

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

**MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA
RECHERCHE SCIENTIFIQUE**

UNIVERSITE SAAD DAHLEB de BLIDA



Faculté des Sciences

Département d'Informatique

Domaine : Mathématique et Informatique

Filière : Informatique

Option : Ingénierie des logiciels

En vue d'obtenir le diplôme de Master 2

**Conception et réalisation d'un système
d'information décisionnel orienté
processus commercial et marketing**

Promotrice : Mlle OUKID.L

Réalisé par :

Encadreur : Mme BELLI.H

BENKALI Hakima

Soutenu le : 26/09/2016

BOUFRIDI Nawel

Devant le jury: Mme I. Chikhi

M M. Bala

MA-004-358-1

Année universitaire 2015/2016

ملخص

تعتبر منظومة قطاع النقل البحري تطويرية، كثيفة ومتعددة الأنشطة و التي تشكل العديد من العوامل التي تؤثر على اداءها. ونظرا لأهمية هذا القطاع تخضع شركات الشحن البحري للوائح والمعايير الدولية، و يجب أن تستوفي متطلبات الأداء ليس فقط فيما يتعلق بهذه المعايير، ولكن أيضا تبعا لنسبة الأرباح المتوقعة و وضعها و مكانتها في السوق.

GEMA هي شركة جزائرية للشحن البحري يتمثل دورها الرئيسي في ضمان تسجيل السفن ، و من أجل ضمان مكانتها الريادية، تريد GEMA أن تنجز لوحة قيادة متطورة لتحسين الإدارة الاستراتيجية.

النظم الحالية لهذه الشركة تتعامل مع بيانات غير منظمة، مع عدم وجود مؤشرات القبض رسميا مما يعيق الوصول إلى المعلومات لاتخاذ القرارات. للقيام بذلك، عهدت الينا ادارة الشركة مهمة تنفيذ نظام دعم القرار الذي من شأنه أن يستجيب لاحتياجات متخذي القرارات للحصول على تحليل البيانات و اعطاء معلومات موثوقة في وقت قصير وتحسين عملية تقديم التقارير.

من خلال هذا العمل سوف نقوم بتصميم نظام استقصاء تخزين و مستودع البيانات من أجل تنفيذ عمليات المبيعات والتسويق، من أجل الوصول الى انجاز لوحة قيادة متكاملة لدعم التحليل واتخاذ القرار.

كلمات البحث :

نظام استقصاء المعلومات، مستودع البيانات، محل البيانات ، نمذجة الأبعاد، لوحة القيادة , ETL, KPI

Résumé

Le secteur maritime est souvent qualifié d'évolutif, intensif, reposant sur une multitude d'activités constituant des facteurs influençant sa performance. En vue de l'importance du secteur, les compagnies maritimes, soumises à des réglementations et des standards internationaux, doivent répondre aux besoins en matière de performance non seulement à l'égard de ces standards, mais aussi vis-à-vis de leurs attentes internes en matière de profits et de leur position au niveau du marché.

GEMA (Société Générale Maritime) est une compagnie maritime Algérienne dont la vocation principale est d'assurer la consignation des navires. Pour garantir toujours sa présence sur le podium, GEMA doit mettre en place un tableau de bord commercial pour améliorer son management stratégique.

Les systèmes d'information automatisés existants au sein de l'entreprise, manipulent un volume important de données, néanmoins, ces dernières ne sont pas suffisamment organisées pour générer des indicateurs de performance structurés, afin de prendre des décisions au moment opportun.

A cet effet, les décideurs nous ont confié la mission de mettre en place un système décisionnel commercial, qui permettra de répondre aux besoins stratégiques de l'entreprise. Le nouveau système à concevoir doit permettre une analyse en ligne OLAP des données, d'avoir des informations fiables dans les brefs délais et d'améliorer le processus d'élaboration des rapports commerciaux.

Dans ce travail, nous proposons une conception et une réalisation d'un entrepôt de données (Data Warehouse) orienté processus commercial et marketing. Nous proposons également un tableau de bord pour l'analyse et l'aide à la décision.

Mots-clés :

Système décisionnelle, entrepôt de données, magasin de données, ETL, KPI, modélisation dimensionnelle, tableau de bord.

Abstract

The maritime sector is evolutionary, intensive, based on myriad activities; it's constituted of factors affecting their performance. In view of the importance of this sector, shipping companies are subject to regulations and international standards must satisfy the performance requirements not only with regard to these standards, but also against their internal expectations for profits and their position in the market.

GEMA is an Algerian shipping company its main purpose is to ensure the recording of ships. To always ensure its presence in the podium, GEMA wants to achieve a dashboard to improve its strategic management. Current systems of this company manipulate unorganized data with the lack of indicators which makes difficult to access for information to make decisions. For this purpose, leaders have entrusted us the implementation of a decision-making system that will respond to requests for data analysis, to have reliable information as soon as possible and to improve the reporting process.

In this work, we will design and implement a Data Warehouse oriented sales and marketing process. And then we'll make a dashboard for the analysis and decision support.

Keywords:

Decision System, Data Warehouse, Data Mart, ETL, KPI, dimensional modeling, dashboard.

Remerciement

Tout d'abord, nous remercions le bon Dieu, notre créateur de nous avoir donné la force, la volonté et le courage afin d'accomplir ce modeste travail.

Nous adressons le grand remerciement à notre promotrice pour ses conseils et ses orientations, ainsi qu'à notre encadreur, pour les conseils qu'elle nous a prodigué, pour la patience, la confiance qu'elle nous a témoigné et qui ont été déterminants dans la réalisation de notre travail.

Nous tenons à exprimer notre profonde gratitude à tous nos enseignants durant les années des études qui ont contribué à notre formation.

Nos remerciements s'étendent également aux messieurs les membres de jury pour l'honneur qu'ils ont fait en acceptant de juger notre travail.

Ces remerciements ne seraient pas complets sans nos pensées pour nos familles, pour nous avoir soutenues et nous avoir permis de réaliser nos études dans les meilleures conditions.

Finalement, nous tenons à remercier toute personne qui a participé de près ou de loin pour la réalisation de ce projet.

Dédicace

J'ai l'immense honneur de dédier ce modeste travail à mes très chers parents qui ont toujours été là pour moi, et qui m'ont donné un magnifique modèle de labeur et de persévérance. J'espère qu'ils trouveront dans ce travail toute ma reconnaissance et tout mon amour.

A mes frères Khaled et Djamel Eddine ainsi qu'à leurs épouses Asma et Fazia

A ma sœur Zahida.

Je le dédie particulièrement à mes grands-parents et mes chères tantes à qui je souhaite une longue et heureuse vie.

A toute ma famille et plus particulièrement mes cousins Anis, Foued, Nassim et Amine.

A ma binôme Nawel et toutes mes amies avec qui j'ai partagé tant de bons moments.

A mon encadreur Mme Belli Hassiba pour son aide si précieuse

Sa patience et sa disponibilité.

Benkali Hakima

Dédicace

Je dédie ce modeste travail à :

Mon très cher père, ma fierté, pour son précieux soutien, pour son encouragement, pour avoir cru en moi, pour son sourire réconfortant.

Ma très chère mère, mon inspiratrice depuis toujours, pour m'avoir tendu chaleureusement ses bras, pour avoir évincé mes moments de doute, pour ses prières et sa patience.

Toute ma famille paternelle BOUFRIDI et maternelle HAKEM.

Mon binôme Hakima ainsi qu'à toute sa famille.

Mes cher(e)s ami(e)s qui m'ont beaucoup aidé.

Tous ceux qui me sont cher(e)s et que je n'ai pas cité ici

BOUFRIDI Nawel

Table des matières

ملخص.....	2
Résumé.....	3
Abstract.....	4
Remerciements.....	5
Dédicace.....	6
Dédicace.....	7
Liste des figures.....	12
Liste des tableaux.....	13
Liste des abréviations.....	14
Introduction générale.....	15
1. Contexte général.....	17
2. Problématique.....	17
3. Objectifs.....	18
4. Organisation du mémoire.....	18
Chapitre I : Synthèse bibliographique.....	20
Introduction.....	21
1. Apparition de la Business Intelligence.....	21
1.1. Définition de la BI.....	21
1.2. Objectifs de la BI.....	21
1.3. Comparaison entre les systèmes transactionnels et les systèmes décisionnels.....	22
2. Les entrepôts de données.....	23
2.1. Historique des Data Warehouse.....	24
2.2. Définition du Data Warehouse.....	25
2.3. Architecture d'un entrepôt de donnée.....	25
2.3.1. Système source.....	26
2.3.2. Zone de préparation des données.....	26
2.3.3. Zone de présentation.....	26
2.3.4. Outils d'accès.....	26
2.4. La stratégie de conception d'un entrepôt de données.....	26
2.5. Le processus d'alimentation E.T.L.....	28
2.5.1. Extraction des données.....	28
2.5.2. Transformation des données.....	29

2.5.3. Chargement des données.....	29
3. La modélisation des entrepôts de données.....	29
3.1.Définition de la modélisation dimensionnelle.....	29
3.2.Les composants du schéma dimensionnel.....	30
a) Table de faits et les mesures.....	30
b) Grain.....	30
c) Table de dimension.....	30
d) Hiérarchie.....	31
e) Le cube.....	31
3.3.Schémas de modélisation.....	32
a) Schéma en étoile.....	32
b) Schéma en flocon de neige.....	32
c) Schéma en constellation.....	33
4. Le concept OLAP.....	33
4.1.Définition de la technologie OLAP.....	34
4.2.Quelques operateurs d'analyse.....	34
4.3.Architecture des serveurs OLAP.....	36
4.3.1. ROLAP.....	36
4.3.2. MOLAP.....	37
4.3.3. HOLAP.....	37
5. Les outils de restitutions.....	38
5.1.Définitions des outils de restitutions.....	38
5.2.Quelques outils de restitutions.....	38
5.2.1. Reporting.....	39
5.2.1.1.Définition du reporting.....	39
5.2.1.2.Objectifs du reporting.....	39
5.2.2. Les tableaux de bord.....	39
5.2.2.1.Définition du tableau de bord.....	39
5.2.2.2.Indicateurs de performance.....	40
5.2.2.3.Démarche de construction d'un tableau de bord.....	41
5.2.2.4.Objectifs du tableau de bord.....	41
5.2.3. Datamining.....	41
5.2.3.1. Définition du Datamining.....	41
5.2.3.2.Objectifs des outils de Datamining.....	42

Conclusion	42
Chapitre II : Etude de cas de l'entreprise	43
Introduction	44
1. Présentation de GEMA.....	44
1.1.Les sociétés filiales et participations de GEMA.....	45
1.2.Statut juridique.....	45
1.3.Domaine d'activité de GEMA.....	45
1.4.GEMA et le marché de consignation.....	46
1.5.Certification de GEMA.....	47
2. Etude de l'existant.....	47
2.1.La dotation d'investissement informatique.....	47
2.2.La dotation d'applications intranet.....	48
2.3.La dotation de moyens de communication.....	48
2.3.1. En interne.....	48
2.3.2. Avec la douane.....	49
2.3.3. Avec les entreprises portuaires.....	49
2.4.L'existant décisionnel.....	49
2.5.Critiques.....	52
Conclusion	52
Chapitre III : Conception de la solution	53
Introduction	54
1. Conception de la solution.....	54
1.1.Analyse des besoins.....	55
1.2.Conception du Data Warehouse.....	55
1.2.1. Modélisation de notre solution.....	56
a) Le fait physique.....	57
b) Le fait compte d'escale réalisé.....	59
c) Le fait BIC.....	60
d) Le fait A40.....	62
e) Le fait transfert.....	63
f) Le fait nouveaux armateurs.....	65
g) Le fait réalisation des clients.....	66
h) Le fait pavillon.....	67
i) Le fait créances armateurs.....	69

1.2.2. Matrice de bus.....	71
1.2.3. Détails des dimensions du système dimensionnel.....	72
1.2.4. Sources du système.....	73
1.3. Conception de l'ETL.....	73
1.3.1. ETL sur les dimensions.....	74
1.3.2. ETL sur les tables de faits.....	74
1.3.3. Traitements effectués.....	74
1.3.4. La périodicité de chargement de données.....	74
1.4. Conception des cubes dimensionnels.....	75
Conclusion.....	75
Chapitre IV : Réalisation de système.....	76
Introduction.....	77
1. Moyens matériels.....	77
2. Moyens logiciels.....	77
3. Architecture technique.....	78
4. Réalisation de la solution.....	79
a. Construction de la zone d'entreposage.....	79
b. Chargement des données.....	79
5. Scénario d'utilisation.....	80
6. La sécurité du système.....	82
Conclusion.....	82
Conclusion générale.....	86
Bibliographie.....	87
Annexes.....	89

Liste des tableaux

Tableau 1 : Comparaison entre le système opérationnel et le système décisionnel.....	23
Tableau 2 : Approches d'implémentation d'un DW.....	28
Tableau 3 : Avantages et inconvénients des modèles en étoile et en flocon.....	34
Tableau 4 : L'évolution du chiffre d'affaire de GEMA.....	48
Tableau 5 : Composition de la table de fait « physique ».....	60
Tableau 6 : Composition de la table de fait « Compte_Escale_Realise».....	61
Tableau 7 : Composition de la table de faits « BIC ».....	63
Tableau 8 : Composition de la table de faits « A40».....	64
Tableau 9 : Composition de la table de faits « transfert ».....	66
Tableau 10 : Composition de la table de fait « Nouveaux_Armateur».....	67
Tableau 11 : Composition de la table de faits « realisationClient ».....	69
Tableau 12 : Composition de la table de faits « pavillon».....	70
Tableau 13 : Composition de la table de faits « creance_Armateur ».....	71
Tableau 14 : Matrice en bus du système dw_gemahouse.....	73
Tableau 15 : Tableau récapitulatif des dimensions et leurs attributs.....	75

Liste des abréviations

ALC : Algérie Ligabue Catering

AVICAT : Avitaillement et Catering des navires

BI : Business Intelligence

CE : Compte d'Escale

CNIS : Centre National de l'Informatique et Statistique

DCM : Direction Commerciale et Marketing

DOSI : Direction de l'Organisation des Systèmes d'Information

DW : Data Warehouse

EDI : Echange de Données Informatisées

ETL : Extraction, Transformation, Loading

FILTRANS : Filiale Transite

FN : Forme Normale

GATMA : Groupe Algérien de Transport Maritime

GEMA : Générale Maritime

HOLAP: Hybrid On-Line Analytical Processing

KPI: Key Performance Indicator

OLAP : On-Line Analytical Processing

OLTP : On-Line Transactional Processing

MOLAP: Multidimensional On-Line Analytical Processing

PGI: Progiciel de Gestion Intégré

RMQ : Responsable Management Qualité

ROLAP: Relational On-Line Analytical Processing

RTC : Réseau téléphonique commuté

SGBD : Système de Gestion de Base de Données

SGBDR : Système de Gestion de Base de Données Relationnelle

SGP : Société de Gestion des Participations

SIGAD : Système d'Information et de Gestion Automatisée des Douanes

SMQ : Système de Management de la Qualité

SOGRAL : Société d'Exploitation de la Gare Routière d'Alger

SPA : Société Par Action.

SQL : Structured Query Language

TD : Table de Dimension

TF : Table de Fait

Introduction Générale

1. Contexte général

C'est dans un environnement fortement complexe et hautement concurrentiel qu'évolue la majeure partie des entreprises. Ce climat de forte concurrence exige de ces entreprises une surveillance très étroite du marché afin de ne pas se laisser écarter par les concurrents et cela en répondant, le plus rapidement possible, aux préoccupations du marché, de leurs clientèles et de leurs partenaires.

Dans ce cadre, les décideurs d'une entreprise ont pour objectif de gérer le présent et l'avenir tout en optimisant les ressources, aussi bien matérielles qu'humaines, dont ils disposent.

Ainsi, ils doivent prendre des décisions, de plus en plus rapides et pertinentes, basées sur une compréhension claire de leurs environnements interne et externe qui deviennent une arme décisive. Donc le système décisionnel est tenu de fournir la bonne information au bon moment et à la bonne personne.

Les années 80 ont vu la naissance et le développement des systèmes d'informations opérationnels permettant une gestion efficace des flux d'informations. Cependant, ces systèmes d'informations n'ont pas pour finalité d'être un outil décisionnel au sein de l'entreprise. D'où, la notion de l'informatique décisionnelle (Business Intelligence) est apparue. Il s'agit d'un ensemble d'outils et de méthodes permettant d'exploiter au mieux les données circulants dans les systèmes d'informations opérationnels au sein de l'entreprise. [1]

Autrement dit, l'informatique décisionnelle permet de répondre aux préoccupations d'ordre décisionnel et stratégique de l'entreprise, en évitant ainsi une prise de décision aléatoire.

1. Problématique

GEMA est une grande entreprise spécialisée dans le domaine de la consignation des navires et des conteneurs. Elle a pour missions principales :

- La mise en place d'une stratégie commerciale adéquate avec l'évolution du marché ;
- La mise en place d'une stratégie de prospection clients pour augmenter le chiffre d'affaire et fidéliser ses clients.

Sur le plan système d'information automatisé, GEMA est dotée d'un système opérationnel intégré de consignation et de comptabilité ainsi qu'un réseau intranet. Ce dernier permet de consolider en temps réel, les données physiques des navires et les données commerciales financières des clients. Néanmoins, ce système consolidé ne permet pas la prise de la bonne décision à cause des contraintes suivantes :

- Inexistence d'indicateurs pertinents définis pour la prise de décision et qui sont générés directement à partir des données consolidées;
- Difficulté dans la définition des indicateurs de mesure de performance commerciaux ;
- Retard dans l'élaboration des rapports d'activités, en vue du volume important des informations manipulées (leur élaboration peut aller jusqu'à 4 mois) ;
- Prise de décision inadéquate avec la stratégie commerciale de l'entreprise.

Cependant, la mise en place d'un système décisionnel commercial est jugée nécessaire pour la prise de la bonne décision au moment propice.

2. Objectifs du projet

Afin de pallier aux problèmes précédemment cités, GEMA nous a mandaté de mettre en place un système d'information décisionnel. Ce dernier permet d'exploiter les données du système consolidé pour améliorer le processus commercial de prise de décision au sein de l'entreprise.

Les objectifs attendus de notre travail sont :

- ❖ Mettre en place des indicateurs de performance pertinents permettant de mesurer le fonctionnement du processus commercial ;
- ❖ Réduire la durée d'obtention d'informations significatives et nécessaires à la prise de décision ;
- ❖ Offrir un outil d'analyse de données aux décideurs avec la possibilité de générer des rapports à la demande en temps réel;
- ❖ Renforcer la fiabilité des informations sur lesquelles se base la prise de décisions ;
- ❖ Centraliser les données afin de faciliter l'accès aux informations, sans négliger l'aspect sécurité ;
- ❖ Réaliser un tableau de bord pour la représentation des données de l'entrepôt ;
- ❖ Intégrer la solution dans l'intranet de l'entreprise.

3. Organisation du mémoire

Nous avons organisé notre mémoire en quatre chapitres principaux:

- Chapitre I: synthèse bibliographique

Cette partie introduit les notions de base des systèmes décisionnels, avec une comparaison entre ces derniers et les systèmes transactionnels. Nous présentons par la suite les entrepôts de

données, l'un des outils qui permettent de mettre en place un système décisionnel dans une entreprise ; nous parlons après du concept OLAP en présentant ses opérations et ses objectifs. Nous passons par la suite à la présentation des différents outils qui permettent la modélisation des entrepôts de données. Et nous finissons par aborder quelques outils de restitution les plus répandus : le reporting, le tableau de bord ainsi que le data mining.

- **Chapitre II: étude de cas de l'entreprise**

Dans cette partie nous présentons notre organisme d'accueil GEMA avec une étude de l'existant, cotés technique et décisionnel, les applications et systèmes utilisés actuellement dans l'entreprise, dont les systèmes qui constitueront la source de notre système à concevoir, suivi d'une critique des points négatifs repérés afin d'atteindre les objectifs.

- **Chapitre III : conception de la solution**

Dans cette partie nous détaillons la démarche suivie pour la conception de notre système, nous commençons d'abord par une analyse des besoins afin de tirer les besoins des décideurs de GEMA pour pouvoir passer à la conception du système à réaliser.

- **Chapitre IV : réalisation et déploiement du système**

Ce dernier chapitre détaille l'implémentation de notre solution, il contient le résultat de notre implémentation, avec la mise en évidence les raisons de nos choix techniques, ainsi qu'une brève présentation de quelques interfaces réalisées.

Chapitre I : synthèse bibliographique

Introduction

Dans ce chapitre, nous allons aborder l'essentiel à connaître sur les systèmes décisionnels (apparition, définitions, objectifs...), les datawarehouse (historique, définition...) ainsi que la modélisation dimensionnelle et les outils de restitutions.

1. Apparition de la Business Intelligence (BI)

Aujourd'hui les sources d'informations sont éparpillées, volumineuses et hétérogènes or ceci constitue la principale difficulté que les décideurs de l'entreprise doivent affronter pour le pilotage et la prise de décisions. La Business Intelligence (informatique décisionnelle) est apparue afin de transformer les données brutes en un système organisé en mettant en œuvre divers outils et méthodes qui permettent d'avoir une vue globale de l'entreprise et de fournir, de manière optimisée, les informations nécessaires à la prise de décision. [2]

1.1. Définition de la BI

Nombreuses sont les définitions du terme business intelligence :

Selon Fernandez : « [...] La Business Intelligence désigne un ensemble d'outils technologiques, méthodiquement assemblés, et déployés en parfaite cohérence avec la stratégie d'entreprise préalablement élaborée. La BI n'a d'autre finalité que de délivrer les informations pertinentes à chaque manager afin qu'il puisse prendre le plus efficacement possible les meilleures décisions selon son contexte d'action, ses prérogatives et ses objectifs tactiques et stratégiques. » [1].

En d'autre terme, c'est l'exploitation des données de l'entreprise au moyen d'outils et de méthodes ainsi qu'une bonne stratégie qui permettent l'extraction, la restitution et la présentation des données pertinentes en vue de faciliter la prise de décisions. C'est-à-dire faciliter la compréhension du fonctionnement actuel de l'entreprise pour permettre l'anticipation des actions dans le but de faciliter le pilotage. L'informatique décisionnelle est donc dédiée aux décideurs et managers de l'entreprise.

1.2. Objectifs de la BI [3]

Le principal objectif de la BI est de s'assurer que les décisions prises soient d'une certitude absolue, pour permettre cela elle doit aussi assurer les objectifs suivants :

- Présenter l'information de manière cohérente : l'objectif est de fournir aux entreprises une vision à jour et complète de leur activité.
- Aider à améliorer la prise de décision : l'élimination des intuitions lors du pilotage de l'entreprise en ce qu'il s'agit des actions futures. Les outils de BI permettent d'avoir des informations historiques, des synthèses, des prédictions liées à plusieurs scénarios, ce qui permet aux décisions d'être prises sur la base de faits plutôt que sur des intuitions.
- Analyser et obtenir des réponses aux questions métiers dans un laps de temps réduit.
- Consolider les données issues de plusieurs départements ce qui améliore l'efficacité.
- Contrôler l'accès aux informations confidentielles.

1.3. Comparaison entre les systèmes transactionnels et les systèmes décisionnels

Les systèmes « opérationnels » ou « de gestion », sont dédiés aux métiers de l'entreprise pour les assister dans leurs tâches de gestion quotidiennes et donc directement opérationnels. La tendance est l'utilisation de P.G.I. (progiciels de gestion intégrée) qui regroupent tous les logiciels de gestion de l'entreprise (finances, ressources humaines, logistique, ventes, etc.) en un unique progiciel paramétrable aux règles de l'entreprise, organisé autour d'une base de données, réduisant ainsi les coûts de communications entre applications. [4]

Les systèmes « décisionnels » sont dédiés aux managers de l'entreprise pour les aider au pilotage de leurs activités. Ils offrent aux décideurs une vision transversale de l'entreprise. La tendance pour réaliser un système décisionnel est à la mise en place d'un entrepôt de données [4].

En d'autres termes, le système décisionnel gère des données analytiques et veille à assurer la disponibilité des informations nécessaires à la prise de décisions, contrairement aux systèmes transactionnels qui s'occupent de la gestion des transactions quotidiennes de l'entreprise.

Bien que les systèmes transactionnels et décisionnels aient le point commun de regrouper les données de l'entreprise est d'en fournir l'accès aux utilisateurs, ils présentent des différences, présentées dans le tableau ci-après :

Critères de comparaison	Systèmes transactionnels	Systèmes décisionnels
Portée de l'interaction utilisateur	Transaction	Base de données entière
Quantité de données affectées par utilisateur	Enregistrements individuels	Groupes d'enregistrements
Temps de réponses	Secondes	Secondes à minutes
Mode d'utilisation	Stable	Dynamique
Natures des données	Primitive (au plus bas niveau de détails)	Agrégées
Mode d'accès a la base de données	Accès prédéfinie	Accès indéfinis ou dynamique
Volatilité des données	Elevées	Faible
Priorités	Hautes performance, grandes disponibilité	Grande souplesse, grande autonomie

Tableau 1 : Comparaison entre le système opérationnel et le système décisionnel [5]

A la lumière de ce tableau nous pouvons constater les principales différences entre les systèmes transactionnels et les systèmes décisionnels, cependant nous ne pouvons juger ni la performance de ces derniers ni d'affirmer que l'un est meilleur que l'autre car un système décisionnel doit avoir son système transactionnel représentant la source qui à partir de laquelle sont construites les informations utiles à la prises de décisions.

2. Les entrepôts de données

Le but de la BI est d'aider à la prise de la bonne décision et de permettre des analyses précises, complexes et de grande envergure et cela a partir des sources de données c'est

pourquoi les décideurs ne doivent pas avoir qu'une vue verticale que fournissent les systèmes transactionnels mais aussi une vue transversale et cela grâce aux systèmes décisionnels.

La question qui se pose c'est : comment organiser ces différentes données dans un ensemble cohérent afin de procéder à toutes les analyses nécessaires pour construire les indicateurs indispensables au pilotage de l'entreprise ? [5]

La réponse à la question précédemment posée est l'utilisation d'un processus d'entreposage de données (datawarehousing)

2.1. Historique des Data Warehouse [6]

L'idée de constituer une base de données orientée sujet, intégrée, contenant des informations datées, non volatiles et exclusivement destinées aux processus d'aide à la décision fut dans un premier temps accueillie avec une certaine perplexité. Beaucoup n'y voyaient que l'habillage d'un concept déjà ancien : l'infocentre. Une copie des données du système opérationnel était prise de façon périodique, et stockée dans une nouvelle base de données. Avec pour objectif de répondre aux besoins des décideurs sans altérer le bon fonctionnement du système opérationnel.

Cependant durant les années 80, un intérêt croissant pour les systèmes décisionnels a vu le jour grâce à l'émergence des SGBDR, la simplicité du modèle relationnel et la puissance offerte par le langage SQL.

Le Data Warehouse est devenu alors une nouvelle source d'information, alimentée avec des données recueillies et consolidées des différentes sources internes et externes.

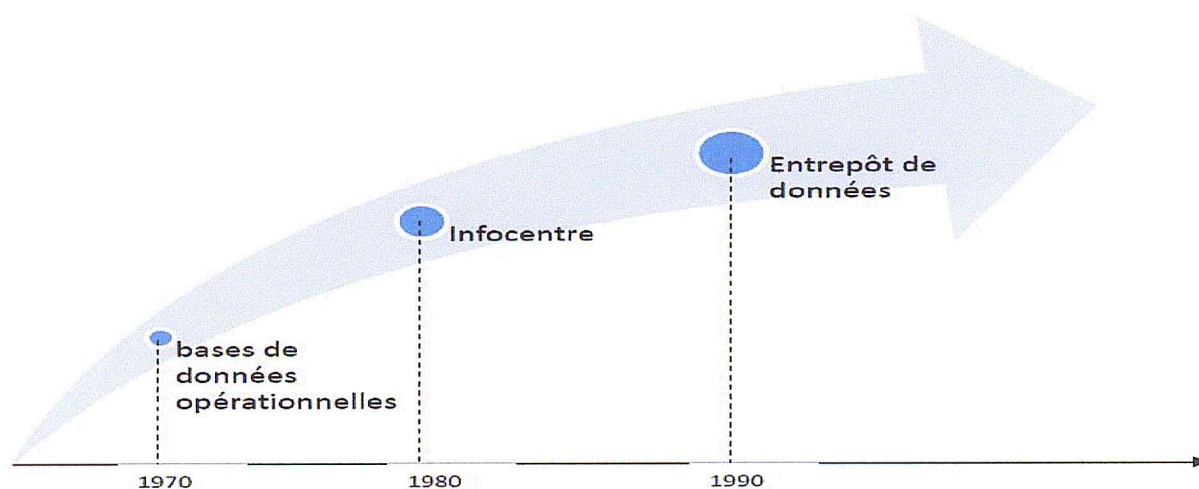


Figure 1: Evolution des bases de données décisionnelles. [6]

2.2. Définition du Data Warehouse

D'après Bill Inmon¹ qui est considéré comme le père des Data Warehouse « Le data Warehouse est une collection de données orientées sujet, intégrées, non volatiles et historiées, organisées pour le support d'un processus d'aide à la décision » [7]

- Les DW sont **orientées sujet**: les données doivent être orienté métier et donc triées par thème ;
- Les DW sont composées de données **intégrées** : un traitement préalable est nécessaire dans un souci de rationalisation et de normalisation ;
- Les données du DW sont **non volatiles**: une donnée entrée dans un entrepôt de données ne doit pas être supprimé ;
- Les données du DW sont **historiées** est donc datées ;
- les données du DW sont **organisées** de manière à permettre l'exécution des processus d'aide à la décision (Reporting, Data Mining...).

A partir de cette définition on peut déduire qu'un entrepôt de données est une structure informatique (support physique) dans laquelle un gros volume de données consolidées (sélectionnées, préparées, et datées) est centralisé à partir de différentes sources de données de l'entreprise et qui est conçu de sorte que les décideurs aient un accès rapide et optimal aux informations stratégique nécessaire à la prise de décision et au pilotage.

2.3. Architecture d'un entrepôt de donnée [8] :

La figure ci-dessous donne une vue globale sur l'architecture d'un entrepôt de donné :

¹ **Bill Inmon** : reconnu comme le père des DW.

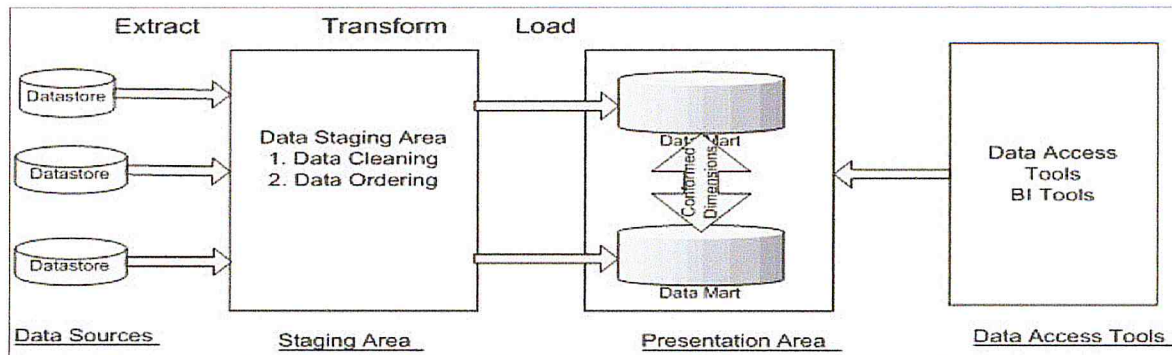


Figure 2: Architecture d'un entrepôt de donnée. [9]

2.3.1. Système source :

Cette zone comprend les différentes sources de données qui alimenteront par la suite le Data Warehouse. Elles proviennent du système opérationnel et des applications liées à l'activité, ces sources peuvent s'agir des bases de données transactionnelles (le plus souvent), des fichiers textes ou tabulés...etc. C'est à partir de ce système source que l'on effectuera l'extraction des données.

2.3.2. Zone de préparation des données :

Cette zone comprend les processus qui permettent de préparer les données provenant du système source et de les rendre intégrables et exploitables au sein du Data Warehouse. A savoir :

- Le stockage des fichiers plats des SGBDR et autre
- Le traitement des données, comprend principalement : le nettoyage, la transformation, la combinaison, la suppression des doublons, la mise en conformité, l'archivage, et l'exportation vers les Data Marts.

2.3.3. Zone de présentation :

Dans cette zone on trouve les sources de données interrogeables de l'entreprise, l'entrepôt de donnée (le Data Warehouse) et les magasins de données (les Data Marts). Les données sont stockées et organisées d'une façon à répondre aux requêtes émises directement par l'utilisateur.

2.3.4. Outils d'accès :

Il s'agit des outils orientés des utilisateurs finaux, à travers lesquels se font l'interaction et l'exploitation des données du Data Warehouse. Ils servent à représenter analytiquement les données de l'entrepôt. Il peut s'agir des outils les plus simples tels que les outils de requêtage, jusqu'à d'autres outils plus sophistiqués, les applications de modélisation et datamining.

2.4. La stratégie de conception d'un entrepôt de données.

Il existe plusieurs approches pour mettre en place un DW. Le tableau ci-après résume les deux approches les plus connues :

	Inmon (top down)	Kimball ² (bottom up)
Schéma [10]		
Principe	Le Datawarehouse constitue un répertoire centralisé pour les données de toute l'entreprise. Les data Marts qui contiennent des données dimensionnelles sont par la suite créés et alimentés par le Datawarehouse.	Les data Marts sont construits en premier pour chaque département de l'entreprise, et sont ensuite rassemblés pour former un Datawarehouse.
Outils	Relationnelle	Dimensionnelle
Avantages	<ul style="list-style-type: none"> • une vue entrepreneuriale des données • coopération de toute l'entreprise pour la réalisation du Data Warehouse • contrôle centralisé • stockage unique de la donnée. 	<ul style="list-style-type: none"> • Implémentation plus simple, plus rapide • moins couteuse • Moins de risque d'échec
Inconvénients	<ul style="list-style-type: none"> • L'implémentation requiert beaucoup de temps, et coût élevé. • Probabilité élevée d'échec • Requiert des compétences transversales 	<ul style="list-style-type: none"> • Chaque data Mart a une vue étroite de la donnée • Possibilité de redondance de la données dans chaque data Mart.

Tableau 2 : Approches d'implémentation d'un DW. [10]

² **Ralph Kimball** : est un informaticien connu comme auteur d'ouvrages concernant les sujets liés au décisionnel.

On peut également trouver l'approche hybride, qui est conseillée par les professionnels de la BI. Elle consiste à faire la conception totale de l'entrepôt de données (i.e. : concevoir toutes les dimensions, tous les faits, toutes les relations), puis créer des divisions plus petites et plus gérables et les mettre en œuvre. Cela équivaut à découper notre conception par éléments en commun et réaliser les découpages un par un.

Cette méthode tire le meilleur des deux précédentes sans avoir les contraintes. Il faut juste noter que cette méthode implique, parfois, des compromis de découpage (dupliquer des dimensions identiques pour des besoins pratiques). [11]

2.5. Le processus d'alimentation E.T.L

L'outil d'alimentation permet de paramétrer des règles de gestion, propres à l'entreprise et à son secteur d'activité. Ces règles visent elles aussi à assurer la cohérence entre les données et à ne stocker dans l'entrepôt de données que des informations préalablement mises en relation les unes avec les autres.

Définition :

« **Extract-Transform-Load** » est connu sous l'acronyme ETL (ou parfois : Datapumping). Il s'agit d'une technologie informatique middleware permettant d'effectuer des synchronisations massives d'information d'une banque de données vers une autre.

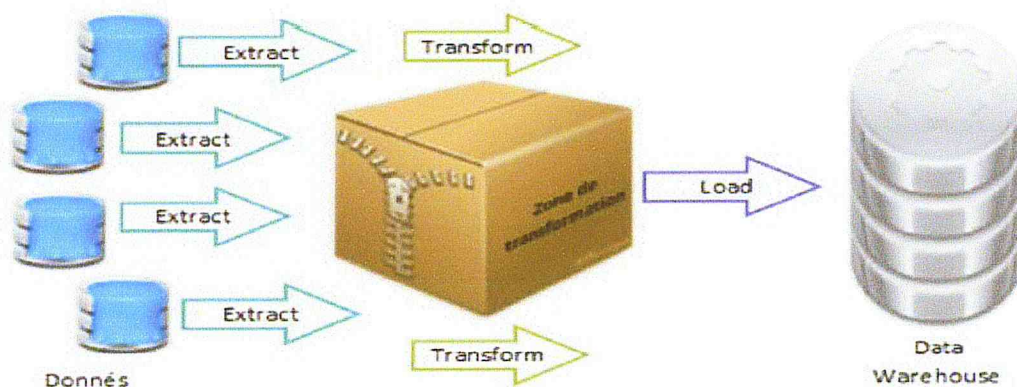


Figure 3: Le système d'alimentation ETL. [12]

2.5.1. Extraction des données [13]

L'extraction des données est la première des étapes des systèmes ETL. Le but de cette étape est la lecture et l'extraction des données du système source. Cette étape s'avère être

critique. En effet, dans le cas où le système source doit fonctionner en permanence (24h/24 et 7 jours sur 7), il faut que l'extraction, opération coûteuse en ressources, doit être faite le plus rapidement possible et souvent durant un laps de temps précis

2.5.2. Transformation des données [13]

Le nettoyage des données est une discipline sur laquelle de nombreux éditeurs travaillent actuellement. Les outils de nettoyage permettent de supprimer les doublons, il s'agit à ce stade d'appliquer des filtres prédéfinis sur les données afin d'attribuer des valeurs cohérentes aux variables mal ou non renseignées ou encore d'harmoniser les formats (exemple pour la date : jj/mm/aaaa). Aussi, des données du système opérationnel doivent être agrégées ou calculées avant leur chargement dans la base décisionnelle.

2.5.3. Chargement des données [13]

Le chargement est la dernière phase de l'alimentation du Data Warehouse. C'est une phase délicate notamment lorsque les volumes sont importants. Pour obtenir de bonnes performances en chargement, il est impératif de maîtriser les structures du SGBD (tables et index) associées aux données chargées afin d'optimiser au mieux ces processus.

3. La modélisation des entrepôts de données

Une fois avoir recensé les différents business process, on pourra commencer à construire le modèle dimensionnel. Pour cela il faudra bien évidemment repérer les tables de faits et celles de dimensions.

3.1. Définition de la modélisation dimensionnelle

La modélisation d'un entrepôt de données est la modélisation multidimensionnelle proposée et popularisée par « **Ralph Kimball** » dans les années 90 « *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* » [14].

Donc, elle consiste à représenter les données comme des points dans un espace multidimensionnel. Les données sont vues comme des sujets d'analyse (**les faits**) étudiés selon plusieurs axes (**les dimensions**). Chaque dimension est liée à un ou plusieurs points de vue définissant ainsi le degré de granularité des données (**hiérarchies**).

Autrement dit, le modèle multidimensionnel se compose de **faits** contenant les **mesures** à analyser et de **dimensions** contenant les **paramètres** de l'analyse.

3.2. Les composants du schéma dimensionnel

Un schéma dimensionnel est composé de :

a) Table de faits et les mesures

La table de faits capture les données qui mesurent les activités de l'organisation. Elle comporte généralement une multitude de lignes, représentant souvent des centaines de millions d'enregistrements lorsqu'elle couvre plusieurs années d'historiques. La caractéristique essentielle de la table de faits c'est de contenir des données numériques « faits » qui peuvent être synthétisées afin de fournir des informations sur l'historique des activités de l'organisation. Chaque table de faits a un index à plusieurs parties qui contient, comme clés étrangères, les clés primaires des tables de dimension associés. Les tables de faits ne doivent pas contenir d'informations descriptives ou des données autres que celles des champs de mesure numérique et des champs d'index correspondants aux tables de dimensions.

Le fait modélise le sujet de l'analyse. Un fait est une « observation du marché », il est formé de mesures correspondant aux informations de l'activité analysée.

Les mesures d'un fait sont numériques et généralement valorisées de manière continue [15], on peut les additionner, les dénombrer ou bien calculer le minimum ou le maximum.

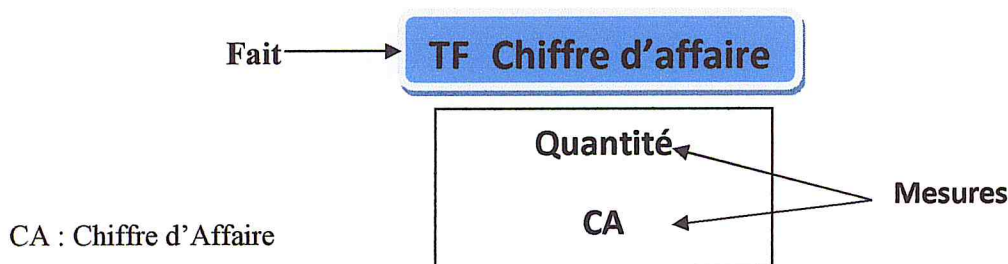


Figure 4 : Exemple de table de faits.

b) Grain

Il désigne la signification d'une ligne dans une table des faits. Les tables dimensions et faits doivent être cohérentes avec le grain défini ce qui crée une cohérence totale dans tous le schéma dimensionnel réalisé. [3]

c) Table de dimension

Les tables de dimensions contiennent le contexte associé aux mesures. Une dimension représente l'axe d'analyse et souvent reflète le critère à base duquel l'information est

recherchée ou analysée. Les tables de dimensions comprennent une multitude d'attributs liés à des caractéristiques métiers. [5]

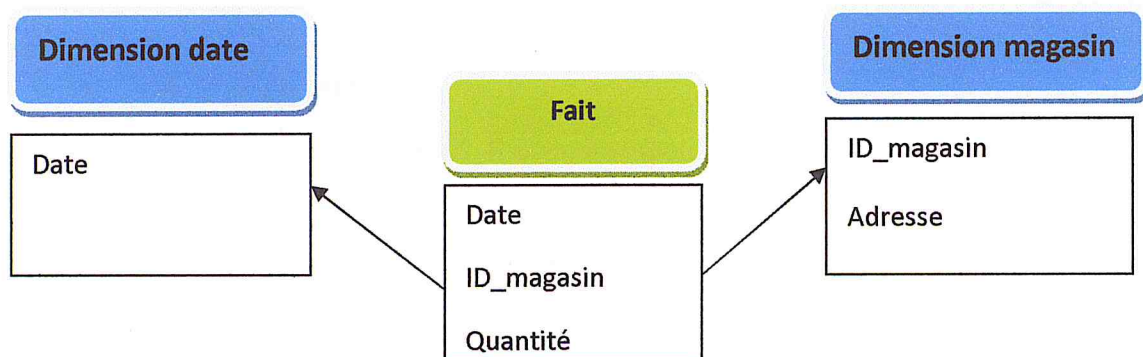


Figure 5 : Exemple de table de dimension.

d) Hiérarchie

C'est une notion qu'on trouve dans les tables de dimension. Une hiérarchie organise les paramètres d'une dimension selon une relation « est-plus fin » conformément à leur niveau de détails. Donc c'est une structuration logique des données selon des niveaux ordonnés. [16]

e) Le cube

La base de données multidimensionnelles s'appuie sur un hyper cube (cube à n dimensions). Les cubes sont définis par un ensemble de dimensions, qui implémentent des axes d'analyse identifiés dans les données et de mesures, qui représentent les quantités numériques à analyser.» [17]

Exemple : Sur le schéma ci-dessous on distingue les dimensions : produit, temps et région. Les faits sont la vente. Pour chaque combinaison des dimensions on peut accéder à la valeur numérique associée au fait vente.

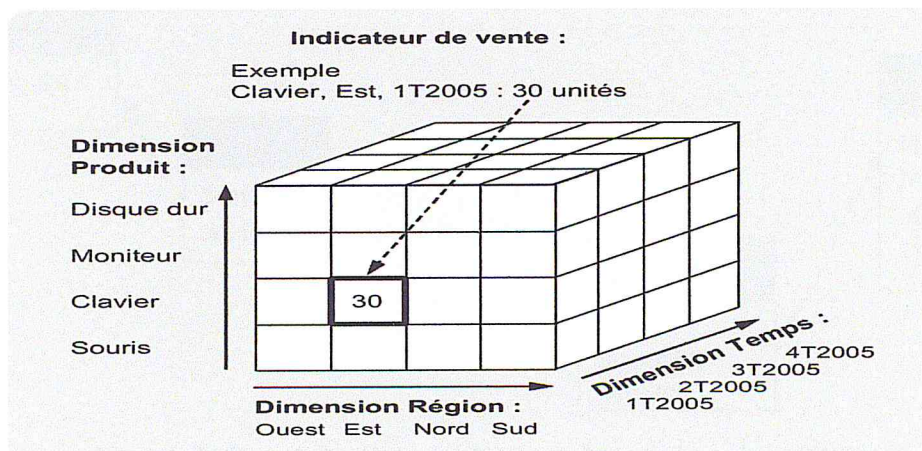


Figure 6: Exemple d'un hyper-cube. [18]

- Avantages et inconvénients des schémas en étoile et en flocon

Le tableau ci-dessous présente les avantages et les inconvénients de chacun des deux modèles :

	Modèle en étoile	Modèle en flocon
Avantages	<ul style="list-style-type: none"> - Facilité de navigation ; - Performances : nombre de jointures limité ; - Gestion des agrégats. 	<ul style="list-style-type: none"> - réduction du volume ; - permettre des analyses par pallier (Drill down/Roll up) sur la dimension hiérarchisée ; - Dimensions normalisées (en 3 FN)
Inconvénients	<ul style="list-style-type: none"> - Dimensions dé-normalisées (non en 3 FN) ; - Redondances dans les dimensions ; - Alimentation complexe. 	<ul style="list-style-type: none"> - navigation difficile ; - nombreuses jointures.

Tableau 3 : Avantages et inconvénients des modèles en étoile et en flocon. [5]

c. Schéma en constellation

Il s'agit de fusionner plusieurs modèles en étoile qui utilisent des dimensions communes. Un modèle en constellation comprend donc plusieurs faits. [5]

4. Le concept OLAP

La différence d'objectifs entre les SGBD et les Data Warehouse a suscité l'utilisation de deux systèmes différents :

- OLTP (On Line Transaction Processing) pour les SGBD.
- OLAP (On Line Analytical Processing) pour les Data Warehouse.

Le terme OLAP (On-Line Analytical Processing) désigne une classe de technologie conçue pour l'accès aux données et pour une analyse instantanée de ces dernières, dans le but de répondre aux besoins de Reporting et d'analyse. [14]

- **Permutation (Switch) :** Consiste à inter-changer la position des membres d'une dimension. Cette opération est illustrée dans la figure suivante :

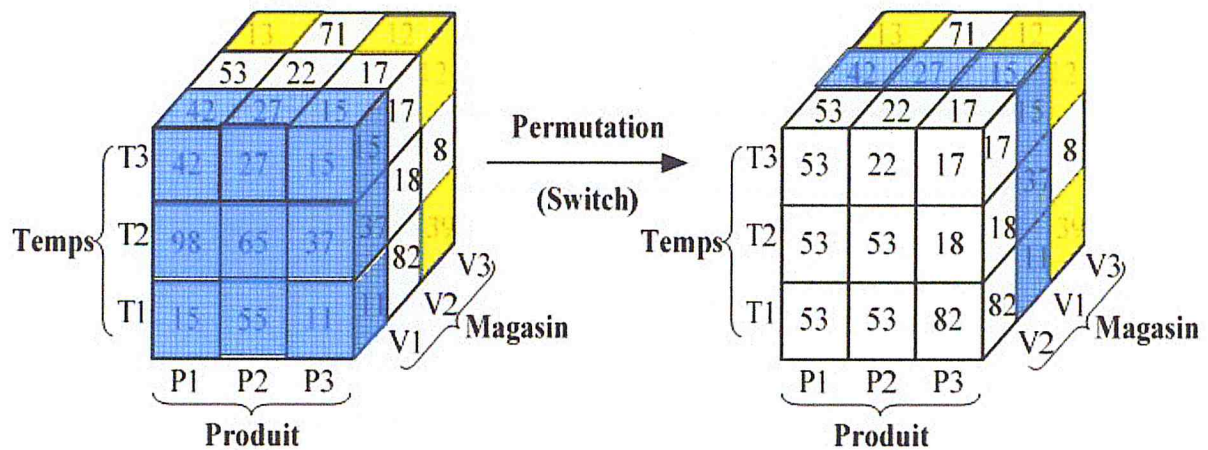


Figure 11 : L'opération Switch.

- **Extraction d'une Tranche du cube (Slice) :** Consiste à sélectionner une dimension, il s'agit de couper une tranche du cube afin d'observer les données de la dimension. Cette opération est illustrée dans la figure suivante :

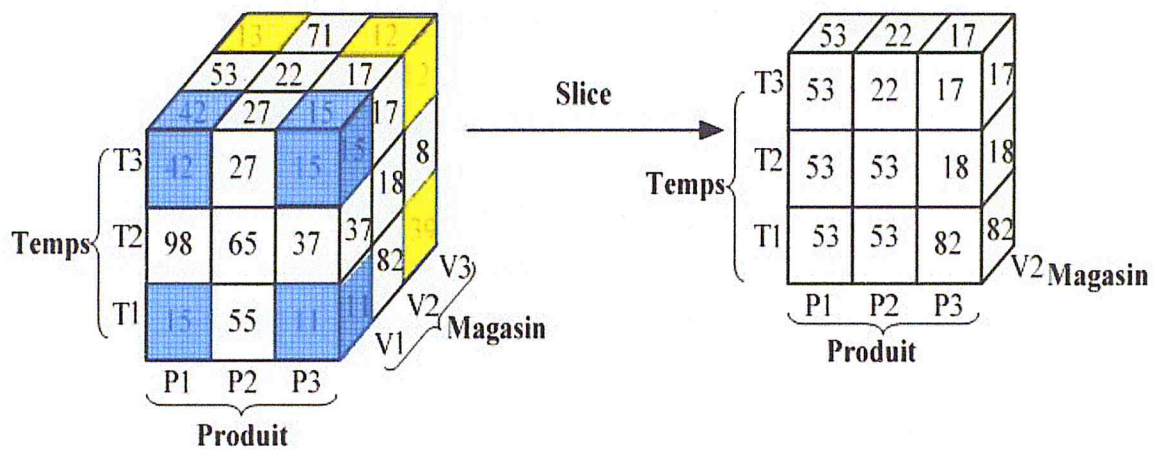


Figure 12 : L'opération Slice.

- **Forage vers le haut (Drill-up ou Roll-up ou Scale-up) :** Consiste à représenter les données du cube à un niveau de granularité supérieur conformément à la hiérarchie définie sur la dimension. Une fonction d'agrégation est utilisée (somme, moyenne, etc.) en paramètre à l'opération, indique comment sont calculés les valeurs du niveau supérieur à partir de celles du niveau inférieur.
- **Forage vers le bas (Drill-down ou Roll-down ou Scale-down) :** Elle fait l'inverse de la précédente (Roll-up), elle consiste à représenter les données du cube à

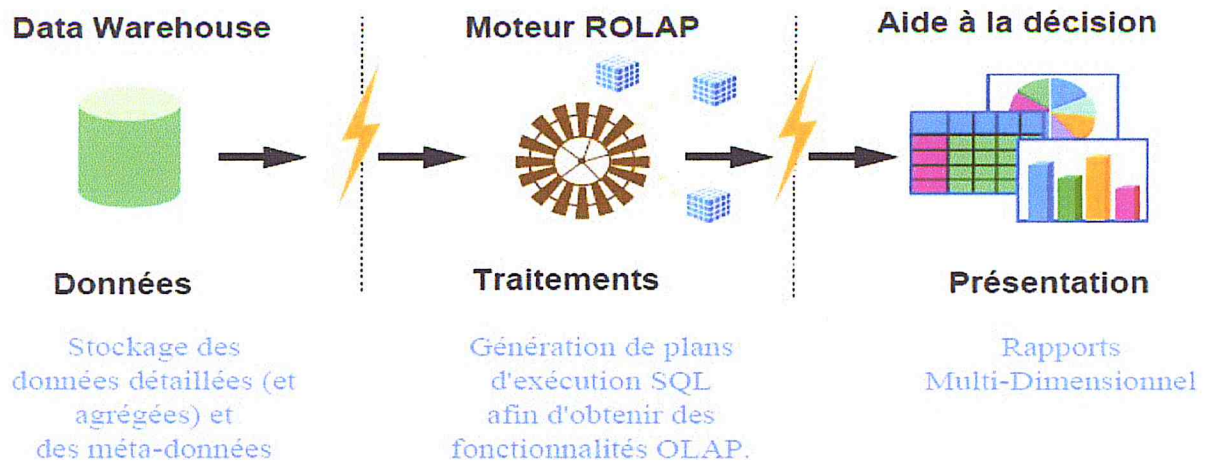


Figure 14: Architecture d'un produit ROLAP. [23]

4.3.2. MOLAP (Multidimensional OLAP) [23]

Cette approche se base sur un modèle structurant les données dans des cubes de données. Un cube de données est composé de cellules qui représentent les mesures (les attributs du fait).

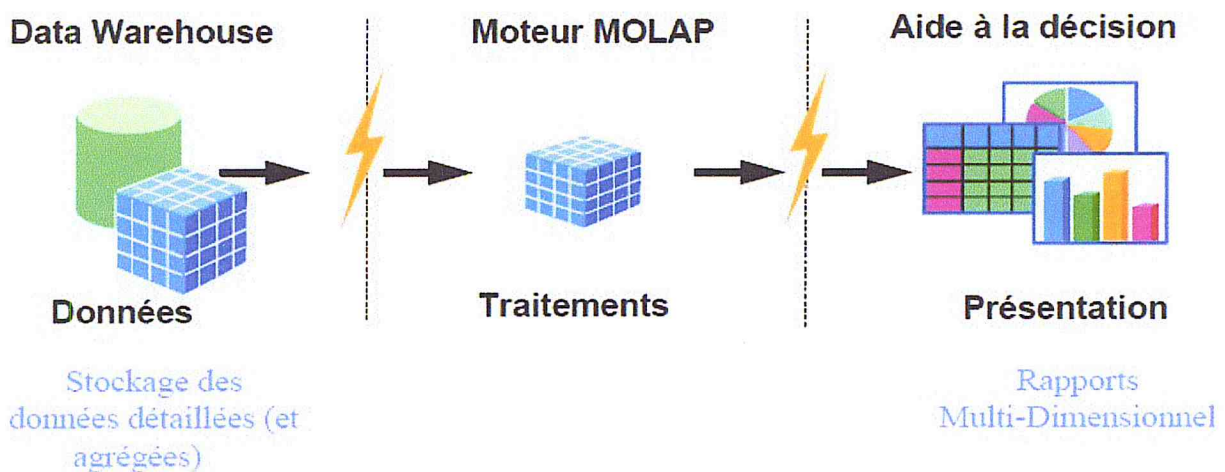


Figure 15 : Exemple d'un produit MOLAP. [23]

4.3.3. HOLAP (Hybrid OLAP) [23]

Ce modèle propose de cumuler les avantages des deux modèles ROLAP et MOLAP. D'un côté garantir la performance en matière d'analyse de la base MOLAP et la capacité du stockage de ROLAP.

Son principe est :

- Stocker les données fréquemment utilisées dans des bases de données multidimensionnelles (cubes).
- Stocker les données non fréquemment utilisées (agrégées) dans des bases de données relationnelles.

un niveau de granularité de niveau inférieur, donc sous une forme plus détaillée (descendre dans la hiérarchie d'une dimension).

Ces deux dernières opérations sont illustrées dans la figure suivante :

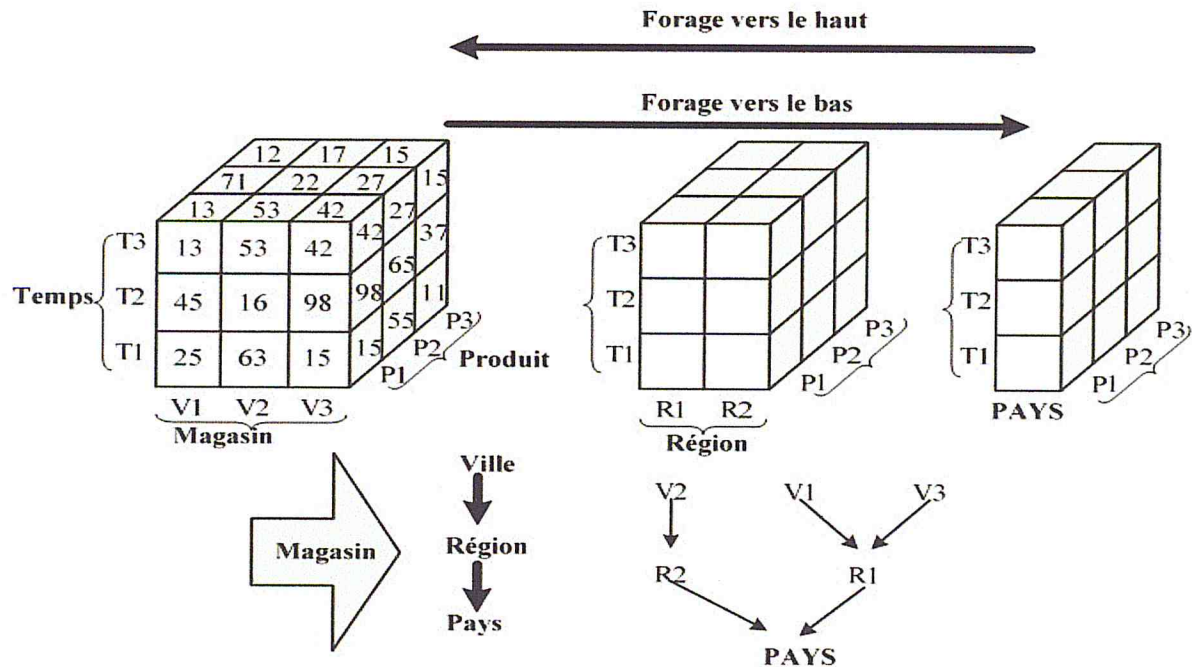


Figure 13: Application du forage vers le bas et vers le haut sur le cube "Ventes". [21]

4.3. Architecture des serveurs OLAP

Lors de la création d'un cube OLAP, il faut décider de la manière de stocker les données automatiques ainsi que les agrégats (regroupements). Donc l'architecture OLAP définit la façon dont seront stockées physiquement les données pour permettre des analyses multi dimensionnelles. [22]

Il y a plusieurs versions d'OLAP, mais les trois les plus répandues sont :

4.3.1. ROLAP (Relational OLAP) [23]

Le modèle ROLAP initialement proposé par Kimball et Ross en 2002. C'est l'approche la plus utilisée. Elle consiste à utiliser un système de gestion de base de données relationnelle. Le modèle multidimensionnel est traduit comme suit :

- Toute dimension est transformée en table.
 - Tout fait est transformé en relation où :
- La clé primaire est la concaténation des clés étrangères référençant les dimensions ou un sous ensemble de ces clés étrangères.
 - Les attributs sont les mesures et les clés étrangères. [15]

5. Les outils de restitutions

Le monde aujourd'hui est interconnecté d'une façon incroyablement complexe, ce qui rend crucial le choix d'un outil approprié en analyse décisionnelle. Les entreprises, dans un effort louable de vouloir aller de l'avant, collectent de grands volumes de données sur la démographie et la fidélité de leurs clients ainsi que le comportement de ses concurrents.

Malheureusement, ces données resteront inexploitable pour les prises de décision, alors que c'était leur objectif premier. Afin de pouvoir donner un sens à toutes leurs données, les entreprises ont besoin de sélectionner les bons outils qui pourront les collecter, les traiter et les présenter de façon utile et dans des délais courts.

5.1. Définitions des outils de restitutions

Le principal objectif de l'entrepôt de données est de rendre les informations aussi accessibles que possible afin que chaque utilisateur final obtienne les renseignements qui l'intéressent. Pour ce faire, il est nécessaire de construire une couche entre les utilisateurs et les informations afin de les aider à trouver ce qu'ils recherchent.

Donc les outils de restitutions ou les outils d'analyse sont la partie publique d'un entrepôt de données [14]. Ils représentent ce que voient les utilisateurs, les outils avec lesquels ils travaillent au quotidien. Sous ce terme sont regroupées toutes les applications qui s'appuient sur les données du Data Warehouse pour les restituer aux utilisateurs finaux.

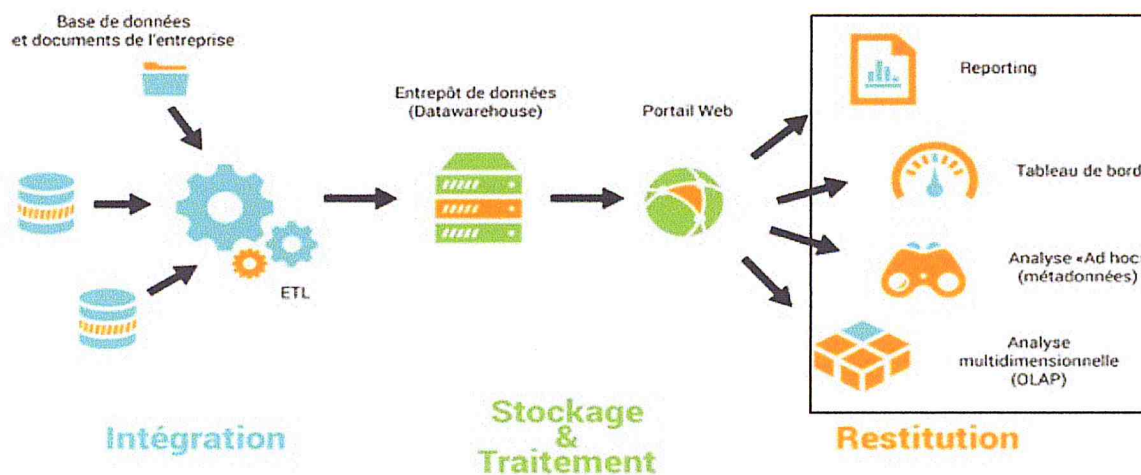


Figure 16 : Exemples de quelques outils de restitutions. [20]

5.2. Quelques outils de restitutions

Les outils de restitution constituent le bout de la chaîne décisionnelle où les utilisateurs exploitent les informations pertinentes représentées et restituées du Data Warehouse pour prendre des décisions, nous allons citer quelques outils de restitutions :

5.2.1. Reporting

5.2.1.1. Définition du Reporting

Alain Fernandez définit le reporting comme suit : « *le reporting est destiné à assurer la réalisation, la publication et la diffusion de rapports d'activité selon un format prédéterminé. Ils sont essentiellement destinés à faciliter la communication de résultats ou d'un suivi d'avancement* ». [1] C'est alors la préparation des rapports d'activité de l'entreprise qui reflète le degré de réalisation des objectifs et qui donne une vue globale sur une activité, un service ou sur l'ensemble de l'entreprise.

Les outils du reporting permettent de formuler des requêtes SQL préétablies pour avoir une représentation graphique et/ou un tableau, qui permettront alors d'élaborer un rapport qui convient à la demande.

5.2.1.2. Objectifs du Reporting

Le but de ces rapports et bilans réguliers est de faire un point ponctuel sur la stratégie de l'entreprise et ainsi permettre d'évaluer les moyens mis en œuvre [24]. De plus, ils fournissent également une aide à la décision pour les choix stratégiques et économiques de l'entreprise.

5.2.2. Les tableaux de bord [25]

5.2.2.1. Définition du tableau de bord

Le Tableau de Bord est un instrument d'aide à la décision, très largement utilisé dans le domaine de la gestion des entreprises. Il présente aux décideurs des indicateurs significatifs sur la qualité des soins et des services et leur permet de prendre des décisions pertinentes.

On peut le définir aussi comme :

- ❖ Outil de communication entre responsables.
- ❖ Outil de prise de décision.
- ❖ Outil de motivation pour les collaborateurs.
- ❖ Instrument de veille permettant de déceler les opportunités et les risques.

La structure et le contenu du tableau de bord dépendent de :

- ❖ L'entreprise : son activité, sa taille ;
- ❖ L'environnement : marché, concurrence ;
- ❖ Le responsable auquel il s'adresse : son champ de responsabilité et d'intervention.



Figure 17 : Exemple d'un tableau de bord-JasperSoft-

Un tableau de bord utile contient généralement 4 éléments essentiels : [25]

- ❖ Un tableau rassemblant les indicateurs pertinents ;
- ❖ Un graphique pour présenter l'information la plus représentative des données du tableau ;
- ❖ Un commentaire clair, précis et concis donnant des indications sur les actions achevées, en cours, et à venir ;
- ❖ Un encart de références avec les coordonnées de l'émetteur et les sources utilisées.

5.2.2.2. Indicateurs de performance

Les indicateurs de performance représentent un élément très important d'un tableau de bord. « Les indicateurs de performance représentent un ensemble de mesures se concentrant sur ceux de la performance organisationnelle qui sont les plus critiques pour la réussite actuelle et future de l'organisation ». [26]

En d'autres termes, les indicateurs de performance en anglais Key Performance Indicator (KPI) constituent un ensemble de mesures concernant la performance de l'organisation, et que, en dépassant un certain seuil, alarmera le décideur pour agir.

L'intérêt d'avoir des indicateurs de performance est que le décideur ne consulte pas les données entières mais celles dont il a besoin pour l'analyse afin de pouvoir prendre une décision en temps réel, seuls les indicateurs les plus importants s'affichent en forme d'un graphique (en général codés en couleur).

5.2.2.3. Démarche de construction d'un tableau de bord

La construction d'un tableau de bord comprend les étapes suivantes :

1. La définition des finalités et des objectifs assignés au tableau de bord.
 - a. Est déterminée par ses utilisateurs, en fonction des enjeux et des objectifs de l'entité.
 - b. Nécessite de porter un regard sur l'organisation au travers d'une analyse des activités.
2. L'identification et la mobilisation des acteurs.
 - a. Nécessité de la participation active des utilisateurs futurs et ceux qui vont en assurer l'alimentation au quotidien.
 - b. L'identification des acteurs et leur rôle est assurée par le pilote du projet ; En général, son utilisateur principal.
3. La définition des critères et indicateurs pertinents.
4. La collecte et le traitement des données.
5. Le choix de la périodicité du tableau de bord : le tableau de bord doit être établi à une fréquence suffisante pour permettre au responsable de réagir à temps. [25]

5.2.2.4. Objectifs d'un tableau de bord [25]

Le tableau de bord est un outil dédié à l'aide à la décision, il permet ainsi de :

- Réduire l'incertitude.
- Faciliter la communication.
- Dynamiser la réflexion.
- Maîtriser le risque.

5.2.3. Datamining

5.2.3.1. Définition du datamining

« Le Data Mining est un terme générique englobant toute une famille d'outils facilitant l'exploration et l'analyse des données contenues au sein d'une base décisionnelle [...] » [1]

Le rôle des outils du datamining est la manipulation de gros volumes de données historiques pour en dégager des éléments exploitables, offrant ainsi une analyse plus poussée permettant de découvrir du sens caché dans les données (la détection de corrélations et de tendances...).

Le Datamining est basé sur des algorithmes statistiques et mathématiques.

5.2.3.2. Objectifs des outils de Datamining

Le datamining répond à deux objectifs, qui sont :

- ❖ Exploiter autant que possible le capital d'informations disponible et faire sortir les informations cachées. [27]
- ❖ Constituer des modèles pour découvrir des tendances ou pour anticiper l'avenir. [28]

Conclusion

L'informatique décisionnelle est considérée comme un outil puissant qui aide les décideurs à mettre en place leurs stratégies, ainsi la mise en place du Data Warehouse implique le passage par des phases pertinentes, qui détermineront, selon leurs qualités de réalisation, l'efficacité du Data Warehouse dans le processus décisionnel :

1. Etude préalable, qui inclut : les besoins de l'entreprise en termes d'informations, les objectifs du Data Warehouse, les données qu'il doit contenir, les résultats attendus, le choix des faits et leurs contextes associés.
2. Détermination de l'architecture et la modélisation dimensionnelle selon l'étude, les circonstances constatées et les choix effectués d'après l'étude.
3. Mise en place de l'ETL qui alimentera le Data Warehouse.

Un Datawarehouse bien conçu et correctement implémenté, influera également sur la performance des outils BI qui utiliseront les données qui y sont stockées.

Dans ce chapitre nous avons présenté une étude bibliographique concernant les systèmes décisionnels et les Data Warehouse.

Dans le chapitre suivant nous allons présenter notre organisme d'accueil ainsi que l'étude de l'existant.

Chapitre II : étude de cas de l'entreprise

Introduction

Après avoir établi la première partie représentée par la synthèse bibliographique sur les systèmes d'informations décisionnels et les entrepôts de données, nous allons maintenant parler de la source de notre projet la Générale Maritime, en commençant par une présentation de l'entreprise, et nous faisons ensuite une étude de l'existant au sein de GEMA.

1. Présentation de GEMA [29]

La Générale Maritime (GEMA) créée en 1987 et certifiée ISO 9001-2008 est une société par action (SPA) dotée d'un Capital social de 327.900.000 de Dinars détenu entièrement par le groupe algérien de transport maritime «GATMA».

Dans le cadre de la mise en œuvre de sa stratégie de développement, GEMA a filialisé les activités Transit et Avitaillement par la création de deux sociétés à capitaux détenus majoritairement par la société mère GEMA dénommées respectivement FILTRANS et AVICAT.

Cette démarche est intervenue après la libéralisation de l'activité consignation³ (métier de base de GEMA) et consacrée par le décret N° 91522 du 22 décembre 1991.

Il s'en est suivi une deuxième phase qui consistait à engager la filiale AVICAT dans un processus de privatisation qui a abouti à la création de deux sociétés à capitaux mixtes : ALC (Algérie Ligabue Catering) et COMARPEX Algérie.

Pour l'exercice de ses activités, GEMA est représentée par un siège social et neuf agences, le siège social se trouve au : 02, rue Jawaharlal Nehru Alger BP 95 Alger. Algérie

Les neuf agences se trouvent sur Alger, Ténès, Oran, Mostaganem, Ghazaouet, Béjaia, Skikda, Jijel et Annaba.

Dans l'amélioration de son créneau, GEMA a créé une nouvelle filiale spécialisée dans la formation en gestion maritime appelée GEMAFORM.



Figure 18 : Logo de GEMA.

³ Ce point est défini dans l'annexe A

1.1. Les sociétés filiales et participations de GEMA

❖ La filiale FILTRANS

Cette filiale intervient dans le secteur de la logistique en tant que commissionnaire en transport international, manutentionnaire, transporteur terrestre en plus de la gestion d'entrepôts de marchandises sous douane et de conteneurs vides. Elle est dotée d'un capital social de 200 millions de DA détenu à 90% par GEMA et 10% par GATMA

GEMA dispose également de participations⁴ dans les sociétés suivantes :

- ❖ Algerian ligabe catering (SPA)
- ❖ Comarpex Algérie (SPA)
- ❖ SOGRAL (SPA)
- ❖ La Société d'Investissement Hôtelière (SPA)

1.2. Statut juridique

Le consignataire de navire doit être considéré comme un « commissionnaire mandataire » de l'armateur. Le consignataire joue le rôle d'intermédiaire entre l'armateur et le client importateur ou exportateur.

1.3. Domaine d'activité de GEMA

L'entreprise GEMA intervient dans le marché de la consignation de tous types de navires dans tous les ports de commerce d'Algérie, la gestion des conteneurs, des marchandises réfrigérées, des marchandises en vrac, des colis lourds. Elle porte aussi aide et assistance pour des études de marché à toute compagnie désirant investir le domaine des transports maritimes en Algérie. Le Groupe GEMA intervient à travers ses filiales dans les activités de transit, transport terrestre, entreposage de marchandise, parc sous douane et avitaillement des navires. Dans son portefeuille clientèle, GEMA possède des clients importateurs, des clients exportateurs, des transitaires et des armateurs étrangers qui la désignent comme agent consignataire en Algérie

⁴ Le détail des participations se trouve dans l'annexe A

La figure ci-dessous montre globalement une opération de consignation d'un navire entre GEMA et un armateur.

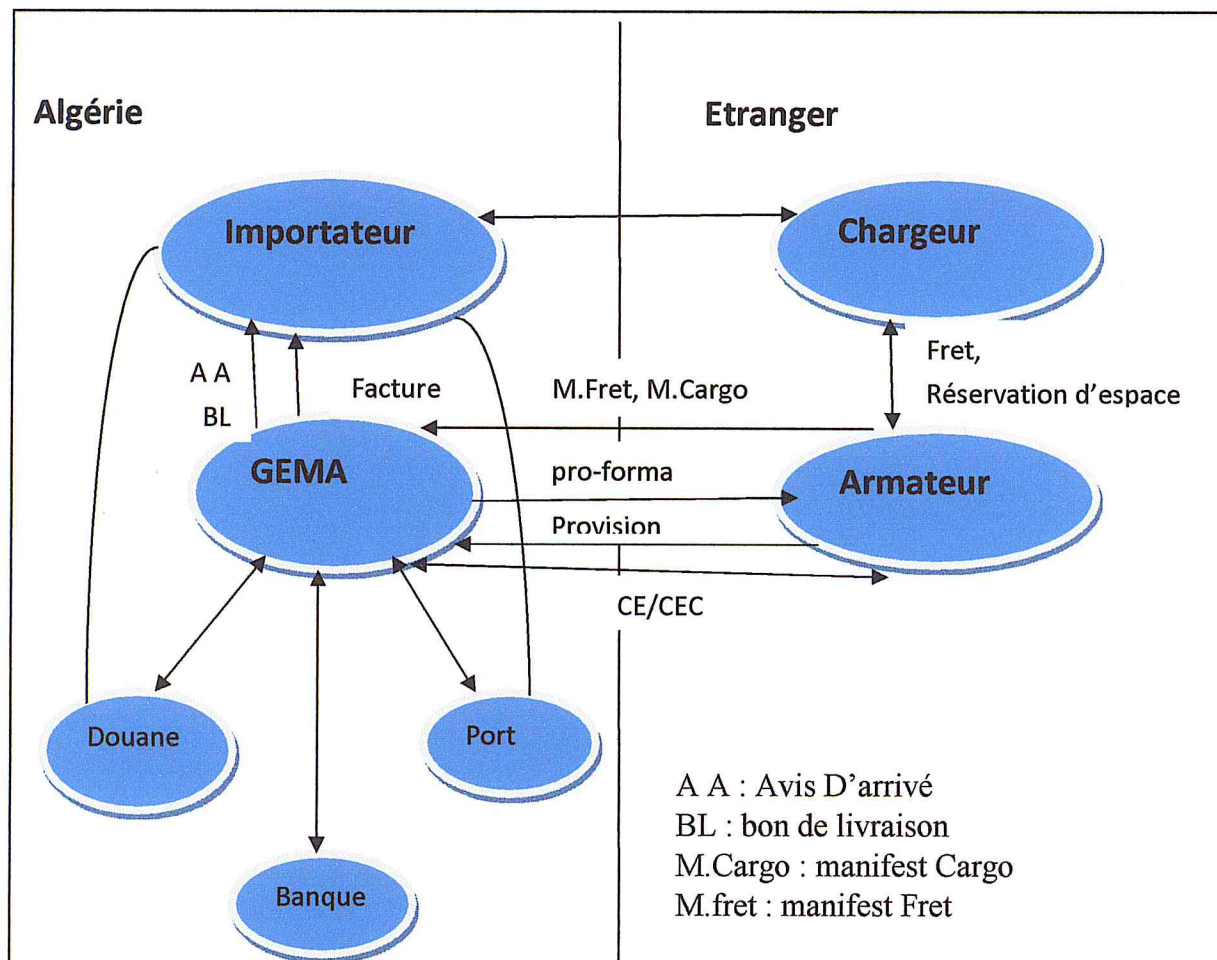


Figure 19 : L'opération de consignation.

1.4. GEMA et le marché de consignation

Un savoir-faire acquis depuis de nombreuses années, de nombreux services offerts, des moyens humains et matériels importants placent sans conteste GEMA comme leader dans les services annexes au transport maritime.

Le chiffre d'affaire enregistré par l'entreprise est en constante évolution représenté dans le tableau ci-dessous :

	chiffre d'affaires	Résultats net	CA/Résultat
Exercice 2010	344 946 DA	169 555 DA	49%
Exercice 2011	373 081 DA	218 347 DA	59%
Exercice 2012	400 303 DA	257 651 DA	64%

Tableau 4 : L'évolution du chiffre d'affaire de GEMA.

1.5. Certification de GEMA

Pour uniformiser ses procédures et améliorer sa prestation, GEMA s'est engagée à la fin de l'année 2005, dans la mise en place d'un système de management de la qualité (SMQ). L'entreprise a été certifiée ISO 9001/2000 par l'organisme SGS au cours de l'année 2009. Depuis, l'entreprise a passé avec succès toutes les étapes de suivi et validité du Système. Comme elle a réussi son accès à la version 2008 et au renouvellement du certificat chaque année.

Ce système a permis une organisation du fond documentaire (procédures, formulaires, instructions de travail) et la mise en place d'une batterie d'outils de management et de contrôle tels les revues de processus, revue de direction, les audits internes, les mesures de performance et enquêtes satisfaction clients.

Cette démarche qualité inscrite dans un processus d'amélioration continue vise aussi à promouvoir la prestation de GEMA et à renforcer sa position par rapport à la concurrence.

2. Etude de l'existant

L'organisation actuelle de GEMA a pour but de répondre aux exigences de la rapidité d'obtention de l'information.

A cet égard, GEMA dispose des systèmes suivants :

2.1. La dotation d'investissement informatique

En 1996, et dans le cadre de la mise en place d'un système intégré de consignation et de comptabilité, GEMA a acquis des serveurs SIEMENS de type RM200, RM300 à architecture RISC avec un système d'exploitation propriétaire UNIX version SINIX 2.43 muni de moteur de base de données INFORMIX ONLINE et d'un applicatif métier de consignation et de

comptabilité acquis avec son code source, (développé avec un outils de développement JAM 7) d'où la possibilité de concevoir des modules complémentaires liés à l'activité .

En 2006, l'entreprise GEMA s'est lancée dans un projet de migration¹, dont l'objectif principal était l'élimination du contrat de maintenance avec SIEMENS pour le matériel et le système d'exploitation.

L'opération de migration est concrétisée par l'acquisition de serveurs FUJITSU TX200-S6 quadri-processeurs muni d'un système d'exploitation « OPEN SOURCE LINUX REDHAT ENTREPRISE 5 ».

La migration a touché aussi le moteur de base de données ainsi que l'outil de développement JAM 7 qui est devenu PANTHER BUILD 5.1.

Un contrat de maintenance a été établi uniquement pour les serveurs jusqu'en 2013. Dans la même année, l'entreprise a mis fin complètement au contrat de maintenance avec SIEMENS et a opté pour la solution des serveurs secours.

Ainsi une étude du projet migration 2, a été réalisée dont l'objectif consiste à convertir leur application métier de consignation et de comptabilité développée en PANTHER BUILD 5.1 en application web développée en PANTHER WEB (produit à acquérir auprès de l'entreprise américaine « PROLIFICS »).

2.2. La dotation d'applications intranet

Dans le but de préserver sa part dans le marché de consignation et faire face à la concurrence, en 2003, l'entreprise a décidé de mettre en place un intranet d'entreprise et développer des applications web en interne, dont le serveur web se trouve au niveau du siège de GEMA. Les applications web développées sont les suivantes :

- Codification/consultation armateur (réalisée)
- Saisie /consultation cotation devise (réalisée)
- Saisie/consultation situation portuaire navire (réalisée)
- Saisie/consultation situation financière armateur (en cours de déploiement)
- Saisie/consultation conteneur (en cours de déploiement)
- Saisie/consultation des archives de GEMA (en cours de déploiement)
- Prospection client sur intranet (en cours de réalisation)

L'accès à l'intranet de l'entreprise se fait via l'adresse www.gemaweb.net (accès sécurisé)

2.3. La dotation de moyens de communication

2.3.1. En interne :

Pour maintenir le fonctionnement des systèmes implantés au niveau des agences maritimes, l'entreprise s'est dotée de moyens de télémaintenance et de moyens de communication, notamment :

Le teleservice, qui permettait d'accéder aux anciens serveurs SIEMENS, à partir d'un serveur SIEMENS central (accès RTC) pour régler les éventuels problèmes techniques.

Le teamviewer, mis en place pour régler des problèmes qui surgissent au niveau des agences et qui utilise l'internet.

Une étude de mise en place d'un réseau VPN (VIRTUEL PRIVATE NETWORK) est en cours de réalisation.

Tout échange d'information entre la direction générale et les agences, les armateurs ou les clients, se fait par la messagerie électronique (mails) de GEMA

2.3.2. Avec la douane :

Une solution EDI (échange de données informatisées) propriétaire a été mise en place en 1998 entre l'entreprise GEMA et le CNIS, pour le transfert de manifeste électronique dans les délais. En 2000 cette solution a été remplacée par la solution RTC, qui s'agit de saisir directement dans le système SIGAD via une connexion téléphonique dédiée.

2.3.3. Avec les entreprises portuaires :

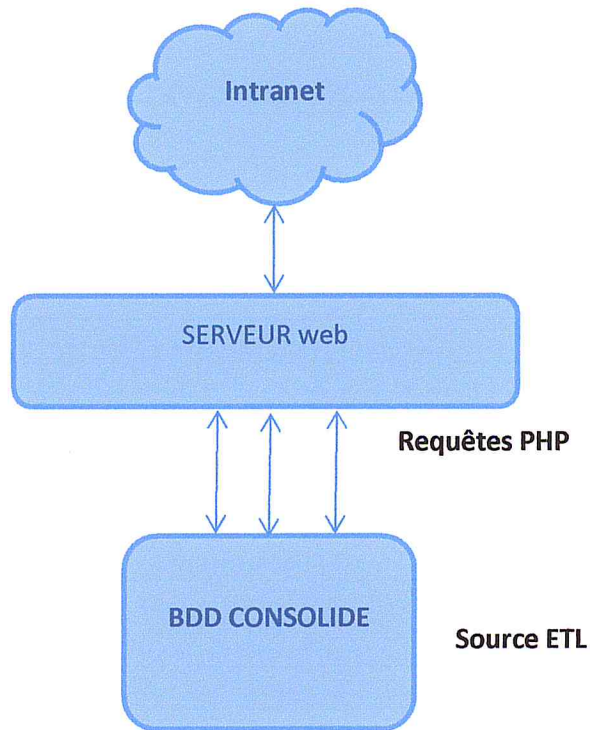
Des EDI à l'import et à l'export permettent de transférer les manifestes électroniques vers l'entreprise portuaire.

Ainsi que GEMA possède son portail électronique www.gema-groupe.com, qui contient toute les informations et l'actualité concernant l'entreprise (contact, réseau des agences)

2.4. L'existant décisionnel

Actuellement GEMA ne dispose pas d'un système décisionnel proprement dit, pour prendre des décisions stratégiques, les décideurs de GEMA se basent sur la base des données consolidés de l'intranet et des tableaux de statistiques conçus en extra des systèmes d'informations automatisés.

Niveau consolidé



Niveau opérationnel

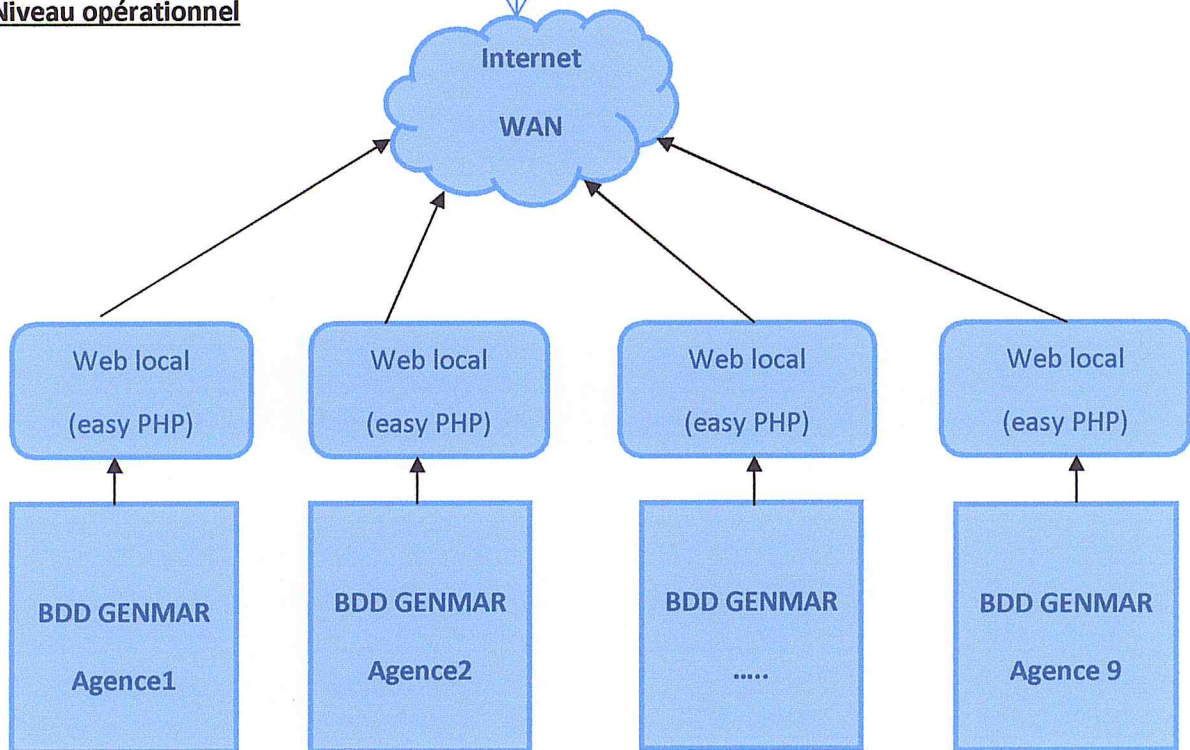


Figure 20: processus de consolidation des données

La base de données « CONSOLIDE » est alimentée via l'intranet de l'entreprise à partir des données opérationnelles qui se trouvent dans la base de données GENMAR dans chaque agence.

Techniquement, les données consolidées sont extraites grâce au web local au niveau de chaque agence de l'entreprise et alimente le web de l'entreprise qui est une partie de l'ETL.

L'organigramme suivant a pour objet de décrire les étapes de la consolidation de l'information d'exploitation afin de produire l'information décisionnelle et analytique.

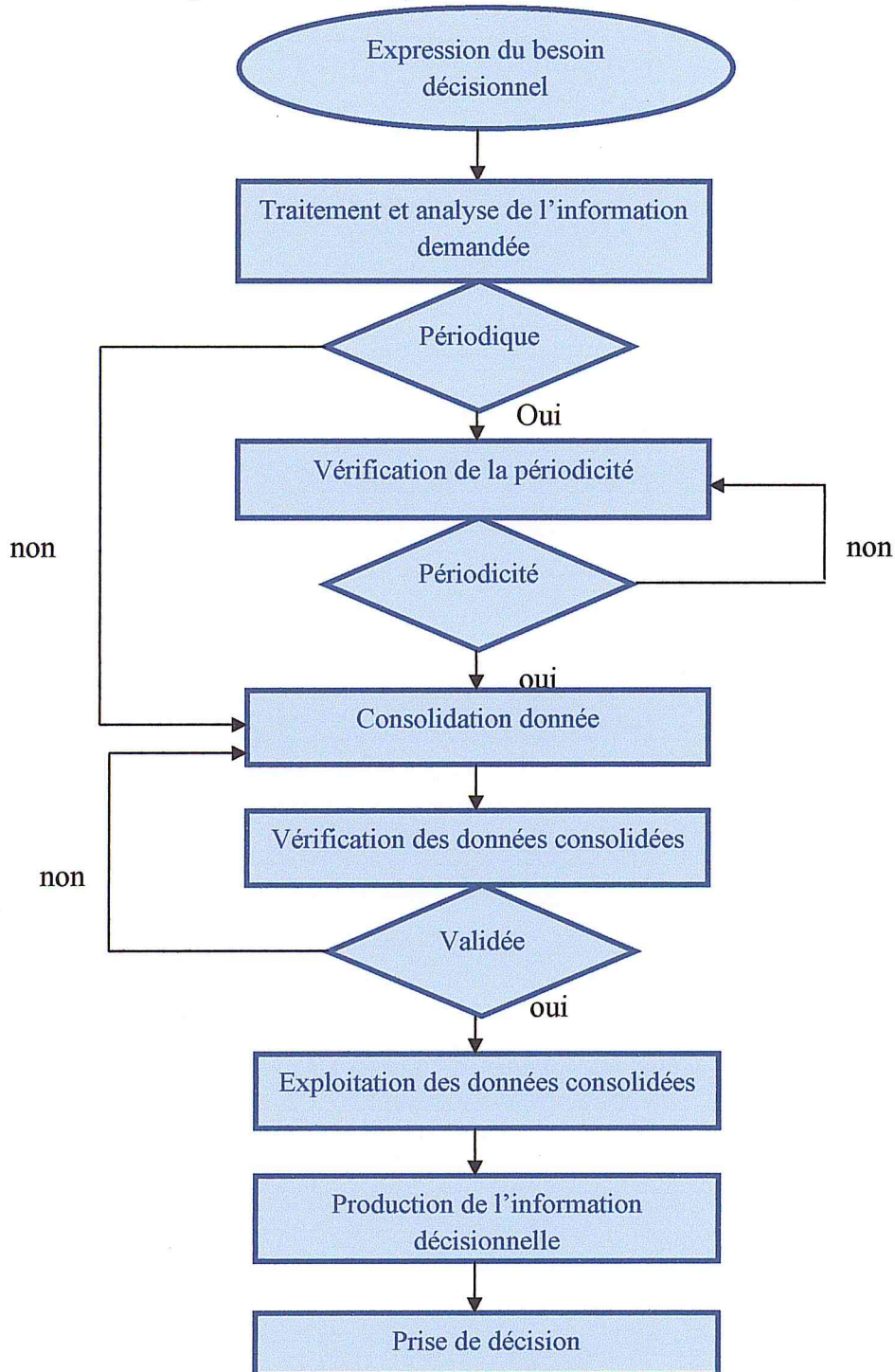


Figure 21: Processus de prise de décision au sein de GEMA.

2.5. Critiques

A partir de notre étude de cas de l'entreprise GEMA, nous sommes sorties avec un ensemble de constatations :

- Inexistence d'indicateurs pertinents définis pour la prise de décision et qui sont générés directement à partir des données consolidées;
- Difficulté dans la définition des indicateurs de mesure de performance commerciaux ;
- Retard dans l'élaboration des rapports d'activités, en vue du volume important des informations manipulées (leur élaboration peut aller jusqu'à 4 mois) ;
- Prise de décision inadéquate avec la stratégie commerciale de l'entreprise.

Conclusion

Par le biais de ce projet, l'entreprise GEMA envisage la mise en place d'un entrepôt de données afin d'assurer la prise de la bonne décision en temps voulu.

Partant de ce constat, nous allons dans ce qui suit présenter la conception de notre système afin de répondre aux besoins de l'entreprise.

➤ **Déclarer le degré de granularité**

Cela revient à spécifier exactement la signification de chaque ligne dans la table de fait.

➤ **Choisir les dimensions appliquées pour chaque table de faits :**

Cela revient à répondre à la question suivante : comment décrire les données résultantes du processus modélisé ?

➤ **Identifier les faits pour chaque table des faits**

Cela revient à répondre à la question suivante : qu'allons-nous mesurer ?

En utilisant les besoins collectés et mentionnés précédemment, nous pouvons concevoir les schémas en étoile constituant notre Data Mart. Ces derniers sont interconnectés entre eux grâce à l'ensemble des tables de dimensions communes entre les tables de fait.

1.2.1. Modélisation de notre solution

Suite à l'étude des besoins réalisée auparavant, nous avons abouti à concevoir neuf (9) schémas principaux :

Les faits définis dans les schémas sont vus par rapport à plusieurs axes d'analyse, qui constituent les dimensions dans ces schémas. Le choix de ces axes a été fait en repérant les aspects qui influent le résultat et la performance de l'activité dans laquelle s'inscrit le « fait » en question :

- **Periode** : représente l'axe temporel de l'analyse, et permet de de situer le fait produit dans la société par rapport à une période donnée.
- **Agence** : l'entité opérationnelle intermédiaire entre le client et l'armateur, assure les opérations maritimes pour un navire et la marchandise transportée
- **Armateur** : personne physique ou morale possédant ou exploitant un navire.
- **Type escale** : une escale est une opération d'accostage d'un navire dans un port consigné par un consignataire. Le type escale est défini par rapport au nombre d'opérations réalisées périodiquement (ligne régulière LR : pour armateurs conventionnés et le nombre d'escale est programmé dans le temps .

Escale en tramping TR : pour armateurs occasionnel , opération maritime aléatoire)

- **Sens du BIC** : un BIC est un transfert inter agence, il se fait dans le cas où un armateur est créditeur dans une agence et débiteur dans une autre, donc son sens est soit émis, soit reçu.
- **Etat A40** : A40 c'est un document intermédiaire entre le consignataire et la douane , il sert comme accord pour effectuer un transfert de fond entre consignataire et

armateur, on parle d'état A40 accepté dans le cas où les services de douane accordent le transfert de fonds, rejeté dans le cas contraire.

- **Nature compte d'escale :** compte d'escale est une facture récapitulative de tous les frais et débours concernant une escale réalisée pour un navire, adressée par GEMA à l'armateur.

Chaque escale doit avoir obligatoirement un compte d'escale initial CE, et un compte d'escale complémentaire CEC après clôture du compte d'escale initial et réception de nouvelles factures fournisseur concernant l'escale.

Etat du transfert : un transfert est une opération d'envoi de fonds de l'entreprise consignataire qui se trouve en Algérie vers un armateur qui se trouve à l'étranger suite à un A40 validé par les services de douanes.

Un transfert passe par deux phases : phase « Transfert en cours de réalisation TER », phase « Transfert réalisé TR »

- **Client :** c'est l'importateur, l'exportateur, le transitaire ou l'armateur, il représente.
- **Pays :** c'est le pays d'origine de l'armateur.
- **Pavillon :** C'est la nationalité du navire, il détermine la loi applicable en haute mer, il permet de bénéficier de la protection réservée aux nationaux.

a) Le fait « physique »

Description

La mesure physique de l'activité de consignation est représentée par le nombre de navires réalisés, le tonnage transporté (débarqué, embarqué) ainsi que le nombre de conteneurs.

L'indicateur physique (nombre de navires, tonnage, nombre de conteneur) réalisé dans une période, est primordial pour mesurer le fonctionnement de l'entreprise.

Grain

Une ligne dans la table de fait « physique » représente le nombre de navires réalisés, le tonnage ainsi que le nombre de conteneur (ce qui constitue l'activité physique) dans une période, pour un armateur, une agence et un type d'escale.

Dimensions

- Période
- Armateur
- Agence
- Type Escale

Mesures

- Le nombre de navires
- Le tonnage

Détails de la table de fait « physique »

Type de l'attribut	Nom de l'attribut	Désignation
Clés étrangères vers les Dimensions	Code_arm	Code d'armateur
	Code_ag	Code agence
	Code_type_Escale	Code type escale
	Jour	Date
Mesures	Nombre_Navire	Le nombre de navires réalisé
	Tonnage	Le tonnage réalisé

Tableau 5 : Composition de la table de fait « physique ».

Schéma en étoile

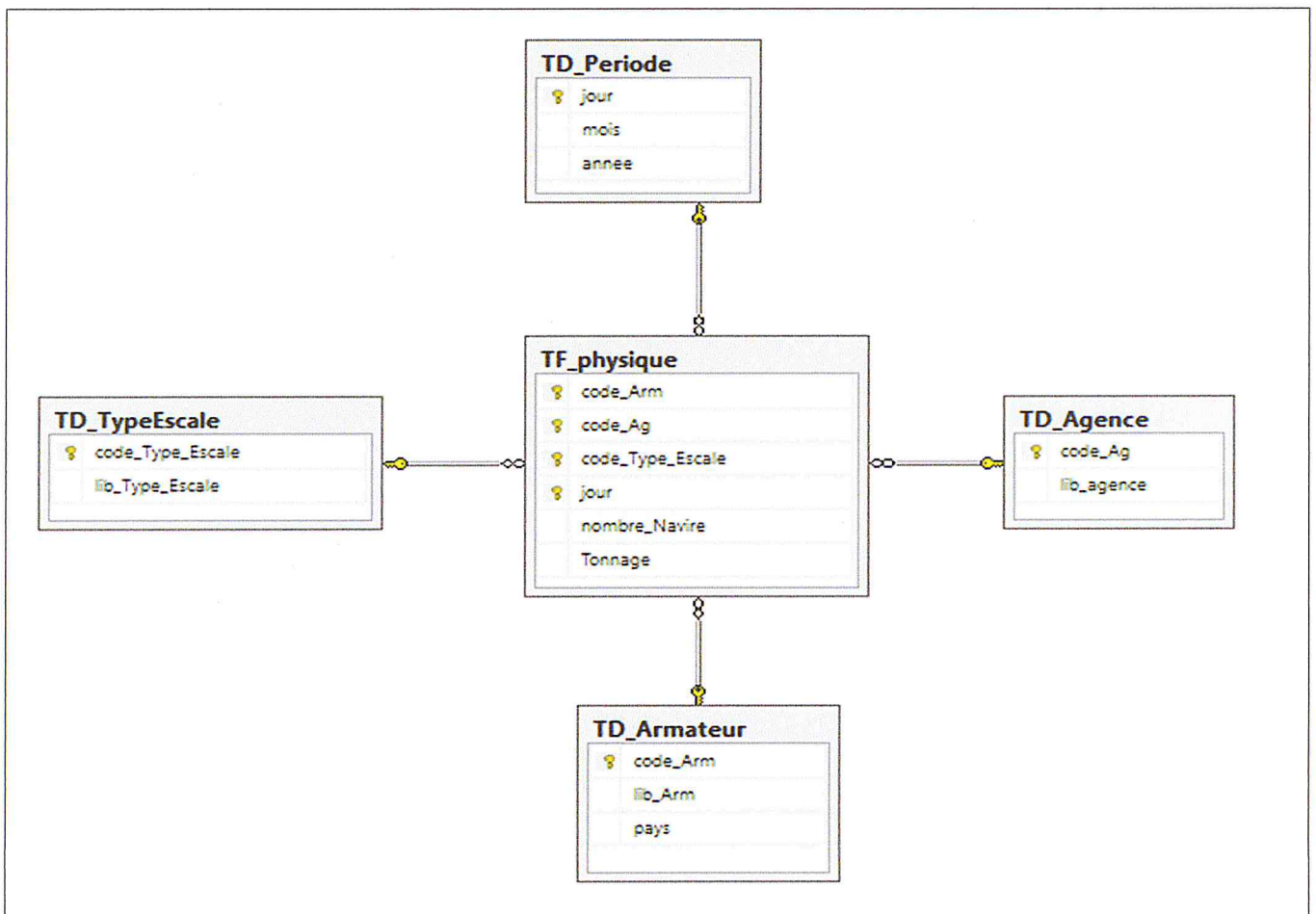


Figure 22: Schéma en étoile de la table de fait « physique ».

b) Le fait « compte d'escale réalisé »

Description

Sur le plan financier, l'entreprise mesure son chiffre d'affaire par rapport au nombre de compte d'escale ou compte d'escale complémentaire réalisé.

Grain

Une ligne dans la table de fait « Compte_Escale_Realise » représente le nombre de compte d'escale ainsi que les montants débit et crédit dans une période, pour un armateur, une agence et nature du compte d'escale

Dimensions

- Armateur
- Agence
- Periode
- Nature compte d'escale

Mesures

- Le nombre de compte d'escale
- Le montant débit du compte d'escale
- Montant crédit du compte d'escale

Détails de la table de fait « Cerealise »

Type de l'attribut	Nom de l'attribut	Désignation
Clés étrangères vers les Dimensions	Code_arm	Code d'armateur
	Code_ag	Code agence
	Jour	Date
	Code_natureCE	Code de la nature du compte d'escale
Mesures	Nombre_CE	Nombre de compte d'escale
	montantDebit	Montant débit du compte d'escale
	montantCredit	Montant crédit du compte d'escale

Tableau 6 : Composition de la table de fait « Compte_Escale_Realise».

Schéma en étoile

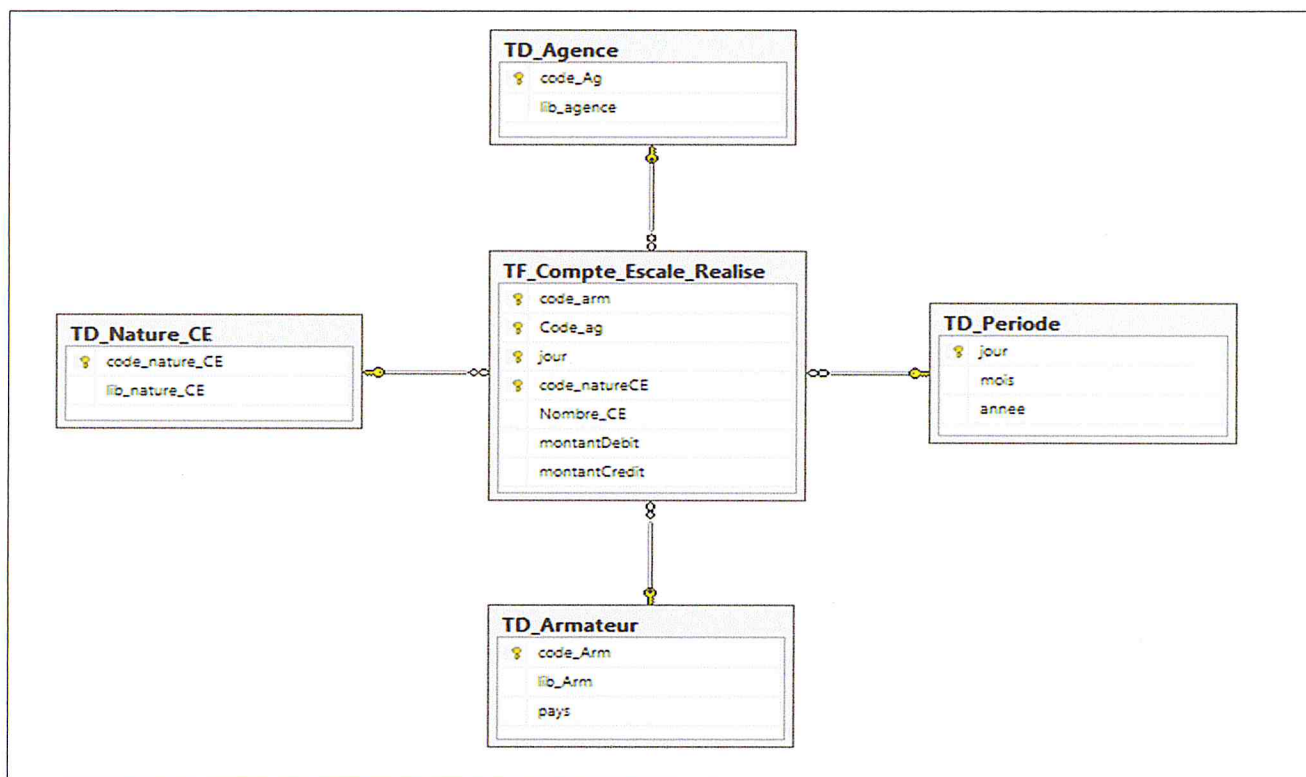


Figure 23 : Schéma en étoile de la table de fait « Compte_Escale_Realise ».

c) Le fait BIC

Description

Le BIC c'est le transfert inter agence pour un armateur qui est créancier dans une agence et débiteur dans une autre.

Le nombre de BIC élevé réalisé dans une agence a une période donnée reflète la non maîtrise des pro-forma.

Grain

Une ligne dans la table de fait « BIC » représente le nombre de BIC ainsi que le montant du BIC dans une période, pour un armateur, une agence et un sens du BIC défini.

Dimensions

- Armateur
- Agence
- Periode
- Sens BIC

Mesures

- Nombre de BIC
- Montant du BIC

Détails de la table de fait

Type de l'attribut	Nom de l'attribut	Désignation
Clés étrangères vers les Dimensions	Code_arm	Code d'armateur
	Code_ag	Code agence
	Jour	Date
	Code_sensBIC	Code sens BIC
Mesures	Nombre_BIC	Nombre de BIC
	MontantBIC	Montant du BIC

Tableau 7 : Composition de la table de faits « BIC ».

Schéma en étoile

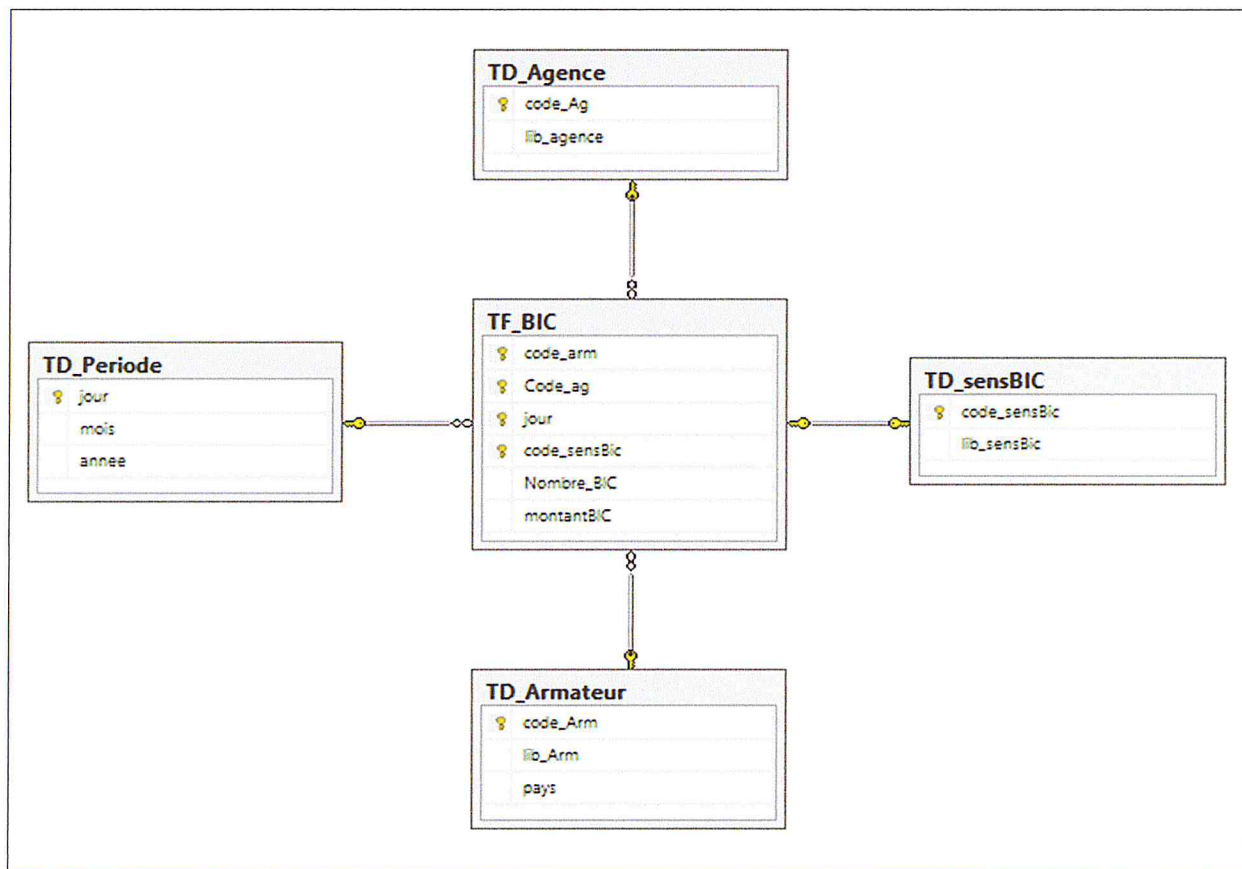


Figure 24 : Schéma en étoile de la table de fait « BIC ».

d) Le fait A40

Description

Il précède un transfert, c'est un document visé par la douane pour qu'elle donne son avis pour pouvoir faire le transfert, il permet de maîtriser le nombre de A40 accordé ou rejeté par la douane.

Grain

Une ligne dans la table de fait « A40 » représente le nombre de A40 ainsi que le montant du A40 dans une période, pour un armateur, une agence et un type du A40 défini.

Dimensions

- Armateur
- Agence
- periode
- Etat A40

Mesures

- Le nombre de A40
- Le montant de A40

Détails de la table de fait

Type de l'attribut	Nom de l'attribut	Désignation
Clés étrangères vers les Dimensions	Code_arm	Code d'armateur
	Code_ag	Code agence
	Jour	Date
	Code_etat_A40	Code etat A40
Mesures	Nombre_A40	Nombre d'A40
	MontantA40	Montant d'A40

Tableau 8: Composition de la table de faits « A40».

Schéma en étoile

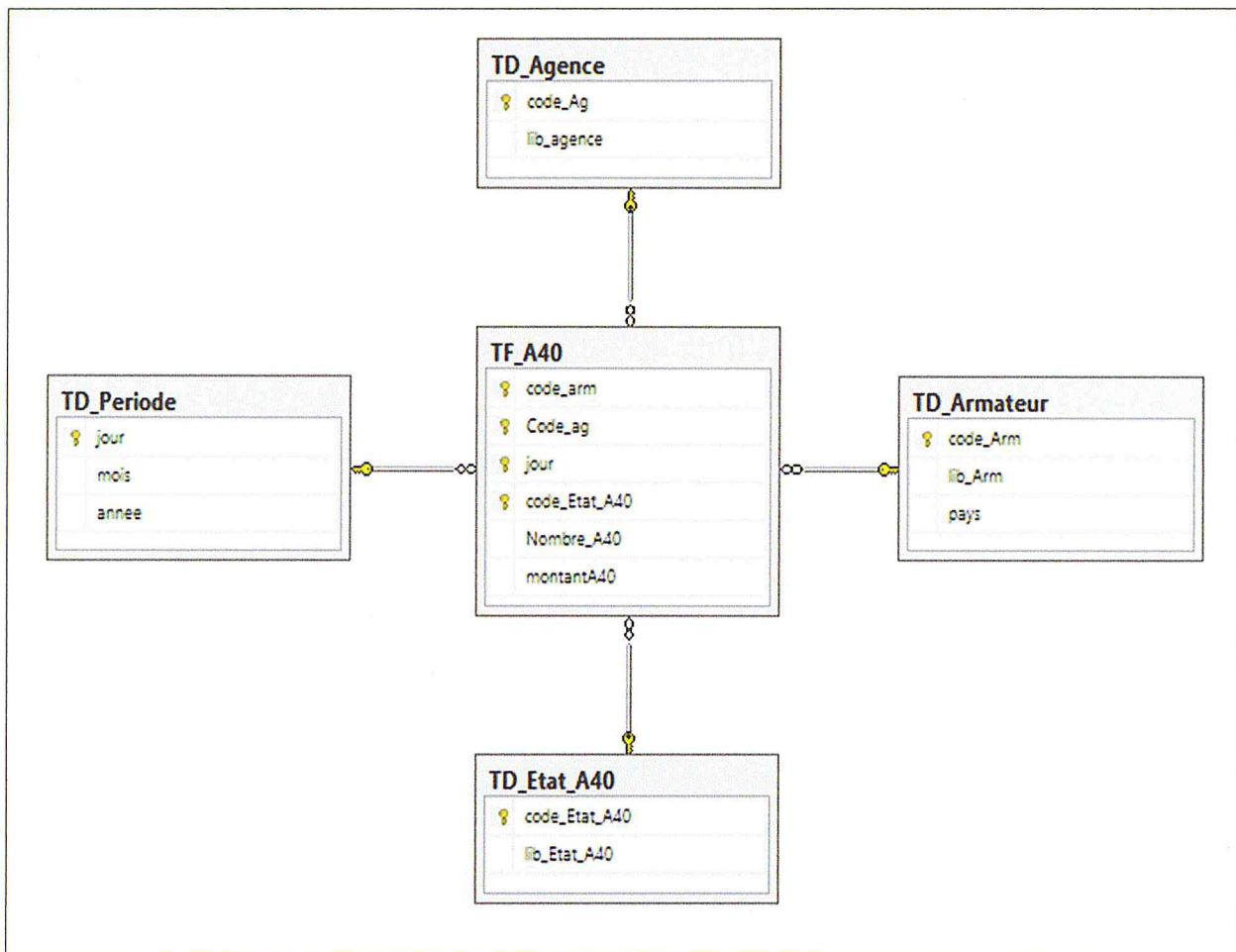


Figure 25 : Schéma en étoile de la table de fait « A40 ».

e) Le fait transfert

Description

C'est un indicateur décisionnel pertinent, il permet de vérifier la maîtrise de réalisation des pro-forma. Un nombre de transfert réduit ainsi qu'un montant faible du transfert impliquent une bonne maîtrise dans le calcul des pro-forma.

Grain

Une ligne dans la table de fait « Transfert » représente le nombre de transfert ainsi que le montant du transfert dans une période, pour un armateur, une agence et l'état du transfert définis.

Dimensions

- Armateur
- Agence
- Periode
- Etat transfert

Mesures

- Nombre de transfert
- Montant du transfert

Détails de la table de fait

Type de l'attribut	Nom de l'attribut	Désignation
Clés étrangères vers les Dimensions	Code_arm	Code d'armateur
	Code_ag	Code agence
	Jour	Date
	Code_etatTransfert	Code etat du transfert
Mesures	nombre_transfert	Nombre de transfert
	montantTransf	Montant du transfert

Tableau 9: Composition de la table de faits « transfert ».

Schéma en étoile

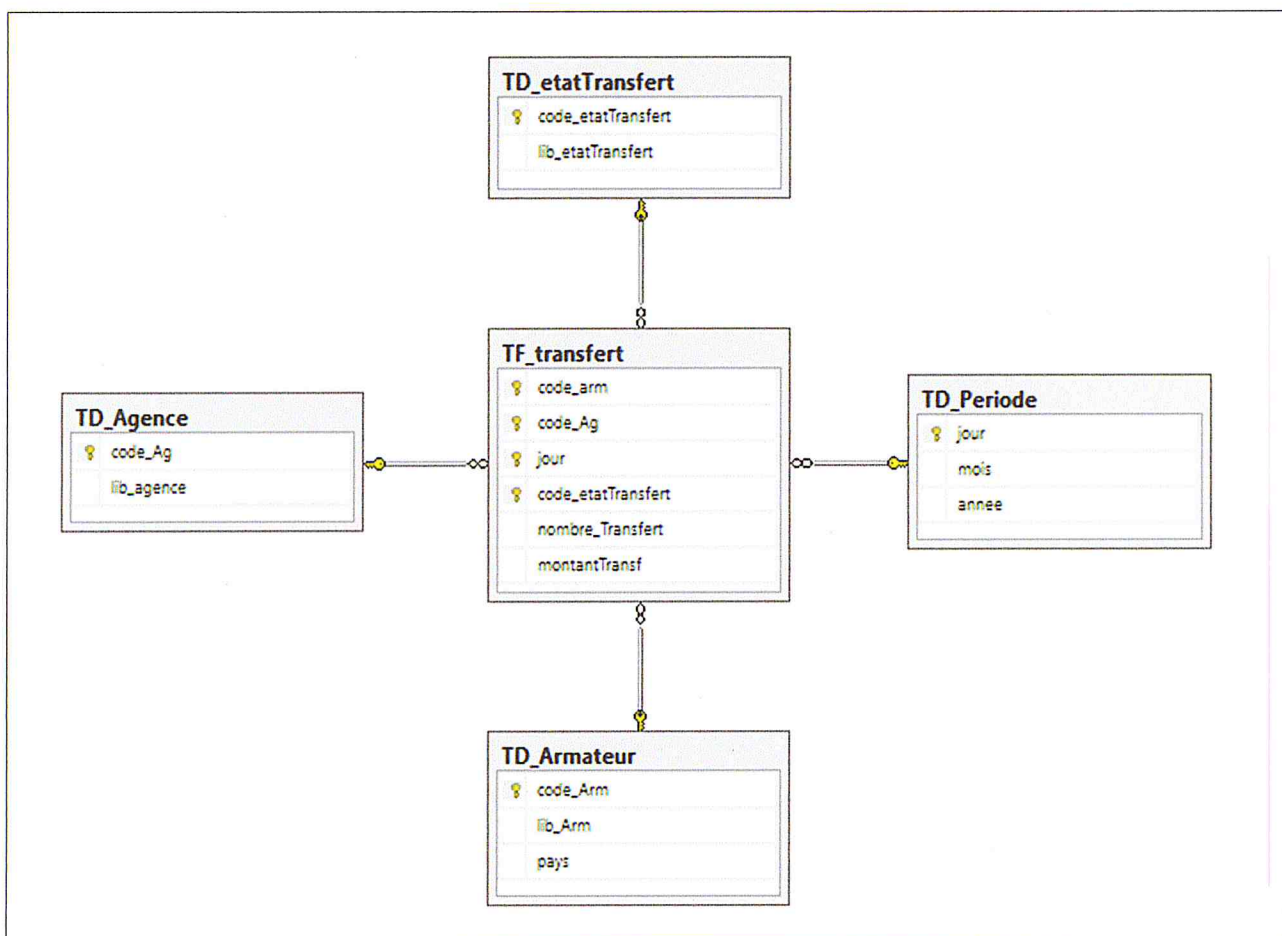


Figure 26: Schéma en étoile de la table de fait « transfert ».

f) Le fait nouveaux Armateurs

Description

Il est important de connaître le nombre de nouveaux armateurs dans une période, pour mesurer l'effort prospection engagé qui constitue l'un des objectifs de GEMA.

Grain

Une ligne dans la table de faits « Armateur» représente le nombre de de nouveaux armateurs dans une période pour une agence et un pays définis.

Dimensions

- Agence
- Periode
- Pays

Mesures

- Le nombre de nouveaux armateurs

Détails de la table de fait

Type de l'attribut	Nom de l'attribut	Désignation
Clés étrangères vers les Dimensions	Code_ag	Code agence
	Jour	Date
	Code_pays	Code pays
Mesures	Nombre_NouveauArm	Nombre de nouveaux armateurs

Tableau 10 : Composition de la table de fait « Nouveaux_Armateur».

Schéma en étoile

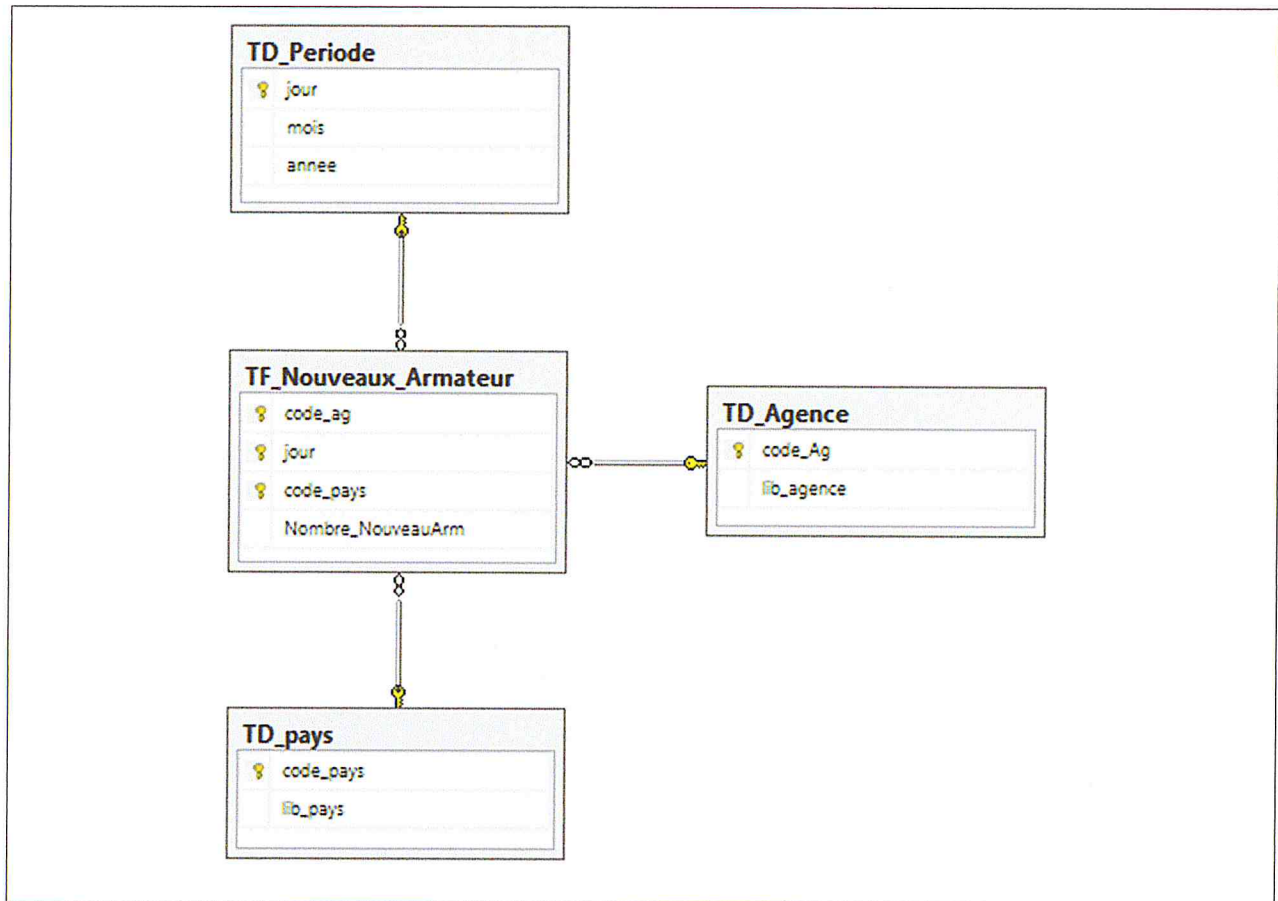


Figure 27 : Schéma en étoile de la table de fait « Nouveaux_Armateur ».

g) Le fait Réalisation des Clients

Description

Il est important pour le décideur de cibler les meilleurs clients par rapport au chiffre d'affaire réalisé afin de fournir des efforts pour de meilleures prestations (faire des remises par exemple)

Grain

Une ligne dans la table de faits « realisationClient » représente le nombre de navires réalisés ainsi que le tonnage réalisés dans une période, pour client et une agence définis.

Dimensions

- Agence
- Periode
- Client

Mesures

- Nombre de navire réalisé
- Nombre de tonnage réalisé

Détails de la table de fait

Type de l'attribut	Nom de l'attribut	Désignation
Clés étrangères vers les Dimensions	Code_ag	Code agence
	Jour	Date
	Code_Client	Code client
Mesures	Nombre_navire_Realise	Nombre de navires réalisés
	Tonnage_Realise	Tonnage réalisé

Tableau 11 : Composition de la table de faits « realisationClient ».

Schéma en étoile

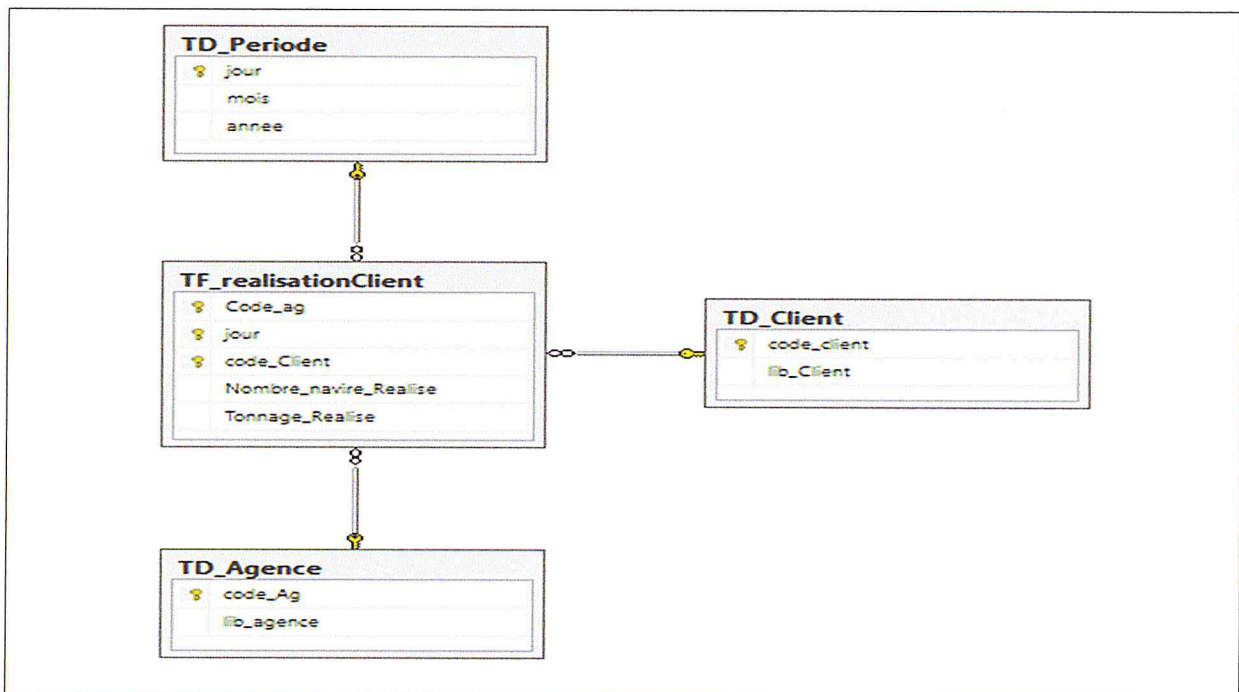


Figure 28 : Schéma en étoile de la table de fait « realisationClient ».

h) Le fait Pavillon

Description

C'est important de savoir les navires traités par pavillon c'est pour le paiement d'une taxe car chaque pavillon a sa propre taxe.

Grain

Une ligne dans la table de faits « Pavillon » représente le nombre de navires dans une période, pour une agence et un pavillon donnés.

Dimensions

- Agence
- Periode
- Pavillon

Mesures

- Le nombre de navire

Détails de la table de fait

Type de l'attribut	Nom de l'attribut	Désignation
Clés étrangères vers les Dimensions	Code_ag	Code agence
	Jour	Date
	Code_Pavillon	Code pavillon
Mesures	Nombre_Navire	Nombre de navires

Tableau 12 : Composition de la table de faits « pavillon».

Schéma en étoile

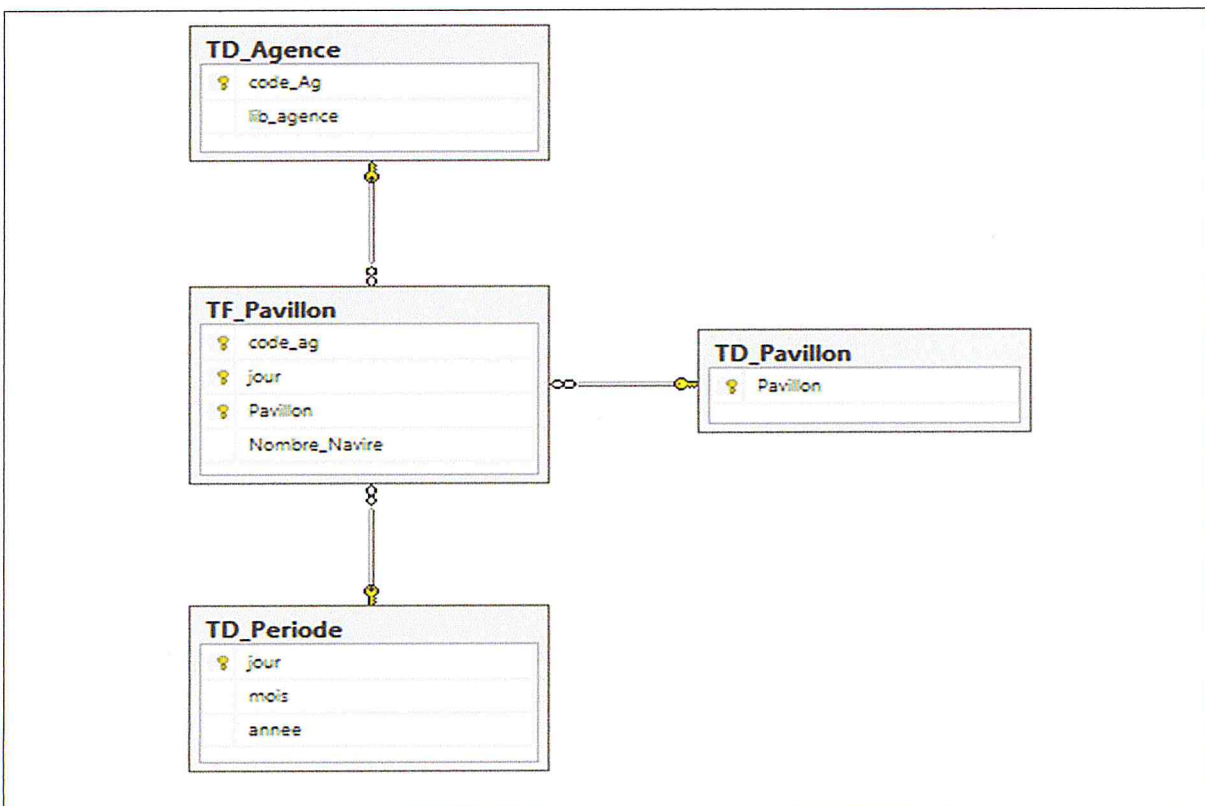


Figure 29 : Schéma en étoile de la table de fait « Pavillon».

i) Le fait Créances Armateurs

Description

Il est important de connaître la valeur des créances des armateurs pour GEMA (compte d'escale ou factures non réglées).

Grain

Une ligne dans la table de faits « creanceArm » représente le montant des créances dans une période, pour un armateur et une agence.

Dimensions

- Armateur
- Agence
- Periode

Mesures

- Montant de créances

Détails de la table de fait

Type de l'attribut	Nom de l'attribut	Désignation
Clés étrangères vers les Dimensions	Code_arm	Code d'armateur
	Code_ag	Code agence
	Jour	Date
Mesures	Montant_Creance	Montant de créances

Tableau 13 : Composition de la table de faits « creance_Armateur ».

Schéma en étoile

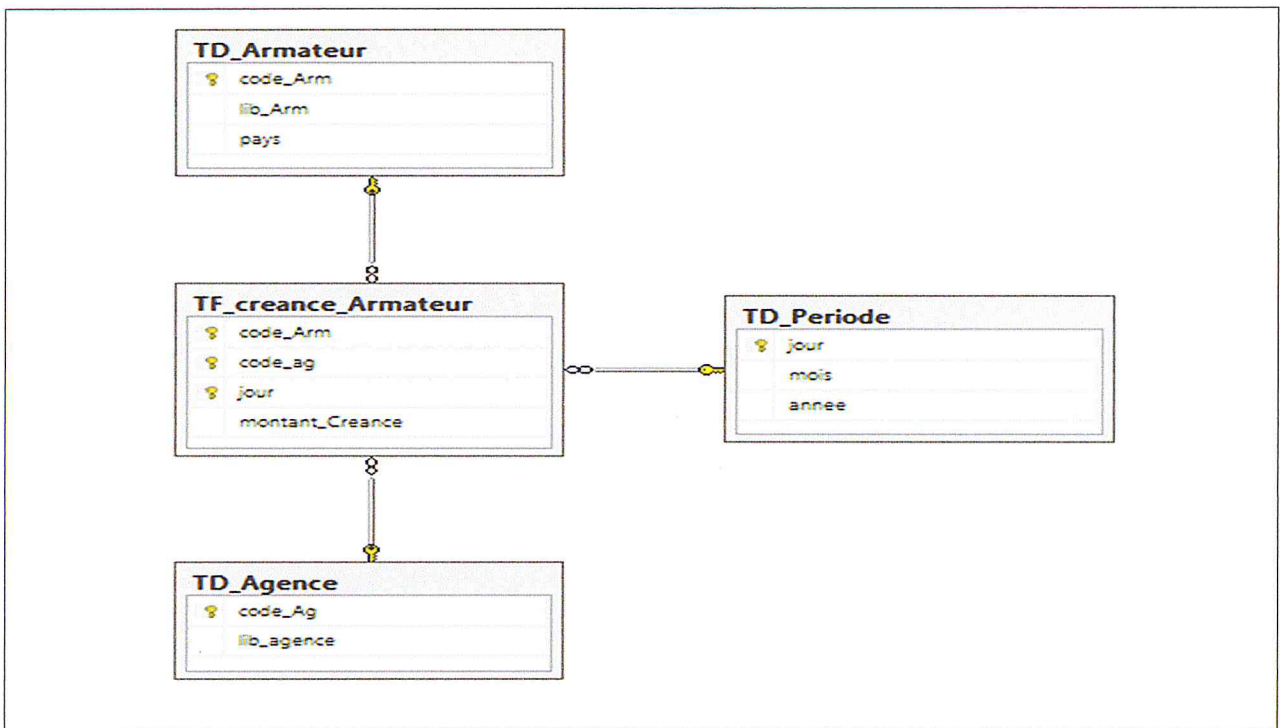


Figure 30 : Schéma en étoile de la table de fait « creance_Armateur ».

1.2.2. Matrice de bus

La matrice de bus montre l'utilisation des dimensions par les différents processus de la société :

	Physique	Compte_escale _realise	BIC	A40	Transfert	Nouveaux_Armateur	realisationClient	pavillon	Creance_Armateur
Periode	X	X	X	X	X	X	X	X	X
Armateur	X	X	X	X	X				X
Agence	X	X	X	X	X	X	X	X	X
TypeEscale	X								
Nature_CE		X							
sensBIC			X						
Type_A40				X					
etatTransfert					X				
Pays						X			
Client							X		
Pavillon								X	

Tableau 14 : Matrice en bus du système dw_gemahouse.

1.2.3.Détails des dimensions du système dimensionnel

Nom de l'attribut	Désignation	Exemple de valeurs
Dimension Periode	La dimension utilisée pour l'analyse temporelle	
Jour	Date complète	2015-11-29...
Mois	Le mois et l'année	2015-11...
Année	L'année	2015...
Dimension Agence	Cette dimension représente l'agence	
Code_ag	Code d'agence	ALG, BEJ...
Lib_ag	Libellé d'agence	Alger, Bejaia...
Dimension Armateur	Cette dimension représente l'armateur, un des types de clients de GEMA	
Code_arm	Code d'armateur : les trois premiers chiffres représente le pays de l'armateur, et le reste représente le numéro séquentiel dans ce pays	250075, 724108...
Lib_arm	Libellé d'armateur	BIRON SA, LUMAR BV
Pays	Pays	France, France...
Dimension typeEscale	Cette dimension représente le type d'escale	
Code_type_Escale	Code du type d'escale	LR, TR
Lib_type_Escale	Libellé du type d'escale	Ligne régulière, ligne en tramp
Dimension sensBIC	Cette dimension représente le sens du transfert inter agence	
Code_sensBIC	Code du sens du BIC	Emis, reçu
Lib_sensBIC	Libellé du sens du BIC	Emis, reçu
Dimension etat_A40	Cette dimension représente l'état du document A40	
Code_etat_A40	Code de l'état d'A40	ACC, REJ
Lib_etat_A40	Libellé de l'état d'A40	Accepté, rejeté
Dimension etatTransfert	Cette dimension représente l'état du transfert	
Code_etatTransfert	Code d'état du transfert	TRE, TER
Lib_etatTransfert	Libellé d'état du transfert	Transfert réalisé, transfert en cours de réalisation
Dimension nature_CE	Cette dimension représente la nature du compte d'escale	
Code_nature_CE	Code de la nature du compte d'escale	CE, CEC

Lib_nature_CE	Libellé de la nature du compte d'escale	Compte d'escale , Compte d'escale complémentaire
Dimension Pavillon	Cette dimension représente le pavillon par lequel passe un navire	
Code_pavillon	Code du pavillon	FRANÇAIS, TURKISH...
Dimension Client	Cette dimension représente le client importateur de GEMA	
Code_client	Code du client	OAIC, ONAB
Lib_client	Libellé du client	OAIC, ONAB
Dimension Pays	Cette dimension représente un pays	
Code_pays	Code du pays : suite de chiffre donné pour chaque pays.	705, 458...
Lib_pays	Libellé du pays	SLOVENIE, MALAISIE...

Tableau 15 : Tableau récapitulatif des dimensions et leurs attributs.

1.2.4. Sources du système

La source de notre système provient principalement du système consolidé qui est considéré comme une zone de stockage alimenté à partir du système opérationnel « GENMAR ». Cette source servira à alimenter les tables de faits, de calculer les mesures, de charger et de mettre à jour les dimensions.

1.3. Conception de l'ETL

L'alimentation est la procédure qui permet de préparer les données sources en vue de leur intégration puis de leur exploitation au sein du DW. La conception de cette opération est une tâche cruciale et complexe car elle détermine la pertinence et la qualité des données du système opérationnel.

Cette procédure commence par le traitement des dimensions en effectuant un chargement initial pour remplir les tables correspondantes. Ceci sera suivi de chargements incrémentaux avec des mises à jour des dimensions (ajout de nouvelles lignes non existantes).

Il existe plusieurs solutions libres (ETL) dans le marché mais on a opté pour une solution écrite à la main pour certaines raisons :

- Bien apprendre les concepts appris en cours « Systèmes décisionnels » en les appliquant.
- Développer ses capacités à programmer en SQL.
- Gérer le moindre détail par nous-mêmes.

1.3.1. ETL sur les dimensions

Le chargement des dimensions nécessite un chargement initial depuis la source après avoir effectué les différents traitements. Une fois les tables remplies, le chargement incrémental s'effectue pour ajouter de nouvelles lignes après avoir vérifié leur absence dans la table de dimension.

1.3.2. ETL sur les tables de faits

Le chargement incrémental responsable du remplissage de la table de faits se fera périodiquement où se feront la transformation, la correspondance avec les dimensions et l'insertion des mesures.

1.3.3. Traitements effectués

Afin de garantir une bonne transformation des données au DW, une stratégie de traitement a été élaborée et qui se résume dans les points suivants :

- Conversion de type de données.
- Traitement des champs vides et nuls.
- Elimination des doublons.
- Corrections des valeurs non significatives.

Les problèmes qui étaient résolues par les traitements mentionnés sont :

- Champs vides et valeurs manquantes ou incomplètes.
- Redondance de données difficilement repérable dans la source due à des utilisations inappropriées des fonctionnalités du système source.
- Données contradictoires dues à des erreurs d'alimentation de la source.

1.3.4. La périodicité de chargement de données :

La périodicité de l'alimentation du Data Warehouse, dépend essentiellement de la fréquence de la mise à jour ou de l'insertion des tables « Opérationnelles » de la base de données. De ce fait, il faut prendre en considération trois facteurs importants :

- ✓ La quantité de données à charger.
- ✓ Le temps nécessaire pour le chargement.
- ✓ Le moment idéal pour lancer le chargement.

Le chargement périodique des données est assuré par l'ETL, c'est pourquoi ce paramètre joue un rôle décisif quant au choix de l'outil ETL. Pour notre projet, nous avons opté pour une fréquence d'alimentation quotidienne.

1.4. Conception des cubes dimensionnels

La conception des cubes dimensionnels, facilite la navigation dans les données et exploite de manière efficace les informations contenues dans le Data Warehouse.

Le choix des cubes à construire, des dimensions participantes, des mesures utilisées et des hiérarchies définies est lié directement aux besoins recensés et à la structure des étoiles conçues.

Conclusion

La conception nous a permis de pré-visualiser le système à réaliser, ceci facilitera l'implémentation de notre système.

Dans la prochaine étape il s'agit de réaliser la solution tout en respectant les étapes des méthodes suivies expliquées dans ce chapitre. Les détails de l'implémentation ainsi que les résultats feront l'objet du chapitre suivant.

Chapitre IV : Réalisation du système

Introduction

La réalisation vient couronner les phases précédentes, donnant une forme concrète à la conception. Cette partie décrit l'environnement existant (logiciel et matériel), l'environnement mis en place pour exécuter notre système, les outils utilisés, ainsi que des aperçus du système conçu et les résultats fournis.

La phase de réalisation comporte quatre étapes qui sont :

- La mise en place du data warehouse (Data Mart).
- La construction de l'ETL.
- La construction et le chargement des cubes OLAP.
- La mise en place des outils de restitution (tableau de bord).

1. Moyens matériels

Pour la réalisation du projet, nous avons utilisé nos propres ressources caractérisés par :

- Windows 7 64bits, 4Go de RAM, Processeur Intel Pentium, CPU 2,00GHz.
- Windows 10 64bits, 4Go de RAM, Processeur Intel Core I7, CPU 2.20GHz.

2. Moyens logiciels

Pour le choix des outils et solutions de développement On a opté pour des solutions **Open source** car celles-ci sont gratuite ou à faible coût, libre de redistribution et surtout bénéficie d'un support Professional garanti.

2.1. Microsoft SQL Server Integration Services

SSIS est une plateforme qui permet d'effectuer l'ETL depuis Plusieurs types de sources vers plusieurs formes d'entrepôts de données (Excel, fichier plat...) tout en passant par la transformation des données via une interface facile à utiliser. [30]



2.2. PHP Storm

PhpStorm est un éditeur pour PHP, HTML et JavaScript, édité par JetBrains.

Il permet d'éditer du code PHP 5.3, 5.4, 5.5, 5.6 et 7.0.

Il possède :

- Une coloration syntaxique ;
- Affichage des erreurs à la volée ;



- Auto-complétion intelligente du code ;
- Réusinage du code.

Il intègre :

- L'envoi des fichiers via FTP ;
- Un gestionnaire de version.

Il permet aussi de visualiser l'architecture de bases de données de différentes sources (MySQL, SQLite, ...). [31]

2.3. WampServer

WampServer est une plateforme de développement Web de type WAMP, permettant de faire fonctionner localement (sans se connecter à un serveur externe)



des scripts PHP. WampServer n'est pas en soi un logiciel,

mais un environnement comprenant deux serveurs (Apache et MySQL), un interpréteur de script (PHP), ainsi que phpMyAdmin pour l'administration Web des bases MySQL. [32]

3. Architecture technique

L'architecture adoptée pour la réalisation de notre système est l'architecture web 3 tiers illustrée dans la figure ci-dessous

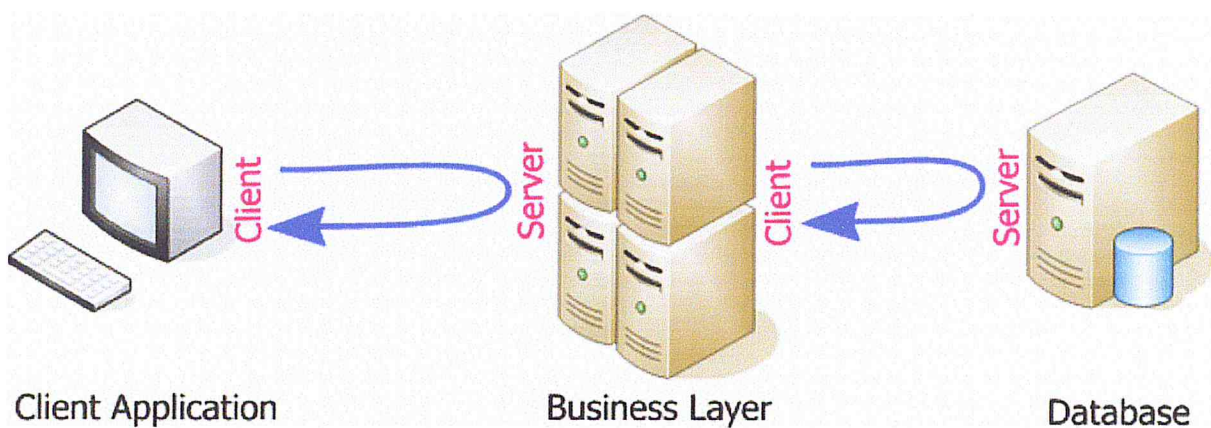


Figure 31 : Architecture technique du système.

Les avantages d'une architecture 3-tiers :

- ✓ Les requêtes clients vers le serveur sont d'une plus grande flexibilité que dans celles de l'architecture 2-tiers basées sur le langage SQL.
- ✓ Cette flexibilité permet à une entreprise d'envisager dans le cadre d'une architecture 3-tiers une grande souplesse pour l'introduction de toutes nouvelles technologies.
- ✓ D'un point de vue développement, la séparation qui existe entre le client, le serveur et le SGBD permet une spécialisation des développeurs sur chaque tiers de l'architecture.
- ✓ Plus de flexibilité dans l'allocation des ressources; la portabilité du tiers serveur permet d'envisager une allocation et ou modification dynamique aux grés des besoins évolutifs au sein d'une entreprise.

4. Réalisation de la solution

4.1. Construction de la zone d'entreposage

Elle s'agit de l'implémentation des Data Marts qui constituent le DW. Nous avons créé une base de données relationnelle conforme aux étoiles déjà conçues à l'aide de PHPMYADMIN.

Dans cette BDD, chaque fait du schéma en étoile correspond à une table appelée table de fait et chaque dimension correspond à une table appelée table de dimension. La table de dimension contient une clé primaire et un ensemble d'attributs. Tandis que la table de fait possède comme attributs les mesures d'activités et les clés étrangères vers les tables de dimension.

4.2. Le chargement des données

Le chargement consiste à alimenter le Data Mart.

Le processus d'alimentation s'est déroulé en trois phrases :

1. L'extraction des données.
2. Le nettoyage et la transformation des données.
3. Le chargement des données dans les tables de faits et les tables de dimensions.

Pour réaliser ce processus on a utilisé SSIS (SQL Server Integration Services). SSIS permet la mise en place de l'ETL.

Les figures ci-dessous illustrent la partie ETL de notre projet :

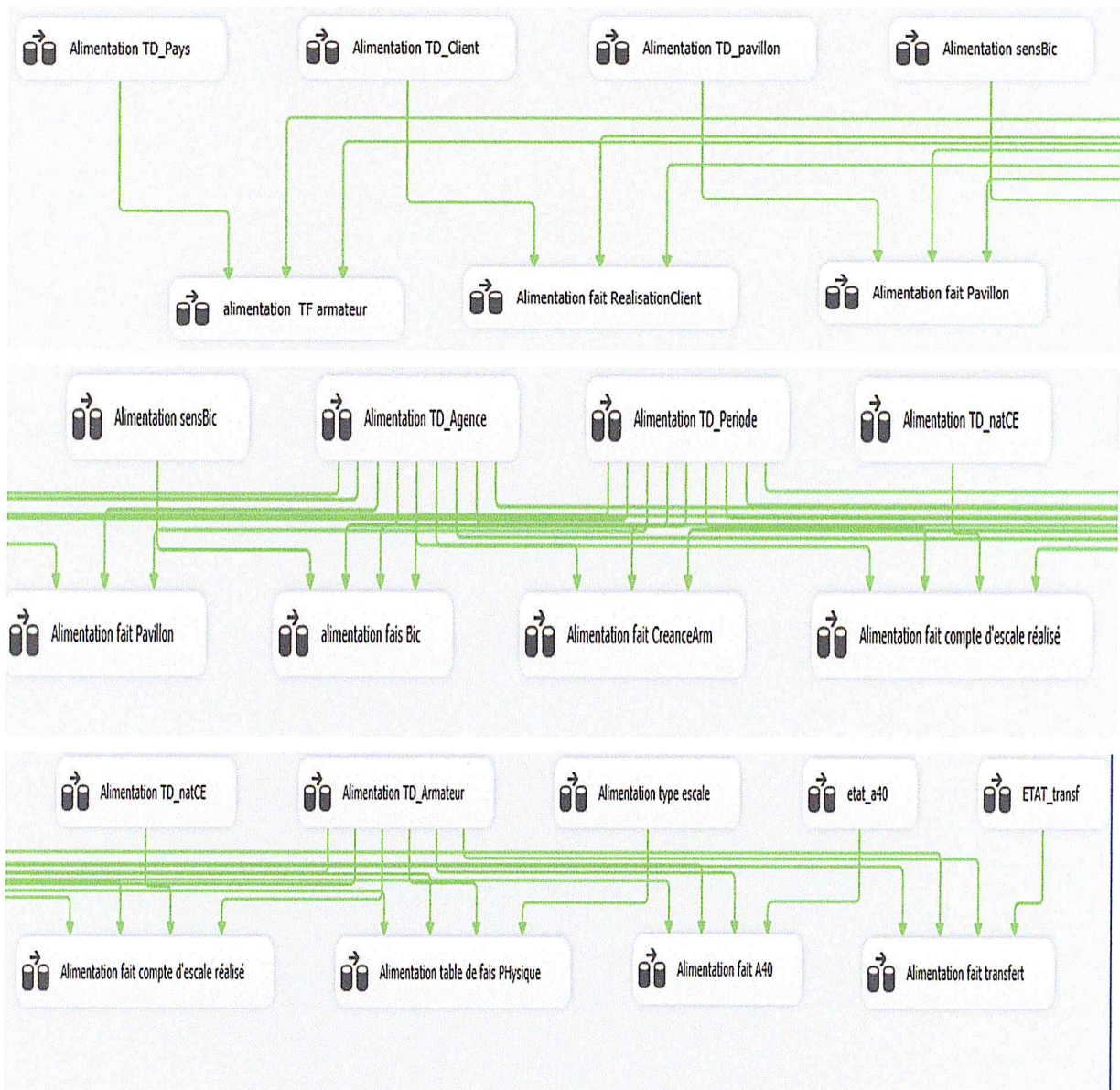


Figure 32: processus ETL de la solution

5. Scénario d'utilisation :

Dans ce qui suit nous allons démontrer avec des captures d'écran un scénario d'utilisation simple qui décrit ce que peut faire l'utilisateur actuellement avec notre solution.

Authentification

Premièrement l'utilisateur entre son nom d'utilisateur et mot de passe.

La figure de la page suivante représente une interface qui permet à l'utilisateur de se connecter au système.



Figure 33: Authentification

Résultats graphiques

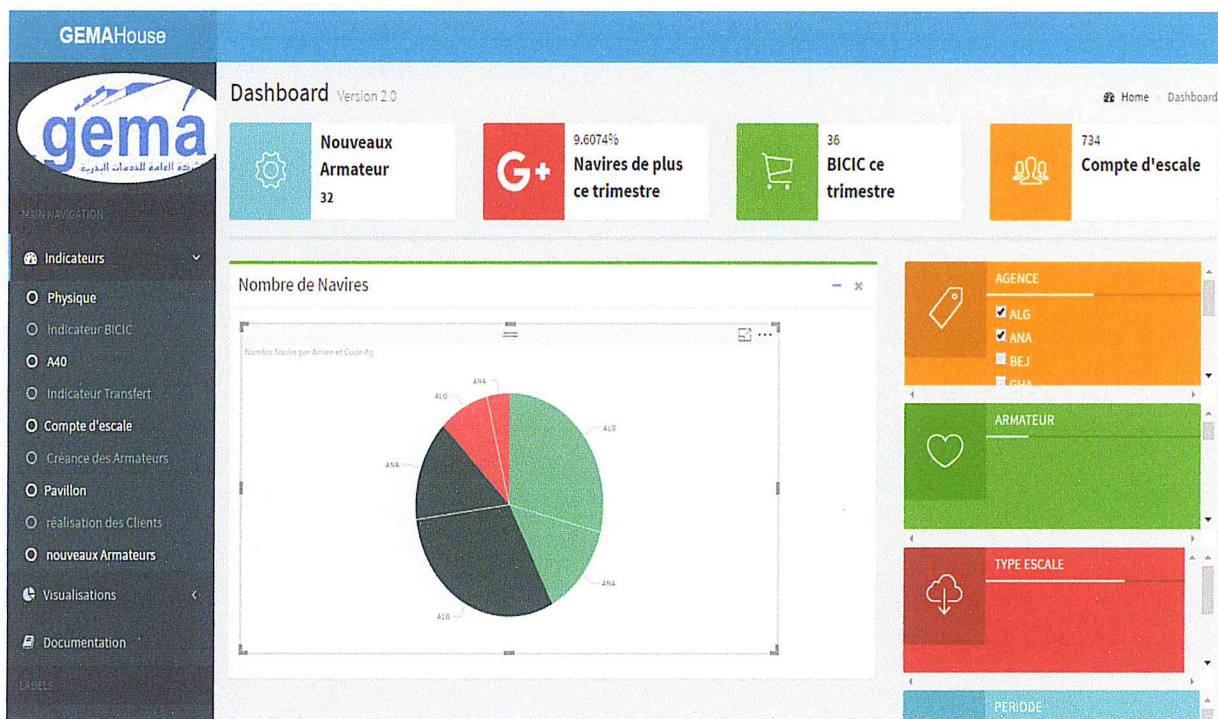


Figure 34: Exemple d'affichage graphique (en camembert) des données.

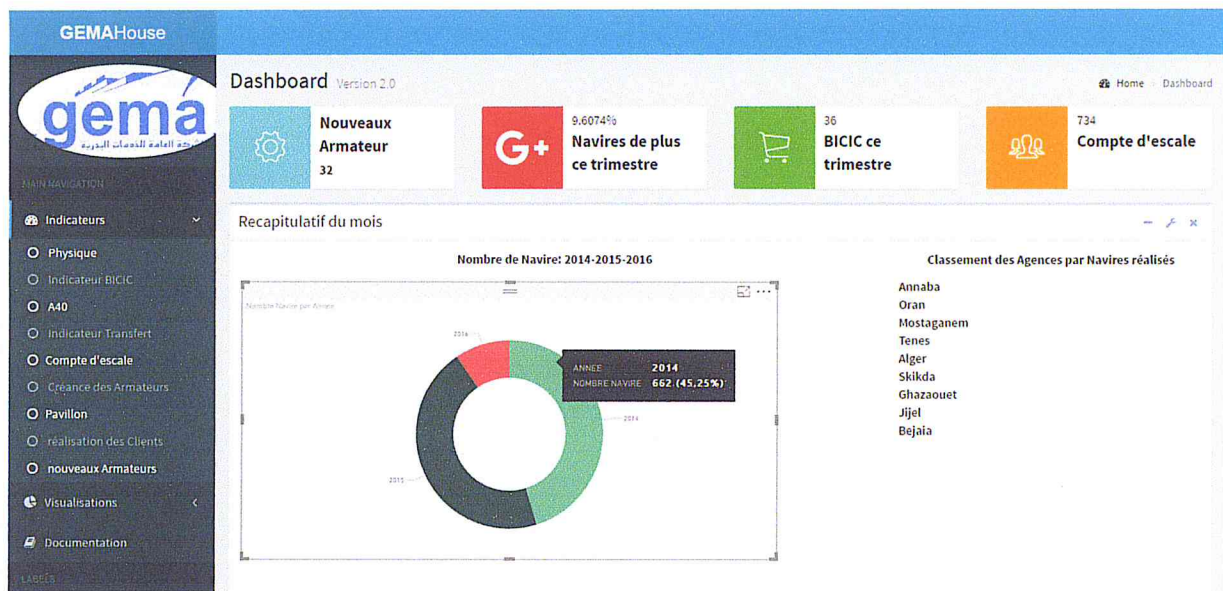


Figure 35 : Exemple d'affichage graphique des données

6. La sécurité du système

Vu l'importance capitale des informations que fournit notre système décisionnel, l'aspect sécurité ne doit en aucun cas être négligé :

- Le réseau intranet de l'entreprise est sécurisé.
- La gestion des habilitations du portail de restitution est assurée, cela permet de limiter l'accès aux décideurs.

Conclusion

Nous avons décrit à travers ce dernier chapitre la réalisation et le déploiement de notre système. Nous avons décrit le matériel, les outils utilisés ainsi que l'architecture technique et l'aspect sécuritaire de notre système.

Conclusion générale

L'informatique décisionnelle permet l'exploitation des données de l'entreprise dans le but de faciliter la prise de décision. Elle permet aussi la compréhension de la situation actuelle de l'entreprise et d'anticiper sa situation future.

A cet effet, et afin d'apporter un réel support à la prise de décision, GEMA a initié le projet : Réalisation d'un Data Warehouse pour la mise en place d'un système décisionnel.

Ce projet fut une nécessité, plus qu'une option dans un marché de consignment aussi riche en opportunités que fort par sa concurrentialité. GEMA avait et a pour priorité de maintenir son taux de croissance et augmenter sa part de marché.

Le travail qui nous a été confié était de contribuer à la conception et au déploiement de ce système. La méthodologie utilisée suit l'approche du cycle de vie dimensionnel proposée par Ralph Kimball. Cette approche se constitue de trois phases : définition des besoins, conception du système et déploiement du système.

Au début, et pour répondre aux besoins des utilisateurs, il faut évidemment connaître et déceler leurs besoins, dans cette quête, les entretiens ont été l'outil principal de la collecte, complétés par des questionnaires.

Nous avons entamé la partie de conception de notre magasin de données. Nous avons proposé une modélisation dimensionnelle en étoile. Cette dernière est basée sur le concept de fait et de dimension. Notre modèle offre comme avantage majeur d'être simple et compréhensible par les utilisateurs finaux.

Une fois les modèles dimensionnels conçus, nous sommes passé à l'alimentation de l'entrepôt de données, l'étape la plus coûteuse en temps, où il a fallu créer plusieurs mappings d'extraction, transformation et chargement de données, par le biais d'outils professionnels certifiés.

Tout comme l'alimentation de l'entrepôt, le chargement des cubes et dimensions de nos Data marts, s'est fait avec des outils professionnels. Les cubes étant chargés et prêts à être interrogés.

Avec la mise en place de cette solution, nous avons réussi à centraliser les données nécessaires à la prise de décision pour un accès plus rapide et simple

Comme un projet informatique n'est jamais complètement terminé, nous pouvons citer les perspectives et développements suivants :

- Suivre le déploiement actuel et effectuer un recueil des remarques des utilisateurs afin d'apporter les corrections nécessaires.

- Réalisation de Data marts pour les autres services (exemple : Ressources Humaines) dans le cadre du projet de GEMA pour la mise en place du SID.
- Utilisation des méthodes et algorithmes de Data Mining pour une meilleure exploitation des données.

Avant de terminer, nous tenons à dire, que ce stage nous a été véritablement bénéfique, nous avons pu mettre en pratique certaines de nos connaissances théoriques acquises au cours de notre cursus universitaire, dans un environnement professionnel, GEMA. Tout cela nous a permis d'acquérir beaucoup d'expérience, qui nous sera très utile par la suite, que ça soit dans le domaine de l'informatique décisionnelle ou autre

Bibliographie

- [1] Fernandez, Alain. 2013. Les nouveaux Tableaux de bord des managers. Eyrolles, 2013.
- [2] Bernard, Duverneuil. Business Intelligence. 2009
- [3] Kimball, Ralph et al. The datawarehouse toolkit. Wiley. 2013.
- [4] Elisabeth, METAIS (page consultée le 22 Février 2016) systèmes informatiques – systèmes d'aide à la décision, Encyclopedia Universalis [en ligne]. Adresse URL=<https://www.universalis.fr/encyclopedie/systemes-informatiques-systemes-d-aide-a-la-decision/>
- [5] Bala, M. « chapite 1 », cours de SID, 2016
- [6] FILALI ABDERRAHMANE, KEDJANE SOFIANE Conception et réalisation d'un Data Warehouse pour la mise en place d'un système décisionnel, mémoire de fin d'étude 2010
- [7] Bill, Inmon. Building the Data Warehouse, John Wiley & Sons, Inc. New York, 2002. 428p
- [8] <http://slideplayer.fr/slide/1159105>
- [9] stat4u. (Page consultée le 10 Mars 2016) Data Warehouse [en ligne]. Adresse URL=<http://datawarehouse4u.info>
- [10] [http:// searchbusinessintelligence.techtarget.in/tip/inmon-vs-kimball-approach-is-suitable-for-your-data-warehouse](http://searchbusinessintelligence.techtarget.in/tip/inmon-vs-kimball-approach-is-suitable-for-your-data-warehouse)
- [11] https://www.dwfacile.com/approches_dm.html
- [12] <http://www.imc.com/services/enterprise-data-warehousing/etl-processmanagement>
- [13] Boussaid, O. Le processus ETL. 2015
- [14] Kimball, Ralph et al. Le data warehouse, guide de conduite de projet, Edition Eyrolles, 2005.
- [15] R. Kimball et M. Ross. Entrepôts de données, guide pratique de modélisation dimensionnelle, Edition Vuibert, 2002.
- [16] Lydie, Soler. Les entrepôts de données, U.F.R. d'informatique, 2008.

- [17] Microsoft. (Page consultée le 20 Mars 2016) [en ligne]. Adresse URL=<https://msdn.microsoft.com/fr-fr/library/ms175641.aspx>
- [18] Michael Tranchant (page consulté le 30 Mars 2016) Qu'est-ce que l'informatique décisionnelle ? [en ligne]. Adresse URL = <http://business-intelligence.developpez.com/tutoriels/quest-ce-que-la-bi/>
- [19] Karim SLAIMI. (Page consultée le 11 Mars 2016). Informatique décisionnelle(BI) : OLAP et la modélisation dimensionnelle [en ligne]. Adresse URL= http://www-igm.univ-mlv.fr/~dr/XPOSE2009/informatique_decisionnelle_olap/index.html
- [20] C. Vangenot. Datwarehouse. Laboratoire de base de données
- [21] Cécile Favre. "Evolution de schémas dans les entrepôts de données : mise à jour de hiérarchies de dimension pour la personnalisation des analyses". Thèse de doctorat, Université Lumière-Lyon2, Lyon, France, 2007.
- [22] Yazid GRIM, (page consultée le 05 Mars 2016). «OLAP, les fondamentaux», [en ligne]. Adresse URL= <http://grim.developpez.com/articles/concepts/olap/>
- [23] Nakache, « Data Warehouse et Data mining », conservatoire national des arts et métiers Lille, 1998.
- [24] M. Poletto, « Quelles sont les bonnes pratiques dans la mise en place d'une solution décisionnelle et comment la maintenir en condition opérationnelle ?, » doctorat professionnelle, Département Informatique, Ecole Supérieure D'Informatique (CESI Exia) France, 2012.
- [25] Chikhi, I. »chapitre2 », cours de SID, 2015.
- [26] Parmenter, David. Key Performance Indicators: Developing, implementing ans using Winning KPIs. 2011.
- [27] FRANCO J, DE LINGNROLLES S. Piloter l'entreprise grâce au data warehouse. 2^{ème} Edition, France : Eyrolles, 2001.
- [28] SAPORTA G. Introduction au Data Mining et à l'apprentissage statistique. Paris : Chaire de Statistique Appliquée & CEDRIC, CNAM, 2012.
- [29] Manuel Qualité (document de GEMA) disponible sur le site : www.gema-group.com

[30] SQL Server Integration Services (page consultée le 20 aout 2016). Adresse URL=
[https://msdn.microsoft.com/fr-fr/library/ms141026\(v=sql.120\).aspx](https://msdn.microsoft.com/fr-fr/library/ms141026(v=sql.120).aspx)

[31] PHP Storm (page consultée le 20 aout 2016). Adresse URL=
<https://fr.wikipedia.org/wiki/PhpStorm>

[32] Wamp Server (page consultée le 20 aout 2016). Adresse URL=
<http://www.wampserver.com/>

Annexes

Annexe A

1. Organigramme de GEMA

L'organigramme suivant donne une vision claire et précise de la structure de l'entreprise

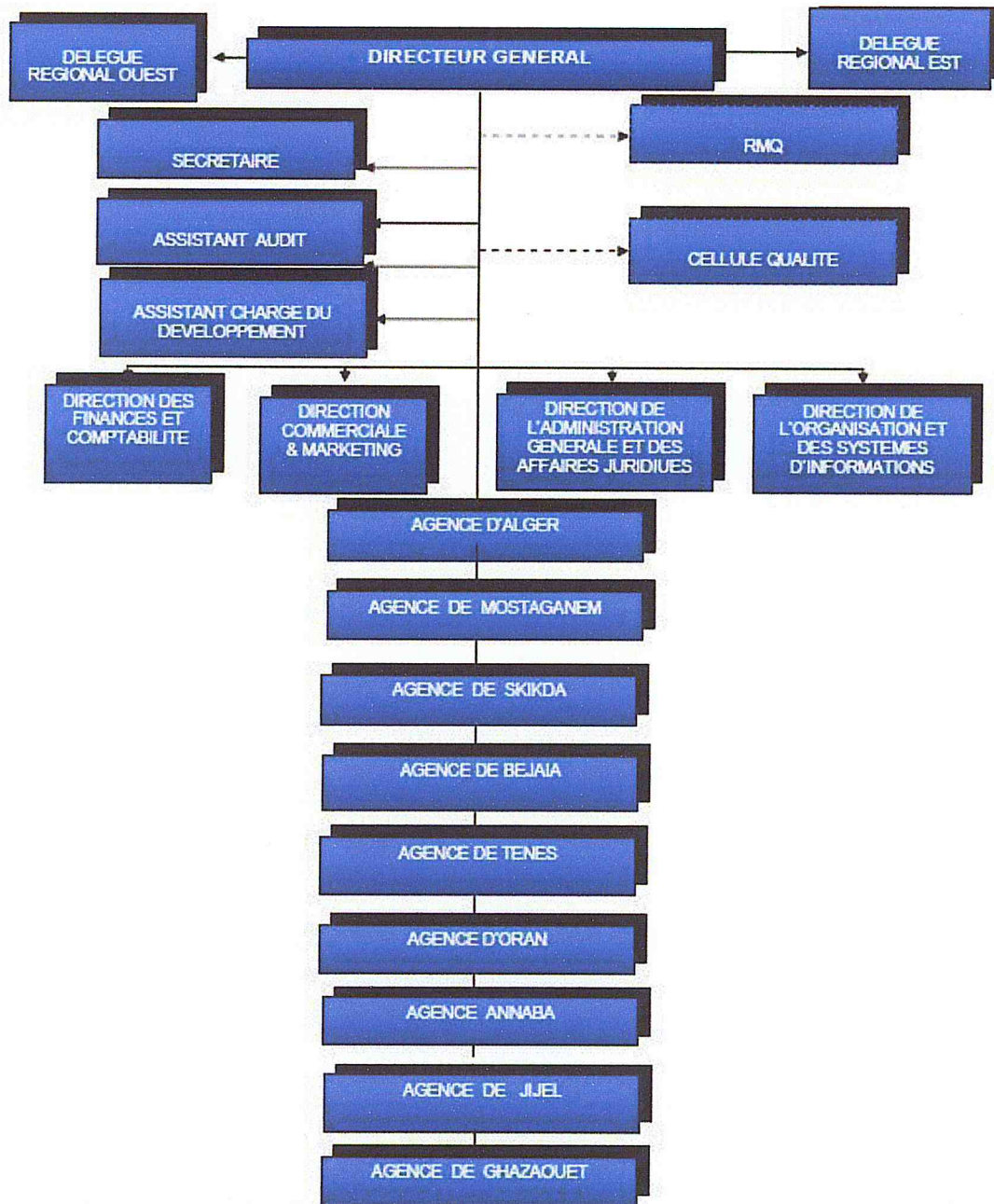


Figure 36: organigramme de GEMA

2. Les sociétés participations dans GEMA

❖ Algerian ligabe catering (SPA)

Société de catering et restauration collective, dotée d'un capital social de 100 millions de DA détenu à hauteur de 35% par GEMA.

❖ **Comarpex Algérie (SPA)**

Société d'Avitaillement des navires et ravitaillement des collectivités, de l'importation et l'exportation des produits alimentaires. Elle est dotée d'un capital social de 100 millions de DA, détenu à hauteur de 21% par GEMA.

❖ **SOGRAL (SPA)**

Société d'exploitation de la gare routière d'Alger dotée d'un capital social de 90 millions de DA, détenu à hauteur de 36% par GEMA.

❖ **La Société d'Investissement Hôtelière -S.I.H - (SPA)**

Société de gestion hôtelière (hôtel Sheraton), dotée d'un capital social de *16.042* millions de DA, détenu à hauteur de 6,42% par GEMA.

3. Définition de la consignation et des activités liés au domaine maritime

Le code maritime Algérien définit l'agent consignataire comme un mandataire de l'armateur. Ainsi, l'agent consignataire est mandaté par un Armateur pour assister le navire en escale dans ses rapports avec les autorités et administrations locales, avec les transitaires, les chargeurs et les réceptionnaires. Son rôle consiste, également, à assister le navire pour tous les événements pouvant surgir lors de l'escale.

Sur le plan commercial et financier et pour compte de ses mandants, il établit le pro forma des comptes d'escales, encaisse les frets, confectionne les comptes d'escales, recrute le fret et défend la notoriété de ses armateurs.

Il peut également, à la demande du réceptionnaire ou dans le cadre d'une convention, assister ce dernier et agir pour son compte en mettant à disposition tous les moyens humains et matériel nécessaire pour assurer le bon déroulement de l'escale (tramp régulier).

4. Politique de tarification de GEMA

Les services rendus par GEMA font l'objet d'une tarification sur une pro forma communiquée aux clients. Le calcul se fait en fonction de la fréquence d'escale du navire et de sa jauge nette.

La politique de GEMA en matière de tarification se base sur la hiérarchisation de la clientèle en ce sens que les clients assidus bénéficient de meilleurs tarifs.

Annexe B

Le questionnaire effectué lors de la collecte des besoins

Questionnaire pour notre encadreur :

1. Qui vont être les utilisateurs finaux de notre système ?
2. A quoi doit permettre notre système ? (objectifs qui dépendent des utilisateurs finaux)

Questionnaire pour les utilisateurs finaux :

- a) Pour connaître l'environnement concurrentiel et la gestion des clients :
 1. Quelles sont les principaux concurrents de GEMA ?
 2. Quelle est la part de marché de GEMA ?
 3. Quelles méthodes utilisez-vous pour faire face à la concurrence ?
 4. Que fait GEMA pour attirer plus de clients ? Et par quel moyen ?
 5. Quelle stratégie suit GEMA pour la fidélisation des clients ?
- b) Pour connaître notre organisme d'accueil GEMA :
 1. Quelles sont les missions de cette direction ?
 2. Quels sont les objectifs de votre secteur d'activités ?
 3. Quels types d'analyses régulières effectuez-vous actuellement ?
 4. Quels problèmes affrontez-vous principalement aujourd'hui ? (au niveau des services commercial et marketing)
 5. Quels types d'analyses que vous souhaitez effectuer et vous ne pouvez pas le faire maintenant ?
 6. Qu'est-ce qui vous empêche d'atteindre les objectifs ?
 7. Quelles sont les données que vous souhaitez historier ? autrement quelles sont toutes les informations dont vous devez/souhaitez disposer pour faire vos planifications ?
 8. Quels types de décision vous prenez et qui nécessite une bonne connaissance de la situation ?
 9. Quel est votre manière d'accès à l'information ?
 10. Est-ce que la procédure actuelle de prise de décision est satisfaisante ? Si non, Quels sont les manques ?
 11. De quelle manière envisagez-vous une meilleure exploitation de l'information ?
 12. Quels sont vos attentes du nouveau système ?