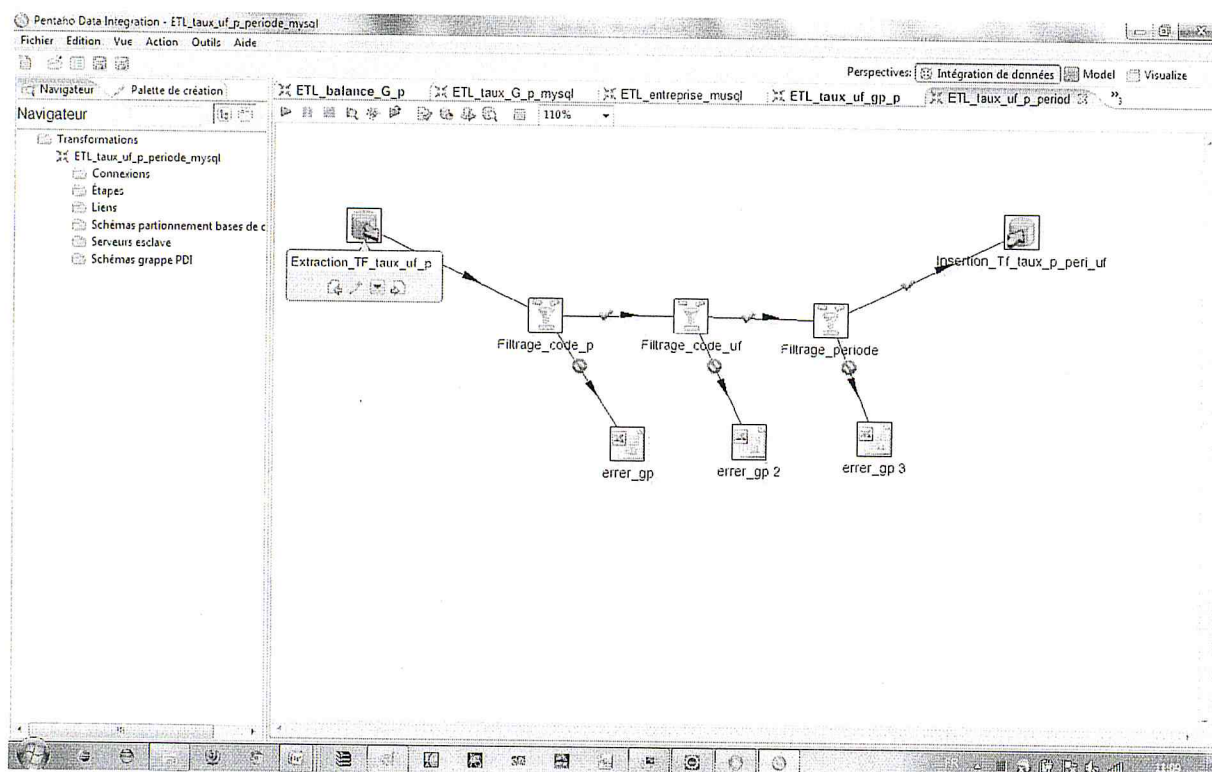


Figure 32 :Etapes de la tâche chargement de la table fait taux par pays et utilisation final

### 3.3 Architecture d'une application Web JEE:

L'Architecture d'une application web JEE est devisée en trois couches ou bien en trois niveaux, basés sur le modèle MVC (Model-View-Controller), l'architecture des trois couches est le modèle le plus générale dans les architectures multicouches, les trois niveaux sont :

- **La couche présentation** : La couche de présentation est la partie visible de l'application qui permet à un utilisateur d'interagir avec le système. Est décrit en langage HTML puis interprété par le navigateur. Dans notre cas on utilise le JSF.
- **La couche métier** : Elle correspond à la partie fonctionnelle de l'application, celle qui implémente la « logique », et qui décrit les opérations que l'application opère sur les données en fonction des requêtes des utilisateurs, effectuées au travers de la couche présentation. En pratique, on trouve au niveau de la couche métier :
  - ❖ Des entitybeans dont la persistance est assurée par la couche de mapping.

- ❖ Des statelessbeans qui proposent des méthodes pour manipuler les entitybean (CRUD).
  - ❖ Des message-drivenbeans qui assurent les traitements asynchrones.
  - ❖ Les API JNDI, pour accéder au service de nommage, et JavaMail, pour envoyer des mails aux clients
- **Couche Accès aux données** : Elle consiste en la partie gérant l'accès aux gisements de données du système. Ces données peuvent être propres au système, ou gérées par un autre système. La couche métier n'a pas à s'adapter à ces deux cas, ils sont transparents pour elle, et elle accède aux données de manière uniforme.

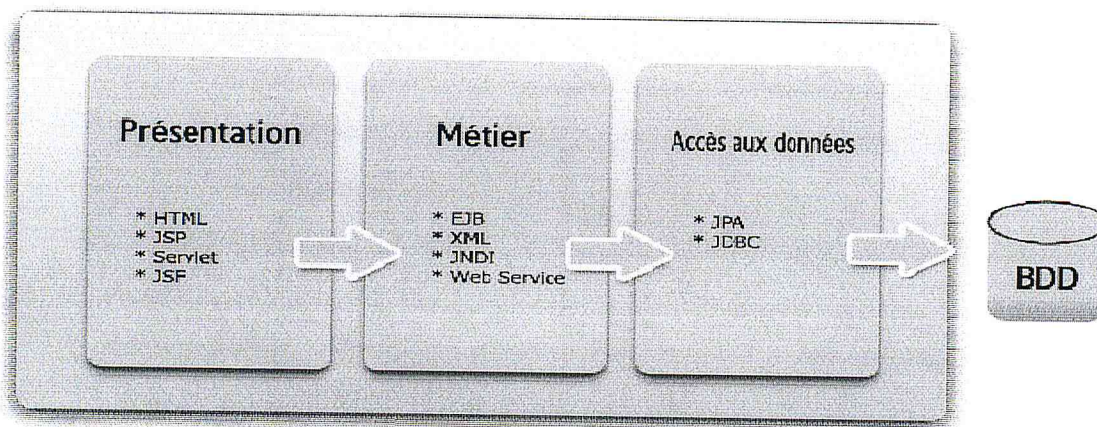


Figure 33 : Architecture en couches d'une l'application JEE

### 3.4 Hibernate :

Hibernate est un framework de mapping objet/relationnel. Ce framework est de plus en plus utilisé. Concrètement, cela veut dire qu'Hibernate nous permet de manipuler les données d'une base de données relationnelle sous forme d'objet. Pour faire cela, Hibernate utilise des fichiers pour relier la base aux objets. Le fait de manipuler directement les données d'une base sous forme d'objet est beaucoup plus pratique, cela nous permet de nous défaire de toute la couche SQL. De plus, cela permet de définir clairement la limite entre la persistance et la couche métier, ce qui se révèle très utile dans le cas d'une application trois-tiers.

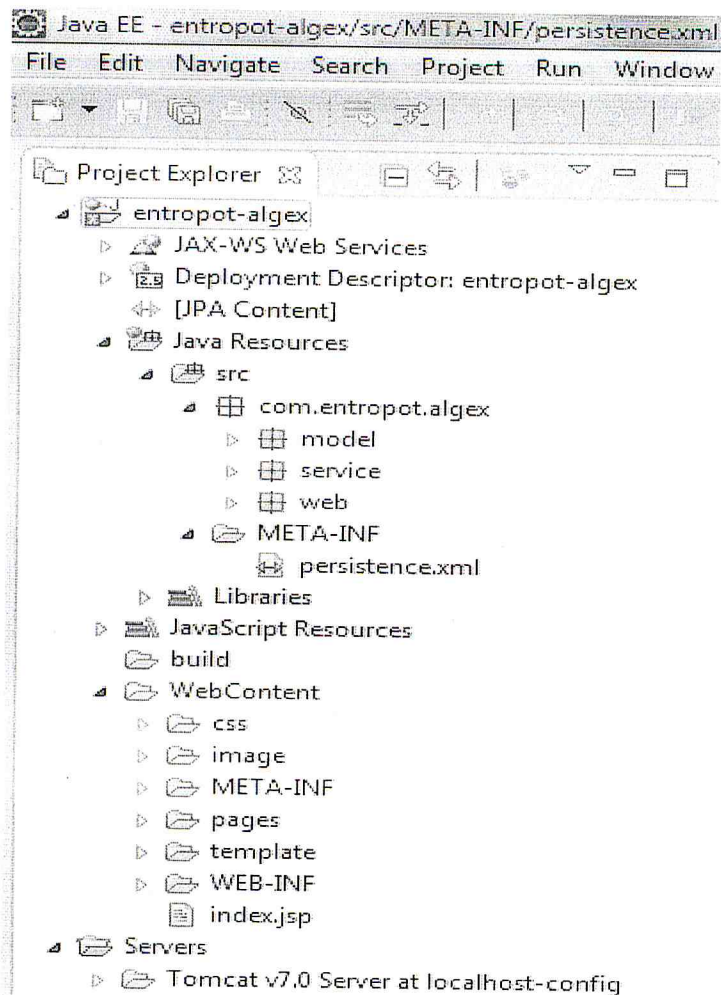
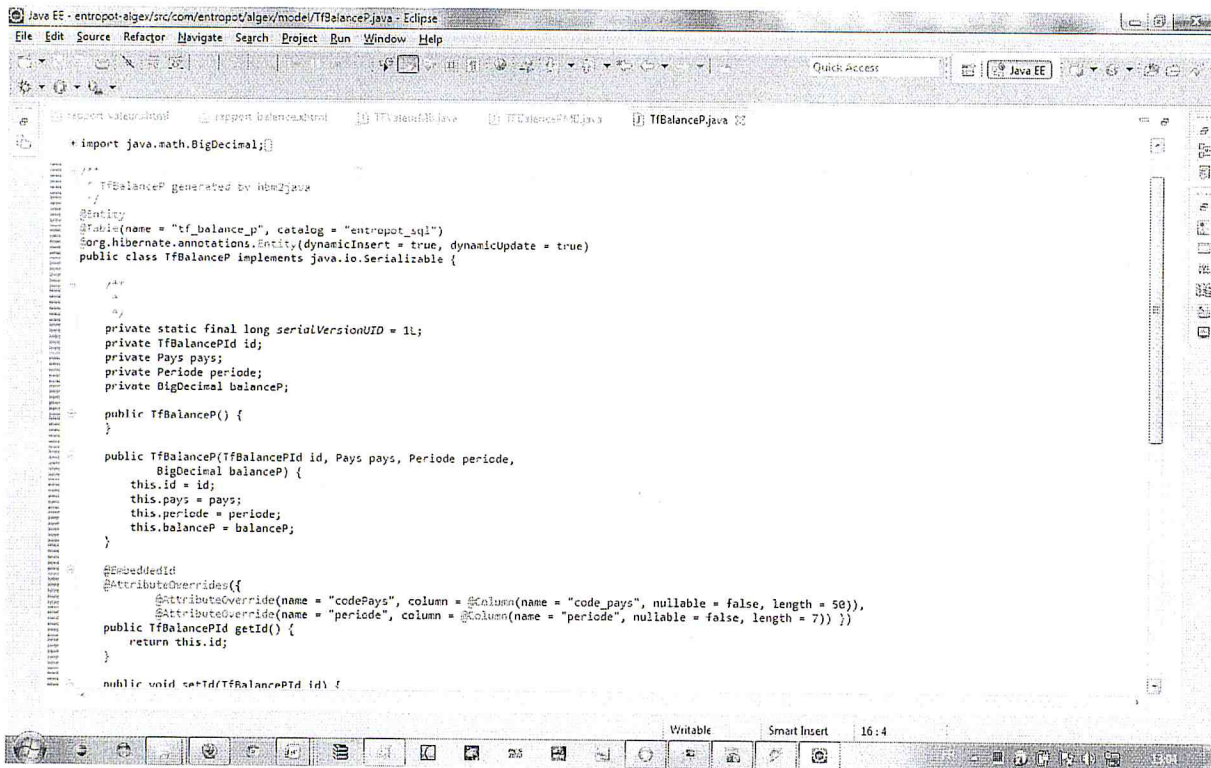


Figure 34 : Structure de l'application java EE

### 3.5 Serveur d'application TOMCAT

TOMCAT est un conteneur libre de servlet et JSPJava EE. Il a été créé par la fondation apache, il implémente les spécifications des servlet java et des JSP.

JEE (Java Enterprise Edition) est une plateforme qui permet de faciliter le développement d'applications d'entreprise.



```
import java.math.BigDecimal;

/**
 * TfBalanceP generated by hbm2java
 */
@Entity
@Table(name = "tf_balance_p", catalog = "entropot_sql")
@org.hibernate.annotations.Entity(dynamicInsert = true, dynamicUpdate = true)
public class TfBalanceP implements java.io.Serializable {

    private static final long serialVersionUID = 1L;
    private TfBalancePid id;
    private Pays pays;
    private Periode periode;
    private BigDecimal balanceP;

    public TfBalanceP() {
    }

    public TfBalanceP(TfBalancePid id, Pays pays, Periode periode,
        BigDecimal balanceP) {
        this.id = id;
        this.pays = pays;
        this.periode = periode;
        this.balanceP = balanceP;
    }

    @EmbeddedId
    @AttributeOverrides({
        @AttributeOverride(name = "codePays", column = @Column(name = "code_pays", nullable = false, length = 50)),
        @AttributeOverride(name = "periode", column = @Column(name = "periode", nullable = false, length = 7)) })
    public TfBalancePid getId() {
        return this.id;
    }

    public void setId(TfBalancePid id) {
    }
}
```

Figure 35 :Exemple de code java

## 4. Fonctionnalités du système

Après avoir présenté la plate-forme (zone de réparation et de présentation de données) de notre système, nous allons présenter dans cette partie des différentes fonctionnalités de notre application (zone de restitution de données).

### 4.1 Interface administration

Cette interface dédiée uniquement à l'administrateur de données permet de créer, modifier ou supprimer des utilisateurs. On peut aussi attribuer à chaque utilisateur un ensemble de droit, avec lesquelles l'utilisateur peut consulter, modifier ou supprimer des rapports publiés par d'autres utilisateurs.

The screenshot shows a web browser window with the URL `localhost:3000/entrepot-algex/pages/Utilisateur/index.jsf`. The page title is "GT - Utilisateur" and the logo "ALGEX" is visible. The main content area is titled "Nouvel Utilisateur" and contains a form for creating a new user. The form fields include:

- #ID: (empty)
- Nom: Benacer
- Prénom: Abdallah
- Phone: 0550299039
- Nom utilisateur: admin
- Mot de passe: (masked with dots)
- Confirmer le Mot de passe: (masked with dots)
- Activé:

Below the form are buttons for "Ajouter" and "Annuler". To the right of the form, there are role selection options: "AGENT" and "ADMINISTRATEUR". The footer of the page contains the text "Algex Entrepot SQL - 2013 ©", "Mercredi 12 Juin 2013 13:01:14", and "Copyright ALGEX-ENT-SQL © 2013".

Figure 36 : Console administrateur

## 4.2 Interface utilisateur

La page d'accueil donne un accès aux utilisateurs finaux et cela se fait toujours suite à une authentification de l'utilisateur, la figure (fig.36) représente la page d'accueil des utilisateurs qui fait office d'authentification afin d'élaborer différentes vues d'analyses multidimensionnelles pour leurs activités, mais aussi de consulter les rapports d'états publié au part avant.

The screenshot shows the user login interface. On the left is the ALGEX logo. On the right is a login form with the following fields and controls:

- Nom utilisateur: user
- Mot de passe: (masked with dots)
- Se souvenir?
- Buttons: "Se connecter" and "Effacer"

Figure 37 : Console utilisateur

### 4.3 Les rapports

L'utilisateur final peut aussi à travers l'interface utilisateur consulter les rapports élaborés au part avant par l'administrateur en utilisant JSF Primefaces.

Le composant JSF Primefaces permet de mettre à disposition des interfaces clients riches dans nos applications Web JEE.

La figure (fig.34) représente un rapport statique (la balance par groupe pays et période) publié par l'administrateur.

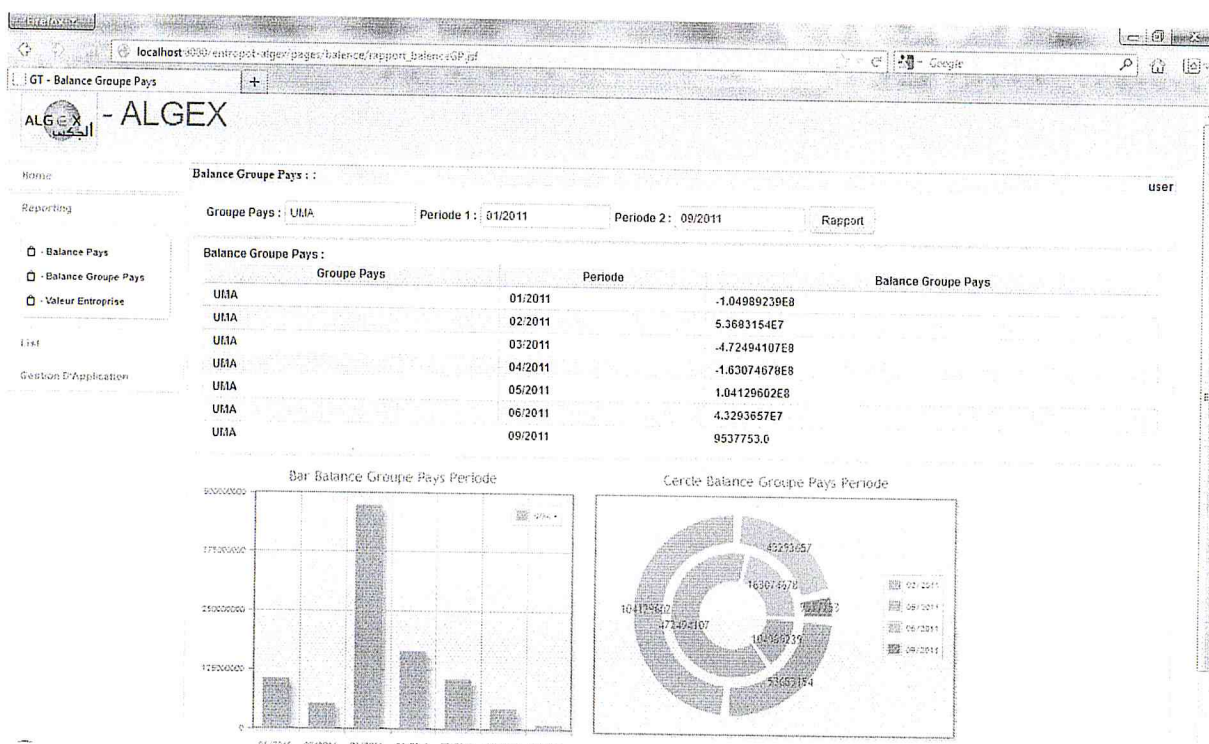


Figure 38 :Rapport statistique de balance par groupe pays

La figure (fig.37) représente un rapport statique (la balance par pays et période) publié par l'administrateur.

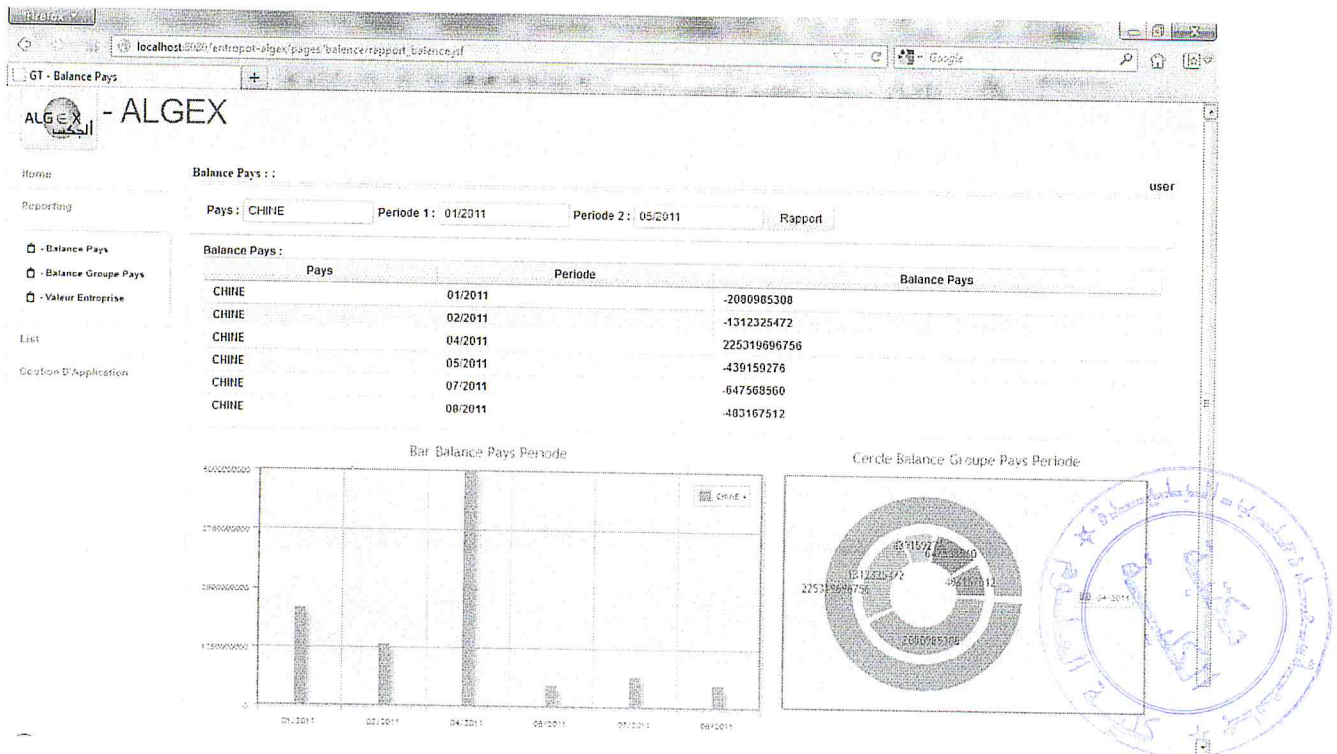


Figure 39 : Rapport statistique de balance par pays

## 5. Sécurité du système

La qualité d'un système suppose au préalable une protection suffisante et cohérente vis-à-vis des risques possible. Le système mis en place doit être capable de se défendre contre toute attaque. Ceci est possible par l'élaboration d'un plan de sécurité rigoureux qui protégera la capitale information de l'entreprise contre de tout individu malsain.

- **Niveau utilisateur**

Pour sécuriser l'accès utilisateurs, les directions d'ALGEX disposent d'un mécanisme d'autorisation et d'authentification. Pour ce qui concerne la sécurité des machines, des antivirus sont installés pour pallier aux attaques de virus et autre logiciel espion, des mises à jour sont effectuer régulièrement.

- **Niveau application**

Le mécanisme de sécurité de notre application et de notre page web est basé sur le mécanisme d'authentification (login et password) et d'autorisation (les droits de publication et de modification des vues et des rapports).

- **Niveau base de données**

C'est la sécurité du serveur de la base de données, basée aussi sur le mécanisme d'autorisation et d'authentification (l'accès à la base) et sur la sauvegarde périodique des données au niveau de l'entrepôt de données.



## **Conclusion**

Nous avons présenté dans ce chapitre notre réalisation, ainsi que les outils avec lesquels nous avons implémenté notre travail. Nous avons réalisé un tableau de bord permettant d'afficher les statistiques, et permettant d'afficher les rapports.

## Bibliographie

- [Kimball, 2003]** Ralph Kimball et Margy Ross. Entrepot de donnees Guide pratique de modelisation dimensionnelle. Edition Vuibert Informatique, 2003
- [El GoIli, 2008]** Ines Gam El GoIli. Ingenierie des Exigences pour les Systemes d'Information Decisionnels : Concepts, Modeles et Processus La methode CADWE, Universite Paris I - PANTHEON — SORBONNE, Octobre 2008.
- [Kimball, 1996]** Ralph Kimball. The data Warehouse toolkit. Edition Jhone Wiley, 1996
- [Bill Inmon, 2000]** B.Inmon, The DataWarhouse TollKit, Eyrolles, 1996
- [Goglin, 2001]** Jean-Francois Goglin. Construction du datawarehouse. Edition Hermes-Lavoisier, Mai 2001.
- [L05]** Kimball Ralph, « Entrepot de donnee, guide pratique de concepteur de data warehouse >; 2000.
- [Kimball, 2000]** Ralph Kimball, Laura Reeves, Margy Ross et Warren Hornthwaite, Concevoir et deployer un Data warehouse Guide de conduit de projet. Edition Eyrolles 2000.
- [Kimball, 1997]** Ralph Kimball. Entrepot de donnees. Edition International Thomson Publishing, 1997.
- [Benitez, 2001]** E. Benitez, C. Collet, et M. Adiba. Entrepots de donnees caracteristiques et problematiques. Revue TSI, 20(2), 2001.
- [Trinidad, 2005]** Maria Trinidad Serna Encinas. Entrepôts de donnees pour l'aide a la decision medicale : conception et experimentation. These pour obtenir le grade de Docteur de l'Universite Joseph Fourier, Juin 2005.
- [Gar, 2000]** Support de cours : « Explosion de l'informatique dedsionnelle » par G.Gardarin.

## Web graphie

- [1] : [www.piloter.org](http://www.piloter.org) cours complet sur le BI
- [2] : [www.algex.dz](http://www.algex.dz) Site officiel d'Algex
- [3] : [www.wikipedia.org](http://www.wikipedia.org) Site wiki