

**République Algérienne Démocratique et Populaire**  
Ministre de l'Enseignement Supérieur et de la Recherche Scientifique

**Université Mouloud Mammeri de Tizi-Ouzou**



**FACULTE DE GENIE ELECTRIQUE ET D'INFORMATIQUE**  
**Département Informatique**

## **Mémoire de Fin d'Etude De MASTER**

**Domaine : Mathématiques et Informatique**

**Filière : Informatique**

**Spécialité : Ingénierie des Systèmes d'Information**

Présenté par:

**KAHINA ARAOUR**

### **Thème**

**Conception et réalisation d'un système d'aide à la  
décision pour le service marketing de BMT-SPA  
pour le port de Bejaia.**

*Mémoire soutenu publiquement le 25/09/ 2016 devant le jury composé de:*

**Présidente : M<sup>elle</sup> Yesli Yasmine.**

**Encadreur : M<sup>me</sup> Bouarab-Dahmani Farida**

**Examinateur : M<sup>r</sup> Yacine Younes.**

**Examinateur : M<sup>r</sup> Radja Hakim.**

**Promotion 2015/2016**

# **Remerciements**

*De prime abord, je tiens à remercier le Bon Dieu tout puissant de m'avoir donné énormément de patience, de courage et de volonté pour réussir mon mémoire.*

*Je tiens à remercier tous les enseignants du département informatique d'avoir assuré ma formation durant mon parcours universitaire. En particulier ma promotrice Mme Bouarab Farida chercheuse chevronnée de l'Université Mouloud Mammeri, de m'avoir fait l'honneur de m'encadrer et m'encourager et aussi pour ses critiques constructives, sans lesquels ce travail n'aurait pu aboutir.*

*J'adresse mes vifs remerciements à mon promoteur Mr Boumerzoug Moussa pour son attention et la confiance qu'il m'a accordé, aussi de m'avoir donné l'opportunité de travailler sur un sujet d'une telle envergure, ainsi qu'à toute l'équipe du service marketing de BMT.*

*Je tiens aussi à remercier vivement Mlle HADDADI Lynda pour son aide, son soutien, ses conseils, qui m'y étaient vraiment bénéfiques, malgré sa surcharge de travail.*

*Enfin, Je remercie les membres du jury d'avoir accepté d'évaluer mon modeste travail, ainsi que toutes les personnes qui ont contribué de près ou de loin à l'élaboration de ce mémoire.*

## Dédicaces

*Je dédie ce travail,  
À ma formidable mère :*

*À celle qui a été l'école de ma vie,  
Qui n'a jamais cessé de veiller sur moi,  
En m'inculquant les notions d'une femme kabyle ;*

*À mon adorable père :*

*Qui était comme un précieux dictionnaire,  
Pour les connaissances qu'il m'a prodigué,  
Pour son soutien, ses sacrifices qu'il a fait pour ma réussite ;*

*À mon cher frère :*

*À mon inspiration, mon modèle de réussite,  
À celui qui m'a toujours encouragé ;*

*À ma petite sœur :*

*À celle que je serai un spécimen d'une femme,  
Kabyle digne, brave, intelligente ;*

*À ma chère tante Salîha :*

*Ainsi que mes adorées sassa, kami, surtout Wardia,  
À celles que je ne remercier assez jamais,  
À tous les membres de ma famille en particulier mon cousin M'hand ;  
À celui qui m'a tellement aidé et soutenu*

*À tous mes amis surtout Lîla B, Chouchou pour laquelle  
Je souhaite Un prompt rétablissement ;*

*À tous les militants de l'identité berbères...*

*À la mémoire du rebelle Matoub LOUNES*

# Glossaire des termes maritimes

## A

---

**Accorage:** Opérations de transport, de chargement ou de déchargement des marchandises, faite aux moyens d'accions entre le navire en mer et le quai, et vice versa.

**Accostage :** Manœuvre d'approche finale du navire à quai conçu pour permettre le stationnement des navires, leur amarrage et la manutention.

**Amarrage :** Immobilisation d'un navire au moyen d'aussières (câbles) à un quai ou une bouée.

**Armateur :** Personne qui arme un navire en lui fournissant matériel, équipage et tout ce qui est nécessaire à la navigation.

**Arrimage :** Opération qui consiste à fixer solidement les marchandises à bord du navire.

## B

---

**Bassin :** Partie d'un port, limitée par des quais, des digues, destinée à recevoir les bâtiments pour les opérations de chargement et de déchargement.

**Bollard :** Gros fût métallique à tête renflée implanté sur un quai pour l'amarrage des navires.

**Booking note (Engagement de fret) :** Accord, signé par les parties, par lequel un chargeur s'engage à remettre une marchandise à un armateur ou à son agent et celui-ci s'engage à l'embarquer.

## C

---

**Cabotage :** Se dit de navires navigants sur des zones définies, en principe limités à un pays. Le cabotage se distingue ainsi de la navigation au long cours et de la navigation au bornage.

**Capitainerie :** Service dépendant de l'autorité portuaire et chargé de coordonner les mouvements de navires dans le port et de la police.

**Chargeur :** Personne qui expédie de la marchandise sur un navire. Personne physique ou morale (propriétaire ou non de la marchandise) qui conclut le contrat de transport et remet, dans la plupart des cas, la marchandise au transporteur.

**Consignataire** : Le consignataire de navire est la personne physique ou morale qui est chargée, soit à titre permanent, soit à titre occasionnel, d'effectuer, au nom et pour le compte de l'Armateur et/ou du transporteur maritime, tant au port d'embarquement qu'au port de débarquement du navire qui lui est consigné.

**Conteneur** : Engin de transport généralement de forme parallélépipédique conçu pour faciliter le transport sans rupture de charge et muni de dispositifs rendant sa manipulation aisée. Les conteneurs normalisés mesurent 20 et 40 pieds de long, soit une capacité respective de 30 et 60 m<sup>3</sup>.

## D

---

**Défense à quai** : Protection destinée à amortir les chocs entre un navire et un quai auquel il est accosté.

**Dépotage** : Manutention consistant à sortir des marchandises d'un conteneur.

**Despatch Money**: Prime payée par l'armateur pour tout temps gagné dans les opérations d'embarquement ou de débarquement d'un navire.

**Digue** : Construction établie devant un port, n'ayant qu'une fonction de protection contre les vagues et courants de marée.

**Duc d'Albe** : Faisceau de poteaux de bois, de tubes d'acier ou bloc de ciment ancrés dans le fond des bassins, sur lequel un navire peut s'amarrer ou s'appuyer.

## E

---

**E.V.P (Équivalent Vingt Pieds)** : en anglais TEU ou Twenty Equivalent Unit. Unité de mesure uniformisée permettant de répertorier les conteneurs selon leur taille et de décrire les capacités des navires porte-conteneurs ou des terminaux. Un conteneur ISO de 20 pieds correspond à 1 EVP.

**Empotage**: Manutention consistant à mettre des marchandises dans un conteneur.

**Escale** : En parlant d'un navire, action de se ravitailler quelque part, d'y faire des opérations commerciales, d'y débarquer ou d'embarquer du fret ou des passagers.

## F

---

**Fret (freight)** : Marchandise que transport le navire, chargement commercial d'un navire, mais aussi le prix du transport proprement dit (taux de fret).

## G

---

**Gare maritime** : Est un quai aménagé pour l'embarquement et le débarquement des passagers.

## H

---

**Handling charges** : Frais consécutifs aux différentes manipulations subies par les conteneurs dans les dépôts ou terminaux.

## M

---

**Manifeste** : Document douanier qui récapitule toutes les marchandises embarquées ou débarquées lors d'une escale.

**Mouillage** : Opération consistant à jeter l'ancre en laissant filer la chaîne de façon à faire crocher l'ancre dans le fond.

## N

---

**Navire** : Bateau destiné à la navigation maritime.

**Numéro d'escale** : Numéro séquentiel attribué à un navire par la direction capitainerie, dès son entrée au port.

## P

---

**Pilotage**: Assistance fournie par un pilote au capitaine de navire pour entrer ou sortir un navire du port.

**Port** : Lieu aménagé pour le transport de personnes ou de marchandises entre la terre et la mer ou le cours d'eau.

**Portique**: Engin destiné à la manutention en bord à quai, il est équipé d'un chariot translatant perpendiculairement au quai et auquel est suspendue la charge.

## R

---

**Rade** : Plan d'eau ayant un accès à la mer et pouvant servir de mouillage.

**Radoub** : Bassin aménagé pour exécuter à sec les réparations sur les navires.

## S

---

**Shift** : est un travail portuaire sans interruption, qui fait référence à des mouvements ou des positions changeantes de marchandises d'un endroit à un autre.

## T

---

**Terminal à conteneur** : est une infrastructure portuaire spécialisée dans le chargement et déchargement des conteneurs transportés par les porte-conteneurs.

**Terre-plein** : est une étendue artificielle de terre acquise sur la mer par terrassement, à l'aide souvent de déchets et de sable, et qui a un niveau altimétrique supérieur au niveau de la mer.

**Transitaire** : est un auxiliaire très important du commerce extérieur. Il intervient dans la chaîne du transport.

# Table de matière

## Sommaire

Partie 01 .....	6
Chapitre 01: Le Système Décisionnel .....	6
Introduction .....	7
1.1. Définition de la décision et ses types: .....	7
1.2. Définition du Système décisionnel .....	9
1.3. Place d'un système décisionnel dans l'entreprise.....	9
1.4. Notion de décideur.....	10
1.5. Contexte du Système décisionnel .....	11
1.6. Architecture d'un Système décisionnel .....	12
1.7. Système décisionnel et fonction marketing .....	13
1.8. Le défi du Système décisionnel .....	14
Conclusion.....	16
Chapitre 02: Les entrepôts de données .....	17
Introduction .....	17
2.1. Définition d'un entrepôt de données .....	17
2.2. entrepôt de données vs Base de données relationnelle .....	19
2.3. Composants de base d'un entrepôt de données: .....	20
2.4. Data Mart.....	22
2.5. entrepôt de données VS magasin de données .....	23
2.6. Objectif d'un entrepôt de données .....	24
2.7. Structure d'un entrepôt de données .....	24
2.8. Conduite et construction d'un entrepôt de données.....	25
Conclusion.....	32
Chapitre 03 : Modélisation dimensionnelle. ....	33
Introduction .....	33
3.1. Définition de la modélisation dimensionnelle .....	33
3.2. Différentes modèles de modélisation dimensionnelle .....	35
3.3. Caractéristiques des dimensions .....	36
3.4. Cube de données .....	37
3.5. Concept OLAP .....	37

3.6. Architecture des systemes OLAP .....	39
Conclusion.....	41
Synthèse de la première partie .....	42
Partie 02.....	44
Chapitre 4: Cas pratique de la BMT-spa.....	44
Introduction .....	44
4.1. Presentation de l'organisme d'accueil .....	44
4.2. Situation géographique .....	45
4.3. Organigramme de BMT: [document interne de l'entreprise] .....	46
4.4. Différentes structures de BMT .....	47
4.5. Les activités de BMT .....	50
4.6. Objectifs de BMT .....	51
Conclusion.....	52
Chapitre 5: Analyse de l'existant. ....	53
Introduction .....	53
5.1. Situation informatique de BMT .....	53
5.2. Etat du système décisionnel actuel de BMT .....	59
5.3. Critique de l'existant .....	62
5.4. Obstacles rencontrés lors du recueil d'information .....	68
Conclusion.....	69
Chapitre 6: Analyse de besoins.....	70
Introduction .....	70
6.1. Présentation de la démarche à suivre.....	70
6.2. Présentation de la démarche d'identification des besoins .....	71
6.3. Récapitulatif sur les besoins recueillis.....	72
6.4. Vue générale de la solution proposée .....	74
Conclusion.....	74
Synthèse de la deuxième partie .....	75
Partie 03.....	77
Chapitre 07: Conception d'Entreposage. ....	77
Introduction .....	77
7.1. Conception de la zone d'entreposage .....	77
7.2. Matrice de bus de données.....	78

7.3. Modélisation dimensionnelle des activités .....	79
Conclusion.....	87
Chapitre 8: Conception de la zone d'alimentation de données. ....	88
Introduction .....	88
8.1. Etude de source de donnée .....	88
8.2. Processus d'alimentation de l'ED.....	90
Conclusion.....	94
Synthèse de la troisième partie .....	95
Partie 04.....	98
Chapitre 9 : Réalisation et mise en œuvre .....	97
Introduction .....	97
9.1. Outils utilisés .....	97
9.2. Réalisation de la solution.....	99
9.3. Architecture technique de la solution .....	105
Conclusion.....	105
Bibliographies .....	108
Webographie .....	109
Annexe 1: schéma descriptif de la réglementation de facturation.....	110
Annexe 2 : Suivi de procédure tarifaire pour chaque type de prestation .....	111
Annexe 3: Diagramme de GANTS .....	115
Annexe 4 : Listes des questions pour entretiens .....	116
Annexe 5 : Etude comparative entre les outils ETL/ talend et pentaho. ....	114

# Liste des figures

Figure 1 : Pyramide des catégories des décisions .....	8
Figure 2 : Les systèmes d'information selon les niveaux hiérarchiques.....	10
Figure 3 : Type de décideurs dans les organisations.....	11
Figure 4 : Architecture d'un système décisionnel .....	13
Figure 5 : Entrepôt de donnée orienté objet .....	18
Figure 6 : Entrepôt de donnée intégrée. ....	18
Figure 7 : Entrepôt de donnée non volatile. ....	18
Figure 8 : Entrepôt de donnée historisées. ....	19
Figure 9 : Architecture de l'entrepôt de données .....	20
Figure 10 : Schéma des data marts. ....	22
Figure 11 : Structure de données d'un entrepôt de données .. .	24
Figure 12 : Architecture en bus dimensionnel .....	26
Figure 13 : Architecture du DW selon Inmon.....	27
Figure 14 : Schéma de système d'alimentation ETL .....	29
Figure 15 : Exemple de table des faits .....	34
Figure 16 : Table de dimension.....	34
Figure 17 : Modèle en étoile (Star schéma). ....	35
Figure 18 : Schéma de modèle en flocon. ....	36
Figure 19 : Un cube de données. ....	37
Figure 20 : Architecture ROLAP .....	40
Figure 21 : Architecture MOLAP .....	40
Figure 22 : Situation géographique de BMT.....	45
Figure 23 : Organigramme de BMT-SPA .....	46
Figure 24 : Organigramme de la direction de Marketing.....	50
Figure 25 : Structure réseau de BMT-spa. ....	54
Figure 26 : Processus métiers de prise de décision .....	61
Figure 27 : Diagramme d'activité d'élaboration des rapports.....	62
Figure 28 : La table VESSEL_CALL. ....	64
Figure 29 : La table facture. ....	64
Figure 30 : La table CNTR et la table ARR_BAYPLAN. ....	65
Figure 31 : La table VESSEL_CALL .....	66
Figure 32 : La table ARRBAYPLAN. ....	66
Figure 33 : La table balance et la table edition1.....	67
Figure 34 : Cycle de vie d'un projet décisionnel . ....	70
Figure 35 : Vue de la solution proposée.....	74
Figure 36 : Schéma en étoile de l'activité suivi des clients .....	82
Figure 37 : Schéma en étoile de l'activité suivis des compagnies maritimes. ....	84
Figure 38 : Schéma en étoile fait suivis des navires. ....	86
Figure 39 : Exemple de table vide.....	91
Figure 40 : Suppression des champs vides.....	92
Figure 41 : L'ensemble des doublons dans la table ARR_BAYPLAN. ....	92

Figure 42 : La table ARR_BAYPLAN après l'élimination des doublons.....	93
Figure 43 : Mapping de la dimension client.....	101
Figure 44 : Mapping da la table fait suivi client. ....	101
Figure 45 : Authentification.....	102
Figure 46 : Créer un nouveau rapport.....	103
Figure 47 : Rapport d'activité de suivis de compagnie maritimes sous format d'un tableau.	103
Figure 48 : Statique sur quantité d'EVP entre les compagnies maritimes .....	104
Figure 49 : Architecture technique du système proposé .....	105

## **Liste des tableaux**

Tableau 1 : Tableau comparatif entre une BDD relationnelle et un entrepôt de données.....	19
Tableau 2 : Tableau comparatif entre un DW et un DM.....	23
Tableau 3:Tableau comparatif entre Kimball et Inmon. ....	28
Tableau 4 : Comparaison entre OLTP vs OLAP.....	39
Tableau 6 : Description des fonctionnalités du CIM.....	58
Tableau 7:Tableau descriptif des fonctionnalités de GestCom. .....	58
Tableau 8:Tableau descriptif de la procédure d'extraction de données. ....	59
Tableau 9:Tableau descriptif de la procédure d'élaboration de rapports. ....	60
Tableau 10:Tableau sur les récapitulatif des besoins des utilisateurs. ....	73
Tableau 11:Matrice de l'architecture en bus dimensionnel. ....	79

# Liste des abréviations

- **BDD** : Base De Données
- **BI** : Business Intelligence.
- **ED** : Entrepôt de données
- **BMT** : Bejaia Méditerranéen Terminal
- **ETL** : Extraction/ Transformation/ Loading
- **DW** : Data Warehouse
- **DM** : Data Marts.
- **HOLAP** : Hybrid On-Line Analytic Processing
- **IDE** : Integrated Development Environment
- **MOLAP** : Multidimensionnel On-Line Analytic Processing
- **MHT** : Montant Hors Taxe.
- **OLTP** : On-Line Transaction Processing
- **OLAP** : On-Line Analytic Processing
- **ROLAP** : Relational On-Line Analytic Processing
- **SGBD** : Système de Gestion de Base de Données
- **SI** : Système d'Information
- **SID** : Système d'Information Décisionnel
- **TB** : Tableau de bord
- **TOS** : Talend Open Studio
- **CA** : Chiffre d'affaire.
- **CTMS** : Container Terminal Management System.
- **WIMAX** :World Wide Interoperability for Microwave Access

# Introduction générale

---

## Contexte général de l'étude

Face à la mondialisation, les entreprises se trouvent confronter à un marché fortement concurrentiel qui évolue rapidement. Pour acquérir de nouvelles parts de marché, et faire face à leurs concurrents, le vrai challenge auquel se confrontent ces entreprises n'est plus le stockage des informations mais plutôt l'exploitation efficace d'une masse de données à des fins analytiques.

Cependant, les décideurs doivent prendre des décisions stratégiques en analysant un grand volume de données historisées. Ces données dans la plupart de temps sont non organisées, prévenant de multiples systèmes opérationnels ce qui rend la difficulté de **prise de décision** en temps voulu.

Les systèmes opérationnels s'avèrent inadaptés à une telle activité [Kimball 1996]. Pour pallier à ce problème, les systèmes décisionnels ont été développés, consciente de son importance, les entreprises s'engageants afin de se munir d'un tel système qui leur permettent d'avoir une vue synthétique de leurs activités. Ainsi avoir l'opportunité d'optimiser le pilotage de leurs activités et d'anticiper sur les évolutions de leurs marchés ainsi que sur le comportement de leurs clients.

La plupart des systèmes décisionnels reposent sur un espace de stockage centralisé, appelé entrepôt de données, dont le rôle est de stocker et d'intégrer des données utiles aux décideurs et de conserver l'historique pour des fins analytiques. Et dans le but d'exploiter ces données des applications finales comme le Reporting ou les tableaux de bord sont indispensables. Ils permettent de mettre à la disposition de tous les décideurs les données nécessaires pour la prise de décision.

C'est dans ce contexte que la BMT-spa (Bejaia Méditerranéen Terminal) une filiale de l'entreprise portuaire de Bejaia(EPB) souhaite améliorer sa procédure de prise de décision en mettant en place un projet de telle envergure qui est le Business Intelligence afin de rester l'un des meilleurs terminaux à conteneurs en Algérie.

# Problématique

BMT est constitué de divers intervenants interdépendants entre eux (consignataires, transitaires, douanes, armateurs...etc.) et de directions (direction des Operations, direction marketing ...etc.) qui veillent au bon déroulement de ses activités ainsi que la continuité de son service. Les diverses opérations de gestion (manutention/acconage) effectuées sur le navire est la mission principale de BMT. En effet, le souci majeur de cette dernière est de s'assurer du bon déroulement de ces opérations dans les meilleures conditions de délai, et de coût.

Dans le but de fidéliser ses clients et d'accroître son chiffre d'affaire, la direction de marketing compte renforcer sa stratégie de travail afin d'améliorer sa prestation du service, faire face à ses concurrents et être leader dans son marché. Cependant, cette entreprise est exposée quotidiennement à certaines difficultés que nous avons pu constater lors de notre étude, principalement au niveau de la direction marketing. Ces difficultés peuvent se résumer en :

- **Non fiabilité du système d'information actuel de la BMT** : le système existant ne répond pas à tous les besoins internes des utilisateurs et ne couvre pas toutes leur demandes quotidiennes (par exemple : la saisie incorrecte des noms des importateurs par la direction des Operations ce qui engendre la difficulté de classement de ces derniers par le service de marketing).
- **Hétérogénéités de données** : les données sont stockées dans plusieurs bases de données éparpillées dans de différents systèmes transactionnels ce qui rend difficile l'accès à l'information en temps voulu et une mauvaise exploitation de cette dernière cela engendre de divers problèmes d'ordre technique (tels que : l'incohérence de donnée) et induit également mauvaise crédibilité de l'information.
- **Le grand volume de données** : les données traitées régulièrement rendent les matériels lents ce qui cause une perte de temps considérable ainsi qu'une diversité de leurs sources depuis les différents services engendrant la difficulté dans l'élaboration des rapports.
- **Une mauvaise qualité de donnée** : la saisie de données qui passe par plusieurs intervenants engendre une redondance de celle-ci. Le mode de contrôle est semi-automatique et influence sur l'exactitude de l'information.
- **Difficulté de regrouper des données** : et d'établir des analyses et des rapports d'activité.

- **La lenteur des procédures reporting :** l'élaboration des rapports d'activités ou d'analyses qui se fait actuellement est quasi manuelle prend beaucoup de temps lors de la génération des rapports d'activités.
- **Un risque d'erreurs élevé :** l'absence de d'intégrité et de système automatique de vérification de donnée augmente le risque d'erreurs.

En effet, cet ensemble d'anomalies constatées lors de notre étude ne permettent pas aux décideurs d'obtenir les informations nécessaires à la prise de décision au moment opportun. Ce qui engendre une perte financière considérable et une stratégie peu performante.

## Objectifs à atteindre

Afin de répondre à la problématique précédente, la BMT-spa décide de se lancer dans un projet de développement d'un système d'aide à la décision, afin de pouvoir satisfaire les besoins de ces décideurs. Dans le cadre de ce projet, notre travail consiste en premier lieu à la conception et la réalisation d'un entrepôt de données qui exploite les données stockées dans des sources différentes. Dans un second lieu, nous allons concevoir et réaliser un reporting qui renferme les indicateurs de performance les plus importants et qui facilite l'exploitation et la navigation dans un entrepôt de données réalisé.

Ce projet permet à l'entreprise de :

- ✓ Savoir la part de marché de chaque compagnie maritime.
- ✓ Savoir la part de marché de chaque client réalisée selon son chiffre d'affaire, ou le nombre d'EVP (Equivalent vingt pieds) ainsi que sa région, son secteur d'activité, et le type prestation.
- ✓ Présenter les données d'une manière cohérente afin d'assurer la lisibilité et une meilleure qualité de données.
- ✓ Consolider et regroupement de données afin de faciliter la procédure d'analyse.
- ✓ Améliorer le rendement en écourtant le temps d'escale.

# Organisation du mémoire

Notre mémoire est organisé comme suit :

## Partie 01 : Etude bibliographique

Cette première partie traite l'aspect théorique de notre travail. Elle est composée de trois chapitres :

- **Chapitre 01 : Systèmes décisionnels.** L'objet de ce chapitre est de donner un panorama du concept de système décisionnel, où nous détaillerons ses objectifs, son architecture, ainsi qu'une comparaison entre ce dernier et le système transactionnel.
- **Chapitre 02 : Entrepôt de données .** Dans ce chapitre, nous allons aborder la notion d'un entrepôt de données, où nous allons définir et discuter de ses objectifs. Nous allons aussi présenter les différents aspects relatifs à la structure de l'ED, ses composants de base ainsi que les différentes approches d'entreposage de données.
- **Chapitre 03 : Modélisation dimensionnelle.** Ce chapitre traite une méthode de conception logique qui vise à présenter les données sous une forme standardisée, qui est la modélisation dimensionnelle.

## Partie 02 : Analyse

Cette partie est consacrée à une présentation de l'environnement professionnel dans lequel nous avons mené notre stage, qui est la BMT (Bejaia Méditerranéen Terminal) .Il est divisé en trois chapitres :

- **Chapitre 04 : Cas pratique de la BMT-spa.** Dans ce chapitre, nous allons faire une présentation de son organisme d'accueil, son environnement, sa structure ainsi que ses activités.
- **Chapitre 05 : Etude de l'existant.** Nous allons, dans ce chapitre, nous concentrer plus particulièrement au service marketing qui constitue le cadre de notre étude, pour mieux comprendre son processus métier, ainsi que l'existant informatique au sein de cette direction.
- **Chapitre 06 : Etude des besoins.** Dans ce chapitre, nous allons nous intéresser aux techniques de collecte d'information utilisées, la démarche suivie afin d'identifier les besoins exprimés par les utilisateurs.

### Partie 03 : Conception

Cette partie est consacrée à la conception de notre entrepôt de données, que nous avons effectué en deux étapes :

- **Chapitre 07 : Conception de la zone de l'entreposage.** À travers ce chapitre, nous allons proposer une modélisation dimensionnelle des données (Faits/Dimensions), répondant aux besoins des utilisateurs et permettant d'exploiter au mieux les données stockées dans un entrepôt de données.
- **Chapitre 08 : Conception de la zone d'alimentation.** Nous présenterons, dans ce chapitre, la conception de la zone d'alimentation de notre entrepôt, qui permet de préparer les données de l'entreprise avant leur intégration dans l'DW.

### Partie 04 : Réalisation

Cette partie est consacrée à la réalisation et déploiement de notre solution :

- **Chapitre 9 : Mise en œuvre de la solution :** Nous allons, dans ce chapitre, présenter les détails de la réalisation de notre solution, qui consiste à donner une forme concrète à la phase de conception. Cette partie décrit l'architecture technique de notre solution, les étapes de réalisation, l'aspect sécurité du système, pour terminer avec quelques perspectives que nous jugeons utiles.

## Chapitre 01:Le système décisionnel

"Chaque fois que vous voyez une entreprise qui réussit, dites-vous c'est par ce qu'un jour quelqu'un a pris une **décision** courageuse."

**Peter Drucker**

" Un système décisionnel est un système qui doit pouvoir faire conjuguer les données à tous les temps. "

**Benjamin EPE**

## Introduction

Au cours de cette dernière décennie, la concurrence à laquelle se livrent les entreprises est marquée par de forte turbulence, nécessitant des réactions rapides. Cependant, l'entreprise doit s'adapter, si possible anticiper, et mettre au point une stratégie susceptible pour mieux atteindre les objectifs. Sachant que celles -ci produisent et manipulent régulièrement de très importants volumes de données. Ces données sont stockées dans plusieurs systèmes de productions.

Alors ; il y en un réel besoin de consolider ces dernières pour pouvoir optimiser le patrimoine informationnel dans l'intérêt d'une analyse pour la bonne prise de décision. Face à ce besoin, est né le concept « Système décisionnel ».

Selon [Inmon, 2005] « ces systèmes regroupent aussi un ensemble d'informations et d'outils mis à la disposition des décideurs pour supporter de manière efficace le processus de prise de décision, ils sont dédiés au management de l'entreprise afin de l'aider au pilotage de ses activités ».

Dans ce chapitre nous allons présenter des notions générales sur le système décisionnel, son aide au pilotage de l'entreprise et particulièrement en fonction du marketing.

### 1.1. Définition de la décision et ses types:

Pour mieux comprendre la définition de la décision, nous présentons quelques définitions :

Pour le sociologue et l'économiste américain [Simon ,1980] ; définit la décision comme: "une opération intellectuelle, débouchant sur un choix, et se matérialisant par une action."

Par contre pour [Hilfer 2008] ; il définit la décision comme étant : "L'acte volontaire par lequel, après examen des divers alternatives, on tranche, on prend parti. "

Dresser une typologie des décisions revient à cerner les principaux facteurs qui agissent sur la prise de décision. D'après la typologie présentée par [Ansoff,1965], il existe trois types de décisions prises dans une entreprise :

### **1.1.1. Décision Stratégique**

Les décisions stratégiques sont les décisions les plus importantes dans la mesure où elles concernent les relations de l'entreprise avec le milieu économique et, porte essentiellement sur des choix de marché et de produits.

Exemple : Prévision des ventes de nouveaux produits à long terme....

### **1.1.2. Décision tactique ou administrative**

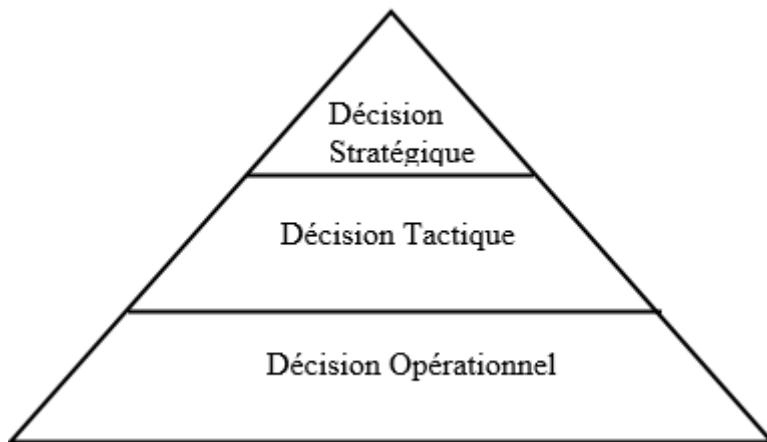
Les décisions tactiques permettent à l'entreprise de déterminer la manière d'utiliser ces ressources afin d'atteindre les objectifs fixés dans le cadre des décisions stratégiques.

Exemple : organiser des formations pour les salariés....

### **1.1.3. Décision opérationnelle**

Il s'agit des décisions relatives à la gestion courante de l'entreprise et concernent l'utilisation optimale des ressources allouées dans le cadre du processus de productivité.

Exemple : établissement des budgets, gestion de stock ....



**Figure 1 : Pyramide des catégories des décisions [Ansoff,1965].**

Ces décisions ne sont pas indépendantes, mais hiérarchisées. Les décisions opérationnelles qui sont les plus nombreuses, concrétisent les décisions tactiques, issues elles-mêmes des décisions stratégiques.

## 1.2. Définition du système décisionnel

Selon [Inmon, 1996] : « Les SID sont apparus comme une nécessité pour faire face à l'augmentation des volumes de données et à la difficulté grandissante des décideurs à disposer de résultats tangibles sur les différentes branches d'activité. Des outils ont été conçus spécifiquement pour extraire les données provenant de sources hétérogènes, les stocker dans un ensemble homogène orienté métier puis enfin restituer les résultats dans des interfaces d'analyse et de Reporting. »

Et d'après [Polletto, 2012] : « L'informatique décisionnelle désigne les méthodes, les outils et les moyens qui permettent de collecter, consolider et modéliser les données d'une entreprise afin d'offrir une aide à la décision et de permettre au corps exécutif de l'entreprise d'avoir une vue d'ensemble de l'activité. »

Le système décisionnel ou (Business Intelligence BI) : "se définit par l'ensemble des moyens, outils et méthodes qui supportent le processus de collecte, consolidation, modélisation, analyse et restitution des informations. " [Cigref, 2009].

Pour [Kimball, 2002] le Système décisionnel : " a pour objectif d'accompagner un ou plusieurs décideurs dans le processus de prise de décision. Il permet aux acteurs concernés de spécifier leurs besoins par des processus de collecte, d'analyse et d'échange d'informations."

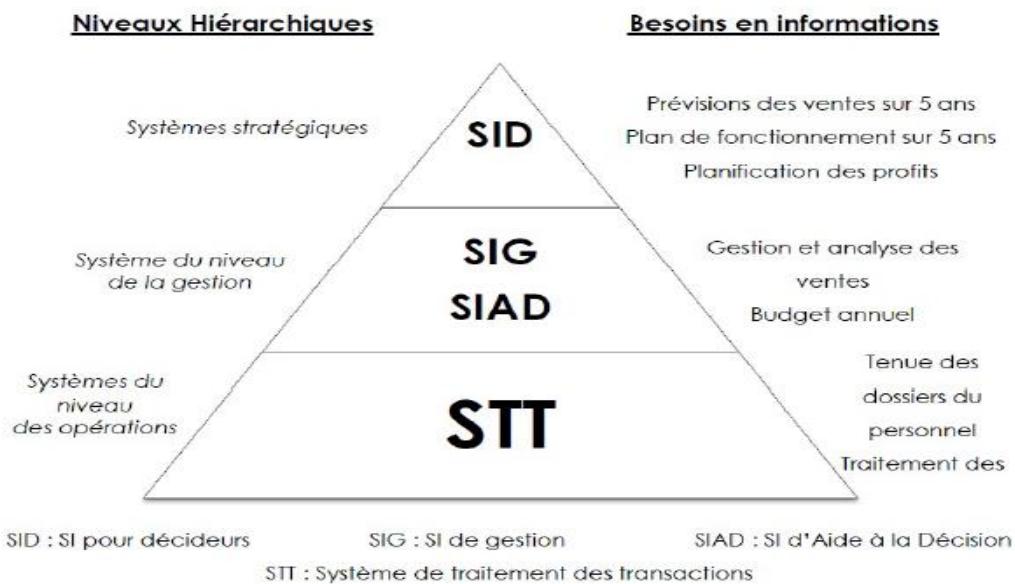
Donc le système décisionnel est un système d'information qui regroupe les données d'aide à la décision et facilite leur exploitation en fournissant les outils adéquats afin d'améliorer la performance de l'entreprise.

## 1.3. Place d'un système décisionnel dans l'entreprise

Il convient tout d'abord de rappeler la définition d'un système d'information, qui met en lumière son importance dans les entreprises en fait dans la mesure où c'est : « un ensemble organisé de ressources : matériel, logiciel, personnel, données, procédures...permettant d'acquérir, de traiter, stocker, communiquer des informations »

Les systèmes d'information avec ses différents types opérationnels ou décisionnels sont construits pour supporter efficacement les divers intérêts organisationnels de l'entreprise.

La figure ci-dessous illustre parfaitement la place qui revient au système décisionnel au sein d'une entreprise qui se place entre le système opérationnel et le système de pilotage.



**Figure 2 : Les systèmes d'information selon les niveaux hiérarchiques.**

## 1.4. Notion de décideur

Jusqu'à les années 80-90, les organisations étaient organisées de manière de pyramide (structure hiérarchique) .les décideurs étaient prises au sommet de la pyramide et les ordres étaient transmis de manière descendante et unilatérale à tous les niveaux opérationnel.

Face à cela, [Sébastien Fantini, 2010] classe les décideurs selon trois catégories qui sont :

### 1.4.1. Décideurs stratégiques

- Horizon de travail : long terme.
- Périmètre de travail : tous les services, tous les territoires.
- Leur rôle : ces décideurs impulsent une politique, définissent les valeurs de l'organisation et donnent les moyens aux ambitions de l'organisation.

Exemple : la direction générale dans une entreprise.

### 1.4.2. Décideurs tactiques

- Horizon de travail : Moyen terme.
- Périmètre de travail : un service ou un territoire.

- Leur rôle : ces décideurs tactiques sont les relais des caps stratégiques, fixés par les décideurs stratégiques.
- Les décideurs tactiques fixent les objectifs de leur direction, élaborent et choisissent la meilleures tactiques pour atteindre les objectifs.

Exemple : la direction des finances, direction de ventes.

#### **1.4.3. Décideurs opérationnels**

- Horizon de travail : court terme.
- Périmètre de travail : un service.
- Leur rôle : faire face à la réalité du terrain, gérer le quotidien.

Exemple : un commercial, un acheteur.

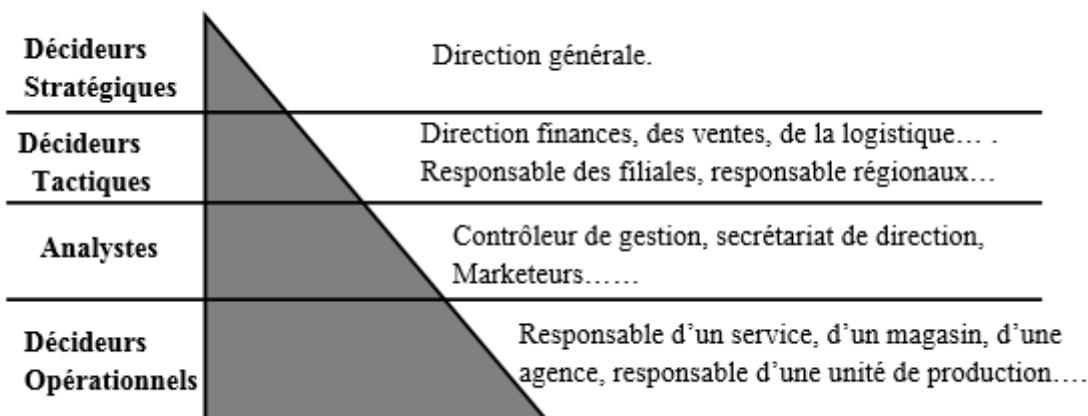


Figure 3 : Type de décideurs dans les organisations. [Sébastien Fantini, p13]

#### **1.5. Contexte du système décisionnel**

De récentes études montrent que la Business Intelligence ou système décisionnel est l'une des préoccupations principales au sein des DSI (Direction de Systèmes d'Information) de grandes entreprises. En effet dans le contexte d'hyper-concurrence, la BI représente une opportunité pour les entreprises afin d'optimiser le pilotage de leur activités, et d'anticiper les évolutions du marché, les comportements des clients ...etc.

Les domaines d'utilisation de la BI touchent la plupart des métiers de l'entreprise :

- Finance, exemple : le reporting financier et budgétaire.
- Vente et commerce, exemple : l'analyse des points de ventes, l'analyse de la profitabilité et l'impact des promotions.
- marketing, exemple : la segmentation client, les analyses des comportements.
- Logistique, exemple : l'optimisation de la gestion des stocks, le suivi des livraisons.
- Ressources humaines, exemple : l'optimisation de l'allocation des ressources.

## **1.6. Architecture d'un système décisionnel**

En 1998 , E.F Codd suggère l'emploi d'un système qui permet d'améliorer le processus de prise de décision par la consultation et l'analyse de grandes masses de données contenue dans les bases opérationnelles des entreprises, pour le but de fournir une réponse rapide à des requêtes analytiques de nature multidimensionnelle [Codd , 1998]. « Ces requêtes représentent des analyses qui s'appuient sur un outil de centralisation des données appelé entrepôt de données » [Kimball, 2002].

Tout système décisionnel est composé de trois zones importantes qui sont : (voir la Figure 4 ci-dessous) :

### **1.6.1. Zone de préparation**

Dans un premier niveau, il est nécessaire de recueillir et d'intégrer les données stockées dans les diverses sources de données, qui sont hétérogène et repartis à l'origine. Les systèmes sources à savoir les systèmes de production de l'entreprise se composent de bases de données opérationnelles. De manière générale un effort est nécessaire pour faire des manipulations sur ces dernières via une famille d'outils dénommé "ETL " pour (Extract, Transforme, Load) pour assurer les fonctionnalités d'extraction, transformation et de chargement de données source afin de les unifier, puis les intégrer dans un espace de stockage centralisé.

### **1.6.2. Zone de stockage**

Les données nettoyées et consolidées seront stockées dans une base de données spécialisées nommées un entrepôt de donnée ou « data warehouse » en anglais qui le cœur du système décisionnels pour être enfin destiner à l'analyse.

### 1.6.3. Zone de restitution

Cette dernière phase concerne la restitution des résultats obtenus à partir de la zone de stockage. Les outils de restitution sont la partie visible offerte aux utilisateurs. On distingue à ce niveau plusieurs types d'outils :

- Les outils de reporting qui permettent la mise à disposition de rapports.
- Les outils d'analyse OLAP pour les analyses multidimensionnelles.
- Les outils de Datamining pour la recherche des corrélations.
- Les tableaux de bord afin de piloter la performance.

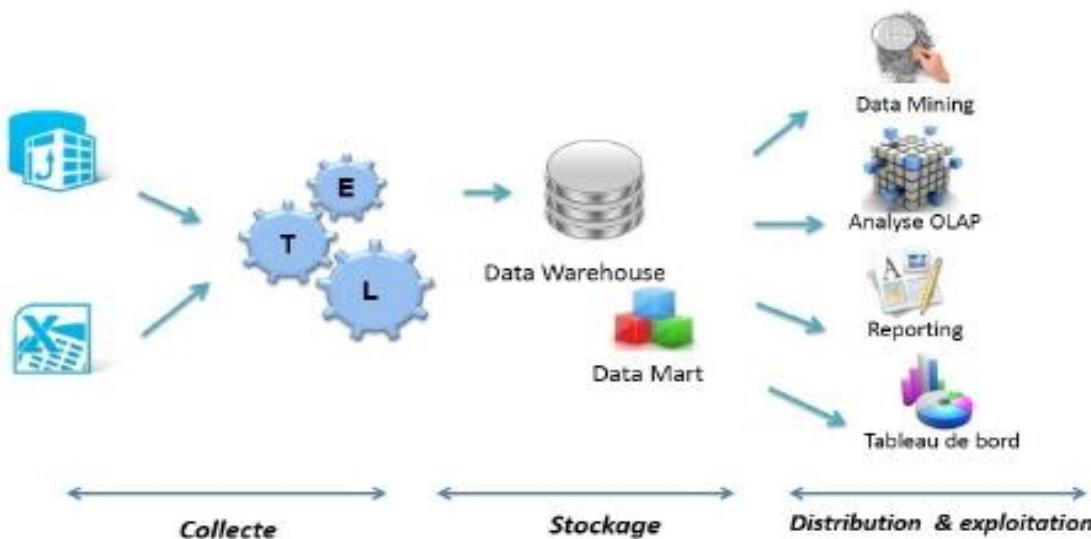


Figure 4 : Architecture d'un système décisionnel.

## 1.7. Système décisionnel et fonction marketing

Le marketing lors de ces dernières années a beaucoup évolué. Il a en particulier été très conditionné par la possibilité offerte par l'informatique de s'adresser directement à chaque client individuellement, qui a permis le développement du fameux **marketing one to one**. C'est dans ce cas que se trouve l'intérêt d'un entrepôt de données afin de gérer les préférences des clients, l'historique des contacts et des transactions. Dans ce cadre la démarche consiste à différencier les clients en fonction de leurs besoins et de leur valeur, d'interagir avec eux afin de personnaliser le produit.

Grâce à de meilleurs systèmes d'information, les entreprises s'ouvrent un monde d'opportunités. La route cependant est longue, il faut commencer par collecter des données sur chaque client, réaliser des analyses, des segmentations, fixer des stratégies en fonction des enjeux, adapter les communications, guider les expériences des clients et coordonner les actions à travers de multiples canaux. Dans cette approche la première difficulté est d'amener l'ensemble des intervenants de l'entreprise à se focaliser sur des objectifs de satisfaction du client, et pour cela de partager les informations et de se coordonner. La difficulté accessoire est de maîtriser toutes les nouvelles technologies nécessaires. **[Michel, 2011]**

## 1.8. Le défi du Système décisionnel

L'entreprise évolue dans un environnement en changement perpétuel. Dans ce contexte elle se trouve dans un cas de prendre de plus en plus des décisions efficaces au moment favorable. Pour prendre des décisions efficaces, elle doit posséder un système d'information fiable et des outils d'aide à la décision performants. L'implémentation d'une telle technologie doit relever deux défis qui inhibent encore aujourd'hui l'adoption de cet outil qui sont :

Dans un premier temps, l'informatique décisionnelle nécessite une forte expertise technique et exige le recours à des logiciels optimisés et fortement intégrés à l'infrastructure existante : systèmes d'exploitation, systèmes de gestion de bases de données avancées.

Le second défi est sans aucun doute lié aux coûts de ces systèmes. Des technologies avancées sont exigées et des solutions de l'informatique décisionnelles ont été conçues pour de grandes organisations à forts budgets. Car les solutions complexes présentant un fort coût d'acquisition et de maintenance décisionnelle sans compromettre le niveau fonctionnel ou la qualité de service.

De ce fait, De nouvelles possibilités apparaissent cependant, au plus grand profit des organisations à budget limite. Des logiciels Open Source sont aujourd'hui disponibles pour toutes les fonctions d'une plateforme décisionnelle délivrant un niveau d'usage, de capacité à monter en charge et de robustesse compatibles avec des besoins critiques.

Voici quelques exemples de logiciels décisionnels qu'ils soient libres ou propriétaires : [web1]

**SAS** : pour Statistical Analysis System (logiciel propriétaire) est le plus important éditeur indépendant du marché du décisionnel. Avec plus de 60 000 sites clients répartis dans 134 pays. Plusieurs applications lui sont propres : tableaux, graphiques...

**Business Object** :(logiciel propriétaire) : il permet l'accès aux applications habituelles du décisionnel (tableaux, graphiques...) ainsi qu'à internet.

**Birt** : qui signifie Business Intelligence and Reporting Tools, est un projet de la Fondation Eclipse créé et animé depuis 2004. Birt figure parmi les outils Open Source de Business Intelligence les plus largement diffusés et utilisés. Le nombre élevé de développeurs dans le monde (un million) témoigne de la facilité d'adoption de cet outil.

**Cognos** : est une compagnie canadienne, filiale d'IBM depuis 2008. Fondée en 1969, la compagnie édite des solutions décisionnelles et propose des services permettant aux entreprises de piloter, superviser et comprendre leur performance. Les solutions qu'elle propose aident les entreprises à mesurer, comprendre et anticiper leur performance opérationnelle et financière et apportent un éclairage étayé sur les meilleures actions à mener. Elles contiennent entre autres : reporting, tableaux de bord, analyse des résultats.

**BI Square Software** : est un éditeur de logiciels de Business Intelligence. Créé en 2010, il fournit des tableaux de bord décisionnels sur l'IT à fin que les entreprises pilotent et optimisent au mieux leur infrastructure informatique (serveurs, stockage, réseau...).

**SpagoBI** : est une solution de Business Intelligence entièrement Open Source. Spago BI est distribué sous licence Mozilla Public License, qui est compatible avec les usages commerciaux. SpagoBI Server est le module principal de la suite, qui offre toutes les fonctionnalités centrales et analytiques de la plateforme. Il fournit un éventail de services administratifs. Il couvre l'ensemble des besoins analytiques, en fournissant plusieurs solutions pour chaque domaine analytique : reporting, tableaux de bord « en temps réel ».

**Pentaho** : est une solution d'informatique décisionnelle Open source développée en Java. L'offre principale de l'Américain Pentaho repose sur sa suite qui rassemble des outils d'intégration de données (ETL), d'analyse, de reporting, de tableaux de bord et de datamining. Il est de plus en plus présent en France, en Europe.

**Talend** : est un éditeur de logiciel open source spécialisé dans l'intégration et préparation de données (ETL) à fin d'offrir à l'entreprise et à son service informatique une nouvelle façon de transformer les données en décisions.

**Remarque** : Dans le cadre de notre travail, nous allons nous intéresser particulièrement à la solution Pentaho et Talend que nous allons décrire plus en détail dans ce qui suit.

## Conclusion

Actuellement, le concept du système décisionnel est un sujet incontournable d'où il se situe au cœur des préoccupations des entreprises. Vu son importance comme un outil servant à identifier les informations nécessaire afin d'entreprendre des meilleures décisions permettant d'augmenter la performance de celles-ci et atteindre les objectifs souhaités.

Dans le prochain chapitre on abordera l'un des concepts les plus importants dans tout projet décisionnel qui est le concept d'un entrepôt de données. Ce dernier est considéré comme la brique essentielle du système décisionnel et le point de départ de toutes les applications d'analyse du Business Intelligence.

# Chapitre 02 : Les entrepôts de données.

## Introduction

Dans le précédent chapitre nous avons présenté brièvement l'entrepôt de données qui est au cœur de tout système décisionnel et la base sur laquelle repose tous les outils d'analyse de données. En fait, il regroupe toutes les données de l'entreprise provenant des différentes sources, internes. Ces données seront nettoyées et organisées d'une façon à être facilement exploitables pour des fins d'analyse.

Dans ce présent chapitre, nous allons voir c'est quoi un entrepôt de données ? Ses caractéristiques ? En quoi diffère-t-il d'une base de données relationnelle ? Et d'autres concepts fondamentaux liés à ce dernier.

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

**Bill Inmon** et **Ralph Kimball** peuvent aisément et raisonnablement se prétendre les fondateurs d'entreposage de données. Ils ne sont peut-être pas les premiers à avoir posé les définitions relatives à un ED mais leurs travaux dans ce domaine et les principes de conception d'un entrepôt que l'un et l'autre ont posé en font des références incontestables.

Selon [Kimball, 2002] : « Le Data Warehouse est un système qui extrait, nettoie, traite et rend conforme des données sources vers un espace de stockage multidimensionnel. Il permet ainsi la mise en œuvre de l'interrogation et de l'analyse à des fins d'aide à la prise de décision ».

Pour **Bill Inmon**, qui est considéré comme le "père des entrepôts de données", définit ED comme : « Le Data Warehouse est une collection de données orientées sujet, intégrées, non volatiles et historisées, organisées pour le support d'un processus d'aide à la décision ». [Inmon, 2005].

- **Données orientées sujet:** les données d'un entrepôt sont organisées autour des thèmes majeurs de l'entreprise tels que : la vente, la production, la comptabilité, ...etc.

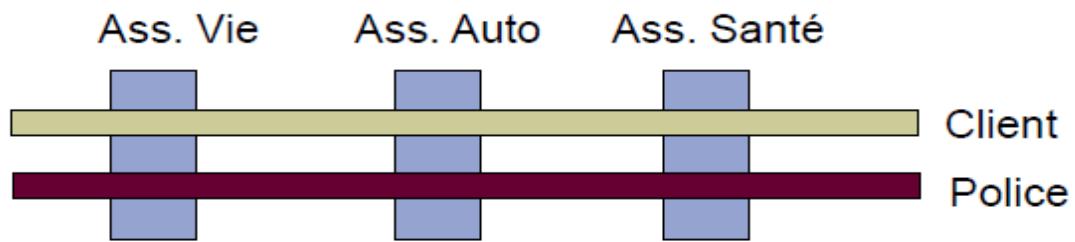


Figure 5 : Entrepôt de données orienté objet. [web2]

- **Données intégrées:** les données proviennent de divers systèmes opérationnels hétérogènes et des applications de production.

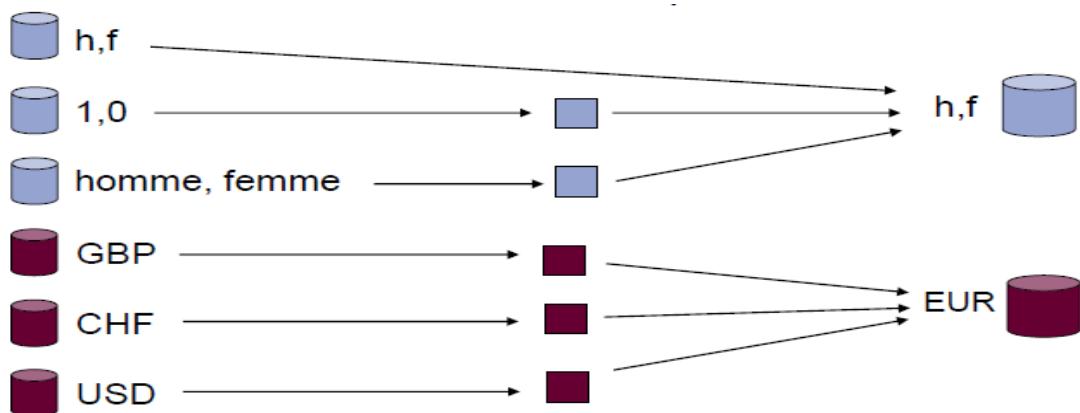


Figure 6 : Entrepôt de données intégrée. [web2]

- **Données non volatiles:** les données stockées dans l'entrepôt ne sont pas modifiées par les utilisateurs.

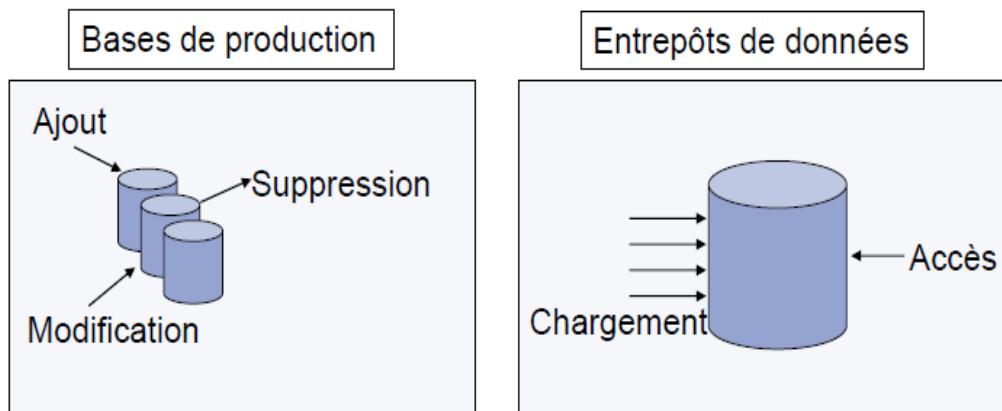


Figure 7 : Entrepôt de données non volatile. [web2]

- **Données Historisées :** les données persistent dans le temps.

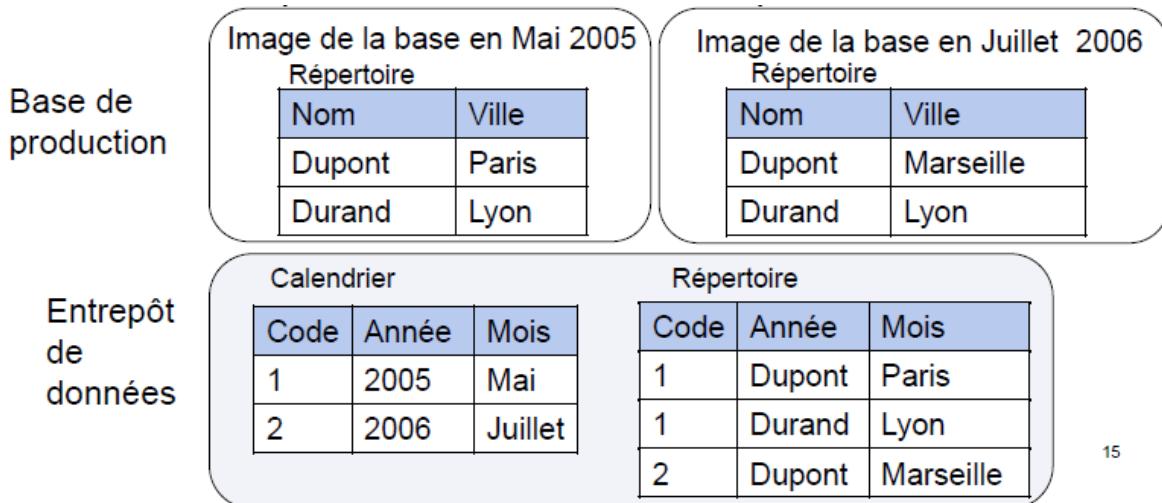


Figure 8 : Entrepôt de données historisées. [web3]

## 2.2. Entrepôt de données vs base de données relationnelle

Une question qui se pose naturellement : Pourquoi ne peut utiliser une base de données relationnelle pour réaliser un système décisionnel ? La réponse est qu'une base de données relationnelle n'est pas adaptée à ce genre de systèmes. Le tableau 2 propose une comparaison entre un DataWarehouse et une BDD (Base de données) :

Caractéristique	Base de données	Entrepôt de données
<i>Opération</i>	Gestion courante, Production	Analyse, support à la décision
<i>Modèle de données</i>	Entité/relation	étoile, flocon de neige
<i>Normalisation</i>	Fréquente	Plus rare dans les <i>Datamarts</i>
<i>Données</i>	Actuelles, Brutes	Historisées, Parfois agrégées
<i>Mise à jour</i>	Immédiate, Temps réel	Souvent différé
<i>Niveau de consolidation</i>	Faible	Elevé
<i>Perceptions</i>	bidimensionnelle	Multidimensionnelle
<i>Opérations</i>	Lectures, Mises à jour, suppressions	Lectures, Analyses croisées, rafraîchissements
<i>Taille</i>	En giga-octets	En téraoctets

Tableau 1 : Tableau comparatif entre une BDD relationnelle et un entrepôt de données.

## 2.3. Composants de base d'un entrepôt de données:

Un **entrepôt de données** (DataWarehouse en anglais) est une structure de base dans un système décisionnel. C'est l'endroit utilisé pour collecter et stocker des informations provenant d'autres sources de données, puis transformées celle-ci en informations manipulées, par les outils d'analyse pour des fins de prise de décision.

La mise en œuvre d'un entrepôt de données ne peut pas se dissocier de ces composantes qui le construisent. Nous allons tout d'abord expliquer chaque composant séparément. L'DW est constitué de quatre zones distinctes, qui sont:

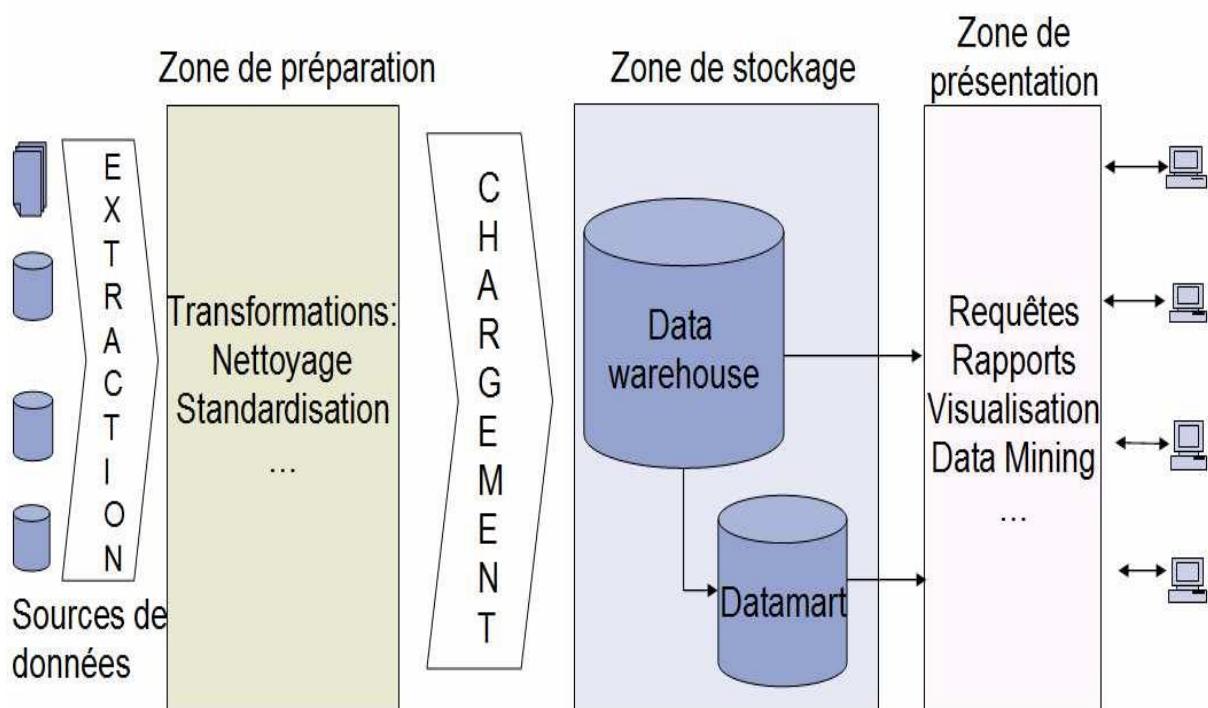


Figure 9 : Architecture de l'entrepôt de données [KIMBALL, 2002].

### 2.3.1. Systèmes sources [Kimball 2002, p10]

Afin de pouvoir alimenter notre entrepôt de données, il nous faut évidemment disposer de systèmes sources qui représentent les diverses applications opérationnelles qui capturent les transactions liées aux activités de l'organisation.

Ces sources de données peuvent être des bases de données internes, des feuilles de calcul, des fichiers plats,....etc.

### **2.3.2. Zone de préparation des données**

C'est à ce niveau que se déroulent les différentes manipulations de données brutes chargées à partir du système source. Elle est constituée d'un ensemble de processus couramment appelés ETL (Extract/Transform /Load) ou **Data StagingArea**. " Cette zone peut être vue comme la cuisine d'un restaurant" [Kimball, 2002].

La zone de préparation est inaccessible par les utilisateurs et elle ne permet pas l'interrogation des données qui sont généralement détruites après chargement dans le DataWarehouse.

### **2.3.3. Zone de stockage des données**

Cette zone représente le lieu de stockage des données. Elle est accessible par les utilisateurs via des outils d'analyses. Le DW est constitué d'un ensemble de Data Marts, les données sont généralement équivalentes à celle du DW mais elles sont présentées de façon adaptée aux besoins spécifiques de la fonction et/ou du domaine utilisateur (par exemple : création d'un DM dédié pour le service Marketing ou Commercial).

### **2.3.4. Outils d'accès aux données**

C'est l'ensemble des outils qui permettent aux utilisateurs d'exploiter les données qui sont stockées dans un entrepôt de données. Il existe plusieurs outils d'accès aux données du DW tels que les outils d'analyse en ligne OLAP, les outils du Datamining qui permettent de découvrir des liens sémantiques entre les données et les outils de reporting ou encore tableau de bord.

### **2.3.5. Métadonnées**

Une méta-donnée représente une donnée sur une donnée. Il s'agit de l'ensemble des informations qui permettent de qualifier une donnée, notamment par sa sémantique, sa règle de calcul, sa provenance, sa qualité,...etc.

Très souvent les données à fédérer dans l'entrepôt de données proviennent de sources très hétérogènes. Cela rend indispensable la présence d'un dictionnaire unique qui sait gérer l'ensemble des fonctions de l'entrepôt de données. Cette cohérence du dictionnaire est décrite au sein des métadonnées du dictionnaire d'un entrepôt de données . Les métadonnées constituent l'ensemble des données qui décrivent des règles ou procès attachés à d'autres données.

## 2.4. Data Mart

### 2.4.1. Définition

" Un data mart est un sous ensemble logique de data warehouse. Au-delà de cette définition relativement simple, on considère souvent un data mart comme la réduction du data warehouse à un seul processus ciblant un groupe métier spécifique. Il est généralement sponsorisé par un département ou une unité en particulier ". [Kimball, 2002]

Il est orienté sujet ou thème et peut être par exemple utilisé pour des applications de CRM (Customer Relationship Management) .

Un Data Mart peut être défini comme une version allégée du DW, il se focalise sur un sujet, un thème ou un métier. Il est plus facile à comprendre et à manipuler, il vise une problématique précise avec un groupe relativement limité d'utilisateurs.

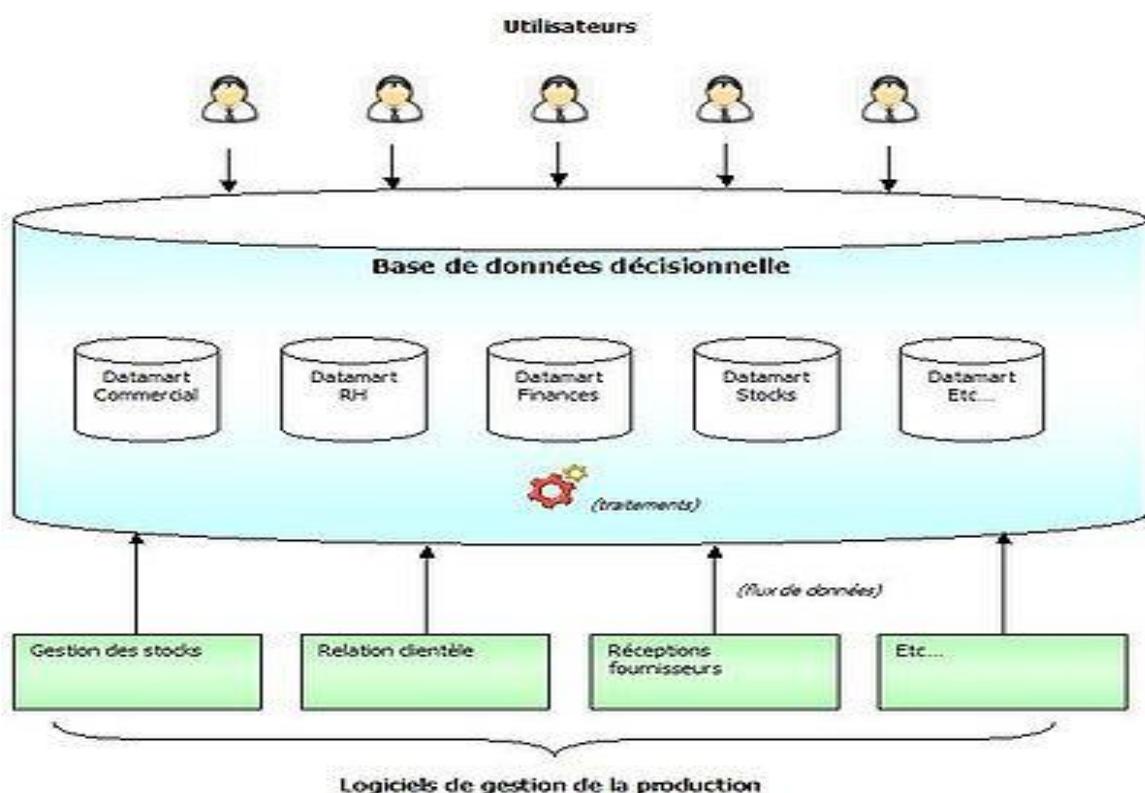


Figure 10 : Schéma des data marts. [web3]

### 2.4.2. Intérêt des magasins de données

- ✓ Les magasins de données offrent un nouvel environnement structuré et formaté en fonction des besoins d'un métier ou d'un usage particulier.
- ✓ Moins de données que les entrepôts de données.
- ✓ Plus facile à comprendre et à manipuler.
- ✓ Amélioration des temps de réponse des requêtes exécutées.

### 2.5. Entrepôt de données VS magasin de données

Le tableau ci-dessous montre les différences entre un entrepôt de données et un magasin de données :

Caractéristiques	Magasin de données	Entrepôt de données
<b>Champ applicatif</b>	Un sujet d'analyse	Plusieurs sujets d'analyses
<b>Temps de développement</b>	Un Mois	Année
<b>Complexité de développement</b>	Faible à moyenne	Grande
<b>Taille de donnée</b>	centaine de Go	centaine de GO jusqu'à TO
<b>Transformation de donnée</b>	Faible à moyenne	Importante
<b>Fréquence de mise à jour</b>	Journalière ou hebdomadaire	Mensuelle
<b>Cible utilisateur</b>	Département	Toute l'entreprise

Tableau 2 : Tableau comparatif entre un ED et un MD.

## 2.6. Objectif d'un entrepôt de données

L'atout principal de l'entreprise réside dans les informations qu'elle possède. Ces informations se présentent généralement sous deux formes : les systèmes opérationnels qui sont l'emplacement de saisie et d'enregistrement de données et l'entrepôt de données qui est l'emplacement de collection et de consolidation de données.

Alors les objectifs fondamentaux d'un entrepôt de données peuvent être cernés comme le dit [Kimball, 2002] comme suit :

- ✓ Assurer que le contenu d'un entrepôt de données soit compréhensible pour l'utilisateur final pour lui faciliter la navigation rapide entre les données. La notion de rapidité fait référence à des délais d'attente égaux à zéro.
- ✓ Rendre les informations de l'entreprise accessible et cohérente via un DW dans la mesure où la modification perpétuelle ne doit pas bouleverser ni les données existantes ni les technologies utilisées.
- ✓ L'ED est considéré comme le socle du système décisionnel vu qu'il recèle les informations propres et crédibles pour faciliter la prise de décision.

## 2.7. Structure d'un entrepôt de données

Un entrepôt de données peut se structurer en quatre classes de données, organisées selon un axe historique et un axe de synthèse :

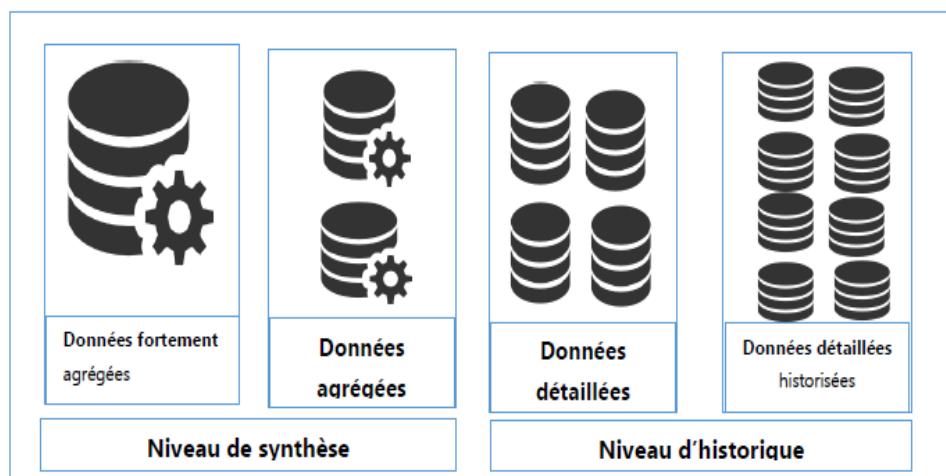


Figure 11 : Structure de données d'un entrepôt de données.

### **2.7.1. Niveau historique**

- **Données détaillées archivées**

Anciennes données rarement sollicitées, généralement stockées dans un disque de stockage de masse, peu coûteux, à un même niveau de détail que les données détaillées.

Exemple : les dossiers d'archives des clients.

- **Données détaillées**

Les données détaillées reflètent les événements les plus récents ; elles sont fréquemment consultées et généralement volumineuses. Exemple : les données provenant du système de production.

### **2.7.2. Niveau de synthèse**

- **Données agrégées**

Données agrégées à partir des données détaillées, ces données sont souvent utilisées car elles constituent déjà un résultat d'analyse et une synthèse de l'information contenue dans le système décisionnel. Elles doivent donc être facilement accessibles et compréhensibles.

Exemple : rapports des ventes.

- **Données fortement agrégées**

Données agrégées à partir des données détaillées, à un niveau d'agrégation plus élevée que les données agrégées.

Exemple : Rapport annuel.

## **2.8. Conduite et construction d'un entrepôt de données**

Pour la construction d'un entrepôt de données on suivra les démarches suivantes :

- Méthodologies conductrices à la mise en place d'un entrepôt de données.
- L'alimentation d'un entrepôt de données.
- La mise en œuvre d'un entrepôt de données.
- L'administration et la maintenance d'un entrepôt de données.

### 2.8.1. Approches de construction d'un entrepôt de données

La plupart des approches existantes pour le développement d'un ED concernent la façon dont les données devraient être structurées, stockées, et contrôlées. **Kimball** et **Inmon** s'entendent sur le fait qu'un entrepôt de données intègre les données de diverses sources de données des systèmes opérationnels.

De ce fait, ces deux gourous d'un entrepôt de données distinguent trois approches, il s'agit :

- L'approche « Bottom-Up », ou DataMarts indépendants de R.Kimball.
- L'approche « Top-Down », ou DataMarts dépendants de B.Inmon.
- L'approche « Middle-Out », ou hybride qui dérive des deux premières approches.

- **Approche Bottom-up**

Selon Kimball, un entrepôt de données n'est rien que l'union des Data Marts. Alors, il décrit le processus d'un entrepôt de données comme suit : « Que chacun construise ce qu'il veut, on intégrera ce qu'il faudra quand il faudra » [KIMBALL, 2002].

Dans cette approche KIMBALL commence à construire un Staging area (Zone de préparation de données) temporaire puis à créer des data marts un par un. Celui-ci les placera au centre de son architecture, sachant que chaque magasin de données est spécifique aux besoins des utilisateurs. Puis un bus de dimension permet de relier ces derniers afin de constituer un entrepôt de données global.

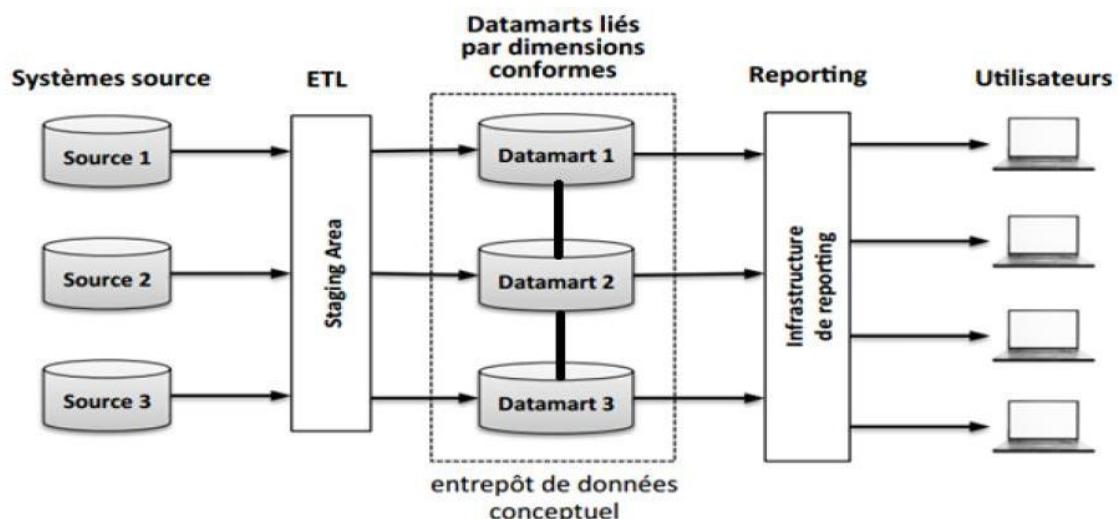


Figure 12 : Architecture en bus dimensionnel [Kimball 2002].

## Avantages de l'approche Bottom-up

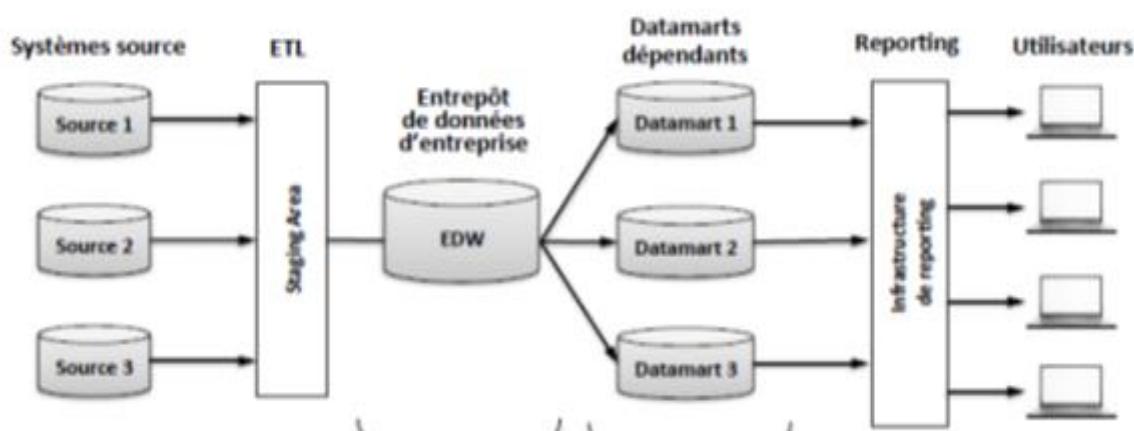
- Architecture la plus simple et la moins coûteuse à développer.
- Intégration des données assurée par les dimensions conforme<sup>1 1()</sup>.
- Approche incrémentale (processus les plus importants d'abord),
- Donne des résultats rapidement.

## Inconvénient de l'approche Bottom-up

- Vision limitée et pas extensible.
- Performances peu optimales des analyses impliquant plusieurs Data Marts.
- Redondances et incohérences entre les DataMarts (car ils sont faits indépendamment les uns après les autres).

### ■ Approche Top-Down

En guise de réponse à l'égard de Ralph Kimball, **Bill Inmon** avait déclaré que le DW est au centre de l'architecture. Il possède sa propre existence physique, orientée stockage, et évolutive par rapport à de nouveaux besoins. Les Data Marts possèdent, eux aussi, leur propre existence physique. Ils viennent s'alimenter à partir d'un entrepôt de données. Cette approche est lourde et complexe car elle consiste à la conception de tout un entrepôt de données, puis à la création des Data Marts départementale. Enfin, les utilisateurs finaux peuvent interroger le ED et DM.



**Figure 13 : Architecture du DW selon Inmon.**

<sup>1</sup>**Conformes :** c'est des tables dimensionnel utilisées par plusieurs table de faits, ils ont donc la même définition dans toutes les faits avec lesquels elles sont rattachées.

## Avantage de cette approche

- Consolidation complète des données de l’entreprise.
- Vision très clair des données de l’entreprise.

## Inconvénients de cette approche

- Peut avoir des redondances de données entre les Data Marts.
- Performance sous optimale des analyses impliquant plusieurs DataMarts

### ▪ Approche Kimball vs approche Inmon

Elément de comparaison	Kimball	Inmon
Processus	Bottom -Up	Top -Down
Structure	DataMarts	DataWarehouse
Schématisation	En étoile	En flocon
La nature du processus de prise de décisions	Tactique	Stratégique
Besoin en intégration de données	Besoin départementaux	Besoin de données pour toute l’entreprise.

Tableau 3:Tableau comparatif entre Kimball et Inmon.

### ▪ Approche hybride

Cette approche est conseillée par les professionnels du BI. Elle consiste en la conception totale d’un entrepôt de données (concevoir toutes les dimensions, tous les faits et 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).

## 2.8.2. Alimentation de l'entrepôt de données avec ETL

Pour alimenter un DW il est intéressant d'extraire les données provenant de système de production ou base de données opérationnel. Pour cela, il faut avoir préalablement savoir identifier les données intéressantes de celles qui ne le sont pas. Pour ce faire, on met en place des interfaces d'extraction, de transformation et d'alimentation nommée « ETL ».

- **Processus Extract-Transform-Load "ETL"**

Un ETL, est une technologie informatique qui permet d'effectuer des synchronisations massives d'information d'une banque de données vers une autre. Cela veut dire extraire des données provenant de sources extérieures, les transformer, puis les charger dans un entrepôt de données ou magasin de données.

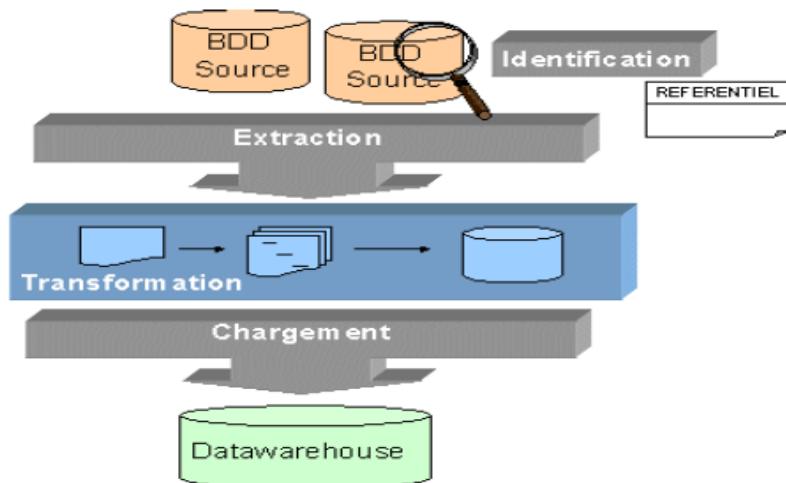


Figure 14 : Schéma de système d'alimentation ETL.

Il est important de savoir que la réalisation de l'ETL constitue 70% d'un projet décisionnel en moyenne. Et ce n'est pas pour rien, ce système est complexe et ne doit rien laisser s'échapper, sous peine d'avoir une mauvaise information dans un entrepôt de données.

Le système d'alimentation de l'ETL se résume en trois phases [figure 14] :

- **Phase d'extraction**

Cette première phase permet d'extraire des données dans divers systèmes sources qui sont hétérogènes à l'origine. L'étape critique est dans le choix des données à extraire et des filtres à appliquer.

### ➤ Phase de transformation

Afin de garantir la pertinence des données extraites et leurs qualités, cette étape exige beaucoup de réflexion. Cela veut dire Après leur extraction, les données subissent une série de transformations et de nettoyage qui sont : cohérence entre les types de données, élimination des doublons...etc. Afin de les rendre présentables, fiables aux décideurs.

### ➤ Phase de chargement (Load)

Une fois les données nettoyées, standardisées, corrigées,...etc. cette dernière étape dans tout processus ETL consiste en l'acheminement de ces données vers leur destination finale (Zone de stockage) qui est l'entrepôt de données ou magasin de données.

## 2.8.3. Déploiement et mise en œuvre d'un entrepôt de données

Une fois que les données se trouvent dans l'entrepôt de données, il ne reste plus qu'à les exploiter. L'utilisateur final doit alors pouvoir interroger les données en ligne, à l'aide d'outils simples et conviviaux, qui leurs permettent de répondre à leurs nouveaux besoins. Parmi les modes d'exploitation des données de l'entrepôt de données, on a :

### A. Reporting

Le terme "Reporting" désigne une famille d'outils de Business intelligence destinés 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 chiffrés ou d'un suivi d'avancement. [web4]

Cet outil permet l'élaboration de comptes rendus pertinents et fournit des bilans analytiques offrant une vision opérationnelle de l'activité et assurant son suivi. Il existe deux types de Reporting :

- **Reporting ad hoc**

Aujourd'hui largement utilisé dans un entrepôt de données, les requêtes ad hoc permettent d'accélérer la récupération des informations vitales pour répondre aux requêtes interactives dans une application à mission critique. L'accès à ce genre de service peut se faire via différentes outils de BI.

Cependant, ces outils offrent à l'utilisateur la possibilité de créer lui-même le rapport qui l'intéresse avec les données qu'il souhaite. L'utilisateur aura accès à des vues métiers spécialement conçus en fonction de ses besoins qui lui permettront de choisir facilement l'information qu'il souhaite. Aucune connaissance en base de données n'est nécessaire.

- **Reporting de masse**

Ce type de rapport par contre est la création à l'avance des modèles de rapports qui seront susceptible d'être publiés aux formats élaborés pour un grand nombre d'utilisateurs.

La mise en forme des documents diffusés est soigneusement travaillée et les modalités de publication sont variées (web, mail, papier, . . .). L'utilisateur (ou destinataire) est un consommateur.

## **B. Tableau de bord**

"un outil d'aide à la décision et à la prévision et un ensemble d'indicateurs peu nombreux conçus pour permettre aux gestionnaires de prendre connaissance de l'état et de l'évolution des systèmes qu'ils pilotent et d'identifier les tendances qui les influenceront sur un horizon cohérent avec la nature de leurs fonctions" [BOUQUIN, 2003]

D'après [Alain Fernandez, 2011]: « Un tableau de bord est un instrument de mesure de la performance facilitant le pilotage proactif d'une ou plusieurs activités dans le cadre d'une démarche de progrès. Le tableau de bord contribue à réduire l'incertitude et facilite la prise de risque inhérente à toutes décisions. Le tableau de bord est un instrument d'aide à la décision. »

Donc, c'est un outil de pilotage constitué d'indicateur clé permettant de mesurer la performance de l'entreprise à une certaine période par rapprochement entre les résultats réalisés et les objectifs tracés, afin de guider les décideurs à mettre en place d'éventuelles actions correctives.

## **C. Data mining**

Ou la fouille de donnée est un ensemble de techniques et de méthodes du domaine des statistiques, des mathématiques et de l'informatique permettant l'extraction, à partir d'un important volume de données brutes, de connaissances originales auparavant inconnues par des méthodes automatiques ou semi-automatiques.

L'utilisation de cet instrument d'analyse et de prospection permet de résoudre des problèmes diverses, allant de la gestion de relation client, en passant par la détection de fraudes ou encore Gestion et analyse de risques tels que les banques.

#### **2.8.4. Maintenance d'un entrepôt de données**

La mise en œuvre d'un entrepôt de données ne signifie pas la fin du projet, un suivi constant doit être planifié pour pérenniser, faire évoluer et croître un entrepôt de données.

Le succès de l'DW dépend de son utilisation massive par les utilisateurs, de ce fait il faut s'investir dans ces domaines suivants :

- **Support aux utilisateurs**

Est utilisé comme assistance de premier niveau qui apporte de l'aide aux utilisateurs finals de l'entrepôt de données afin de leur faire apprécier son rôle dans l'entreprise et ainsi détecter les correctifs nécessaires à apporter.

- **Formation des utilisateurs**

Il est indispensable d'offrir un programme de formation aux utilisateurs d'un entrepôt de données.

- **Management de l'évolution**

L'évolution de l'ED concerne sa mise à jour, en fonction de différents changements des sources ou des besoins des utilisateurs. De ce fait l'implémentation de l'ED devrait suivre cette évolution. Toutefois la maintenance d'un entrepôt de données serait envisageable aux nouveaux besoins ou pour de nouvelles améliorations.

### **Conclusion**

L'entrepôt de données est probablement, après l'internet l'une des tendances les plus récentes qui bouleversent l'environnement des entreprises, d'ailleurs celles-ci l'utilisent de plus en plus cette dernière décennie.

Alors l'ED est considéré comme le cœur du système d'information décisionnel. Vu que l'entrepôt de données est une collection de données, alors quelle est la démarche qui consiste à transformer ces données source en structure d'un entrepôt de données finalisée ?

Pour répondre à cette question nous allons démontrer dans le chapitre suivant les modèles de conception d'un entrepôt de données.

# Chapitre 03 : Modélisation dimensionnelle.

## Introduction

La modélisation dimensionnelle est une méthode de conception logique qui a comme but de présenter les données sous une forme intuitive, qui permet des accès hautement performants. Cette modélisation s'avère plus intéressante que la modélisation entité/association dans le cas des bases décisionnelles.

Dans ce chapitre, nous allons s'intéresser sur la modélisation dimensionnelle, présentée les concepts qui y sont liés (table de faits, dimensions, mesures). Puis nous présentons les différents modèles et formes d'entreposage de données (en étoile, en flocon, et en constellation).

### 3.1. Définition de la modélisation dimensionnelle

Selon [Kimball, 2002] « Le modèle dimensionnel vise à présenter les données sous une forme intuitive dont l'objectif est de se rapprocher de la manière dont les décideurs perçoivent les données d'analyse »

Ce modèle désigne la structure de données la plus adaptable pour modéliser la grande masse de données stockées dans un entrepôt de données.

« Conceptuellement, cette modélisation multidimensionnelle a donné naissance aux concepts de fait et de dimension » [Kimball 2002]

Toutefois, Ce modèle est basé sur une table de fait et d'un ensemble de tables plus petites nommée tables dimensionnelles. Chacune de ces dernières possède une clé primaire unique. Les concepts de la modélisation dimensionnelle sont les suivants :

#### 3.1.1. Table de fait

Un "fait" représente la valeur d'une mesure, sur un sujet d'analyse. Une mesure est un indicateur d'analyse de type numérique.

"La table de faits est (dans la plupart des cas) est la table la plus volumineuse (avec le plus grand nombre de lignes) du modèle."

La table de fait est la table principale du modèle dimensionnel, elle contient les valeurs des mesures et les clés vers les tables de dimensions.

Exemple:

Table de faits des ventes	
Clés étrangères vers les dimensions	Clé date (CE)
	Clé produit (CE)
	Clé magasin (CE)
Faits	Quantité vendue
	Coût
	Montant des ventes

Figure 15 : Exemple de table des faits [web2]

### 3.1.2. Table de dimension

Une **dimension** est un axe d'analyse selon lequel sont visualisées les mesures d'activité d'un sujet d'analyse.

Les tables de dimensions accompagnent une table fait, servent à enregistrer les descriptions textuelles des dimensions de l'activité, contient moins d'enregistrement qu'une table de fait. « Une table de dimension établit l'interface homme / entrepôt, elle comporte une clé primaire », [Kimball, 2002].

Exemple :

Dimension produit
Clé produit (CP)
Code produit
Description du produit
Famille du produits
Marque
Emballage
Poids

Figure 16 : Table de dimension. [web2]

### 3.2. Différentes modèles de modélisation dimensionnelle

Afin de modéliser un DW, la table de faits et les tables dimensionnelles sont assemblées dans une structure de données qui corresponds aux besoins de l'entreprise ; afin de faciliter l'exploitation des données. Pour cela il existe plusieurs modèles de modélisation dont les plus utilisés sont :

#### 3.2.1. Modèle en étoile

Schéma en étoile est un modèle de données d'un ED. Ce modèle est inventé par **Kimball** qui indique que chaque table de fait représente un processus métier, ainsi il place la table de fait au centre de son diagramme celle-ci est relié à des tables dimensionnelles. Ces dernières ne sont pas reliées entre elles. Ce modèle est le plus utilisés car:

- Il offre la lisibilité et il est facile à comprendre par l'utilisateur.
- Sa facilité de navigation dans la base de données.

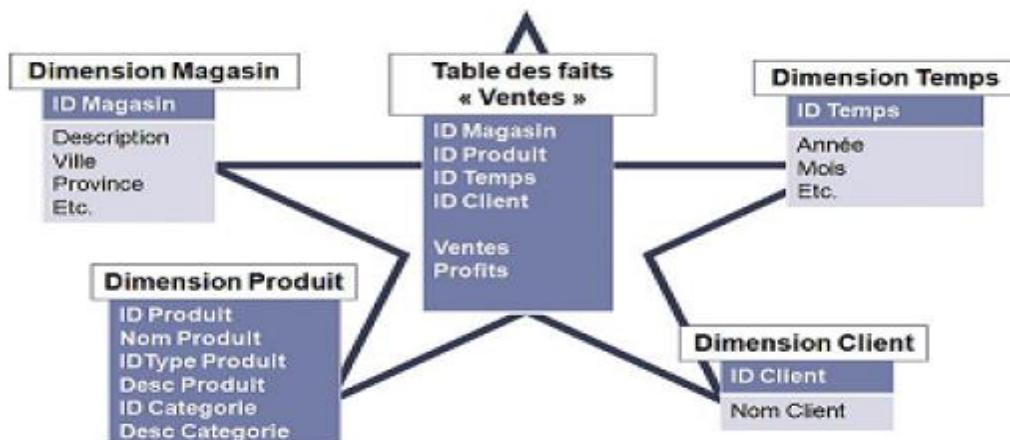


Figure 17 : Modèle en étoile (Star schéma). [web3]

#### 3.2.2. Modèle en flocon

Un modèle en flocon est un modèle proposé par Inmon pour lequel chaque dimension est décomposée en sous hiérarchie et représentée avec plusieurs tables.

L'avantage de cette modélisation permet d'expliquer la structure des hiérarchies en normalisant les dimensions ce qui permet d'économiser l'espace disque de stockage.

Par contre Les représentations en flocon sont déconseillées en général par [Kimball, Ross, 2008, p.56] car :

- Le modèle en flocon est plus complexe et son appréhension par l'utilisateur est difficile.
- Les performances en requête sont diminuées par les jointures.
- Le gain en espace disque est faible (les dimensions sont peu volumineuses relativement aux faits).

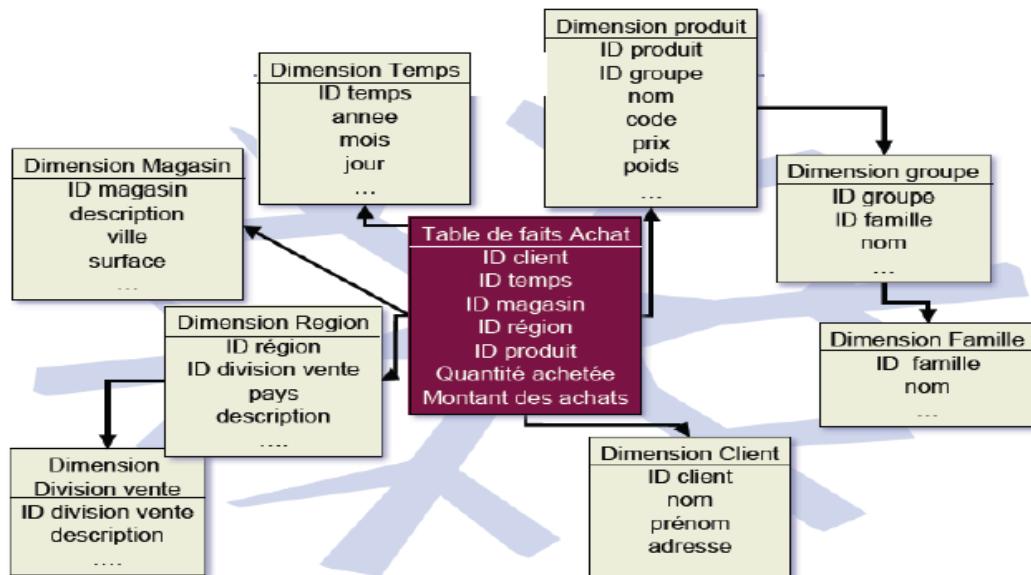


Figure 18 : Schéma de modèle en flocon.

### 3.2.3. Modèle en constellation

Ce modèle est une extension du modèle en étoile. Il consiste à fusionner plusieurs schémas en étoile qui utilisent des dimensions communes. Un schéma en constellation comprend donc plusieurs faits reliés à un ensemble de dimensions qui peuvent être partagées.

## 3.3. Caractéristiques des dimensions

### • Additives des données

« L'additivité est la possibilité de faire la somme des mesures présentées dans les tables de faits qui peuvent être parmi un de ces trois types » [Kimball, 2002].

- ✓ **Mesures additives** : elles peuvent être sommées suivant n'importe quelle hiérarchie d'une dimension. Par exemple, coût, vente, quantité.
- ✓ **Mesures semi-additives** : elles peuvent être sommées selon les hiérarchies de certaines dimensions mais pas toutes les dimensions.
- ✓ **Mesures non additives** : qui, leurs valeurs correspondent à des moyennes, des taux ou autres.

L'utilisateur peut spécifier les différentes fonctions telles que la somme, le minimum, le maximum, la moyenne.

### 3.4. Cube de données

La base de données multidimensionnelles s'appuie sur un hyper cube (cube à n dimensions). L'administrateur définit les dimensions, qui représentent une façon de trier l'information. Cette modélisation permet à l'utilisateur sans l'aide de l'informaticien d'avoir l'information dont il a besoin. L'utilisateur choisit deux ou trois critères à visualiser sous forme de tableau ou de cube.

Il peut également faire pivoter les axes d'analyse pour projeter les informations sous un angle différent. Ainsi, il peut visualiser les informations sous forme de cube ou chaque côté du cube dispose d'une dimension (exemple : produit, région, temps) et les caractéristiques correspondantes (exemple : région (Ouest, Est, Nord, Sud).

Chaque côté du cube La figure ci-dessous illustre parfaitement ça :

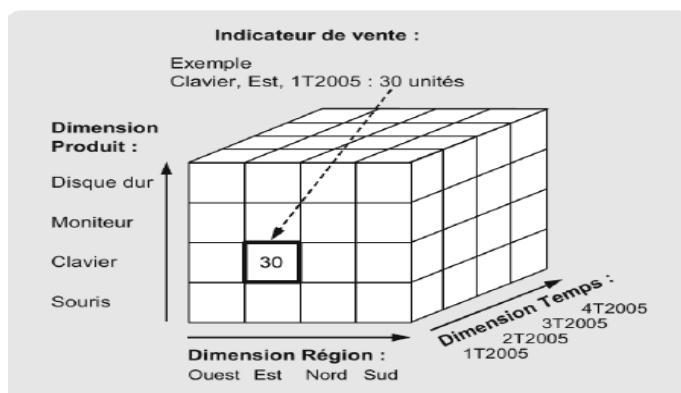


Figure 19 : Un cube de donnée.

### 3.5. Concept OLAP

Pour gérer son patrimoine informationnel les entreprises s'appuient sur de systèmes à savoir :

- OLTP (On Line Transaction Processing) pour les SGBD.
- OLAP (On Line Analytical Processing) pour un entrepôt de données.

#### 3.1.3. Définition de OLAP

Le concept OLAP (On-line AnalyticalProcessing) fut introduit et défini en 1993 par le docteur **E.F Codd** l'un des célèbres concepteurs de la base de données relationnelle.

Le concept de l'OLAP est défini comme une : " activité globale de requêtage et de présentation de données textuelles et numérique connues dans un entrepôt de données ; style d'interrogation et de présentation spécifiquement dimensionnel". [Kimball 2002]

Donc un entrepôt de données repose sur le système OLAP, il est caractérisé par un volume très important de données. Il fonctionne seulement en lecture et a pour objectifs principaux de rassembler, traiter et de présenter des données multidimensionnelles et venant de sources diverses à des fins d'analyse et de décision, offrant ainsi à l'utilisateur une vue orientée métier.

### **3.1.4. Définition OLTP**

Est un type de système qui gère les processus transactionnels en ligne (OLTP : "On-Line Transactional Processing"). Il est généralement composé d'un ensemble d'applications métiers autour d'une base de données (dite « opérationnelles »). OLTP a pour objectif de pouvoir insérer, modifier, intégrer rapidement et par plusieurs utilisateurs simultanément les données liées à la transaction de l'entreprise.

### **3.1.5. OLAP vs OLTP**

Après avoir défini précédemment le concept OLAP et OLTP, et la place de chacun dans l'entreprise, on trouve que le système opérationnel et le système décisionnel présentent quelques différences. Ces différences peuvent être perçues du point de vue des données ainsi que celui de l'utilisation. Le tableau ci-dessous illustre la différence entre ces deux systèmes :

Caractéristiques	OLTP	OLAP
<b>Utilisation</b>	SGBD (production)	Data Warehouse
<b>Opération typique</b>	Mise à jour	Analyse
<b>Type d'accès</b>	Lecture/ Ecriture	Lecture
<b>Niveau d'analyse</b>	Elémentaire	Global
<b>Quantité d'informations échangées</b>	Faible	Importante
<b>Orientation</b>	Ligne	Multidimensionnelle
<b>Taille DB</b>	Faible (max quelques GB)	Importante (>TB)
<b>Ancienneté des données</b>	Récente	Historique

Tableau 4 : Comparaison entre OLTP vs OLAP.

### 3.6. Architecture des systèmes 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). Il y a plusieurs versions d'OLAP, mais les trois les plus répandues sont :

- Architecture ROLAP :(Relationnal OLAP).
- Architecture MOLAP :(Multidimensionnel OLAP).
- Architecture HOLAP :(Hybrid OLAP).

#### 3.6.1. ROLAP (Relationnal OLAP)

« Est un ensemble d'interfaces utilisateurs et d'applications qui donnent une vision dimensionnelle à des bases de données relationnelles » [Kimball, 2002].

Les systèmes ROLAP « **R**elationnel **O**n-line **A**nalys**T**ic **P**rocessing » sont en mesure de simuler le comportement d'une SGBD multidimensionnel en exploitant un SGBD Relationnel. Ainsi L'utilisateur aura l'impression d'interroger un cube multidimensionnel alors qu'en réalité il ne fait qu'adresser des requêtes sur une base de données relationnelles. Par contre il est plus facile et moins cher à mettre en place. Cette technique reste la plus efficace dans la mesure où elle permet de stocker un grand volume de données.

Exemple de moteur ROLAP : Mondrian

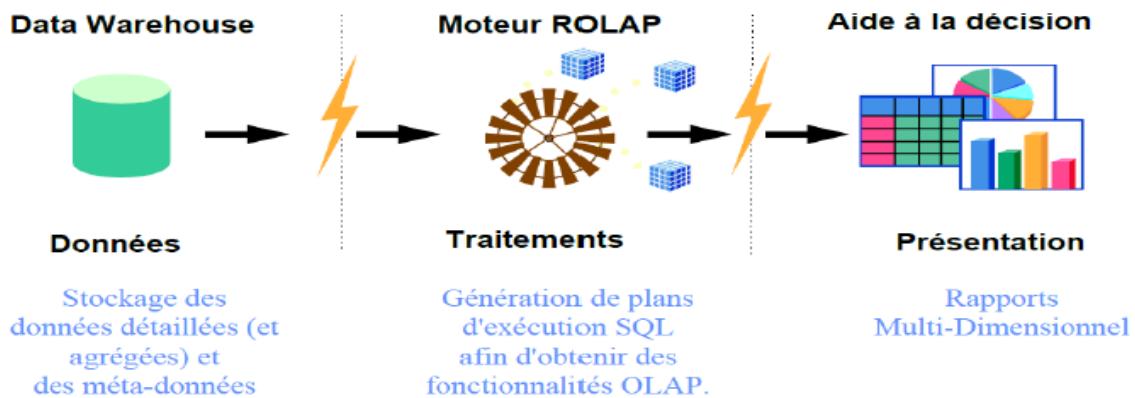


Figure 20 : Architecture ROLAP [NAKACHE, 1998]

### 3.6.2. MOLAP (Multidimensional OLAP)

«Est un ensemble d'interfaces d'applications et de technologies de bases de données propriétaire dont l'aspect dimensionnel est prépondérant » [Kimball, 2007].

Cette architecture s'appuie sur une base de données multidimensionnelle (cube) qui permet de stocker les données fréquemment utilisés nécessitant un temps de réponse minimal. Son inconvénient qu'elle est plus cher à mettre en place, et elle ne supporte pas une grande masse de donnée.

Exemple de moteur MOLAP : Microsoft Analysis Services.

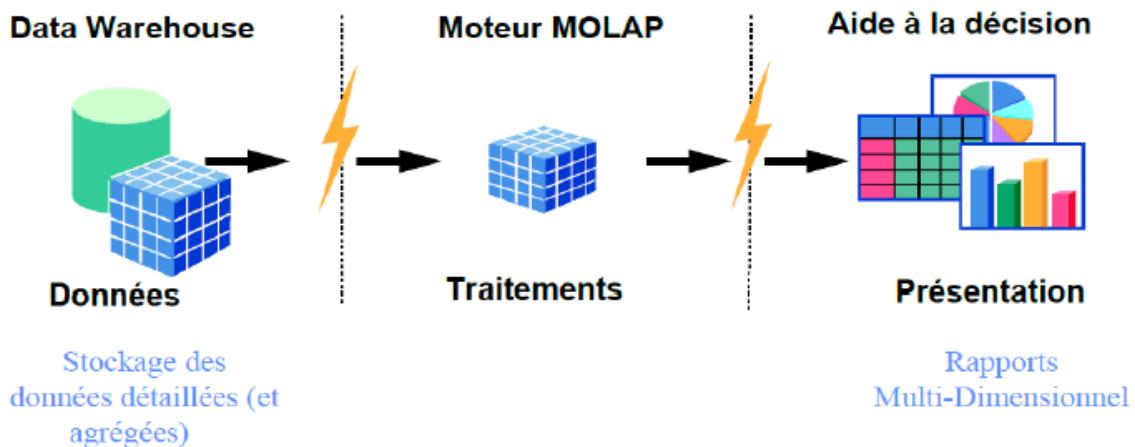


Figure 21 : Architecture MOLAP [NAKACHE, 1998].

### **3.6.3. HOLAP (Hybrid OLAP)**

Cette architecture est une solution hybride entre MOLAP et ROLAP .Cette approche combine les deux approches relationnelles et dimensionnelles. En effet, les données détaillées de base d'un entrepôt de données sont stockées dans une base de données relationnelle et les données agrégées sont stockées dans une base de données multidimensionnelle.

Son avantage qu'elle est bonne au niveau du coût et de performance.

## **Conclusion**

Dans ce chapitre nous avons pu voir que la modélisation dimensionnelle est la structure de données la plus adaptés pour l'implémentation d'un entrepôt de données et celle qui répond parfaitement aux attentes des utilisateurs. Puis nous avons présenté la technologie OLAP et son importance dans l'informatique décisionnelle. Cette technique offre aux utilisateurs la possibilité d'analyser les données suivant les axes propres à leur métier.

## Synthèse de la première partie

Dans cette première partie, nous avons traité l'aspect théorique de notre projet, où nous avons introduit les notions fondamentales concernant les systèmes décisionnels, ainsi que les data warehouse et leur rôles dans les organisations.

Nous avons défini en premier lieu, le concept du système décisionnel, ainsi que son rôle crucial au sein des entreprises. Puis nous avons présenté son architecture qui s'appuie sur les systèmes de production (base de données sources) qui sont peu adaptés à la prise de décision.

Puis nous avons focalisé dans le second chapitre sur le DW qui est au cœur de tout un système décisionnel ou nous avons définis qu'elle une zone de centralisation des informations décisionnelles en assurant l'intégration des données extraites, la pérennité de ces données dans le temps. Après l'étude de l'importance d'ED par rapport à la BDD et nous avons terminé avec la démarche constructive d'un entrepôt de données.

Nous avons ensuite abordé la modélisation dimensionnelle, en recouvrant les différents modèles dimensionnels existants ; à savoir le modèle étoile, flocon, et constellation, qui sont avec le cube de données. Nous avons abordé aussi le concept OLAP sur lequel repose un data warehouse avec ces différentes architectures (ROLAP, MOLAP, HOLAP).

Dans la partie suivante, nous allons présenter la deuxième partie de notre PFE ou nous allons étudier l'organisme d'accueil de BMT-spa avec les différents systèmes dont elle dispose, puis concevoir et déployer la solution répondant à leurs besoins.

---

## **Partie 02: Analyse.**

---

Chapitre 04 : Cas pratique de la BMT-spa

Chapitre 05 : Analyse de l'existant

Chapitre 6: Analyse de besoins.

# Chapitre4:Cas pratique de la BMT-spa.

## Introduction

Dans ce chapitre, nous allons présenter l'environnement professionnel dans lequel notre projet se déroule.

Nous commençons d'abord par une brève présentation de la BMT-SPA, puis nous introduisons la structure générale de son organisation avec ses différentes directions en particulier sa direction marketing pour qui notre projet est destiné, nous détaillons son organigramme ainsi que ses objectifs.

### 4.1. Présentation de l'organisme d'accueil

**BMT (Bejaia Méditerranéen Terminal) - SPA** est une jointe venture entre l'**Entreprise Portuaire de Bejaia**. **EPB** est l'autorité portuaire qui gère le port de Bejaïa. **PORTEK System and Equipment**, une filiale du groupe PORTEK, est un opérateur de Terminaux à conteneurs présent dans plusieurs ports dans le monde et également spécialisé dans les équipements portuaires.

L'activité principale de **BMT** est la gestion et l'exploitation du Terminal à conteneurs. Sa mission principale est de traiter dans les meilleures conditions de délais, de coûts et de sécurité, l'ensemble des opérations qui ont rapport avec le conteneur. Pour se faire, elle s'est dotée d'équipements performants et de systèmes informatiques pour le support de la logistique du conteneur afin d'offrir des services de qualité, efficaces et fiables pour assurer une satisfaction totale des clients.

BMT veille au développement et à la gestion de son terminal à conteneurs où l'intégrité, la productivité, l'innovation, la courtoisie, et la sécurité sont de rigueur. BMT est constamment soucieuse des intérêts de ses clients avec lesquels elle partage le souci de performance et de coût.

Elle met à la disposition de ses clients des ressources humaines et des moyens nécessaires pour optimiser sa productivité et atteindre des niveaux de performance concurrentielle.

## 4.2. Situation géographique

Implanté au centre du pays, au cœur de la méditerranée dans le nord du continent africain, le Port de Bejaia occupe une situation géographique stratégique. Il dessert un hinterland (arrière-pays) important et très vaste. La ville, le Port et le terminal à conteneurs de Bejaia disposent de ce fait de voies de communication reliant l'ensemble des routes du pays, des voies ferroviaires et à proximité d'un aéroport international.



Figure 22 : Situation géographique de BMT.

### 4.3. Organigramme de BMT: [document interne de l'entreprise]

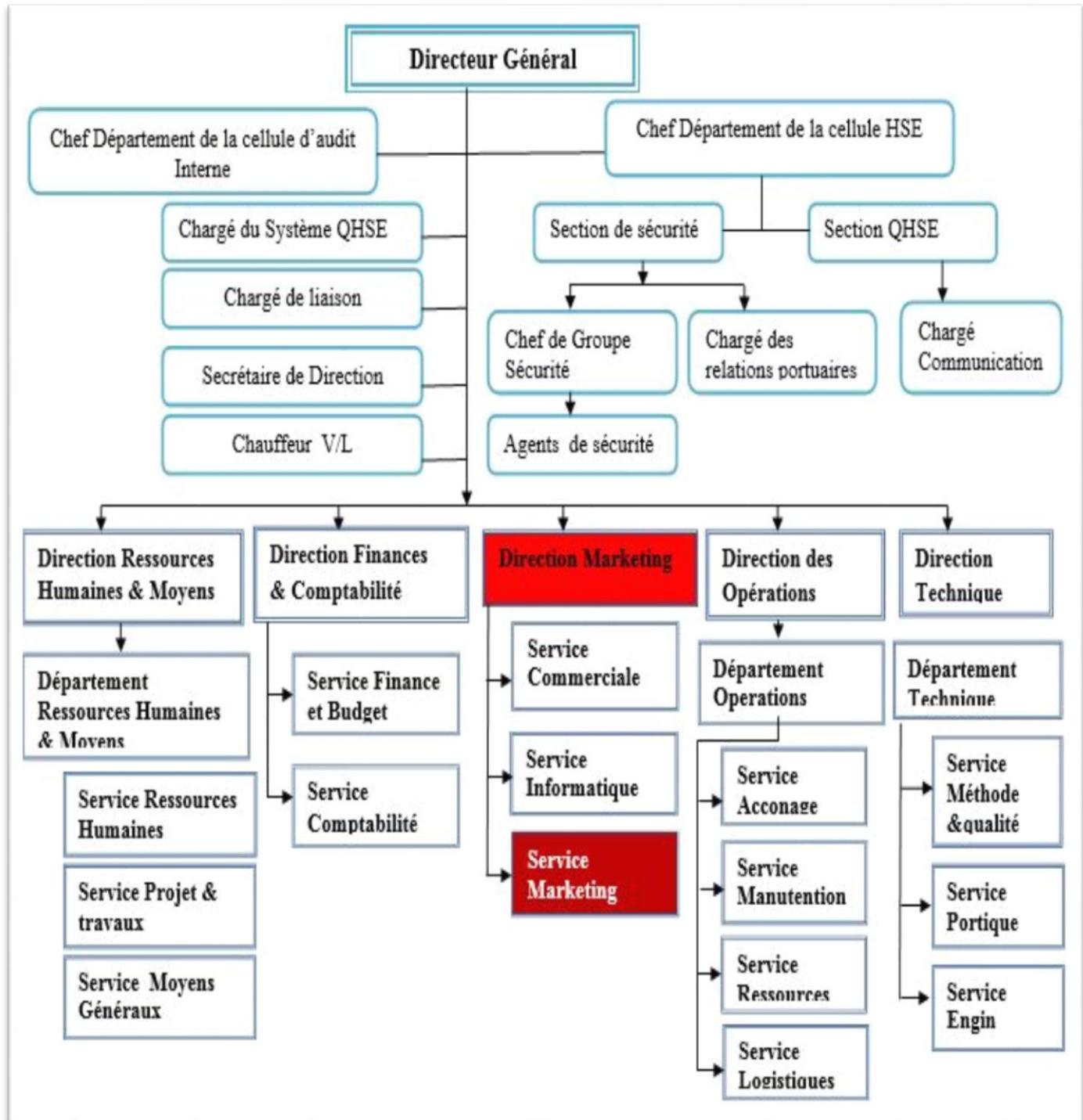


Figure 23 : Organigramme de BMT-SPA

## 4.4. Différentes structures de BMT

### 4.4.1. Direction générale

A sa tête le directeur général qui gère l'entreprise, a le pouvoir de décision, administre l'entreprise, assigne des directives pour les différents structures et fait la liaison entre les directions de l'entreprise.

- **Cellule d'audit et contrôle**

Assure le suivi et l'audit des procédures et la veille réglementaire.

- ✓ Secrétaire de direction.
- ✓ Chauffeur V/L (Véhicule/Légers).

- **Cellule QHSE (Qualité & Hygiène Sécurité d'entreprise):**

Assure la mise en œuvre et suivi du plan QHSE de l'entreprise.

- **Cellule Sécurité :** Assure la sécurité de la marchandise, du parc à conteneurs et la propriété de l'entreprise.

- **Les directions du métier**

- ✓ Direction des Ressources Humaines et Moyens.
- ✓ Direction des Finances et Comptabilité.
- ✓ Direction Marketing.
- ✓ Direction des Opérations.
- ✓ Direction Technique.

### 4.4.2. Direction des ressources humaines et moyens

Sa mission est de mettre en œuvres des systèmes de gestion intégrée à la stratégie de BMT pour atteindre ses objectifs et qui traduisent une adéquation entre les impératifs économiques et les attentes du personnel.

### 4.4.3. Direction des finances et comptabilité

La mission de la Direction des Finances et Comptabilité est :

- ✓ Veiller à l'adéquation de la politique financière de l'entreprise avec les objectifs globaux.
- ✓ Cordonner et suivre les relations avec les institutions financières.
- ✓ Assurer les relations avec les banques, et les administrations fiscales et parafiscales.

- ✓ Assurer le recouvrement des créances de toute nature.
- ✓ Etablir et suivre les budgets et les plans de financement.
- ✓ Elaborer les plans de financement en assurant l'actualisation et l'exécution.
- ✓ Déterminer, rechercher et négocier les financements les plus appropriés en relation avec les établissements concernés.
- ✓ veiller à l'application des règles comptables et à la tenue correcte des livres au sein de la société.
- ✓ Elaborer le bilan et autres états financiers et comptables.
- ✓ Etablir et analyser le bilan de fin d'année

#### **4.4.4. Direction des opérations**

La mission de la Direction des Opérations est de :

- ✓ Assurer la planification des escales, de parc à conteneurs et la planification des ressources, équipes et équipements.
- ✓ Prendre en charge les opérations de manutentions, comme la réception des navires porte-conteneurs et leurs chargements et déchargement.

Suivre les opérations de l'accostage tel que : le suivi des livraisons, dépotages, restitutions des conteneurs vides et le traitement des conteneurs frigorifiques.

#### **4.4.5. Direction technique**

Sa mission est d'assurer une maintenance préventive et curative des engins du parc à conteneurs. Voici ces différents services :

- **Service engins**  
Assure l'entretien des véhicules lourds.
- **Service portiques**  
Assure l'entretien des portiques et de la grue mobile.
- **Service méthodes**  
Assure la mise en œuvre du plan de maintenance des équipements.

### **4.3.6. Direction de marketing**

Cette direction veille à la marque de l'entreprise en se préoccupant en permanence d'entretenir des relations fidèles avec ses clients. Elle est restructurée récemment en trois service que nous allons présenter ci-dessous :

- **Service commercial**

Le service commercial veille à l'élaboration d'une bonne politique commerciale et tarifaire, en appliquant des plans d'action de l'entreprise en termes d'efficacité de facturation de recouvrement afin d'améliorer et d'assurer la pérennité de sa relation avec ses clients.

Pour le déroulement de la procédure de facturation le service commercial suit un schéma de réglementation [**annexe 1**].

Toutefois, à chaque fin de prestation réalisée (manutention, acconage, autres) réalisée au port ; des documents (bon de commande, bon délivré... etc.) sont établis par la direction des opérations puis ils seront transmis au service commercial. Le chef de service commerciale contrôle la conformité de ces dossiers à facturer. En cas d'erreur la signale au service des opérations pour rectification, sinon il applique la procédure de tarif pour chaque type de prestations effectuée [**annexe 2**], puis saisie les données dans l'application GestCom, et la valide .Enfin, faire parvenir la facture finale aux clients.

- **Service Marketing**

Elle vise à faire connaître ses missions, ses programmes, ses orientations et ses performances auprès de ses clients. Elle amène son environnement externe à prendre conscience de l'importance des démarches qu'elle entreprend dans le développement et l'amélioration de la qualité des services.

- **Service informatique**

Il met à la disposition des acteurs de BMT les moyens informatiques (matériels, logiciels) permettant la mise en œuvre du système d'information global. Il gère aussi les ressources informatiques de l'entreprise et veille à ce qu'elle soit un pôle d'aide et de compétences relatif aux NTIC (Nouvelles Technologies de l'Information et de la Communication).

Sa mission est d'assurer :

- ✓ Administration du système CTMS (Container Terminal Management System).
- ✓ la maintenance du parc informatique de l'entreprise et le développement de nouvelles applications aux différentes structures de l'organisation.
- ✓ Assure la veille technologique en matière de la communication et de l'information.

- **Organigramme de la direction Marketing**

La direction marketing est divisé en trois service (commercial + Marketing + Informatique) que nous avons vu ci-dessus, qui ont pour mission d'assumer le rôle de présentation de l'entreprise en Algérie et à l'étranger.

La direction de marketing est organisée comme suit :

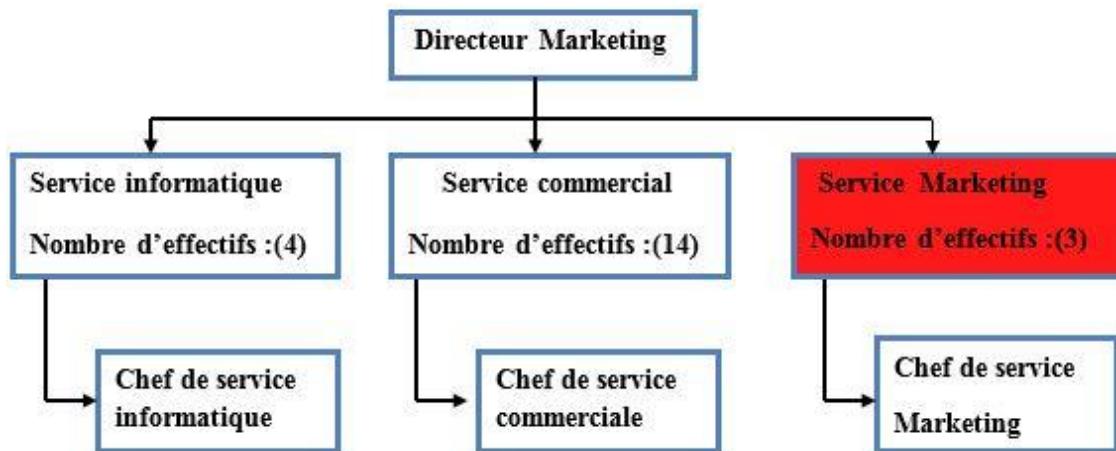


Figure 24 : Organigramme de la direction de Marketing.

- **Missions du service du marketing**

- ✓ Assure la promotion de l'image de marque de l'entreprise et la mise en œuvre du plan d'action de marketing
- ✓ Coordonner et veiller à la bonne exécution des actions marketing,
- ✓ Assumer le rôle de représentation de l'entreprise en Algérie et à l'étranger.
- ✓ Faire des classifications régulières pour leurs clients afin de les segmenter.
- ✓ Veiller à assurer la qualité de service.
- ✓ Augmenter le taux de pourcentage de leur marché.

## 4.5. Les activités de BMT

Bejaia Méditerranéen Terminal (BMT) reçoit annuellement un grand nombre de navires pour lesquels elle assure les opérations de planification, de manutention et d'accostage avec une traçabilité des opérations.

#### **4.5.1. Operations planification**

- Planification des escales.
- Planification du déchargement/chargement.
- Planification du parc à conteneurs.
- Planification des ressources : équipes et moyens matériels.

#### **4.5.2. Operations de manutention**

- La réception des navires porte-conteneurs
- Le déchargement / chargement des conteneurs du navire.

#### **4.5.3. Operations d'acconage**

- Transfert des conteneurs vers les zones d'entreposage.
- Transfert des conteneurs frigorifiques vers la zone « reefers ».
- Mise à disposition des conteneurs aux services de contrôle aux frontières.
- Suivi des restitutions et des mises à quai pour embarquement ou débarquement.
- Sécurité absolue sur le terminal.

### **4.6. Objectifs de BMT**

BMT a pour objectif de faire de son terminal à conteneur une infrastructure moderne à même de répondre aux exigences les plus sévères en matière de qualité dans le traitement du conteneur. La mise à disposition d'une nouvelle technologie dans le traitement du conteneur pour :

- Faire face à la concurrence nationale et internationale.
- Gagner des parts importantes du marché
- Mettre en place des procédures efficaces de gestion et une prestation de service répondant aux normes universelles
- Satisfaction complète de la clientèle et usagers portuaires en matière de transport et de manutention
- Tenir l'engagement d'assurer un service de qualité dans les meilleurs délais
- Offrir un niveau élevé de l'efficacité opérationnel pour les clients
- Obtenir l'excellence dans la gestion des opérations terminales.

## Conclusion

A travers ce chapitre, nous avons introduit l'organisme d'accueil de BMT-spa. Cette présentation nous a permis de mieux comprendre sa structure ainsi le fonctionnement de ces direction en particulier le rôle et les missions de la direction de marketing ou notre stage s'est déroulé.

Dans le prochain chapitre nous évoquons le cadre de notre étude, en analysant les systèmes opérationnels et le système décisionnel utilisés par la direction de marketing particulièrement par le service de marketing.

# **Chapitre 5:Analyse de l'existant.**

## **Introduction**

L'étude de l'existant constitue la phase la plus importante dans un cycle de développement d'un projet informatique. En effet, la connaissance du système actuel est indispensable pour pouvoir porter un jugement juste et avoir toutes les informations nécessaires afin de proposer une solution adéquate qui répond aux besoins des utilisateurs.

Nous allons essayer, à travers ce chapitre, faire une analyse aussi complète que possible sur l'existant informatique à la BMT-SPA, en terme de matériaux, architecture réseau et ainsi que de système de production .en suite, étudier les applications métiers utilisés par le service marketing ou notre projet s'est déroulé ainsi que leur processus de prise de décision actuel. Enfin mettre un diagnostic de cette analyse, ce qui permet ensuite de compléter le chapitre suivant qui est l'étude des besoins.

### **5.1. Situation informatique de BMT**

Vue le volume et la complexité des informations traité quotidiennement par la BMT, ont donc nécessité le recours à l'outil informatique.

La circulation fiable et rapide de l'information et de donnée est une condition primordiale pour le bon acheminement de la gestion des terminaux et assurer la compétitivité commerciale du port de Bejaia.de ce fait, BMT-Spa possède un grand parc informatique que nous allons présenter ci-dessous :

#### **5.1.1. Parc informatique**

##### **A. Réseaux**

Afin de se connecter au réseau Internet, la BMT-spa s'appuie sur la technologie WIMAX. Celle-ci possède trois serveurs d'Active Directory (Windows server 2008), ainsi que deux Switch de marques CISCO répartis sur les services existant et reliant les autres terminaux.

Cette société dispose de deux types de réseaux :

1. Un réseau LAN (Giga Ethernet) basé sur un matériel Cisco relié à une VPN (Virtual Privat Network).
2. Un réseau WAN basé aussi sur un matériel Cisco reliant les sites distants par une fibre optique (2Mb/s).

Le parc informatique disponible ainsi que la configuration du réseau permettent de choisir entre plusieurs types d'architecture :

- **L'architecture Client/serveur** : la présence d'un serveur de données, et d'applications ainsi qu'un nombre important de PC répartis sur les différentes structures favorisent ce type d'architecture.
- **L'architecture 3tiers** : la présence d'un réseau de serveurs de données (oracle 11g), serveurs d'application (apache) ainsi que des serveurs web logique (8.1) ainsi que des PC avec une grande puissance de calcul permet d'adopter cette solution.

La cartographie ci-dessous montre l'architecture réseau de BMT-spa :

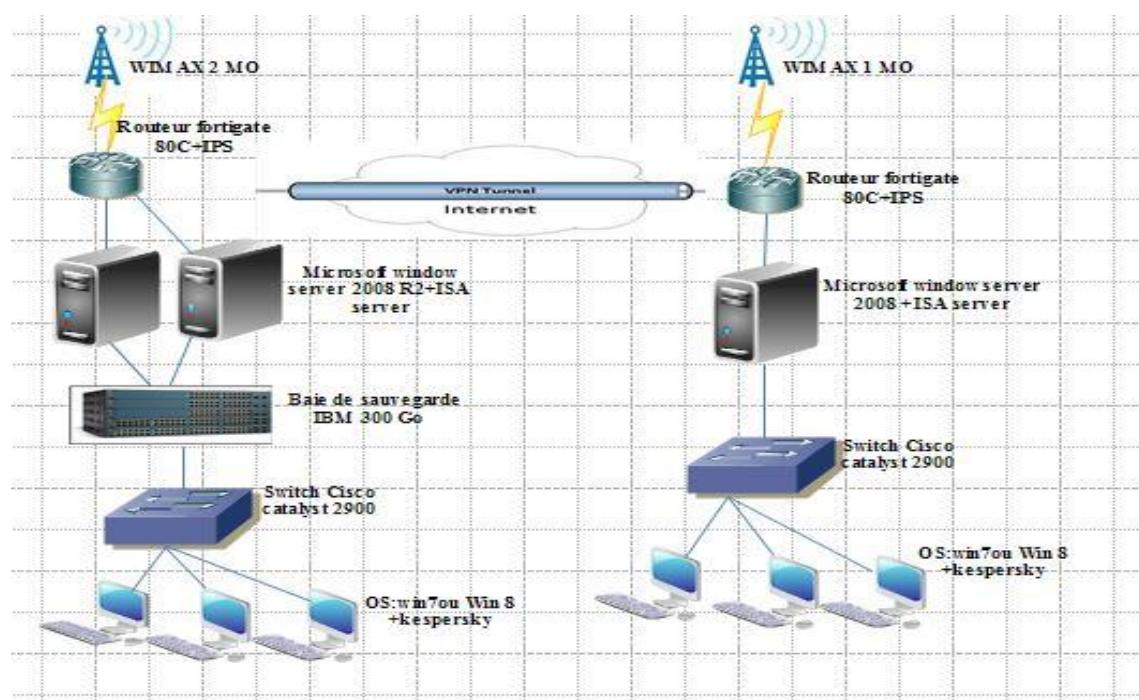


Figure 25 : Structure réseau de BMT-spa.

## B. Ordinateur

La BMT possède un important dispositif informatique en matière d'ordinateur (80% ACER, HP). Ces PC sont dotés de processeurs, des mémoires vives de différentes performances,

disposant tous d'un antivirus Kaspersky endpoint Security 10. Leurs System exploitation est Windows 7 ou Windows 8.

### C. Applications Existantes et environnement de développement

Pour développer, déployer et exécuter des Applications de tous types, BMT utilise la plate-forme .NET, Java SE, ou J2EE, qui repose sur le système d'exploitation Windows. Comme système de gestion de base de données(SGBD), elle utilise Oracle 11G, et MySQL 5.

- **Application utilisées**

BMT dispose d'un système de gestion du Terminal à conteneurs **CTMS** (Container Terminal Management Système) qui est développé par un prestataire de service. Cette application est basé sur environnement client/ serveur, qui gère les activités fonctionnelles et opérationnelles du port en temps réel.

Les activités opérationnelles dont gère l'application « CTMS » renferment les domaines suivant : Navire, conteneurs...Etc. Tandis que les activités fonctionnelles représentent le second aspect du système « CTMS» concernent la gestion des ressources humaines, et du parc auto.

BMT dispose aussi d'une application **GestCom** de gestion de portefeuille client développé en interne par le service informatique, qui permet de suivre les facturations des clients dans un délai précis.

- **Etude détaillée de ces applications**

Nous avons procédé à une étude détaillée des applications utilisées par le service du marketing de l'entreprise BMT-spa.

Le tableau ci-dessous montre la description de ces applications, ces volets et ses utilisateurs finaux.

#### ❖ Description des modules du Système CTMS

CTMS pilote l'ensemble des processus métiers de l'entreprise .il est composé de 4 modules essentielles qui sont :

- ✓ Ship planning.
- ✓ Yard planning.
- ✓ Ressource planning.
- ✓ CIM (Client Interface Management).

Les tableaux ci-dessous illustrent en détail chacune de ces rubriques avec leurs volets

**LA RUBRIQUE SHIP PLANNING**

<b>DESCRIPTION GENERALE</b>	<p>➤ PLANIFICATION NAVIRE :</p> <p>CE MODULE EST CONÇU POUR LA PLANIFICATION DU NAVIRE ET LES DIFFERENTES OPERATIONS LIÉE A LUI (PLANIFICATION D'ESCALE, ACCOSTAGE A QUAI, ARRIMAGE, DEBARQUEMENT EN RADE...ETC.). CE MODULE CONTIENT PLUSIEURS VOLETS.</p>
<b>SES VOLETS</b>	<ul style="list-style-type: none"> <li>▪ <b>VESSEL STATISTIC</b> : COMPORTE TOUS LES DETAILS SUR LE NAVIRE (NOM DU NAVIRE, DATE D'ACCOSTAGE...).</li> <li>▪ <b>VESSEL CALL</b> : PLANIFICATION D'ESCALE DES NAVIRES.</li> <li>▪ <b>VESSEL PLANNING</b> : PERMET D'ENREGISTRER LES OPERATIONS D'ARRIMAGE, ET LA STRUCTURE NAVIRE.</li> </ul>
<b>LES UTILISATEURS FINAUX</b>	CETTE RUBRIQUE EST DESTINÉE A LA DIRECTION DES OPERATIONS.

**LA RUBRIQUE YARD PLANNING**

<b>DESCRIPTION GENERALE</b>	<p>➤ PLANIFICATION DU PARC:</p> <p>CE MODULE EST DEVELOPPE POUR PLANIFICATION DU PARC AFIN D'ASSURER LES OPERATIONS D'ACCONAGE (DECHARGEMENT / CHARGEMENT) DES CONTENEURS DU NAVIRE A QUAI. CELLE-CI COMPORTE PLUSIEURS VOLETS QUI PERMETTENT DE SUIVRE AUTOMATIQUEMENT L'ETAT DU PARC AINSI QUE L'AVANCEMENT DE CES OPERATIONS.</p>
<b>SES VOLETS</b>	<ul style="list-style-type: none"> <li>▪ <b>MACRO PLANNING</b> : CE VOLET EST CONÇU COMME UNE VUE GLOBALE AFIN D'ASSURER LA PLANIFICATION DES CONTENEURS DANS DES BLOCS (LIEU DE LEUR DÉPÔTS).</li> <li>▪ <b>MICRO PLANNING</b> : CE VOLET EST CONÇU COMME UNE VUE DÉTAILLÉE POUR SUIVRE LA PLANIFICATION DE CONTENEURS DANS LEUR BLOC.</li> </ul>
<b>LES UTILISATEURS</b>	CETTE RUBRIQUE EST DESTINÉE A LA DIRECTION DES OPERATIONS

<b>LA RUBRIQUE RESOURCE PLANNING</b>	
<b>DESCRIPTION GENERALE</b>	<p>➤ <b>RESSOURCE PLANNING :</b>      CE MODULE EST CONÇU POUR LA PLANIFICATION DES RESSOURCES (MATERIELS ET HUMAINS) ET LES METTRE EN DISPOSITION AFIN D'ASSURER TOUTES LES OPERATIONS LIÉES AUX NAVIRES (MANUTENTION ET ACCONAGE), LE BON DÉROULEMENT DE PLANIFICATION DU PARC.</p>
<b>SES VOLETS</b>	<ul style="list-style-type: none"> <li>▪ <b>STAFF DEPLOYMENT FOR SHIP OPERATIONS :</b> CE VOLET PERMET DE DEPLOYER TOUTES LES RESSOURCES NÉCESSAIRES POUR ASSURER LES OPERATIONS (AMARRAGE, ACCOSTAGE ET DEBARQUEMENT ET EMBARQUEMENT A QUAI) DU NAVIRE.</li> <li>▪ <b>STAFF DEPLOYMENT FOR YARD OPERATIONS :</b> C'EST DEPLOYER TOUS LES RESSOURCES DISPONIBLE POUR ASSURER LES DIFFÉRENTES OPERATIONS REALISER DANS LE PARC (TELS QUE ENTREPOSAGE, LE SHIFTING ...ETC).</li> <li>▪ <b>STAFF ROSTER:</b> CE VOLET EST CONÇU POUR SAVOIR LE STAFF UTILISE POUR LE SERVICE D'AVANCEMENT DES OPERATIONS.</li> </ul>
<b>LES UTILISATEURS FINAUX</b>	CETTE RUBRIQUE EST DESTINÉE À LA DIRECTION DES OPERATIONS.

Tableau 5 : Tableau descriptif des modules de CTMS.

### ❖ Description de l'application CIM

<b>CIM (CLIENT INTERFACE MANAGEMENT)</b>	
<b>DESCRIPTION GENERALE</b>	<b>➤ GESTION DES CLIENTS :</b>  EST UNE APPLICATION WEB QUI EST CONÇUS POUR SUIVRE LA GESTION DU CLIENT AVEC SON ROLE DANS LES ACTIVITES D'IMPORT ET D'EXPORT. ELLE COMPORTE CERTAINS VOLETS LIEE AUX CTMS TELS QUE YARD PLANNING....
<b>SES VOLETS</b>	<ul style="list-style-type: none"> <li>▪ <b>VESSEL</b> : COMPORTE LES DETAILS SUR LE NAVIRE.</li> <li>▪ <b>IMPORT</b> : COMPORTE LES INFORMATIONS NECESSAIRES SUR L'ACTIVITE D'IMPORTATION.</li> <li>▪ <b>EXPORT</b> : COMPORTE LES INFORMATIONS NECESSAIRES SUR L'ACTIVITE D'EXPORTATION.</li> <li>▪ <b>YARD</b> : COMPORTE LES INFORMATIONS LIEES AUX OPERATIONS DEROULEES AUX PARCS.</li> <li>▪ <b>OPERATION</b> : PERMET LA GESTION DE L'HISTORIQUE DES OPERATIONS FAITES SUR LE CONTENEUR AINSI QUE LEURS MAJ.</li> <li>▪ <b>MONITORING</b> : LES INFORMATIONS DE CONTROLE</li> <li>▪ <b>BILLINGS</b> : COMPORTE LES INFORMATIONS SUR LA FACTURATION DU NAVIRE.</li> </ul>
<b>LES UTILISATEURS</b>	CETTE RUBRIQUE EST DESTINEE A LA DIRECTION DES OPERATIONS.

Tableau 6: Description des fonctionnalités du CIM.

### ❖ Description de l'application GestCom

<b>L'APPLICATION GESTCOM</b>	
<b>DESCRIPTION GENERALE</b>	<b>➤ GESTCOM :</b>  EST UNE APPLICATION CONÇUS SPECIALEMENT AU SERVICE COMMERCIAL POUR LEUR FACILITER LA PROCEDURE DE FACTURATION D'UNE MANIERE AUTOMATIQUE AFIN DE GAGNER DU TEMPS, ASSURER LA FIDELITE DES DONNEES.
<b>SES VOLETS</b>	<ul style="list-style-type: none"> <li>▪ <b>FACTURATION</b> : COMPORTE LES INFORMATIONS DETAILLEES SUR LES FACTURES ETABLIES AVEC LEUR HISTORIQUE.</li> <li>▪ <b>RECOUVREMENT</b> : CE VOLET PERMET DE SUIVRE LE RECOUVREMENT DE LA CREANCE DES CLIENTS.</li> <li>▪ <b>TABLEAU DE BORD SUIVIS CLIENT</b> : REPRESENTE L' INDICATEUR SORTIE DES FACTURATIONS REALISEES AFIN DE LEUR PERMETTRE DE MESURER LEUR CHIFFRE D'AFFAIRE PAR CLIENT ET PAR PRESTATION.</li> </ul>
<b>LES UTILISATEURS</b>	CETTE APPLICATION EST UTILISEE PAR LE SERVICE COMMERCIAL

Tableau 7:Tableau descriptif des fonctionnalités de GestCom.

## 5.2. Etat du système décisionnel actuel de BMT

Actuellement, le service marketing ne dispose d'aucun système d'aide à la décision automatique qui peut répondre à la majorité des besoins des décideurs. Tout processus de prise de décision se base en général sur les rapports dont les données sont:

- 1) Extraite à partir des systèmes transactionnels (CTMS, GestCom).
- 2) Extraite à partir d'autres sources comme les fichiers Excel.

### 5.2.1. Description du déroulement des tâches

Dans le cadre du suivi de l'activité de BMT (manutention et acconage, facture client), un rapport d'activité mensuel, semestriel et annuel est établie par le service marketing. Ainsi tout processus d'analyse à ce niveau s'appuie essentiellement sur des rapports dont les données sont extraites et consolidées à partir des systèmes transactionnels d'une manière manuelle résumé dans les deux phases suivantes :

#### A. Extractions de données

L'extraction de données s'effectue au niveau des systèmes opérationnels de manière périodique, ils se font généralement chaque mois, d'autres se font chaque semestre, ou chaque année auprès des différentes sources de données, celles-ci sont fournies sous format fichiers plats au département marketing.

Le tableau ci-dessous présente d'exemple d'extraction de donnée :

EXTRACTION DE DONNEE	SOURCE DE DONNEE	PERIODE	TYPE DE FICHIER FOURNI
NUMERO DE CONTENEUR DEBARQUE, TYPE DE CONTENEUR (20OU 40), DATE D'ACCOSTAGE, NOM IMPORTATEUR, NOM DE COMPAGNIE MARITIME.	CTMS	MENSUELLE, SEMESTRIELLE APRES ANNUELLE	FICHIER PLAT (EXCEL)
CHIFFRE D'AFFAIRE PAR IMPORTATEUR (CLIENT), PAR PRESTATION, PAR CODE CLIENT, PAR NOM DU CLIENT.	GESTCOM	MENSUEL, SEMESTRIELLE APRES ANNUELLE	FICHIER PLAT (EXCEL)
CODE CLIENT, NOM CLIENT, DATE, VILLE, REGION, PRODUIT, NOMBRE EVP (20OU 40) IMPORTE.	GESTCOM	MENSUEL, SEMESTRIELLE APRES ANNUELLE	FICHIER PLAT (EXCEL)
DATE DE MOUILLAGE, DATE D'APPAREIL-LAGE (DATE QUITTE LE QUAI), TOTAL MOVE (NOMBRE DE MANIPULATION FAITE SUR LE NAVIRE).	FICHIER EXCEL	MENSUEL, SEMESTRIELLE APRES ANNUELLE	FICHIER PLAT (EXCEL)

Tableau 8: Tableau descriptif de la procédure d'extraction de données.

## B. Consolidation de données

La consolidation de données se fait d'une manière semi-automatique, c.à.d. que les rapports d'analyses sont générés par deux Marketeurs en utilisant des outils très classiques comme le logiciels Microsoft Office (Excel et Word). Chaque rapport généré à la fin de mois (voir semestre, puis année) est envoyé vers le chef de service qui fera une vérification sur ces derniers puis il le fera valider et envoyer au directeur de marketing.

Le tableau ci-dessous présente l'établissement de quelques rapports :

RAPPORT	DONNEE SOURCE	TYPE DE FICHIER ETABLIS	PERIODE D'ETABLISSEMENTS RAPPORT
REPARTITION VOLUME (EVP) DEBARQUE PAR LES IMPORTATEURS (TOP 30) ENTRE LES COMPAGNIES MARITIMES	NUMERO DE CONTENEUR DEBARQUER, TYPE DE CONTENEUR (20OU 40), DATE D'ACCOSTAGE, NOM IMPORTATEUR, NOM COMPAGNIE MARITIME.	FICHIER PLAT (EXCEL)	1 JOUR
CLASSEMENT DES CLIENTS PAR EVP PUIS PAR REGION ET PAR SECTEUR D'ACTIVITE POUR CALCULER LA PART DE MARCHE EN POURCENTAGE.	CODE CLIENT, NOM CLIENT, DATE CREATION, VILLE, REGION, PRODUIT, NOMBRE EVP (20OU 40) IMPORTÉ.	FICHIER PLAT (EXCEL)	1 JOUR
CLASSEMENT DES CLIENTS PAR CHIFFRE D'AFFAIRE PUIS PAR REGION ET SECTEUR D'ACTIVITE.	CHIFFRE D'AFFAIRE PAR IMPORTATEUR (CLIENT), PAR PRESTATION, PAR CODE CLIENT, PAR NOM DU CLIENT.	FICHIER PLAT (EXCEL)	1 JOUR
REALISATION SUR EXCEL D'UN TABLEAU DE BORD MANTENTION SUR LES INDICATEURS DE PERFORMANCE (TEMPS AU RADE, NOMBRE DE CONTENEUR /HEURE, TEMPS AU QUAI).	DATE DE MOUILLAGE, DATE D'ACCOSTAGE, DATE DEBUT DES OPERATIONS, DATE FIN DES OPERATIONS, DATE D'APPAREIL-LAGE, TOTAL MOVE.	FICHIER PLAT (EXCEL)	1 JOUR

Tableau 9: Tableau descriptif de la procédure d'élaboration de rapports.

### 5.2.2. Description du processus métiers

Le service marketing s'appuie sur le logiciels bureautique (Excel) afin d'élaborer leurs rapports finaux. Afin de mieux comprendre leurs processus métiers, nous avons pu résumer leur procédures de travail dans un schéma qui illustre toutes ces étapes précédentes voici la figure qui montre cette architecture de prise de décision actuelle

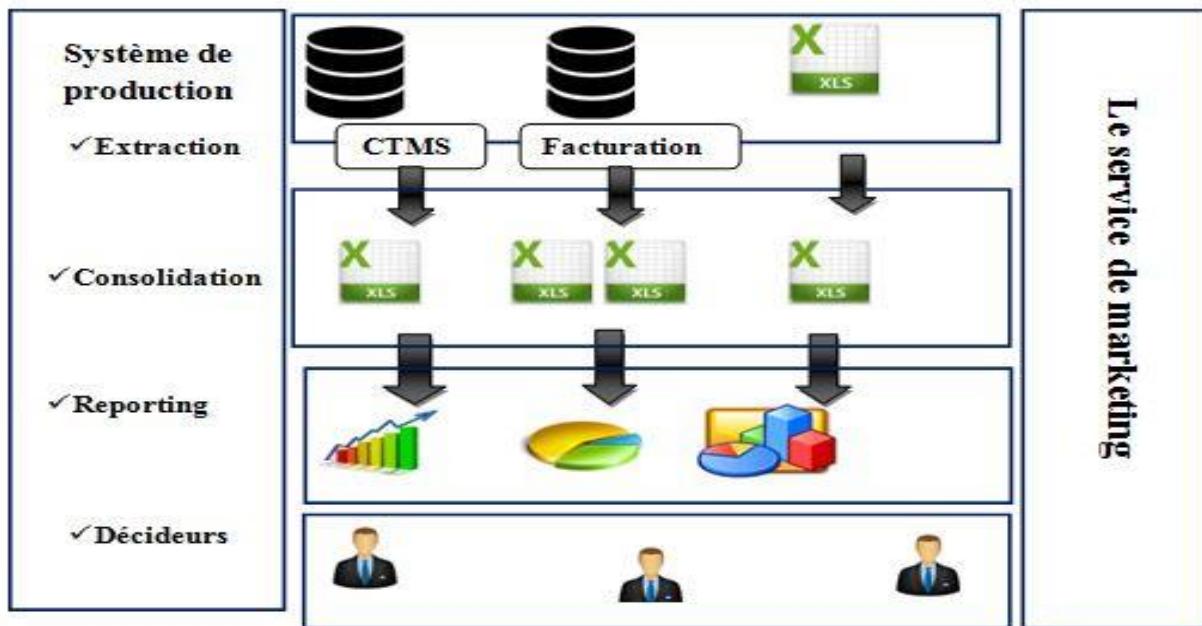


Figure 26 : Processus métiers de prise de décision

### 5.2.3. Description du processus de prise de décision

Dans le cadre du suivi de l'activité de BMT (manutention et acconage, facture client), un rapport d'activité mensuel, trimestriel et annuel est établi par le service marketing.

Les analystes de marketing identifient la source de données à utiliser depuis les systèmes opérationnels existants. Si les données sont disponibles à leurs niveaux, ils élaborent les rapports d'activités que nous avons citées déjà sous un fichier format (.xls). En cas de bug informatique c.à.d. que le matériel se plante ou tombe en panne ; ils font appeler au service informatique afin de régler le problème signalé, puis extraire les données voulues, enfin mettre celle-ci sur Excel afin de pouvoir les exploiter, et les analyser. La procédure d'établissement de rapports finaux et leur validation prend généralement 1 Jours sinon 2 jours si le chef de service est sous charge.

Une fois les rapports sont élaborés, ils seront acheminer vers le directeur de marketing qui sera en tirer la bonne décision.

Cette procédure S'illustre dans le tableau ci-dessous :

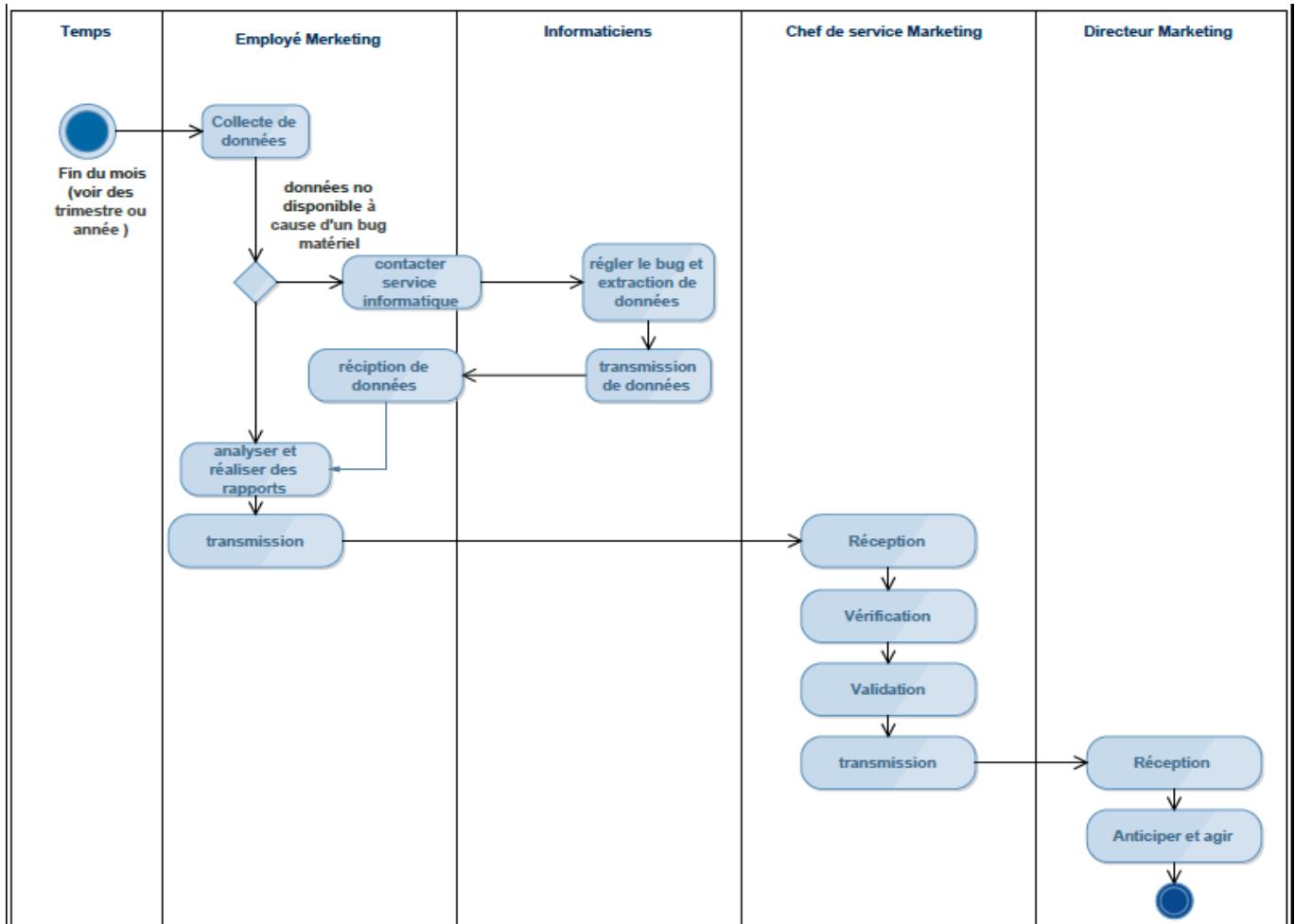


Figure 27 : Diagramme d'activité d'établissement des rapports.

### 5.3. Critique de l'existant

Après avoir fait une étude de l'existant un diagnostic de celui est nécessaire pour dégager les différentes anomalies constatées dans le fonctionnement de leur système actuel.

Les insuffisances que nous avons pu récences sont :

#### 5.3.1. Systèmes opérationnels

Le système actuel du service de marketing présente un certain nombre d'insuffisances, qui peuvent se résumer en :

- ✓ Difficultés dans l'élaboration des rapports d'activité à cause d'insuffisance de données (la saisie incorrecte de certaines données telles, les noms des importateurs ...etc.) qui passe par plusieurs service la direction des opérations puis service commerciale.
- ✓ L'impossibilité d'analyser un grand volume de données à cause des limites imposées par les outils utilisés en termes de fonctionnalités et de volume supporté. (Utilisation d'Excel).
- ✓ Mauvaise codification de données ce qui retarde la procédure de travail.

### **5.3.2. Bases de données utilisés**

La base de données de **CTMS** qui est exploité par le service marketing a été conçue par un prestataire de service. Celle-ci est une BDD centralisée installé dans le serveur oracle 11G. Elle regroupe plus de 200 tables. Après avoir consulté quelques tables qui y été mis à ma disposition nous avons pu remarquer c'est quelques anomalies :

- **Incohérence des données**

La codification appliquée par la douane sur les navires, saisie ensuite sur le manifeste puis enregistrer par Direction des opérations sur l'application CTMS engendre des incohérences de données ce qui influence sur la fiabilité de l'information. Pour cela, le service marketing ne donne pas une grande importance aux codes des informations, d'où l'absence d'uniformisation. Ce problème est considéré comme le souci majeur de tous les concepteurs et les développeurs au sein de l'entreprise.

Dans la table Vessel\_call (escale navire) par exemple, nous avons remarqué qu'un même nom de navire est présenté sous deux ou plusieurs codification différentes (deux codes «2005000021, 2005000029 » pour le même navire DENIZHAN BAYRAKTA) comme le montre la figure suivante :

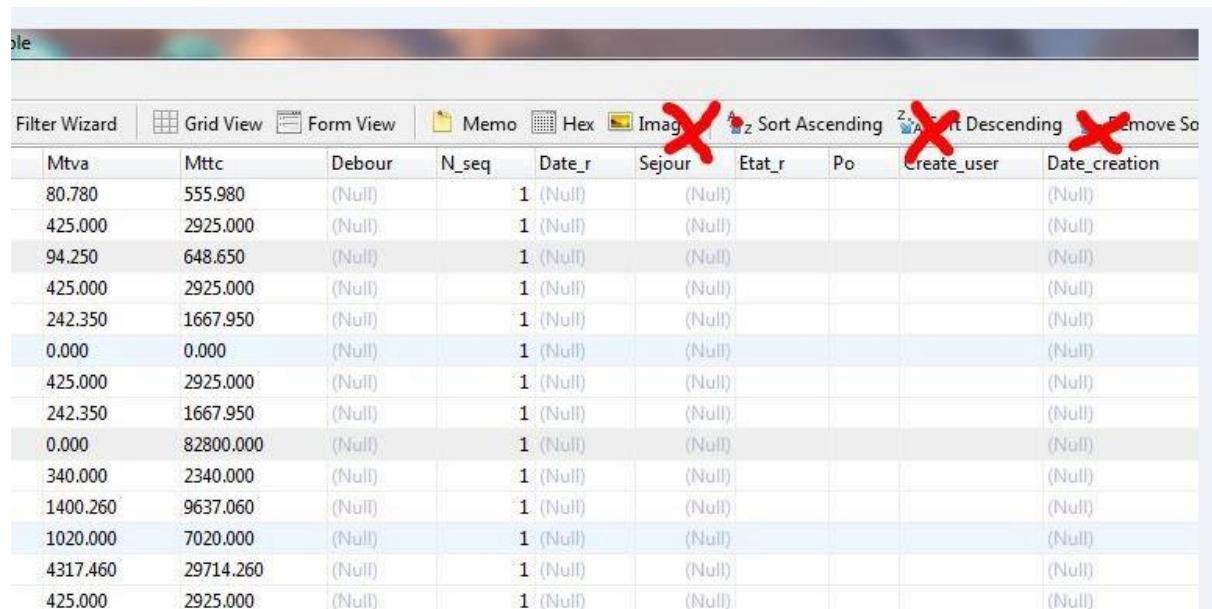
VV_CD *	VSL_NM *	IN_VOY_NBR	OUT_VOY_NBR	VV_STATUS_IND	SHPG_SVC_CD	SHPG_ROUTE_NBR	BERTH_APPL_DTTM
2006000031	NORTH AFRICA 2	NB028R	NB028R	CL	NAF		1 06/02/2006 09:30:22
2005000008	CMAORAN	DM060R	DM060R	CL	NAF		1 03/07/2005 14:09:36
2005000009	MSCHOGGAR	55547	55547	CL	MSCN		1 04/07/2005 17:00:18
2005000021	DENIZHAN BAYRAKTA	DN058R	DN058R	CL	NAF		1 31/07/2005 04:34:15
2005000022	SADAN BAYRAKTAR	SK054R	SK054R	CL	NAF1		1 31/07/2005 05:28:20
2005000029	DENIZHAN BAYRAKTA	56709	56709	CL	NAF		1 14/08/2005 16:23:53
2005000030	KARINA	50923	50923	CL	NAF1		1 15/08/2005 15:33:05
2005000005	BODRUM	380	380	CX	MAEN		1 22/06/2005 10:16:04

Figure 28 : La table VESSEL\_CALL.

- Plusieurs champs vides dans les tables

Dans la table par exemple "facture" qui contient les informations liée à la facturation du client (tels n\_imp, date\_fact), nous avons constaté que plus de 50% de champs sont vides.

La figure ci-dessous illustre cette anomalie :



The screenshot shows a database grid with 15 rows and 10 columns. The columns are labeled: Mtva, Mtcc, Debout, N\_seq, Date\_r, Séjour, Etat\_r, Po, Create\_user, and Date\_creation. Most cells contain '(Null)' or similar empty values, indicating significant data entry errors. Red X marks are placed over the sorting and filtering icons at the top of the grid.

Mtva	Mtcc	Debout	N_seq	Date_r	Séjour	Etat_r	Po	Create_user	Date_creation
80.780	555.980	(Null)		1 (Null)	(Null)				(Null)
425.000	2925.000	(Null)		1 (Null)	(Null)				(Null)
94.250	648.650	(Null)		1 (Null)	(Null)				(Null)
425.000	2925.000	(Null)		1 (Null)	(Null)				(Null)
242.350	1667.950	(Null)		1 (Null)	(Null)				(Null)
0.000	0.000	(Null)		1 (Null)	(Null)				(Null)
425.000	2925.000	(Null)		1 (Null)	(Null)				(Null)
242.350	1667.950	(Null)		1 (Null)	(Null)				(Null)
0.000	82800.000	(Null)		1 (Null)	(Null)				(Null)
340.000	2340.000	(Null)		1 (Null)	(Null)				(Null)
1400.260	9637.060	(Null)		1 (Null)	(Null)				(Null)
1020.000	7020.000	(Null)		1 (Null)	(Null)				(Null)
4317.460	29714.260	(Null)		1 (Null)	(Null)				(Null)
425.000	2925.000	(Null)		1 (Null)	(Null)				(Null)

Figure 29 : La table facture .

- Saisi incorrecte des informations

Nous avons remarqué que les informations saisi dans le CTMS par la direction des Opérations engendrent des erreurs qui jouent sur la fiabilité des données .cela implique des ambiguïtés et retarde la procédure de travail d'analyse pour service du marketing.

Pour les tables par exemple (ARP\_BAYPLAN<sup>2</sup> et CNTR<sup>3</sup>) l'information est mal saisi pour le PLOAD<sup>4</sup> pour un même numéro de conteneur (CNTR\_NBR='MSCU1613003'), comme le montre la figure suivante :

The screenshot shows two separate SQL queries run in Oracle SQL Developer:

- Top Query (CNTR):** `select pload from CNTR where CNTR_NBR='MSCU1613003';` The results show two rows: one for 'PLOAD' and one for 'ESVLC'.
- Bottom Query (ARR\_BAYPLAN):** `select pload FROM ARR_BAYPLAN WHERE CNTR_NBR='MSCU16130034';` The results show a single row for 'PLOAD'.

**Figure 30 : La table CNTR et la table ARR\_BAYPLAN.**

Dans les tables ARR\_BAYPLAN et la table VESSEL\_CALL nous avons constaté que durant la même date, pour un même navire on lui a affecté deux code différents.

<sup>2</sup> ARR\_BAYPLAN : C'est la table de chargement navire.

<sup>3</sup> CNTR : Conteneur.

<sup>4</sup>PLOAD: c'est l'attribut Port Load (chargement du port).

The screenshot shows the Oracle SQL Developer interface. The top menu bar has tabs for 'Query Builder Untitled1\*', 'Editor Untitled3\*', 'Editor Untitled4\*' (which is active), and 'Editor Untitled18\*'. A query is being run in the editor:

```
SELECT VV_CD from VESSEL_CALL where LAST MODIFY_DTTM between '03-07-2007' and '05-07-2007';
```

The results pane is titled 'Results' and contains tabs for 'Result Sets', 'Messages', 'Explain Plan', and 'Pivot & Chart'. The 'Result Sets' tab is selected, showing five tabs labeled 'Set 6', 'Set 7', 'Set 8', 'Set 9' (which is highlighted in yellow), and 'Set 10'. Below these tabs is a message: 'Drag a column header here to group by that column'. The data table has a single column labeled 'VV\_CD \*'. The values listed are:

VV_CD
2007000156
2007000158
2007000161
2007000162

Figure 31 : La table VESSEL\_CALL

The screenshot shows the Oracle SQL Developer interface. The top menu bar has tabs for 'Query Builder Untitled1\*', 'Editor Untitled3\*', 'Editor Untitled4\*', and 'Editor Untitled18\*' (which is active). A query is being run in the editor:

```
SELECT VV_CD from ARRBAYPLAN where LAST MODIFY_DTTM between '03-07-2007' and '05-07-2007';
```

The results pane is titled 'Results' and contains tabs for 'Result Sets', 'Messages', 'Explain Plan', and 'Pivot & Chart'. The 'Result Sets' tab is selected, showing five tabs labeled 'Set 6', 'Set 2', 'Set 3', 'Set 4', and 'Set 5'. Below these tabs is a message: 'Drag a column header here to group by that column'. The data table has a single column labeled 'VV\_CD \*'. The values listed are:

VV_CD
2007000164
2007000164
2007000164
2007000164
2007000164
2007000164
2007000164
2007000164

Figure 32 : La table ARRBAYPLAN.

- Redondance de données**

La redondance de donnée implique des traitements complexes d'analyse de données. Si nous prenons l'exemple des deux tables (CNTR, ARP\_BAYPLAN) nous avons remarqué qu'elles contiennent les même attributs (Last\_Modify\_Dttm, Last\_Modify\_User\_Id) en plus elles offrent les mêmes informations sur les prestations effectuées.

The figure shows two side-by-side tables from a MySQL database. The left table is titled 'ARR\_BAYPLAN' and the right table is titled 'CNTR'. Both tables have three columns: 'CNTR\_SEQ\_NBR \*', 'LAST MODIFY DTTM \*', and 'LAST MO'. The data in both tables is identical, showing a sequence of dates from 14/06/2005 at 01:00:17 to 14/06/2005 at 04:42:58, with each entry labeled 'cim1' in the 'LAST MO' column.

CNTR_SEQ_NBR *	LAST MODIFY DTTM *	LAST MO
869	15/06/2005 04:42:58	cim1
870	14/06/2005 01:00:17	cim1
871	14/06/2005 01:01:18	cim1
872	14/06/2005 01:02:18	cim1
873	14/06/2005 01:03:29	cim1
874	14/06/2005 01:04:37	cim1
875	14/06/2005 01:05:33	cim1
876	14/06/2005 01:08:39	cim1
877	14/06/2005 01:09:59	cim1
878	14/06/2005 01:11:11	cim1
879	14/06/2005 01:12:17	cim1
880	14/06/2005 01:13:27	cim1
881	14/06/2005 01:15:10	cim1
882	14/06/2005 01:16:13	cim1
866	15/06/2005 04:32:57	cim1
867	15/06/2005 04:35:11	cim1
868	15/06/2005 04:37:20	cim1
869	15/06/2005 04:42:58	cim1
870	14/06/2005 01:00:17	cim1
871	14/06/2005 01:01:18	cim1
872	14/06/2005 01:02:18	cim1
873	14/06/2005 01:03:29	cim1
908	15/06/2005 04:57:10	cim1
909	14/06/2005 03:03:09	cim1
910	14/06/2005 03:04:19	cim1
911	14/06/2005 03:05:25	cim1
912	14/06/2005 03:06:49	cim1
913	14/06/2005 03:08:06	cim1

Figure 33 : La table ARR\_BAYPLAN et la table CNTR.

La base de données facturation qui est exploitée par le service marketing été conçue et développé par le service informatique sous le SGBD MySQL 5. cette BDD contient 53 tables. Après avoir consulté la base de donne de facturation, voici quelques anomalies que nous avons recensées :

- Présence de tables vides**

Plus de 40% de tables présentées dans la base de donnée facturation sont vides ou non utilisées. la figure ci-dessous montre quelques exemples de ces tables (balance client et édition) sont vides.

The figure shows two empty tables from a MySQL database. The left table is titled 'balance' and the right table is titled 'edition1'. Both tables have multiple columns, but all rows are empty, showing '(Null)' in every cell.

Code_cl	Nom_cl	Debit	Credit	Solde	Solde_init
(Null)	(Null)	(Null)	(Null)	(Null)	(Null)

N_fact	Cntr	Type	Taille
(Null)	(Null)	(Null)	(Null)

Figure 33 : La table balance et la table edition1.

(La table balance\_cl qui devrait contenir les informations liée aux clients mais tous ces attributs sont vides).

### **5.3.3. Système décisionnel Actuel**

Une entreprise équipée d'une bonne solution décisionnelle a un avantage certain, mais comme nous l'avons constaté, le système d'information actuel du service marketing ne répond pas à toutes les demandes quotidiennes des utilisateurs, en termes d'analyse. Ils se basent sur des statistiques et des rapports établis à partir des systèmes opérationnels et des données enregistrées dans d'autres sources (fichiers Excel), pour prendre des décisions influentes et assez importantes. Cette méthode n'aboutit pas à des résultats exacts et fiables.

L'inefficacité du système décisionnel est considérée comme un sérieux souci, auquel la BMT (Bejaia Méditerranéen Terminal) doit faire face, et c'est essentiellement le point sur lequel se concentre notre étude.

## **5.4. Obstacles rencontrés lors du recueil d'information**

Malgré, la motivation et la collaboration exprimée par le personnel du service marketing et de l'IT lors de notre collecte des informations nécessaires liées à leurs besoins. Cela n'empêche pas que cette tâche ne s'est pas accomplie sans difficultés ce qui a eu pour effet de ralentir et rendre plus complexe, la progression de notre travail. Ces difficultés peuvent être résumées comme suit :

- ✓ Difficulté de planification d'un entretien dans certaines périodes à cause de la charge de l'emploi du temps.
- ✓ Certaines annulations de dernières minutes, ce qui a un peu bousculé l'ordre d'entretiens que nous avons établi.
- ✓ La rétention d'informations sous couvert de confidentialité et de sécurité de données.
- ✓ Indisponibilité de l'encadreur du service informatique durant la période de stage ce qui a trop ralenti le travail.
- ✓ Trop d'attente à cause de réunion ou personnel en déplacement à d'autre direction.
- ✓ Réponse trop expéditive en raison de surcharge de travail.
- ✓ Complexités des bases de données opérationnelles à cause du manque du schéma relationnel.

## Conclusion

Dans ce chapitre, nous avons décrit l'existant informatique au sein de la société BMT-spa afin d'avoir une vision globale sur son état informatique et applicatif. Cela nous a permis d'approfondir l'étude de notre existant au sein du service marketing. Afin de pouvoir proposer une solution qui pourra être proche aux attentes des analyses et des décideurs de ce service. Dans le chapitre suivant nous allons identifier les besoins des utilisateurs finaux de notre nouveau système.

# Chapitre 6:Analyse de besoins.

## Introduction

« Les chances de succès d'un Data Warehouse se trouvent considérablement accrues par la bonne compréhension des utilisateurs et de leurs besoins » [Kimball, 2002].

L'identification des besoins est une étape primordiale dans le succès de tout un projet de système décisionnel. Toutefois, Ralph Kimball place les besoins des utilisateurs au centre de son univers. Vu que, cette procédure nous permet de bien comprendre et de faire une capture des besoins utilisateurs. Il est donc impératif d'être à l'écoute de ces derniers afin de Cerner avec précision leurs besoins et les traduire pour les intégrer lors de la conception.

Comme tout projet informatique il est souvent très recommandé d'établir une bonne démarche de gestion de projet. Dans notre projet, nous nous sommes inspirées de la méthodologie pour les projets décisionnels préconisés par «**Ralph Kimball** » qui représente une succession d'étapes nécessaires à la conception, au développement et au déploiement d'un entrepôt de données efficace. Cette méthode nous a aussi permis de bien choisir les méthodes et outils adaptés. La figure ci-dessous illustre l'ensemble des étapes suivies pour la réalisation de notre projet :

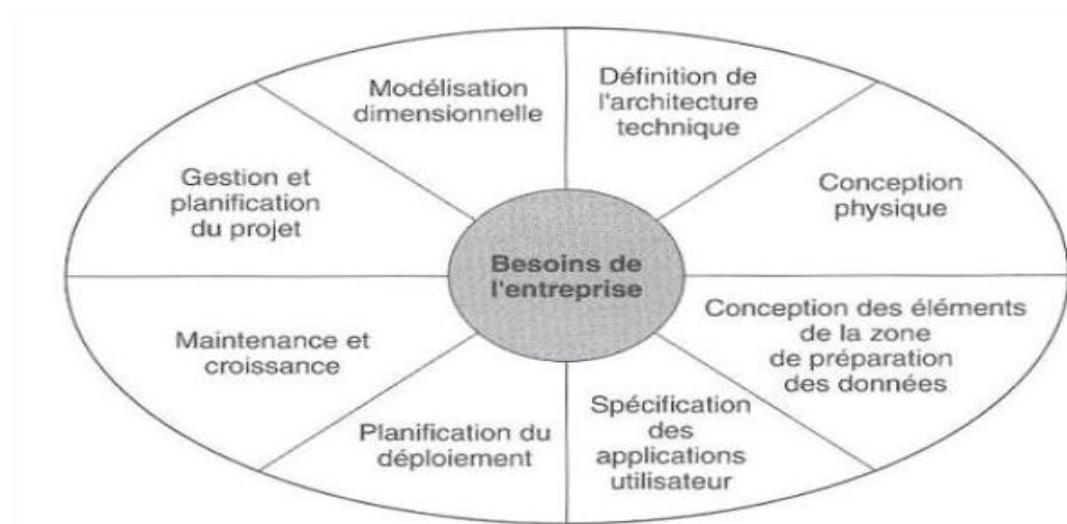


Figure 34 : Cycle de vie d'un projet décisionnel [Kimball, 2002].

### 6.1. Présentation de la démarche à suivre

Comme on peut s'y attendre dans tout un projet informatique, le cycle de vie commence souvent par la planification du projet. Celle-ci aborde la définition du projet ainsi que son

étendue [**Annexe 3**]. Elle se concentre surtout sur les besoins exprimés par les utilisateurs de l'entreprise et leurs niveaux d'exigences.

## **6.2. Présentation de la démarche d'identification des besoins**

Afin de faire une étude complète et d'obtenir le maximum d'informations, nous avons concentré en premier lieu sur le recueil des besoins exprimés par les utilisateurs, puis pour les évaluer, nous avons combiné avec la disponibilité des données sources.

Pour collecter l'ensemble d'informations auprès des utilisateurs, il existe plusieurs méthodes dont on trouve les entretiens, et les questionnaires... (Voir **annexe 4**).

Nous avons opté pour la technique d'entretien grâce à l'ensemble d'avantages qu'elle présente et qui peuvent se résumer en :

- Rapidité et simplicité d'utilisation.
- Courts délais : La préparation de l'entretien ne peut prendre que peu de temps aux personnes interrogées.. .
- La participation des intervenants permet d'augmenter les chances de recueillir l'information voulue.

Cependant, cette méthode paraît évidente et simple mais en réalité c'est une étape rigoureuse, qui nécessite la réflexion, afin de bien comprendre le travail, les objectifs et les enjeux du métier du service marketing. Pour cela, nous allons suivre la démarche ci-dessous :

### **6.2.1. Identification des personnes à interviewer**

Il faut prendre contact avec des personnes susceptibles d'apporter des informations utiles et de fournir les explications recherchées. Ces personnes occupant des rôles différents dans l'entreprise et ayant des responsabilités, pouvons-nous aidés à avoir différents points de vue et recensés les informations nécessaires.

Les utilisateurs de ce nouveau système peuvent être regroupés en (04) catégories qui sont :

- **Décideur stratégique** : représenté par le DG de l'entreprise.
- **Décideurs tactiques** : représenté par directeur de marketing et son chef de service.
- **Administrateur Système** : c'est l'administrateur de base de données.
- **Analystes** : représenté par les Marketeurs et les agents commerciaux.

### **6.2.2. Elaboration des entretiens**

Dans cette étape, nous avons commencé par la définition du contexte général, la date des entretiens. Cette planification ne se limite pas à prendre rendez-vous avec les personnes à interviewer mais aussi à organiser les entretiens afin d'optimiser la collecte d'informations.

Puis, nous avons préparé la liste des questions à poser (**Annexe 4**). Mais tout d'abord nous avons commencé par une petite description de notre projet, après nous avons passé au cœur du sujet en faisant parler l'interviewé de ce qu'il fait, pourquoi il le fait, et la manière dont il prend ses décisions, afin de bien comprendre sa procédure de travail.  
Enfin, après chaque entretien, une identification des points obscurs est nécessaire afin de les clarifier dans les prochains entretiens.

### **6.2.3. Documentation**

Bien que les entretiens représentent la méthode d'identification des besoins la plus utilisée, les informations fournies par les utilisateurs sont insuffisantes pour la construction d'un entrepôt de données. De ce fait, La documentation a été utilisé afin d'analyser les différents documents disponibles et qui peuvent nous donner des informations utiles pour notre projet.

### **6.2.4. Etude des sources de données**

Cette phase consiste à étudier les systèmes opérationnels existants utilisés par le service Marketing, afin de confirmer les informations acquises lors de notre étude de l'existant et de détecter les besoins exprimés par les managers relatifs aux données existantes.

Les bases de données (CTMS, facturation) et les fichiers plats (.xls) utilisée par le service marketing permettent de fournir presque toutes les informations nécessaires pour l'alimentation de l'ED. Lors de notre étude détaillée des bases de données sources ; nous avons pu identifier les tables utilisés : Les tables transactionnelles (ARR\_BAYPLAN, CNTR, facturation ...etc.)

## **6.3. Récapitulatif sur les besoins recueillis**

Cette phase consiste à rédiger un recueil officiel qui regroupe des conclusions issues de fin des entretiens. Afin de pouvoir valider ces conclusions par les utilisateurs.

Ce recueil contient des comptes rendus d'entretiens rassemblés et synthétisés, qui permettront d'avoir une vue globale du projet et du point de vue des décideurs.

### 6.3.1. Besoin fonctionnel

Les besoins exprimés par les utilisateurs portent sur les coûts, les délais, le trafic des conteneurs. Le tableau ci-dessous représente une synthèse sur les besoins fonctionnels détectés et exprimés lors de la collecte d'informations et classés par sujet d'analyse. Ces analyses sont effectuées par mois ou par semestre ou année ainsi, elles peuvent être demandées à n'importe quelle période de l'année

SUJET D'ANALYSE	BESOIN ANALYTIQUE
<b>SUIVI DES CLIENTS</b>	<ul style="list-style-type: none"> <li>● PART DE MARCHE DU CHIFFRE D'AFFAIRE REALISE PAR CHAQUE CLIENT, SA REGION, SON SECTEUR D'ACTIVITE, TYPE DE PRESTATION.</li> <li>● PART DE MARCHE DE CHAQUE CLIENT PAR NOMBRE D'EVP, PAR TYPE DE PRESTATION, SA REGION, SON SECTEUR D'ACTIVITE.</li> </ul>
<b>SUIVIS DES COMPAGNIES MARITIMES.</b>	<ul style="list-style-type: none"> <li>● LA PART DE MARCHE DE CHAQUE COMPAGNIE MARITIME</li> <li>● LA PART DE MARCHE DE CHAQUE COMPAGNIE MARITIME PAR TYPE DE PRESTATION : (EMBARQUEMENT DU VIDE ET DU PLEIN /DEBARQUEMENT DU CONTENEUR PLEIN OU REFEERS (CONTENEURS FRIGORIFIQUE).</li> </ul>
<b>SUIVIS DES NAVIRES</b>	<ul style="list-style-type: none"> <li>● NOMBRE DE CONTENEUR TRAITE PAR HEURE (VESSEL RATE).</li> <li>● TOTAL DE MANIPULATION SUR LES NAVIRE (TOTAL MOVE).</li> <li>● TEMPS D'ATTENTE TOTAL. (TEMPS EN RADE ET TEMPS EN QUAI).</li> </ul>

Tableau 10:Tableau sur les récapitulatif des besoins des utilisateurs.

### 6.3.2. Besoin technique

Le projet d'élaboration d'un système décisionnel constitue un réel besoin pour le service marketing. De ce fait, le service informatique avait tracé certains critères que le futur système doit respecter. Ces critères sont les suivants :

- ✓ Héberger l'entrepôt de donnée sur un système de gestion de base de données : oracle 11g.
- ✓ Avoir l'accès à l'information en temps réel.
- ✓ Utiliser un outil open source de préparation de donnée.
- ✓ Utiliser un outil open source de présentation de donnée, pour faire le reporting.

## 6.4. Vue générale de la solution proposée

Après avoir étudié l'existant et les besoins au niveau du service marketing de la BMT-Spa, nous avons opté pour une solution que nous jugeons la plus adéquate. Cette solution repose sur la mise en place d'un système décisionnel permettant d'atteindre l'objectif principal, qui consiste à assurer un meilleur suivi des clients ainsi la classification de compagnie maritimes.

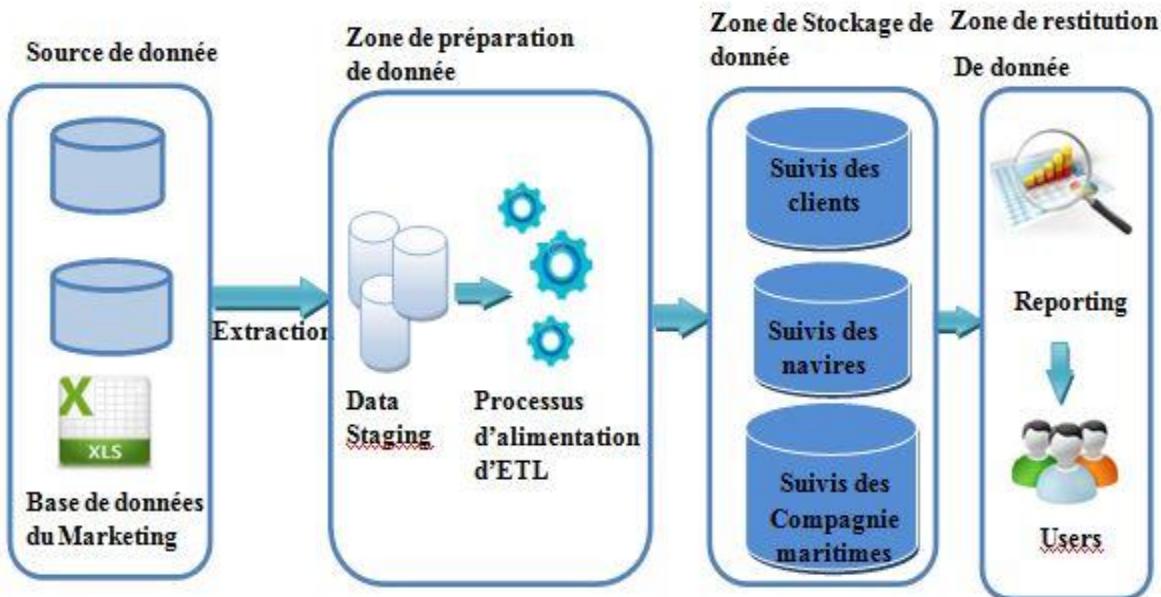


Figure 35 : Vue de la solution proposée

## Conclusion

Tout au long de ce travail, nous avons pu constater que la collecte des besoins est une étape cruciale, qui définit la suite des événements par rapport à notre projet. En effet les résultats de cette étape définiront la manière de construire le DW et ses données. Le chapitre suivant porte sur la conception d'un entrepôt de données à savoir : la conception du modèle dimensionnel, la conception de la zone d'entreposage et d'alimentation, et la conception des cubes dimensionnels.

## Synthèse de la deuxième partie

Dans cette deuxième partie, qualifiée préliminaire, nous avons commencé par une présentation de l'organisme d'accueil de la BMT-Spa, avec ses différentes directions en particulier la direction du Marketing où notre projet s'est déroulé.

Ensuite, nous avons analysé l'existant informatique au sein de la BMT-Spa, puis étudier leurs différentes systèmes transactionnels ainsi l'existant décisionnel ; afin de mieux comprendre leur processus métiers et d'en conclure les dysfonctionnements qui seraient à l'origine des difficultés. Enfin nous avons mis un diagnostic de cette analyse, ce qui permet ensuite de compléter le chapitre suivant qui est l'étude des besoins.

En outre, notre engagement durant le déroulement de ce projet est d'être à l'écoute des utilisateurs afin de bien cerner leurs besoins. L'étude des besoins est une étape très cruciale dans la réussite d'un projet décisionnel, elle permet de comparer les attentes exprimées aux données existantes.

Dans la partie suivante, nous allons présenter la conception de notre solution pour répondre à la problématique posée, tout en se basant sur les résultats précédents

---

## **Partie 03: Conception.**

---

Chapitre 07 : Conception de la zone d'entreposage.

Chapitre 8 : Conception de la zone d'alimentation de  
données.

# Chapitre 07: Conception de la zone d'entreposage.

## Introduction

Après avoir collecté les besoins utilisateur de ce système et analyse les données existantes, nous voilà prêts à lancer la conception logique et physique qui est une pierre angulaire de notre un entrepôt de données .

Sachant que la conception et l'intégration de la nouvelle solution doit être compatible avec l'existant informatique de l'entreprise BMT-spa.

Dans cette partie, nous avons repartis notre un entrepôt de données en deux phases qui sont :

- La conception de la zone d'entreposage : consiste à définir le schéma conceptuel de notre entrepôt de données (tables de fait, tables de dimensions...).
- la conception de la zone d'alimentation : consiste à définir le déroulement du processus ETL afin d'intégrer les données dans le ED.

### 7.1. Conception de la zone d'entreposage

Le modèle conceptuel de DW doit être simplifié au maximum pour comprendre ce que l'entrepôt mémorise. On parle de modèle multidimensionnel. Notre objectif est de proposer une méthode de modélisation faits/dimensions répondant aux mieux aux besoins des utilisateurs pour leur faciliter l'exploitation des données stockées dans un entrepôt de données.

Tout le long de notre conception nous suivrons la démarche de « Ralf Kimball » qui est la plus adaptées à notre cas et qui se résume en Quatre étapes :

#### 7.1.1. Sélection du processus d'activités à modéliser

Un processus est une activité opérationnelle effectuée dans une organisation qui est généralement pris en charge par une source de gestion et collecte de données. Pour notre cas Le processus clé est : le suivi des clients.

### **7.1.2. Choix de grain du processus d'activités**

La définition de la granularité de la table de fait, revient à spécifier exactement ce que représente une ligne de cette table. Et ce, en déclarant l'expression la plus fine qui contiendra le besoin de l'utilisateur de manière plus détaillée.

### **7.1.3. Choisir les dimensions participantes à ce processus**

Les dimensions fournissent les questions suivantes : "qui fait quoi, quand, pourquoi, et comment" aux alentours du contexte d'un événement particulier d'un processus donné. Les tables de dimensions contiennent les attributs descriptifs pour définir ce processus. Après le choix de la granularité, l'identification des dimensions est assez simple. Selon [Kibball, 2002]: « Le temps est la seule dimension qui figure systématiquement dans tout un entrepôt de données , et plus souvent la première dimension dans le classement sous-jacent de la base de données ». C'est pour cela que la table de dimension est appelée quelquefois l' "âme" d'un entrepôt de données .

### **7.1.4. Identifier les faits numériques**

Les faits sont les mesures qui résultent d'un processus d'un événement particulier. Les faits sont généralement des quantités numériques additives telles que le chiffre d'affaire qui répondent à la question "qu'est-ce qu'on mesure ?". Les décideurs sont vivement intéressés par l'analyse de ces mesures afin de déterminer la performance du processus concerné.

## **7.2. Matrice de bus de données**

Pour la mise en œuvre d'un ED, il faut créer une matrice de bus dimensionnel qui est une ressource importante dans la coordination de données.

L'élaboration de cette matrice consiste en premier lieu à recenser l'ensemble des magasins de données qui constituent une collection de faits numériques, puis ensuite identifier les dimensions impliquées dans ces magasins.

<b>Faits Dimensions</b>	<b>Suivis des Clients</b>	<b>Suivis des Compagnies Maritimes</b>	<b>Suivis des navires</b>
Temps	✗	✗	✗
Client	✗		
Activité	✗		
Région	✗		
Prestation	✗		
Compagnies		✗	
Navires			✗
PortDesProv		✗	
Conteneur		✗	✗

Tableau 11: Matrice de l'architecture en bus dimensionnel.

### 7.3. Modélisation dimensionnelle des activités

Après avoir défini les magasins de données que nous allons concevoir et les dimensions de la matrice du bus décisionnel, nous passerons maintenant à la construction du modèle dimensionnel.

#### 7.3.1. Activité de suivi des clients

L'activité "suivis des clients" est le premier processus à modéliser dans la conception de notre DataMarts. Cette activité permet de suivre le comportement du client et sa contribution dans la part de marché du port afin de mieux le fidéliser.

- **Déclaration du grain**

Le grain de cette activité se trouve dans le suivi du nombre d'EVP (Equivalent Vingt Pieds) de conteneurs importé ainsi que le chiffre d'affaire réalisé.

Ce suivi s'effectue selon le type de prestations (débarquement, entreposage...etc.), secteur d'activités (agroalimentaire, électronique...etc.), par région (centre, Soummam...etc.), et ce par période (mois, semestre et annuel).

- **Choix des dimensions**

Les dimensions liée à la table “ Fait\_Suivi\_Clients ”

- Dimension DIM\_Temps

<b>DESIGNATION</b>	<b>DETAILS</b>
<b>ID_TEMPS</b>	IDENTIFICATEUR DE LA DIMENSION TEMPS.
<b>DATE_DU_JOUR</b>	FORMAT COMPLET DE LA DATE.
<b>ANNEECALENDRAIRE</b>	C'EST L'ANNEE DU CALENDRIER
<b>SEMESTRE</b>	NUMERO DU SEMESTRE DANS L'ANNEE.
<b>MOIS_NUM</b>	NUMERO DU MOIS DANS L'ANNEE.
<b>LIBELLE_MOIS</b>	LA DESIGNATION DU MOIS
<b>JOUR</b>	C'EST LE NUMERO DE JOUR DU MOIS
<b>LIBELLE_JOUR</b>	LE NOM DE JOUR DE SEMAINE

- Dimension DIM\_Client

<b>DESIGNATION</b>	<b>DETAILS</b>
<b>ID_DIM_CLT</b>	IDENTIFICATEUR ARTIFICIEL DE LA DIMENSION CLIENT.
<b>NOMCOMPLET_CLT</b>	NOM COMPLET DU CLIENT.
<b>ADRESSE_CLT</b>	ADRESSE DU CLIENT
<b>NUM_TEL</b>	NUMERO DU TELEPHONE.
<b>EMAIL_CLT</b>	EMAIL DU CLIENT.

- Dimension DIM\_Prestation

DESIGNATION	DETAILS
ID_DIM_PRST	IDENTIFICATEUR ARTIFICIEL DE LA DIMENSION PRESTATION.
CODE_PRST	CODE DE LA PRESTATION
DESIGN_PRST	DESIGNATION DE LA PRESTATION

- Dimension DIM\_Secteur\_Activité

DESIGNATION	DETAILS
ID_DIM_ACTVT	IDENTIFICATEUR ARTIFICIEL DE LA DIMENSION ACTIVITE.
DESIGN_ACTVT	DESIGNATION DE L'ACTIVITE.

- Dimension DIM\_Region

DESIGNATION	DETAILS
ID_DIM_RGN	IDENTIFICATEUR ARTIFICIEL DE LA DIMENSION REGION.
LIBELLE_RGN	LIBELLE DE LA REGION.
LIBELLE_VILLE	LIBELLE DE LA VILLE

- **Choix du fait (les mesurables)**

Les mesurables de l'activité de suivis des clients :

- ✓ **Chiffre d'affaire** : c'est le chiffre d'affaire réalisé par le client.
- ✓ **QTE d'EVP** : quantité d'EVP importé par un client.

- Schéma en étoile correspondant

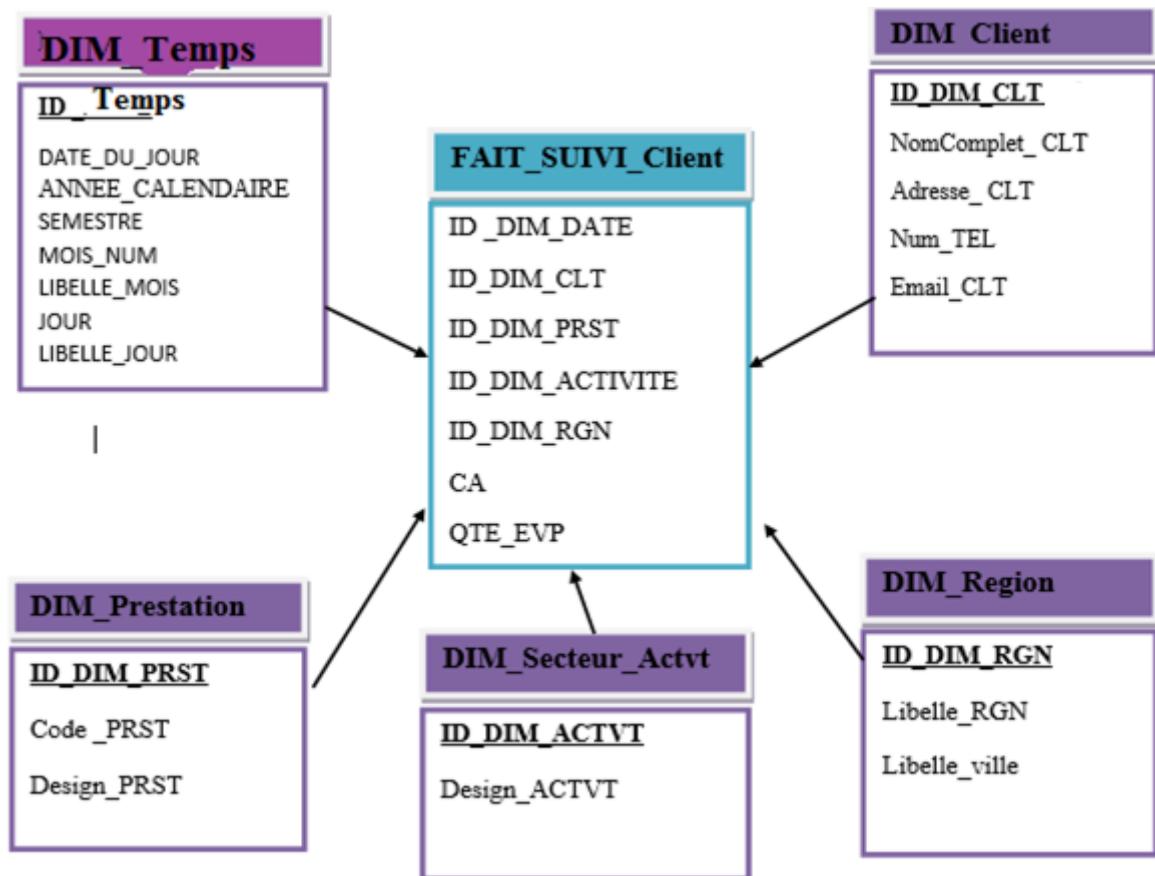


Figure 36 : Schéma en étoile de l'activité " fait suivi des clients".

### 7.3.2. Activité de suivi de compagnies maritimes

C'est l'activité de suivi de chaque compagnie maritime dans le transport maritimes et leurs contributions dans l'efficacité du marché du port de Bejaia. Et le rôle de chaque une d'entre elle de répondre aux attentes de ses clients en assurant une fiabilité de transport de marchandise.

- Déclaration du grain

Le grain de cette activité se trouve dans le suivi du nombre d'EVP (Equivalent Vingt Pied) par chaque compagnie maritime. Ce suivi s'effectue selon nom de la compagnie maritime, port de provenance, port de destination, par conteneur, et ce par période (mois, semestre et année).

- Choix de dimension

- Dimension DIM\_Temps

DESIGNATION	DETAILS
<b>ID_TEMPS</b>	IDENTIFICATEUR DE LA DIMENSION TEMPS.
<b>DATE_DU_JOUR</b>	FORMAT COMPLET DE LA DATE.
<b>ANNEECALENDRAIRE</b>	C'EST L'ANNEE DU CALENDRIER
<b>SEMESTRE</b>	NUMERO DU SEMESTRE DANS L'ANNEE.
<b>MOIS_NUM</b>	NUMERO DU MOIS DANS L'ANNEE.
<b>LIBELLE_MOIS</b>	LA DESIGNATION DU MOIS
<b>JOUR</b>	C'EST LE NUMERO DE JOUR DU MOIS
<b>LIBELLE_JOUR</b>	LE NOM DE JOUR DE SEMAINE

- Dimension DIM\_CMPGN\_MRTM

DESIGNATION	DETAILS
<b>ID_DIM_CMPAGN_MRTM</b>	IDENTIFICATEUR ARTIFICIEL DE LA DIMENSION COMPAGNIE MARITIME.
<b>NOM_CMPAGN_MRTM</b>	NOM DE LA COMPAGNIE MARITIME

- Dimension DIM\_CNTR

DESIGNATION	DETAILS
<b>ID_DIM_CNTR</b>	IDENTIFICATEUR ARTIFICIEL DE LA DIMENSION CONTENEUR.
<b>CODE_CNTR</b>	CODE DU CONTENEUR.
<b>TYPE_CNTR</b>	TYPE DU CONTENEUR SOIT 20 PIEDS OU 40 PIEDS.
<b>CAT_CNTR</b>	CATEGORIE DU CONTENEUR (FRIGORIFIQUE, DANGEREUX...ETC.).

- Dimension DIM\_Port

DESIGNATION	DETAILS
ID_DIM_PRT	IDENTIFICATEUR ARTIFICIEL DE LA DIMENSION PORT
PRT_NOM	NOM DU PORT DESTINATION/PROVENANCE.

- **Choix du fait (les mesurable)**

QTE\_EVP : Nombre d'EVP importé par une compagnie maritime.

- **Schéma en étoile correspondant**

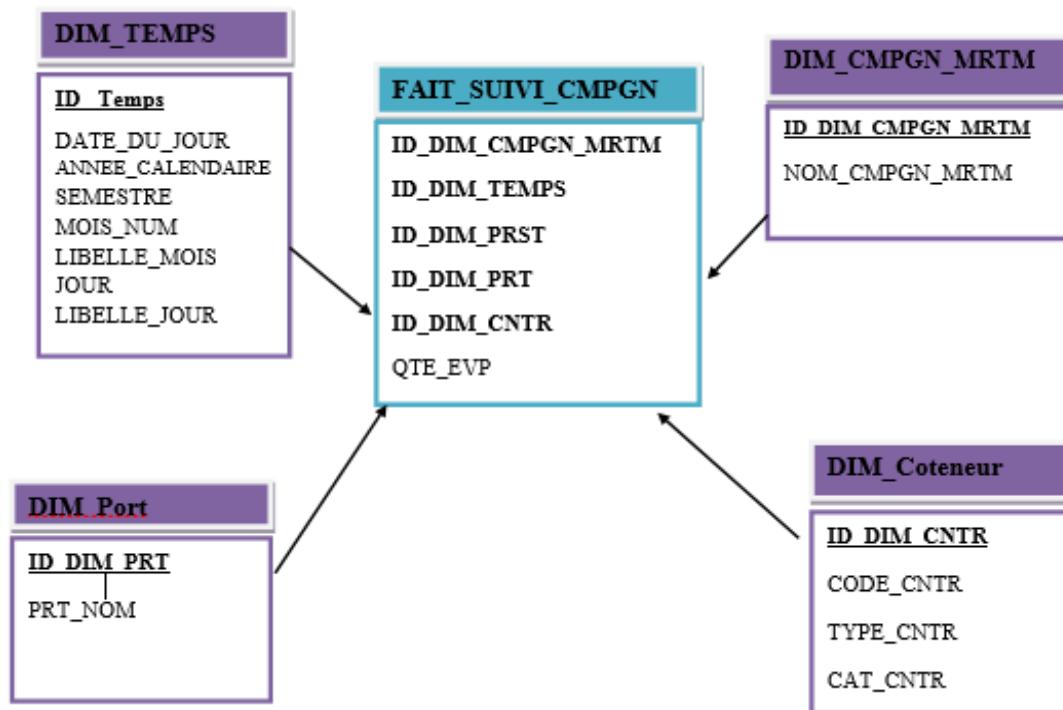


Figure 37 : Schéma en étoile de l'activité "suivi des compagnies maritimes".

### 7.3.3. Activité de suivi de navires

Cette activité permet de gérer les mouvements (Mouillage, accostage, appareillage...etc) des navires, en effectuant le suivi en temps réel des escales.

Cette activité met l'accent sur le suivi de leurs mouvements depuis leurs arrivées en rade jusqu'à leurs accostages et débarquement à quai. Ainsi elle suit la productivité des navires, c.à.d. le temps de traitement réalisée sur ces derniers.

- **Déclaration du grain**

Le grain de cette activité se situe dans le suivis des escales des navires, depuis leur placement en rade jusqu'à leur affectation à quai afin de calculer le temps d'attente totale.

- **Choix de dimension**

- Dimension DIM\_Temps

<b>DESIGNATION</b>	<b>DETAILS</b>
<b>ID_TEMPS</b>	IDENTIFICATEUR DE LA DIMENSION TEMPS.
<b>DATE_DU_JOUR</b>	FORMAT COMPLET DE LA DATE.
<b>ANNEECALENDRAIRE</b>	C'EST L'ANNEE DU CALENDRIER
<b>SEMESTRE</b>	NUMERO DU SEMESTRE DANS L'ANNEE.
<b>MOIS_NUM</b>	NUMERO DU MOIS DANS L'ANNEE.
<b>LIBELLE_MOIS</b>	LA DESIGNATION DU MOIS
<b>JOUR</b>	C'EST LE NUMERO DE JOUR DU MOIS
<b>LIBELLE_JOUR</b>	LE NOM DE JOUR DE SEMAINE

- Dimension navire

<b>DESIGNATION</b>	<b>DETAILS</b>
<b>ID_DIM_NVR</b>	IDENTIFICATEUR ARTIFICIEL DE LA DIMENSION TEMPS.
<b>NOM_NVR</b>	NOM DU NAVIRE.
<b>NOM_CMPGN_MRTM</b>	NOM DE LA COMPAGNIE MARITIME.

- Dimension conteneur

DESIGNATION	DETAILS
<b>ID_DIM_CNTR</b>	IDENTIFICATEUR ARTIFICIEL DE LA DIMENSION CONTENEUR.
<b>CODE_CNTR</b>	CODE DU CONTENEUR.
<b>TYPE_CNTR</b>	TYPE DU CONTENEUR SOIT 20 PIEDS OU 40 PIEDS.
<b>CAT_CNTR</b>	CATEGORIE DU CONTENEUR (FRIGORIFIQUE, DANGEREUX...ETC.).

- Choix du fait (les mesurable)

Temps d'attente totale : c'est le temps d'attente en rade et à quai.

Total de manipulation sur les navires (total move).

Mesure non additive : nombre de conteneur par heure.

- Schéma en étoile correspondant

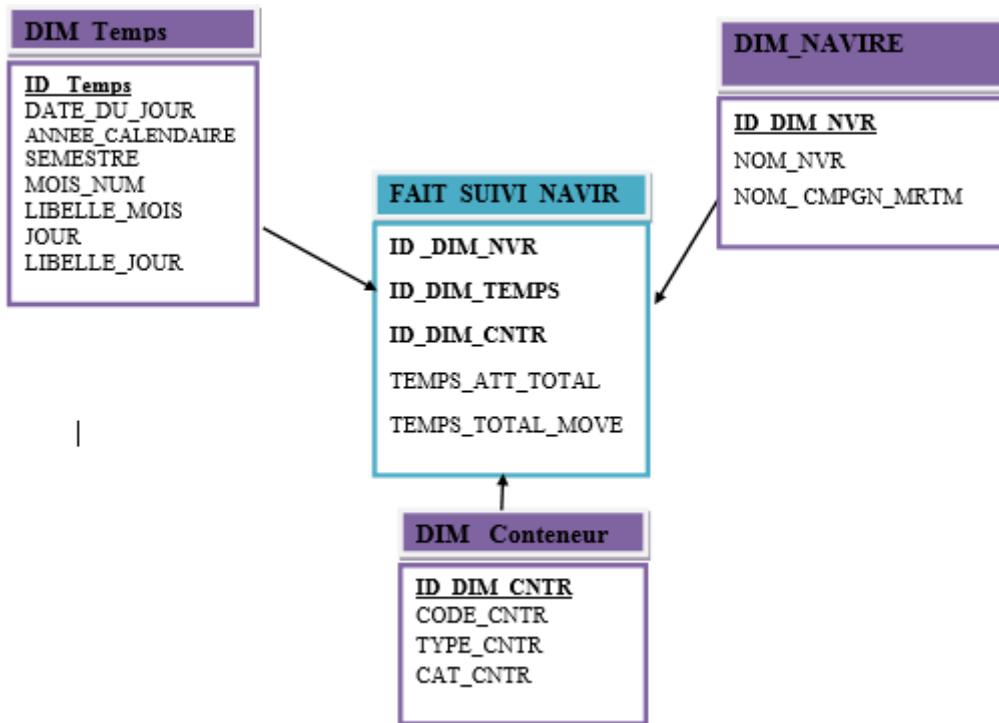


Figure 38 : Schéma en étoile "fait suivis des navires".

## Conclusion

La conception de la zone d'entreposage de données est une étape essentielle du projet décisionnelle. Elle se fait grâce à la modélisation dimensionnelle, qui offre aux utilisateurs des modèles intuitifs, tout en définissant les axes d'analyse et les indicateurs mesurés. Cette étape se base sur l'étude préliminaire, afin de satisfaire les besoins en analysant les besoins exprimés par les utilisateurs. Une fois la conception logique d'un entrepôt de données est établie, nous passons dans le chapitre suivant à la conception de la zone d'alimentation de l'DW.

# **Chapitre 8:Conception de la zone d'alimentation de données.**

## **Introduction**

L'alimentation de données est la procédure qui permet de transférer des données du système opérationnel vers un entrepôt de données en les adaptant. La conception de cette opération est une tâche assez complexe (elle constitue 70% d'un projet décisionnel en moyenne). Il est nécessaire de déterminer quelles données seront chargées, quelles transformations et vérifications seront nécessaires, la périodicité et le moment auxquels les transferts auront lieu.

### **8.1. Etude de source de donnée**

Afin de déceler les tables de données porteuses d'information, nous avons examiné les bases de données opérationnelles utilisées par le service de marketing.

Comme source de données nous avons :

- 1) La base de données CTMS centralisée installé dans le serveur oracle 11G qui est constitué de plus de 200 tables.
- 2) La base de données facturation qui est conçus sur un SGBD MySQL 5 qui contient 54 tables.
- 3) Un fichier Excel (.xls) de suivis des escales.

Cette phase nécessite beaucoup de réflexion et de concentration. Pour la mener à bien, nous avons listé en premier lieu les données nécessaires, ainsi que leurs emplacements. Ensuite, nous avons préparé les transformations nécessaires à appliquer (suppression de champs vide) et les liens logiques entre les données sources et cibles. Voici les tables nécessaires pour le data warehouse :

<b>NOM DE TABLES</b>	<b>BDD SOURCE</b>	<b>DESCRIPTION</b>
FACTURATION	BBD MySQL	IL S'AGIT DE LA TABLE LA PLUS VOLUMINEUSE, ELLE CONTIENT L'ENSEMBLE DES INFORMATIONS SUR LA FACTURE CLIENTS.  CHAQUE LIGNE DE LA TABLE COMPREND UN ENSEMBLE DE CHAMPS SIGNIFICATIFS (N_FACT) QUI EST LE NUMERO DE FACTURATION, CODE_CL QUI EST LE CODE D'UN CLIENT....ETC.
IMPORTATEUR	BBD MySQL	CETTE TABLE REGROUPE TOUS LES CLIENTS DU PORT.  LES IMPORTATEURS SONT IDENTIFIES PAR UN ENSEMBLE DE CHAMPS, SELON UN NUMERO UNIQUE (N_IMP), UN NOM (NOM_IMP), PAR UN EMAIL WEB...ETC.
LIGN_FACT	BBD MySQL	CETTE TABLE REPRESENTE LES DIFFERENTES LIGNES DE FACTURATION, ELLE CONTIENT LES CHAMPS LIEE A LA FACTURE CLIENT ; CEPENDANT ELLE CONTIENT L'IDENTIFIANT DE LA LIGNE FACTURE (ID_LIGN_FACT),LA QUANTITE DE FACTURATION (QTE_FACT)...ETC.
SERVICE	BBD MySQL	LA TABLE SERVICE PRESENTE LES DIFFERENTES PRESTATIONS REALISEES. ELLE CONTIENT LES CHAMPS SUIVANT : DESIGNATION DU SERVICE (DESIGN_S), ETAT DU SERVICE SI ELLE REGLE OU NON (ETAT)...ETC.
CENTRE	BBD MySQL	IL S'AGIT DE LA TABLE QUI DESIGNE LES TYPES DE PRESTATION AINSI QUE LEUR CODE SPECIFIES. ELLE CONTIENT LES CHAMPS SUIVANTS : DESIGNATION DE PRESTATION (DESIG_C), CODE DE PRESTATION (N_COMPTE_C)...ETC.
ARR_BAYPLAN	BBD ORACLE	CETTE TABLE REPRESENTE LA TABLE DE CHARGEMENT NAVIRE. ELLE CONTIENT LES CHAMPS SUIVANT : LE NOM DE LA COMPAGNIE MARITIME (OPER_CD), LE PORT DE CHARGEMENT(PLOAD), LE PORT DE DESTINATION (PDEST), CATEGORIE DE CONTENEUR (CAT_CD), CODE DU CONTENEUR (CNTR_SEQ_NBR), ...ETC.

VESSEL CALL	BDD ORACLE	CETTE TABLE REPRESENTE LES INFORMATION LIEE A L'ESCALE DU NAVIRE. ELLE CONTIENT LES CHAMPS SUIVANT : LE NOM DU NAVIRE (VSL_NM) ; DATE D'ESCALE (VSL_BERTH_DTTM)...ETC.
FICHIER INFO ESCALE	FICHIER EXCEL	CE FICHIER CONTIENT TOUTES LES INFORMATIONS NECESSAIRES LIEES AUX NAVIRES : NOM DU NAVIRE (NAVIRE P22/23), NOM DE LA COMPAGNIE MARITIME (SL), DATE DE DEBUT DES OPERATIONS (DEBUT OPS), DATE FIN DES OPERATIONS (FIN OPS)....ETC.

Figure 39 : Tableau descriptif des sources de données.

## 8.2. Processus alimentation de l'entrepôt de données

Après avoir conçu le modèle de données, comment alimenter l'ED ? Cette étape n'est pas un simple programme d'extraction, transformation et de chargement. Il s'agit plutôt de l'activité la plus importante du projet en termes de charge de travail et de complexité des tâches elle représente 70% de réalisation de projet, car elle permet de passer de la structure de données source à la structure dimensionnelle de l'ED. Ce processus, s'effectue en plusieurs étapes :

### 8.2.1. Extraction des données

Après avoir déterminé les données à charger, la première étape du processus ETL peut commencer. Cette opération peut commencer une fois le plan de préparation de données est établi et les sources de données identifiées précisément.

Dans un premier temps, toutes les données source du système, identifiées comme étant pertinentes, sont extraites et injectées dans une zone de préparation de données temporaires (staging area) ; qui est une zone transparente pour les utilisateurs, et représente le chantier de l'entrepôt de données. [KIMBALL, 2002]. Cette technique d'extraction correspond à l'étape de mise en route d'un entrepôt de données.

### 8.2.2. Transformation des données

Avant de commencer la transformation des données, il était indispensable d'effectuer des vérifications sur la qualité de ces dernières ; ainsi que la cohérence des données existantes dans la source : Vérification des clés primaires des tables (unique et non nulles), des clés étrangères dans les tables référencées, ...etc.

Une fois les données extraites, elles subissent une série de transformations destinées à les modifier en informations nettoyées, purifiées, consolidées et présentables aux utilisateurs. Pour cela, nous avons alors réalisé les différents types de transformations suivantes :

- **Suppression de tables vides**

En premier lieu, la transformation que nous avons faite était la suppression des tables inutiles vus que les bases de données de la BMT sont composée d'un nombre important de tables vides ou non utilisées (plus 40% des tables présentes dans la BDD facturation sont vides ou non utilisées), issues généralement de la duplication de certaines tables. De plus, la base de donnée oracle contient plus **200** tables hors pour notre entrepôt nous n'avons besoin que de **2** tables, et pour la bdd MySQL qui contient 56 tables hormis que nous avons besoin que 5 table, donc toutes les tables inutiles ont été supprimées.

The screenshot shows a Microsoft Access application window. The title bar reads "age @facturation (facturation) - Table". Below the title bar, there are three tabs: "age @facturation (facturation)" (selected), "avoir @facturation (fact...)" (disabled), and "costemer @facturation (...)" (disabled). The menu bar includes "File", "Edit", "View", "Window", and "Help". Below the menu is a toolbar with icons for "Import Wizard", "Export Wizard", "Filter Wizard", "Grid View", "Form View", "Memo", "Hex", and "Image". The main area displays a table with the following columns: Num\_ag, Code\_cl, Nom\_cl, Mttc, Ag\_1, Ag\_2, Ag\_3, Ag\_4, and Ag\_5. A single row is present, showing all fields as "(Null)".

Num_ag	Code_cl	Nom_cl	Mttc	Ag_1	Ag_2	Ag_3	Ag_4	Ag_5
(Null)	(Null)	(Null)	(Null)	(Null)	(Null)	(Null)	(Null)	(Null)

Figure 40 : Exemple de table vide.

- **Traitement des champs inutiles**

Nous avons supprimé plusieurs champs inutiles dans les tables sources dont nous n'avons pas besoin. Par exemple dans la table facture, nous disposons de **19** colonnes, mais pour notre cas nous n'avons besoin que de **8** colonnes : le numéro importateur(N\_imp)...etc.

Mtva	Mttc	Debour	N_seq	Date_r	Sejour	Etat_r	Po	Create_user	Date_creation
80.780	555.980	(Null)	1	(Null)	(Null)			(Null)	
425.000	2925.000	(Null)	1	(Null)	(Null)			(Null)	
94.250	648.650	(Null)	1	(Null)	(Null)			(Null)	
425.000	2925.000	(Null)	1	(Null)	(Null)			(Null)	
242.350	1667.950	(Null)	1	(Null)	(Null)			(Null)	
0.000	0.000	(Null)	1	(Null)	(Null)			(Null)	
425.000	2925.000	(Null)	1	(Null)	(Null)			(Null)	
242.350	1667.950	(Null)	1	(Null)	(Null)			(Null)	
0.000	82800.000	(Null)	1	(Null)	(Null)			(Null)	
340.000	2340.000	(Null)	1	(Null)	(Null)			(Null)	
1400.260	9637.060	(Null)	1	(Null)	(Null)			(Null)	
1020.000	7020.000	(Null)	1	(Null)	(Null)			(Null)	
4317.460	29714.260	(Null)	1	(Null)	(Null)			(Null)	
425.000	2925.000	(Null)	1	(Null)	(Null)			(Null)	

Figure 41 : Suppression des champs vides.

- Traitement des erreurs saisis et élimination des doublons

Dans certain cas, nous avons effectué une analyse syntaxique sur certains attributs afin d'éliminer les erreurs de saisis et la suppression des doublons que nous avons constatés. Et lors de notre étude, nous avons constaté l'impossibilité de corriger certain champs de table tels que la saisie incorrect des noms d'importateurs.

EDIT	ID_DIM_CMPGN_MRTM	NOM_CMPGN_MRTM
1		ZIML
2		ZIML
3		ZIML
4		ZIML
5		ZIML
6		ZIML
7		ZIML
8		ZIML
9		CFSL
10		ZIML

Figure 42 : L'ensemble des doublons dans la table ARR\_BAYPLAN.

EDIT	ID_DIM_CMPGN_MRTM	NOM_CMPGN_MRTM
1		ZIML
2		CFSL
3		CMA LINE
4		MSC
5		MASA
6		MSCA
7		CMAA
8		NASHCO
9		CNAN
10		TRSA

Figure 43 : La table ARR\_BAYPLAN après l'élimination des doublons.

### 8.2.3. Chargement des données

La dernière mission de l'ETL, est de charger les données dans le DW. C'est-à-dire le chargement des dimensions qui est une tâche très simple considérant que tous les traitements sur les données ont été fait et enfin le chargement des tables de faits.

Pour cette étape de chargement nous avons procédé comme suit :

- **Chargement de la dimension temps**

La particularité de la dimension TEMPS est qu'elle est chargée une seule fois durant le cycle de vie de l'ED. Cette dimension est la seule qui n'existe pas dans le système source et qui nécessite un traitement totalement différent des autres dimensions. Nous allons pour cela construire cette dimension temps.

- **Chargement des dimensions non changeantes**

Cette étape consiste à charger les tables dimensionnelles statiques (non changeantes) dans le temps. Dans notre cas c'est la table dimension prestation et région

- **Chargement des dimensions changeantes**

Après avoir chargé toutes les dimensions non changeantes, nous passons au chargement des dimensions changeantes telles que la dimension clients.

- **Chargement de Tables de faits**

Après avoir chargé toutes les dimensions, nous passons au chargement des différentes tables de faits. Cette étape correspond au chargement des différentes mesures liées aux informations de

l'activité à analyser. Il faut assurer que pour chaque clé étrangère dans la table de fait, il existe une correspondance dans la table dimensionnelle.

## Conclusion

Dans ce chapitre nous avons décrit la conception de la zone d'alimentation de l'entrepôt, de l'identification des sources de données jusqu'à la conception du processus d'alimentation lui-même, en passant par l'architecture ETL.

L'objectif était le bon acheminement des données provenant du système de production leurs chargements dans ED, tout en assurant la qualité et l'intégrité des données, en évitant de nuire au système de production (c.à.d. le système existant).

## Synthèse de la troisième partie

Dans cette partie, nous avons présenté la conception de notre solution pour notre entrepôt de données . Nous avons commencé par une représentation des données de l'entrepôt grâce à la modélisation dimensionnelle, qui permet de naviguer dans les données afin de satisfaire les besoins d'analyse.

Ensuite, nous avons poursuivi par la conception de la zone d'alimentation, qui constitue l'activité la plus importante et la plus complexe de tout projet décisionnel, car elle permet de passer de la structure de données sources à la structure multidimensionnelle.

Une fois la conception achevée, nous passons dans la partie suivante à la réalisation et le déploiement de notre solution. Cette partie consiste à donner une forme concrète à la phase de conception.

---

## Partie 04: Réalisation

---

Chapitre 9 : Réalisation et mise en œuvre.

# Chapitre9 : Réalisation et mise en œuvre.

## Introduction

Après avoir présenté la conception de notre entrepôt de données, nous allons décrire dans ce chapitre la mise en place de notre solution et décrire en détails sa réalisation.

Pour cela, il a été nécessaire de recourir à un certain nombre d'outils et mettre en place du nouveau système. Ce chapitre s'intitule autour de la description de l'architecture technique de notre solution ainsi que les étapes de réalisation de différentes zones de données conçus dans chapitre précédent. Pour illustrer notre travail nous avons utilisés des captures d'écrans.

### 9.1. Outils utilisés

#### 9.1.1. Talend Open Studio (TOS)

Pour la zone d'alimentation, nous avons choisi Talend Open Studio (TOS) qui est basé sur l'IDE « Eclipse » et intègre un ensemble de composants implémentés en JAVA.

Après une étude comparative (**ANNEXE 5**), le choix a été porté sur Talend Open Studio dans sa version « 6.2 ». Ce choix se justifie par les avantages offerts par cette solution :

- ✓ Solution logicielle open source, développée par la société Talend en France, qui offre les fonctionnalités nécessaires pour répondre aux besoins d'extraction et chargement des données. Son usage permet de se doter d'un outil puissant sans engager des frais de licence,
- ✓ Solution facile à utiliser avec une interface très agréable et très maniable (beaucoup d'actions se font en glisser-déposer), incluant une documentation complète et riche,
- ✓ TOS est adaptable et polyvalent, admettant une communauté très réactive et participative pour son développement,
- ✓ Traitement des composants : FTP, SMTP...etc.
- ✓ Il possède une très grande gamme de composants (246 composants), permettant de se connecter à quasiment toutes les sources de données (oracle, MySQL, PostgreSQL...Etc.), types de fichiers existants (Excel, XML...Etc.), ainsi que aux progiciels (ERP, CRM...etc.). Talend est un code générateur java toujours visible par l'utilisateur, c'est-à-dire que l'installation de TOS n'est pas obligatoire sur les serveurs de

production. Il est possible d'enregistrer les transformations sous forme de scripts, qui sont alors exécutables à partir de n'importe quelle machine disposant d'une JVM (pour Java).

### **9.1.2. SGBD Oracle Data base 11g Express Edition**

L'entrepôt de données a été implémenté et stockées sous le SGBD payant Oracle Data base 11g Express Edition. Ce choix se justifie par le fait que le service informatique de BMT ait déjà acquis une licence. D'une autre part, La version 11g d'Oracle offre une gamme riche d'outils de configuration et d'administration, aussi elle permet une gestion efficace et sécurisé de données.

### **9.1.3. Pentaho**

Pour la zone de présentation de données nous avons opté Pentaho qui est une plate-forme décisionnelle open source complète qui intègre un serveur web Apache Tomcat 6.0. Ce dernier interagit avec l'ensemble des clients Pentaho pour afficher les différents rapports, tableaux de bord, cubes, . . . etc.

Il permet la distribution de fonctionnalités et de documents décisionnels à un grand nombre de destinataires via une interface web. L'activité principale de **Pentaho** est l'exploitation des données d'une entreprise pour ainsi faciliter la prise de décision. Notre choix se justifie par les avantages offerts par cette suite logicielle, qui peuvent se résumer en :

- Il n'a pas besoin de connaître JAVA pour travailler avec Pentaho : seule la maîtrise du langage SQL est nécessaire, ainsi que des connaissances de base en XML, HTML. Il faut bien sûr s'auto former (ou être formé) aux clients de conception.

## 9.2. Réalisation de la solution

### 9.2.1. Réalisation de la zone d'entreposage

Il s'agit de l'implémentation des DataMarts constituant le DW. De ce fait nous avons créé base de données sous de oracle 11 g (oracle express Edition). Pour les implémenter physiquement, nous avons utilisée l'outil ROLAP qui permet de gérer la structure multidimensionnelle dans un SGDB relationnel. D'abord nous avons créé les tables de dimensions chacun d'entre elles contiennent une clé primaire (clé de substitution) et un ensemble d'attributs. Puis nous avons créé les tables de fait celles-ci possèdent comme attributs les mesures d'activités et les clés étrangères vers les tables de dimension.

- **Construction de la zone d'alimentation (ETL)**

Cette zone représente l'étape la plus importante et la plus complexe en termes de charge de travail. Cette phase d'alimentation des magasins de données se déroule en plusieurs étapes, qui sont : chargement de la dimension temps qui se fait une fois durant le cycle de vie de l'ED, chargement des dimensions non changeantes, chargement des dimensions changeantes et chargement des tables de faits.

Comme nous l'avons mentionné précédemment, nous avons opté pour **Talend Open Studio** pour la réalisation des différents programmes ETL. Cet outil présent de grandes capacités de traitement des gros volumes de données, il offre aussi plusieurs fonctionnalités telles que la récupération de schéma de données après modification, vérification de type de données et signalisation d'erreurs.

Le tableau ci-dessous présente les fonctionnalités de TOS que nous avons utilisé dans le processus ETL :

COMPOSANT	DESCRIPTION
<b>TLOGROW</b>	D'AFFICHER LES DONNEES SUR LA CONSOLE
<b>TMYSQLINPUT</b>	EST LINKE AVEC UNE TABLE MYSQL AFIN DE FOURNIR DES DONNEES EN ENTREE
<b>TMYSQLOUTPUT</b>	EST LINKE AVEC UNE TABLE MYSQL AFIN DE POUVOIR RECEVOIR DES DONNEES

<b>TMySQLConnection</b>	OUVRE UNE CONNEXION A UNE BASE MYSQL
<b>TOracleInput</b>	EST LINKE AVEC UNE TABLE ORACLE AFIN DE FOURNIR DES DONNEES EN ENTREE
<b>TOracleOutput</b>	EST LINKE AVEC UNE TABLE MYSQL AFIN DE POUVOIR RECEVOIR DES DONNEES
<b>TFileInputExcel</b>	LE COMPOSANT <b>TFileInputExcel</b> LIT UN FICHIER EXCEL (.XLS OU .XLSX) OU UN FLUX DE DONNEES ET EN EXTRAIT LES DONNEES LIGNE PAR LIGNE
<b>TMap</b>	PERMET DE REALISER ENTRE AUTRE : <ul style="list-style-type: none"> <li>· JOINTURE (INTERNE/EXTERNE),</li> <li>· TRANSFORMATION DES DONNEES,</li> <li>· CONCATENATION ET INVERSION DE CHAMPS,</li> <li>· FILTRAGE,</li> <li>· GESTION DES REJETS,</li> <li>· QUALITE DE DONNEES.</li> </ul>
<b>TJoin</b>	PROCESSING DE REALISATION DES JOINTURES
<b>TUniqRow</b>	PERMET DE DETERMINER LES DOUBLONS EN FONCTION D'UNE COLONNE DE DONNEES EN ENTREE. EN SORTIE, IL EST POSSIBLE SOIT D'AVOIR LES DONNEES UNIQUES SOIT D'AVOIR LES DONNEES DONT LA COLONNE ETUDIEE A DEJA ETE RENCONTREE.
<b>TFILTERJOIN</b>	PROCESSING PERMETTANT DE LIMITER LES COLONNES EN SORTIE DU COMPOSANT.

*Tableau 12: Tableau descriptif des composants utilisés.*

Maintenant avec ses différents nous avons pu les utiliser afin de réaliser le processus ETL ; extraire les données sources depuis les bases de production et leurs subirent de divers transformation et puis les charger dans un entrepôt de données.

La figure suivante illustre parfaitement l'une des transformations faites sur la table source importateur pour charger la table dimension client.

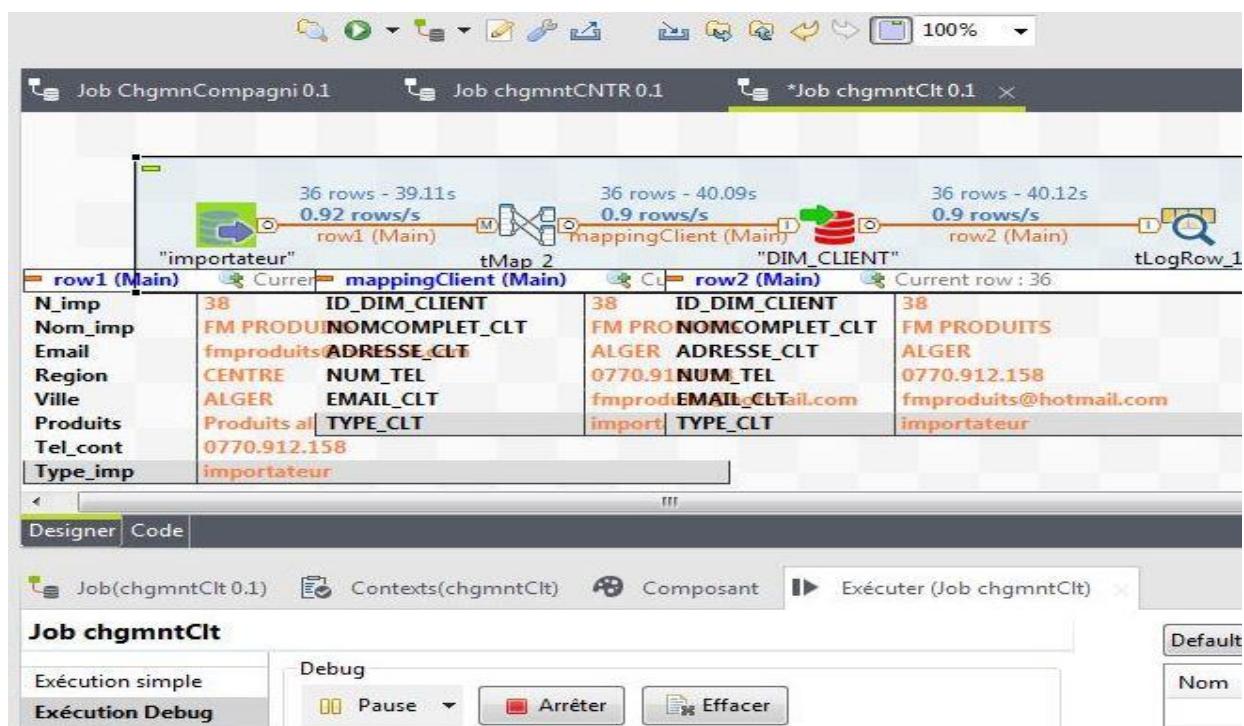


Figure 44 : Mapping de la dimension client.

Talend open studio nous a permis d'exécuter notre job en mode trace debug qui nous permet un suivi du traitement des données et nous fournit l'option de l'insertion ligne par ligne et affiche le résultat d'une manière dynamique. La figure suivante présente le mapping de la table fait clients :

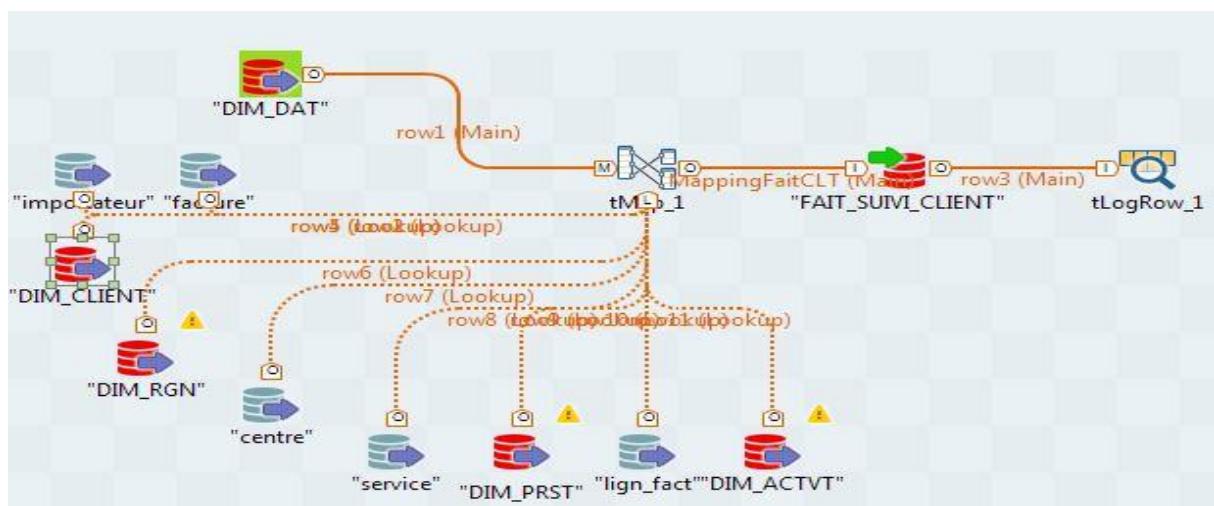


Figure 45 : Mapping da la table fait suivi client.

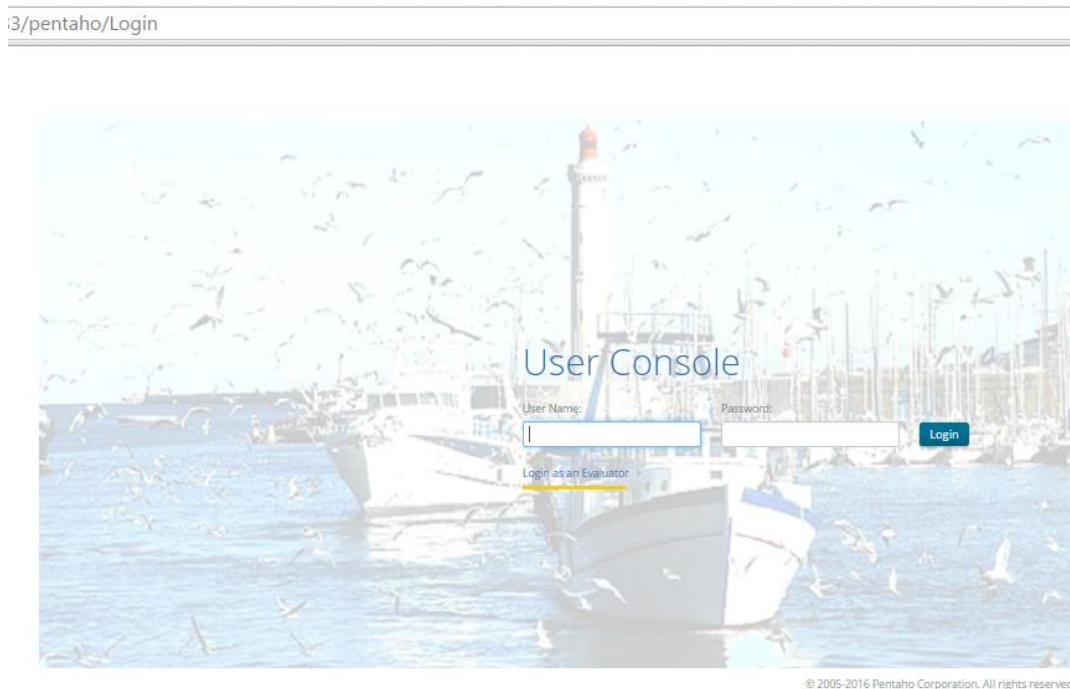
### 9.2.2. La réalisation du reporting

Pour présenter les données contenues dans le DW à l'utilisateur final, nous avons utilisé : le Reporting .à l'aide d'outil Pentaho Reporter Design. La partie Reporting est dynamique puisque l'utilisateur (qui a le privilège nécessaire) a la possibilité d'ajouter de nouveaux rapports ou de les modifier.

Dans ce qui suit nous allons montrer 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 tape son login ainsi que son mot de passe :



**Figure 46 : Authentification.**

- **Création d'un nouveau rapport**

Une fenêtre s'ouvre qui permet à l'utilisateur de créer un nouveau rapport d'activité en lui ouvrant la possibilité d'accès à l'entrepôt de données puis la présentation de données sous différents états en utilisant une représentation graphique, l'utilisateur n'a qu'à choisir le type de présentation (courbe, barres, etc.), il l'obtiendra facilement en spécifiant les axes d'analyses sur lesquelles il veut afficher ces résultats. Voici un exemple de rapport :

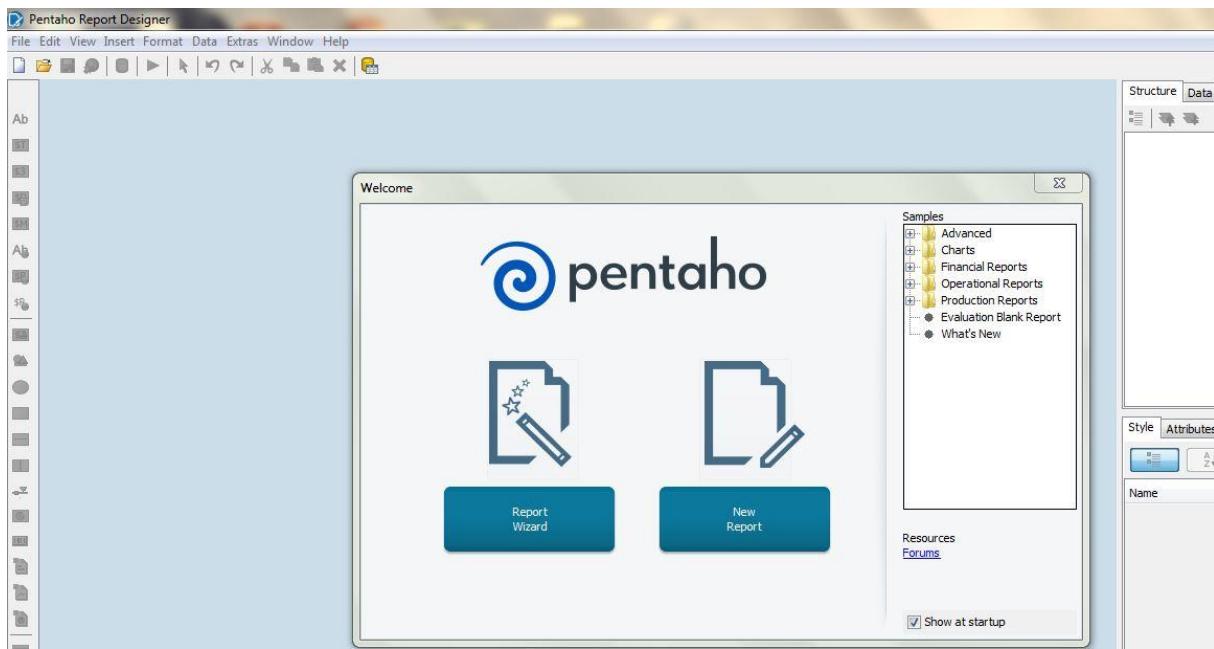
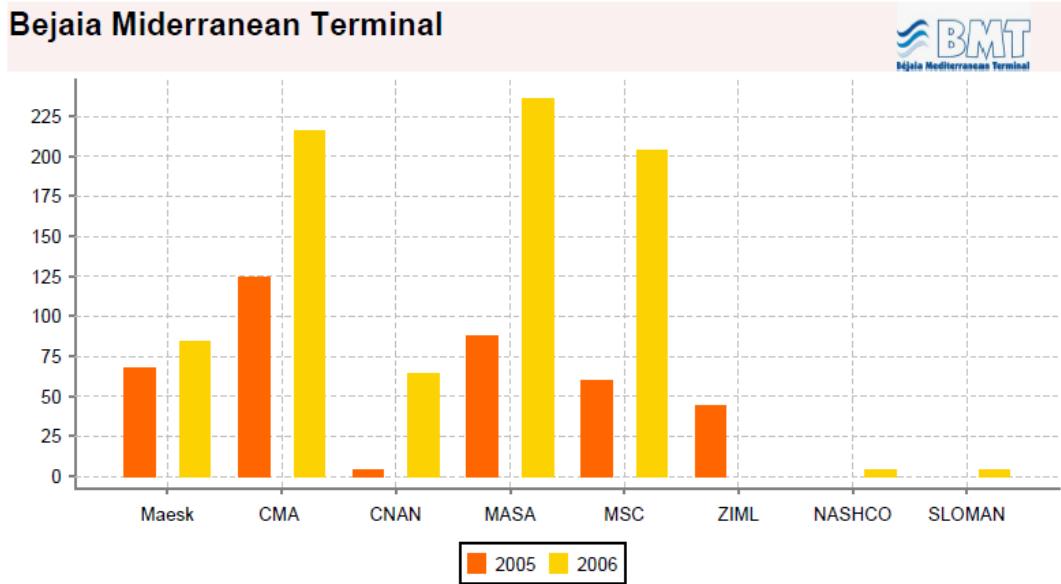


Figure 47 : Créer un nouveau rapport.

- **Etat de sortie du rapport**

Activité de suivi de compagnie maritime				Annee: 2005 à 2006
nom compagnie	Qte_evp	Conteneur	Année	
Maesk	48	20	2 005	
Maesk	20	40	2 005	
CMA	112	20	2 005	
CMA	12	40	2 005	
CNAN	4	40	2 005	
MASA	36	20	2 005	
MASA	52	40	2 005	
MSC	24	20	2 005	

Figure 48 : Rapport d'activité de suivis de compagnie maritimes sous format d'un tableau.



**Activité de suivi de compagnie maritime** Année: 2005 à 2006

Figure 49 : Statique sur quantité d'EVP entre les compagnies maritimes.

- **Exporter le rapport vers une autre Format**

L'utilisateur peut exporter le rapport final créé vers un format choisi comme PDF or EXCEL pour l'imprimer.

### 9.2.3. La sécurité du système

Au niveau de l'application utilisateur, nous avons utilisé la stratégie de sécurité disponible avec Pentaho BI Server qui est basée sur le mécanisme d'autorisation et d'authentification permettant d'affecter des droits d'accès pour chaque utilisateur. Au niveau de la base de données qui est notre entrepôt de données , nous utilisons la gestion des mécanismes offert par le SGBD Oracle 11 g, ce dernier est paramétré pour autoriser ou interdire les accès et la manipulation des données du Au niveau de l'application utilisateur, nous avons utilisé la stratégie de sécurité disponible avec Pentaho BI Server qui est basée sur le mécanisme d'autorisation et d'authentification permettant d'affecter des droits d'accès pour chaque utilisateur. Au niveau de la base de données qui est notre entrepôt de données , nous utilisons la gestion des mécanismes offert par le SGBD Oracle 11 g, ce dernier est paramètre à limiter les accès ainsi que la manipulation des données stockées dans un entrepôt de données.

### 9.3. Architecture technique de la solution

Pour assurer la mise en œuvre de notre entrepôt de données une architecture de notre solution serait envisageable ainsi que les différents outils utilisés. La figure suivante représente l'architecture technique de notre solution

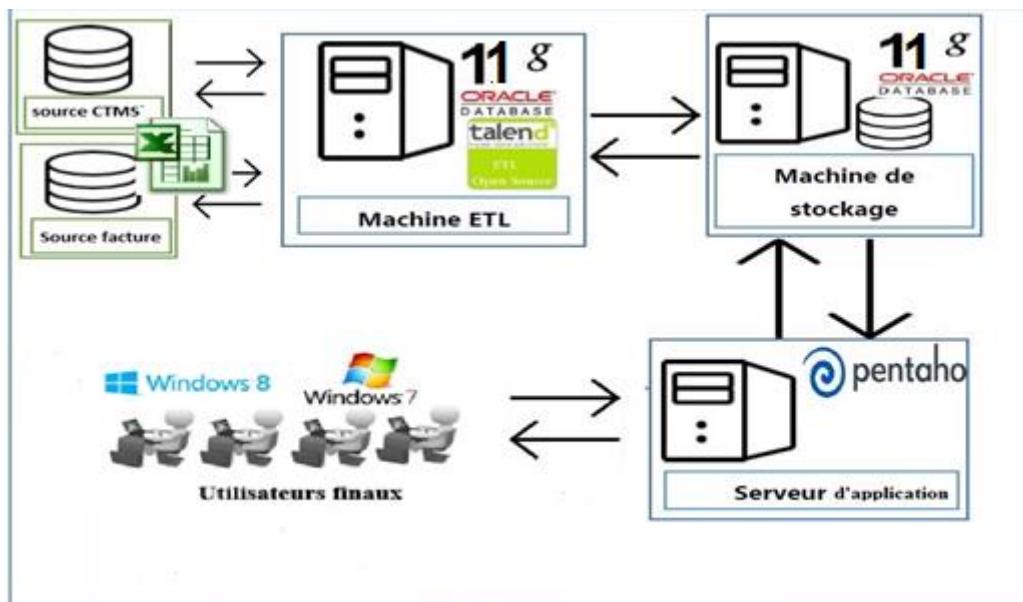


Figure 50 : Architecture technique du système proposé

## Conclusion

Dans cette partie nous avons décrit l'architecture technique qui va recevoir notre solution. Ainsi que l'environnement de développement qui ont été pris en compte lors de la phase de réalisation du DW. Nous avons aussi présenté tous les outils que nous avons utilisés et qui sont Oracle 11g et talend open studio ainsi que Pentaho BI open source qui nous a permis la collecte, l'intégration et l'organisation des données issues de différentes sources et la restitution des données afin de réaliser le reporting.

## Conclusion générale et perspectives

De nos jours, l'information est considérée comme la clé primaire de l'économie et constitue un instrument de compétition, c'est pourquoi les entreprises sont conscientes de la maîtrise de l'information afin de la rendre disponible sous la bonne forme, au moment opportun à la bonne personne qui sera l'exploiter. C'est pour cela qu'il faut accompagner les décideurs des entreprises par d'outils analytiques sophistiqués pour leur permettent de prendre des décisions pertinentes.

Afin de préserver sa place dans un marché extrêmement concurrentiel, la BMT-Spa s'est engagé dans le projet de mise en place d'un système décisionnel dédié au service marketing basé sur l'architecture d'un entrepôt de données qui permet de transformer les données de production en décision stratégiques.

Afin de bien mener notre projet, nous avons adopté une méthodologie itérative de planification et de suivi de projet inspirée de l'approche du cycle de vie dimensionnel de Ralph Kimball, qui consiste en différentes étapes pour la mise en place d'un système décisionnel. Cette méthode utilise principalement l'approche guidée par l'analyse des besoins.

Dans un premier temps, nous avons étudié l'existant de l'entreprise afin de comprendre le processus métier de ses activités puis cerner les besoins de son service marketing. Nous avons utilisé le recueil par entretien comme outil principal, complétés par l'analyse documentaire, puis clarification des points obscurs en faisant des réunions avec notre encadreur afin de mieux comprendre. De ce fait nous put ressortir les sujets d'analyses qui sont : suivi des clients, suivi de compagnies maritimes et en fin suivi des navires.

Nous avons ensuite entamé la conception de notre entrepôt de données. A partir des besoins recueillis, nous utiliser la modélisation dimensionnelle, pour la représentation graphique de chaque activité d'une manière simplifier basé sur la table de fait et table de dimensions.

Une fois les modèles logiques définis, nous avons passé à l'alimentation de la zone de stockage, ce qui nous a consommé énormément de temps (extraction de l'ensemble des données source, vérification de leur qualité, apport de transformations nécessaires et chargement dans la zone de stockage) et ce par le biais d'outil ETL. Une fois l'alimentation terminée, nous sommes passées au chargement de nos DataMarts.

Enfin pour présenter les données stockées dans un entrepôt de données nous avons réalisé un reporting ou nous avons cité l'architecture technique de notre solution ainsi que les différentes technologies utilisés.

Nous pouvons à la fin citer quelques perspectives qui semblent importantes afin d'apporter des améliorations et développements au système, et qui sont les suivantes :

- ✓ Suivre la solution actuelle afin d'apporter les modifications nécessaires pour l'amélioration des performances du système.
- ✓ Intégrer de nouveaux outils performants tels que le CRM (Customer Relationship Management).
- ✓ Réaliser d'autres magasins de données pour les autres services : service des opérations et service de comptabilité.
- ✓ Proposé une nouvelle codification de données ou bien une codification uniformisé pour les noms des importateur basé sur la simplicité.
- ✓ Intégrer de nouveaux outils d'analyses tels que le tableau de bord permettant de suivre et d'anticiper le fonctionnement des activités du service Marketing, afin de mesurer la performance de cette dernière,

Pour finir, nous pouvons conclure que malgré les obstacles rencontré, ce projet de fin d'études est une expérience amplement enrichissante qui nous a permis d'accroître notre savoir, tout en mettant en pratique nos connaissance acquise durant notre cursus cela nous a permet aussi d'explorer de nouvelles technologies et d'outils d'analyse et de développement, et aussi cela nous a permet de dire que c'est une agréable expérience professionnelle, qui nous a fait découvrir le marché du travail afin de nous préparer à continuer notre chemin.

## Bibliographies

[Ansoff, 1965]: H.I.Ansoff, « corporate strategy», Mc Grow-Hill, 1965.

[Bouquin, 2003] : Bouquin Henry ; « Le contrôle de gestion » ; P.U.F ; 2003.

[Codd, 1998]: E.F. Codd, Providing OLAP (On Line Analytical Processing) to user analyst: an IT mandate, rapport technique, 1998.

[Cigref,2009] :Reprit de Matthieu lafare thèse professionnelle HEC : « business intelligence »

[Fernendez, 2011] A. Fernandez, "L'essentiel du tableau de bord: Méthode complète et mise en pratique avec Microsoft Excel »,2011.

[Hilfer ,2008] : Jean-Pierre HILFER, Michel KALIKA et Jaques ORSONI, « Management : Stratégie et Organisation », Vuibert, 2008.

[Inmon, 1996]: W.H. Inmon, «*Building the Data Warehouse*», New York: John Wiley & Sons, 1996.

[Inmon, 2005]: W. H. Inmon, Building the Data Warehouse, John Wiley and sons, New York, 4ème ed. 2005.

[Kimball, 2002] : Kimball R., Ross M., The Data warehouse Toolkit 2ème edition, John Wiley & Sons Inc., New York, 2002.

[Kimball, 2007] :R. Kimball, M. Ross et W. Thornthwaite, « Le Data Warehouse – Guide de conduite de projet »,2007.

[Michel, 2011] :Propos Sur les SI Décisionnels, Michel Bruley Septembre 2011.

[Nakache, 1998] : Didier Nakache; « *Data Warehouse et Data Mining* »; Conservatoire National des Arts et Métiers de Lille; Version 1.1; 15 juin 1998.

[POLETTTO, 2012] : POLETTTO, Maxime. L'informatique décisionnelle. Thèse professionnelle.

[SebastianFantini, 2010]: Business Intelligence avec SQL Server 2008 R2.

[Simon ,1980] : H.Simon . La décision par les ordinateurs, Economica, 1980.

# Webographie

**[WEB1] :**

[http://fr.wikiversity.org/wiki/Prévision\\_décisionnelle/Application\\_de\\_la\\_prévision.](http://fr.wikiversity.org/wiki/Prévision_décisionnelle/Application_de_la_prévision)

**[web2] :**

thèse de Lydie Soler Janvier 2008 Sur les entrepôts de données.

**[web3] :**

[https://fr.wikipedia.org/wiki/Datamart#/media/File:DMT\\_DRs.JPG](https://fr.wikipedia.org/wiki/Datamart#/media/File:DMT_DRs.JPG).

**[web4] :**

[http://www-igm.univ-mlv.fr/~dr/XPOSE2009/informatique\\_decisionnelle\\_olap/md.html](http://www-igm.univ-mlv.fr/~dr/XPOSE2009/informatique_decisionnelle_olap/md.html).

**[web5]:**

<http://www.piloter.org/business-intelligence/reporting.htm>

## ANNEXES

### Annexe 1: schéma descriptif de la réglementation de facturation.

#### 7. Schéma descriptif de la procédure :

Qui ?	Fait quoi ?	Pourquoi ?	Quand ?	Comment ?
Le Chargé de la facturation	Réception des dossiers à facturer	Facturation des prestations fournies	A la fin de la prestation fournie par la direction des opérations	-Fiches de suivi des opérations -bordereau d'envoi
Le chargé de la facturation	Vérification des dossiers à facturer	Vérifier la conformité des dossiers à facturer	A la réception des dossiers	-Fiches de suivi et bordereau d'envoi -bon de commande -CMR
Agents de la facturation	Etablissement des factures	Exigence de la loi	A la validation des dossiers à facturer	Logiciel de facturation
Le Chargé de la facturation	Vérification des factures	Validation des factures	A l'émission des factures	Factures par rapport aux dossiers facturés
Le Chef du Service Commercial	2 ème Vérification des factures	Validation des factures	Après la 1ère vérification	Factures par rapport aux dossiers facturés
Agents de facturations	Dispatching des factures	Classement des factures	Dès validation des factures	Original+2souches pour le client,1souche pour la DFC,1souche pour l'archivage
Agents de recouvrement	Remise des factures	faire parvenir les factures aux clients	A la fin du dispatching	-Au niveau du service -Tournées périodiques
Agents de facturation	Archivage	-Traçabilité -Sauvegarde	A la fin du dispatching	-au niveau de la salle d'archives

## Annexe 2 : Suivi de procédure tarifaire pour chaque type de prestation

Exemple de prestation: manutention.

### 2 - TARIFS DE MANUTENTION AUX NAVIRES

Applicable à partir du 01/01/2014

#### **Vracs solides (céréales)**

Le calcul des frais de débarquement (FD) pour les produits en vrac est basé sur les unités et le poids figurant sur le manifeste ou sur le connaissance. En temps normal, aux forfaitaire de la tonne débarquée est de **180DA /tonne** hors taxe, tous frais fondu.

**NB : Toute fraction sur le tonnage est considérée comme unité complète.**

#### **Tarif des embarquements/débarquements :**

Le débarquement ou le débarquement d'un conteneur est facturé à l'unité (à la boîte) dans le type de conteneur (20 pieds ou 40 pieds) indépendamment de son poids(plein ou vide).

#### **Tarifs des prestations de manutention aux navires :**

##### **tarifs comprennent :**

Le déchargement ou le chargement avec les moyens du navire et/ou portiques de quai. Le transfert du conteneur du navire jusqu'au terminal et vice-versa. Les équipes affectées au déchargement et chargement. Les pointeurs affectés au comptage au niveau du navire. Les engins et moyens mis en place.

##### **tarifs ne comprennent pas :**

Les opérations de pointage et d'engins dans le cadre de la préparation d'un débarquement. Les fins de vacances en normal. Des diverses attentes enregistrées pendant les horaires normaux. Des grues affectées au chargement et au déchargement. Des attentes imputables au navire opérant en heures supplémentaires.

Les tarifs de débarquement des conteneurs pleins y compris les transferts au parc sont de :  
 - Conteneur 20 pieds : **7.560 DA/HT.**  
 - Conteneur 40 pieds : **10.800 DA/HT.**

Le même tarif est appliqué pour une opération de débarquement ou d'embarquement à la fois, sans rapprochement au terminal (**land and reship**).

Les tarifs d'embarquement des conteneurs vides ou pleins à l'export sont :

- Conteneur 20 pieds : **6.300 DA/HT.**
- Conteneur 40 pieds : **9.000 DA/HT.**

A l'export des conteneurs pleins de marchandises d'origine algérienne, les opérations de manutention bénéficient d'une réduction de **30%** sur les tarifs en vigueur.

#### **b) Tarifs ouverture/fermeture des calles :**

Le tarif ouverture ou fermeture des calles est fixé comme suit : **10.000 DA/opération.**

Une opération est soit une ouverture de calle, soit une fermeture.

#### **c) Tarifs shifting à board :**

- Conteneur 20 pieds (plein ou vide) (HT) : **3.100 DA/unité.**

- Conteneur 40 pieds (plein ou vide) (HT) : **4.400 DA/unité.**

#### **d) Travaux de régie aux navires :**

Ce sont des travaux spécifiques commandés par le bord ou son représentant qui peuvent affecter le rendement, ou alors des prestations n'entrant pas dans le cadre d'un travail normal (désarrimage, mauvais chargement, saisissage, déssaisissage, balayage et ramassage dans les calles, ...).

Ces travaux sont facturés pour un shift et par équipe (où équipe est composé de 04 hommes OM/OMS/cariste/conducteur selon la prestation).

- Équipe/shift : **18.000 DA.**

#### **2.3 - Levage pour mise à disposition de conteneurs vides pour l'export :**

- Conteneur 20 pieds **3.500 DA/HT.**

- Conteneur 40 pieds **3.500 DA/HT.**

Une réduction de 30% est appliquée à l'export de marchandise d'origine Algérienne.

#### **2.4 - Levage pour rapprochement au niveau de la zone extra-portuaire :**

- Conteneur 20 pieds **3.500 DA/HT.**

- Conteneur 40 pieds **3.500 DA/HT.**

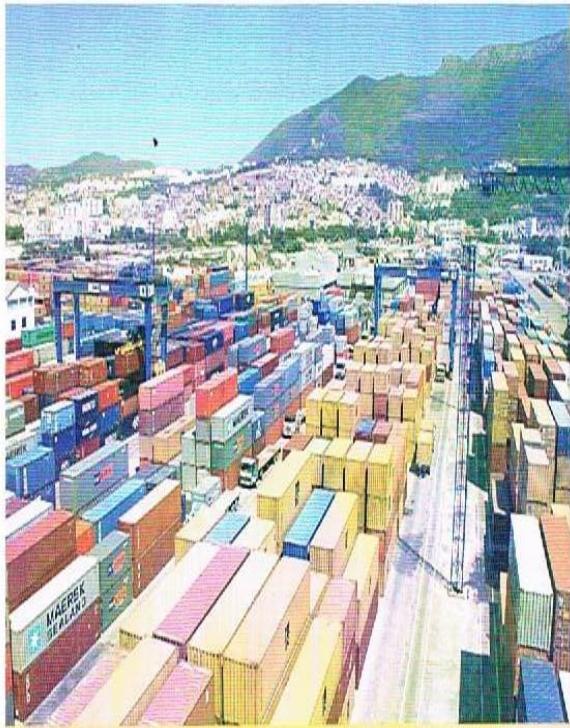
#### **2.5 - Rapprochement des conteneurs vides :**

Le tarif de transport pour le rapprochement des conteneurs vides depuis la zone extra-portuaire d'Iriyahan vers le terminal à conteneurs est comme suit :

- Conteneur 20 pieds **1.350 DA.**

- Conteneur 40 pieds **2.700 DA.**

Type de prestation : Acconage.



**Tarifs des Prestations  
Levage, Entreposage et Autres**



**3 - TARIFS DE LEVAGE (ACCONAGE) AU PARC.**

Applicable à partir du 01/01/2014

**3.1 - Tarifs des prestations d'accionnage (levage) dans le terminal :**

Toute opération dans le parc (visite, dépotage, empotage, livraison, réception, pesée, mise à disposition,...) comporte un ou plusieurs levages (LOLO) :

Lift On (**CHARGEMENT**) – Lift Off (**DECHARGEMENT**)

Et des shifting au niveau du parc. Le chargement et le déchargement d'un conteneur sont des opérations différentes facturées à part sur la base suivante :

OPÉRATIONS	20 pieds   ou   40 pieds	Obs.
LOLO pour visite	3.850 DA	
LOLO pour Dépotage	3.850 DA	
LOLO pour Pesée	3.850 DA	
LOLO pour Livraison	3.850 DA	
LOLO pour Restitution	3.850 DA	
LOLO pour Empotage Interne	3.850 DA	
LOLO Divers	3.850 DA	

Ces tarifs sont harmonisés et unifiés indépendamment du nombre de coup de fourches ou de shifting, en relation directe avec la prestation.

- Toute opération de levage de conteneurs vides pour l'export, bénéficie de 30% de réduction du tarif.

## Annexe 2 : Suivi de procédure tarifaire pour chaque type de prestation

### Prestation entreposage

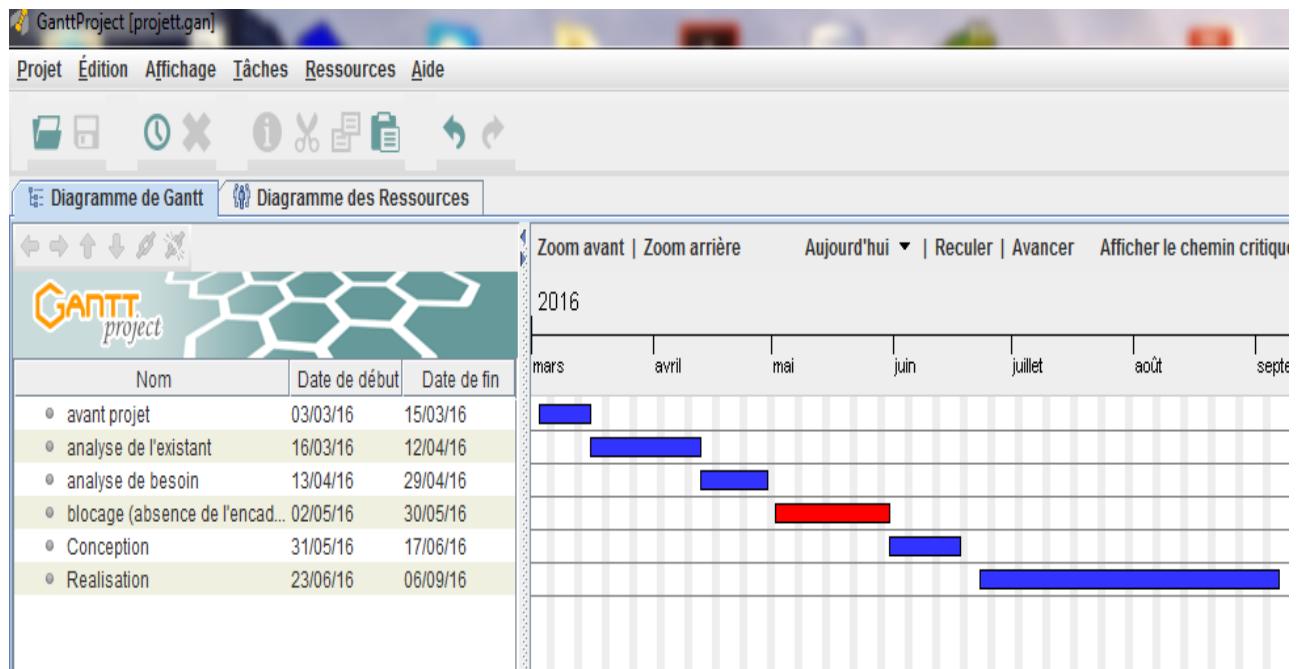
<p><b>5 - AUTRES PRESTATIONS</b> Applicable à partir du 01/01/2014</p> <p><b>.1 - Conteneurs Frigorifiques :</b> La fourniture d'énergie électrique des conteneurs frigorifiques sera facturée comme suit : 3.500 DA / jour/boîte (20 pieds ou 40 pieds). Le tarif inclus la prise en charge des conteneurs frigorifiques dans l'installation préalablement, au niveau du parc frigorifique.</p> <p><b>.2 - Tarifs de location des engins :</b> En fonction de la disponibilité, BMT peut mettre à disposition de ses clients des engins dans l'enceinte du terminal à conteneurs et du port pour une période de temps déterminée, à la location s'effectue sur la base de l'heure et est facturée selon le barème suivant :</p> <table border="1" style="margin-left: auto; margin-right: auto; border-collapse: collapse; text-align: center;"> <thead> <tr> <th>Type d'engins</th> <th>Tarif/Heure (HT)</th> </tr> </thead> <tbody> <tr> <td>Tracteur remorque</td> <td>2.500 DA</td> </tr> <tr> <td>Spreader</td> <td>2.500 DA</td> </tr> <tr> <td>Chariot élévateur (&lt;05t)</td> <td>2.000 DA</td> </tr> <tr> <td>Chariot élévateur à pince rotative</td> <td>3.000 DA</td> </tr> <tr> <td>Chariot élévateur (&lt;10t)</td> <td>4.200 DA</td> </tr> <tr> <td>Chariot élévateur (&gt; 10t)</td> <td>5.000 DA</td> </tr> <tr> <td>Steeker</td> <td>6.000 DA</td> </tr> </tbody> </table> <p>- Le minimum de perception de chaque engin est 1 heure.</p> <p><b>i.3 - Prestations d'empotage ou de dépotage (manuel ou mécanisé) :</b> Des équipes et moyens adéquats seront utilisés pour assurer le dépotage ou empotage des conteneurs. Cette prestation sera facturée comme suit : - 4.000 DA pour conteneurs 20 pieds. - 6.000 DA pour conteneurs 40 pieds.</p> <p>Les moyens et équipes correspondent dans le cas d'un : Empotage ou dépotage manuel à une ou plusieurs équipes de dépotage 1 équipe est constituée de 04 agents polyvalents. Ou bien d'un ou plusieurs chariots élévateurs (Empotage/Dépotage mécanisé)</p>	Type d'engins	Tarif/Heure (HT)	Tracteur remorque	2.500 DA	Spreader	2.500 DA	Chariot élévateur (<05t)	2.000 DA	Chariot élévateur à pince rotative	3.000 DA	Chariot élévateur (<10t)	4.200 DA	Chariot élévateur (> 10t)	5.000 DA	Steeker	6.000 DA	<p><b>5.4 - Prestation de dépotage et empotage partiel dans le cas des visites :</b> - 4.000 DA pour conteneurs 20 pieds. - 6.000 DA pour conteneurs 40 pieds.</p> <p><b>5.5 - Prestation de visite intégrale :</b> Le tarif d'une visite intégrale est le suivant :</p> <table border="1" style="margin-left: auto; margin-right: auto; border-collapse: collapse; text-align: center;"> <thead> <tr> <th>Type</th> <th>durée ≤ 1 Shift</th> <th>Durée &gt; 1 Shift</th> </tr> </thead> <tbody> <tr> <td>20 pieds</td> <td>7.000 DA</td> <td>11.000 DA</td> </tr> <tr> <td>40 pieds</td> <td>9.000 DA</td> <td>13.000 DA</td> </tr> </tbody> </table> <p>Cette prestation inclut l'affectation d'une équipe qui assurera le dépotage et l'empotage et tous travaux de manutention nécessaires à la visite intégrale, avec mise à disposition d'engins dans le cas de dépotage mécanisé.</p> <p><b>5.6 - Nettoyage et préparation des conteneurs vides :</b> a) Le nettoyage des déchets après visites et dépotages des conteneurs (emballages déversés de marchandises) seront assurés par les agents d'assainissement de BMT. La prestation de nettoyage est fixée à : - 450 DA pour conteneur 20 pieds. - 700 DA pour conteneur 40 pieds.</p> <p>b) A la restitution et sur demande du consignataire, il sera procédé à la vérification l'état du conteneur et si besoin à son nettoyage avant la fermeture des portes pour préparer à l'embarquement. Toutefois, il est de la responsabilité de l'armateur ou de son consignataire de faire une vérification finale avant la pose des scellés au moment de l'embarquement. La prestation de nettoyage est fixée à : - 450 DA pour conteneur de 20 pieds. - 700 DA pour conteneur de 40 pieds.</p> <p><b>5.7 - Contrôle et Scellage des conteneurs :</b> Les conteneurs pleins destinés à la visite sont systématiquement scellés après la visite. Le tarif de la pose du scellé est de : 150 DA/HT par conteneur.</p> <p><b>5.8 - Prestation de transport par Route/Rail :</b> Le transport des conteneurs sur site client avec retour du vide sera facturé conformément à une convention entre BMT et son client, qui définit clairement les responsabilités de chaque partie, ainsi que les tarifs en vigueur.</p>	Type	durée ≤ 1 Shift	Durée > 1 Shift	20 pieds	7.000 DA	11.000 DA	40 pieds	9.000 DA	13.000 DA
Type d'engins	Tarif/Heure (HT)																									
Tracteur remorque	2.500 DA																									
Spreader	2.500 DA																									
Chariot élévateur (<05t)	2.000 DA																									
Chariot élévateur à pince rotative	3.000 DA																									
Chariot élévateur (<10t)	4.200 DA																									
Chariot élévateur (> 10t)	5.000 DA																									
Steeker	6.000 DA																									
Type	durée ≤ 1 Shift	Durée > 1 Shift																								
20 pieds	7.000 DA	11.000 DA																								
40 pieds	9.000 DA	13.000 DA																								

### Autre prestation

<p><b>4 - PRESTATIONS D'ENTREPOSAGE DANS LE PARC A CONTENEURS</b> Applicable à partir du 01/01/2014</p> <p>La prestation d'entreposage est calculée comme suit : <b>l'import :</b> 1 à 3 jours gratuits.</p> <table border="1" style="margin-left: auto; margin-right: auto; border-collapse: collapse; text-align: center;"> <thead> <tr> <th>JOURS</th> <th>20 pieds</th> <th>40 pieds</th> </tr> </thead> <tbody> <tr> <td>Du 4 au 10e</td> <td>70 DA</td> <td>140 DA</td> </tr> <tr> <td>Du 11 au 15e</td> <td>250 DA</td> <td>500 DA</td> </tr> <tr> <td>Du 16 au 25e</td> <td>500 DA</td> <td>1000 DA</td> </tr> <tr> <td>Du 26 au 35e</td> <td>800 DA</td> <td>1600 DA</td> </tr> <tr> <td>Plus de 35 jours</td> <td>1200 DA</td> <td>2400 DA</td> </tr> </tbody> </table> <p><b>l'export du vide :</b> conteneurs vides entreposés en zone portuaire ou extra-portuaire 1 à 15 jours gratuits</p> <table border="1" style="margin-left: auto; margin-right: auto; border-collapse: collapse; text-align: center;"> <thead> <tr> <th>JOURS</th> <th>20 pieds</th> <th>40 pieds</th> </tr> </thead> <tbody> <tr> <td>Plus de 15 jours</td> <td>250 DA</td> <td>500 DA</td> </tr> </tbody> </table> <p><b>l'export du plein (Marchandises d'origine Algérienne) :</b> 1 à 10 jours Gratuits.</p> <table border="1" style="margin-left: auto; margin-right: auto; border-collapse: collapse; text-align: center;"> <thead> <tr> <th>JOURS</th> <th>20 pieds</th> <th>40 pieds</th> </tr> </thead> <tbody> <tr> <td>Du 11 au 15e</td> <td>30 DA</td> <td>60 DA</td> </tr> <tr> <td>Du 16 au 25e</td> <td>60 DA</td> <td>120 DA</td> </tr> <tr> <td>Du 26 au 35e</td> <td>85 DA</td> <td>170 DA</td> </tr> <tr> <td>Plus de 35 jours</td> <td>120 DA</td> <td>240 DA</td> </tr> </tbody> </table>	JOURS	20 pieds	40 pieds	Du 4 au 10e	70 DA	140 DA	Du 11 au 15e	250 DA	500 DA	Du 16 au 25e	500 DA	1000 DA	Du 26 au 35e	800 DA	1600 DA	Plus de 35 jours	1200 DA	2400 DA	JOURS	20 pieds	40 pieds	Plus de 15 jours	250 DA	500 DA	JOURS	20 pieds	40 pieds	Du 11 au 15e	30 DA	60 DA	Du 16 au 25e	60 DA	120 DA	Du 26 au 35e	85 DA	170 DA	Plus de 35 jours	120 DA	240 DA	<p><b>Au réexport du plein (Marchandises d'autres origines) : 1 à 5 jours Gratuits.</b></p> <table border="1" style="margin-left: auto; margin-right: auto; border-collapse: collapse; text-align: center;"> <thead> <tr> <th>JOURS</th> <th>20 pieds</th> <th>40 pieds</th> </tr> </thead> <tbody> <tr> <td>DU 6 au 15e</td> <td>70 DA</td> <td>140 DA</td> </tr> <tr> <td>DU 16 au 25e</td> <td>120 DA</td> <td>240 DA</td> </tr> <tr> <td>DU 26 au 35e</td> <td>171 DA</td> <td>342 DA</td> </tr> <tr> <td>Plus de 35 jours</td> <td>500 DA</td> <td>1.000 DA</td> </tr> </tbody> </table> <p><b>CAS PARTICULIERS :</b> - Les conteneurs type flat rack ou autre conteneurs pleins ne donnant pas la possibilité de gerber au dessus, seront soumis aux tarifs d'entreposage majoré de 300%.</p>	JOURS	20 pieds	40 pieds	DU 6 au 15e	70 DA	140 DA	DU 16 au 25e	120 DA	240 DA	DU 26 au 35e	171 DA	342 DA	Plus de 35 jours	500 DA	1.000 DA
JOURS	20 pieds	40 pieds																																																					
Du 4 au 10e	70 DA	140 DA																																																					
Du 11 au 15e	250 DA	500 DA																																																					
Du 16 au 25e	500 DA	1000 DA																																																					
Du 26 au 35e	800 DA	1600 DA																																																					
Plus de 35 jours	1200 DA	2400 DA																																																					
JOURS	20 pieds	40 pieds																																																					
Plus de 15 jours	250 DA	500 DA																																																					
JOURS	20 pieds	40 pieds																																																					
Du 11 au 15e	30 DA	60 DA																																																					
Du 16 au 25e	60 DA	120 DA																																																					
Du 26 au 35e	85 DA	170 DA																																																					
Plus de 35 jours	120 DA	240 DA																																																					
JOURS	20 pieds	40 pieds																																																					
DU 6 au 15e	70 DA	140 DA																																																					
DU 16 au 25e	120 DA	240 DA																																																					
DU 26 au 35e	171 DA	342 DA																																																					
Plus de 35 jours	500 DA	1.000 DA																																																					

## Annexe 3 : Diagramme de GANTS.

Pour optimiser les chances de réussite et garantir le bon déroulement de notre projet, nous avons organisé et planifié le projet en utilisant le logiciel Gantt Project de la manière suivante :



## **Annexe 04: listes des questions pour entretiens.**

### **Questions aux analystes**

- Quels types d'analyse régulière effectuez-vous actuellement ? sur quel type de données ?
- Comment obtenez-vous les données ?
- Quelles sont vos sources de données ? quel sont vos indicateurs de performances ??
- Quels sont les données que vous souhaiteriez historier ?
- Quels sont les principaux problèmes auxquels vous faites face aujourd'hui ?
- Disposez-vous d'un outil d'analyse automatique ?
- Quels sont vos besoins et vos attentes vis-à-vis de ce nouveau système ?

### **Questions à l'administrateur de donnée**

- Quel type d'analyse est généré régulièrement ?
- Quels sont les outils d'accès et d'analyse ?
- Comment les besoins des utilisateurs en matière d'accès aux données et d'analyse sont-ils pris en charge ?
- Quels sont les problèmes connus concernant la qualité des données ?
- Qu'attendez-vous de l'entrepôt de données ?

### **Questions aux décideurs tactiques**

- Quels sont vos objectifs par ordre de priorité ?
- Quelle est votre manière d'accès à l'information? Comment accédez-vous à l'information ou quels sont vos moyens d'accès à l'information?
- Quels sont les insuffisances de la procédure actuelle de prise de décision ?
- Quelles sont vos méthodes pour évaluer la réussite. ? c.à.d. Quels sont vos critères de réussite ?
- Comment savez-vous que vous êtes sur la bonne voie ?
- Quelles sont les problèmes à lesquelles vous affronter aujourd'hui ??
- Comment anticipiez-vous les problèmes que vous rencontrez ??
- De quelle manière envisagez-vous une meilleure exploitation de l'information ?

## Annexe 5 : Etude comparative entre les outils ETL/ talend et pentaho.

Les outils ETL sont généralement utilisés dans les systèmes décisionnels afin de permettre l'**Extraction**, la **Transformation** et le **Chargement** de données depuis des sources diverses (bases de données, fichiers) vers un entrepôt de données. Parmi les outils ETL open source ; nous pouvons citer Talend Open Studio, Pentaho Data Intégration, Clover ETL, Octopus ...etc. Nous avons réalisé une comparaison entre Talend Open Studio et Pentaho Data Intégration., qui sont les plus répondu dans le marché :

	FONCTIONNALITES	PENTaho	TALEND
ACCES AUX DONNEES	APPEL DE PROCEDURES STOCKEES	OUI	SEULEMENT POUR CERTAINS SGBD
	EXECUTION DES REQUETES	OUI	OUI
	PERFORMANCE	OUI, MOYENNE	OUI, ELEVEE
	OUTIL DE CREATION DE REQUETES	NON	OUI
TRAITEMENT DES DONNEES	FONCTIONS DE TRANSFORMATIONS DES DATES ET DES NOMBRES	OUI	OUI
	GENERATION DE DOCUMENTATION TECHNIQUE ET FONCTIONNELLE	NON	OUI
	GESTION DES ERREURS D'INTEGRATION	OUI, POUR CERTAINES ETAPES	OUI
	RECUPERATION SCHEMA DE DONNEE	NON	OUI
GESTION DE LA SECURITE	SECURITE SUR L'ACCES AUX METADONNEES	OUI	OUI
	SECURITE SUR LA CONSOLE D'ADMINISTRATION	OUI	OUI