

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

Université Saad Dahlab, Blida
USDB.

Faculté des sciences.
Département informatique.



**Mémoire pour l'obtention
d'un diplôme d'ingénieur d'état en informatique.**
Option : Système d'Information

Sujet :

**Conception & déploiement d'un
système décisionnel pour le
pilotage des créances clients**

Présenté par : KINOUCHE Sofiane
TAMOURT Yacine

Promoteur : M^{elle} N. BOUSTIA
Encadreur : Mr H. ALIOUI

Organisme d'accueil : NAFTAL SPA.

Soutenue le: date soutenance, devant le jury composé de :

Président
Examinateur
Examinateur

- Numéro/2006-2007-

MIG-004-177-1



Liste des figures

Fig-01- Structure d'un système d'aide à la décision	6
Fig-02- Organisation orienté sujet du data warehouse	10
Fig-03- Intégrité des données du data warehouse	11
Fig-04- Historisation des données dans le data warehouse	11
Fig-05- Non-volatilité des données dans le data warehouse	12
Fig-06- Les composants d'un Data Warehouse	14
Fig-07- Structure d'un data warehouse	16
Fig-08- Architecture d'un produit MOLAP	23
Fig-09- Architecture d'un produit ROLAP	23
Fig-10- Exemple de représentation multidimensionnelle	25
Fig-11- Exemple de fait	25
Fig-12- Exemple de dimension	26
Fig-13- Exemple de modélisation en étoile	26
Fig-14- Exemple de modélisation en flocon	27
Fig-15- Exemple de modélisation en constellation	28
Fig-16- La construction d'un entrepôt de données	28
Fig-17- Le processus d'alimentation ETL	30
Fig-18- Schéma du cycle de vie dimensionnel	32
Fig-19- Création de l'entreprise	37
Fig-20- L'ancienne structure en UNDS de l'entreprise Naftal Spa	37
Fig-21- La nouvelle structure divisionnaire de Naftal Spa	38
Fig-22- L'organigramme général de l'entreprise Naftal Spa	39
Fig-23- Organigramme de la DCSI	42
Fig-24- Schéma de l'environnement de Naftal Spa	45
Fig-25- L'évolution des consommations GPL	49
Fig-26- La représentation sous forme de cube	59
Fig-27- Schéma du modèle dimensionnel de base de l'activité « Etat_Créance »	64
Fig-28- Dimension Temps du schéma activité « Etat_Créance »	65
Fig-29- Dimension Produit du schéma activité « Etat_Créance »	66
Fig-30- Dimension Client du schéma de l'activité « Etat_Créance »	67
Fig-31- Dimension Centre du schéma activité « Etat_Créance »	68
Fig-32- Dimension Type_Echéance du schéma activité « Etat_Créance »	68
Fig-33- Modèle dimensionnel en flocon de l'activité « Etat_Créance »	70
Fig-34- Schéma du modèle dimensionnel de base de l'activité « Chèques_Impayés » ...	71
Fig-35- Modèle dimensionnel en flocon de l'activité « Chèques_Impayés »	73
Fig-36- Schéma du modèle dimensionnel de base de l'activité « Recouvrement »	74
Fig-37- Modèle dimensionnel en flocon de l'activité « Recouvrement »	76
Fig-38- Architecture détaillée du système	77
Fig-39- Concept de fait	82

Fig-40- Concept de dimension	82
Fig-41- Cube MOLAP de l'activité « Etat_Créance »	83
Fig-42- Cube MOLAP de l'activité « Chèques_Impayés »	84
Fig-43- Cube MOLAP de l'activité « Recouvrement »	85
Fig-44- Le processus d'alimentation de notre entrepôt de données	87
Fig-45- La fréquence de chargement des systèmes sources vers l'entrepôt de données ...	88
Fig-46- Les outils d'accès à notre entrepôt de données	89
Fig-47- Disposition physique du système.....	92

Liste des tableaux

Tab-1- Différences entre système de production et data warehouse.....	12
Tab-2- Tableau de synthèse des architectures de data warehouse.....	19
Tab-3- Evolution des consommations GPL.....	49
Tab-4- Les points de ventes.....	49
Tab-5- Liste des tables de l'entrepôt de données.....	79
Tab-6- Liste des attributs de l'entrepôt de données.....	79
Tab-7- Liste des tables des faits choisies avec leurs mesures et dimension d'analyse Sollicitées.....	81

Remerciements

Quelques lignes ne pourront jamais exprimer la reconnaissance que nous éprouvons envers tous ceux qui, de près ou de loin, ont contribué, par leurs conseils, leurs encouragements ou leurs amitiés à l'aboutissement de ce modeste travail.

Nos remerciements vont particulièrement aux gens de Naftal Spa pour leur aide précieuse, à leur tête M^{me} TAJET sans oublier Mr SEDDAR.

Nos vifs remerciements accompagnés de toute notre gratitude à notre promotrice M^{lle} BOUSTIA pour ses précieux conseils, sans elle, il nous aurait été difficile de mener à bien notre projet.

Et enfin, nous tenons à remercier tous ceux qui ont contribué à l'aboutissement et la réussite de ce travail.

Merci

Dédicaces

A mes chers parents qui sont les personnes les plus chères que j'ai dans la vie, pour leur amour et leur soutien.

A mon frère et mes sœurs, pour leurs encouragements. Et à toute ma famille.

A tous les gens de Naftal Spa qui ont contribué à la réalisation de ce travail et à leur tête ma tante M^{me} Tafat.

A mes amis qui ont été toujours là pour m'écouter.

A mon binôme Sofiane.

A tous ceux qui nous ont aidé de prêt ou de loin et à vous.

Je dédie ce modeste travail.

Yacine

Dédicaces

*Aux témoignage d'affection, d'amour et de grande reconnaissance,
aux êtres les plus chers que j'ai dans la vie, qui se sont toujours
sacrifiés et m'ont soutenu avec tout ce qu'ils ont.*

*...Mes parents,
Que Dieu me les garde*

*A ceux qui ont su me consolider durant les moments les plus
difficiles de ma vie, à mes frères et sœurs,*

A tous les miens,

A tous mes amis qui ont toujours partagé avec moi, joies et peines,

A Monsieur Haroun, à qui je tiens à exprimer ma grande gratitude,

*A tous mes enseignants du primaire à l'université, et à leur tête
notre promotrice M^{lle} Boustia pour son aide précieuse et sa patience
avec nous,*

A mon binôme Yacine,

Et à vous,

Je dédie le fruit de mes années d'étude

Sofiane

Résumé

La mise en œuvre d'un vrai système décisionnel est aujourd'hui indispensable pour toute entreprise qui souhaite tirer parti des trésors que recèlent les mines d'informations qui sont dispersées en son sein.

Un système décisionnel efficace intègre les nouvelles technologies pour mieux collecter des données disparates dans toute l'entreprise, de les nettoyer et de les mettre en cohérence et enfin les organiser au regard des axes d'analyse souhaités. Ceci grâce à l'apparition d'un nouveau concept appelé "data warehouse" ou "entrepôt de données".

Notre travail consiste à mettre en œuvre un système décisionnel métier pour pilotage des créances clients au sein de NAFTAL Spa.

L'approche retenue pour la réalisation du système est l'approche de cycle de vie dimensionnel proposée par Ralph Kimball. Notre approche est divisée en trois phases : Identification des besoins, Conception et Déploiement du système.

Abstract

The putting in work of a true system decisional is today indispensable for all enterprise that wishes to pull party of treasures that conceals mines of information that are dispersed in his/her/its breast.

A system efficient decisional integrates the new technologies to collect some disparate data better in all the enterprise, to clean them and to put them in consistency and in short to organize them to the look of axes of analysis wished. This thanks to the apparition of a new concept called "data warehouse".

Our work consists in the creation of a profession's system decisional for piloting of credence's customers within NAFTAL SPA.

The approach kept for the realization of the system is the approach of dimensional life cycle proposed by Ralph Kimball. Our approach is divided in three phases: Identification of needs, Conception and Spreading of the system.

ملخص

إن استخدام نظام قراري أصبح اليوم ضروري لأي مؤسسة تتطلع إلى الاستفادة قدر الإمكان من الثروات الكامنة في الكم الهائل من المعلومات الموزعة بين مختلف مصالحها.

يدمج النظام القراري الفعال، التكنولوجيات الحديثة لتحسين عملية جمع المعلومات غير المتجانسة داخل المؤسسة، ثم تنقيتها، تنسيقها و أخيراً تنظيمها حسب محاور التحليل المنشودة و هذا بفضل ظهور تصور جديد هو مخزن المعلومات (داتا ويرهاوس).

يتمثل عممتنا في وضع نظام قراري عملي للتحكم في ديون الزبائن لمؤسسة نפטال (م.ذ.أ).

الطريقة المنتهجة لتحقيق هذا النظام هي طريقة "دورة الحياة ذات الأبعاد" المقترحة من طرف رالف كيمبل، هذه الطريقة تنقسم إلى ثلاثة أطوار: تشخيص، تصور و عرض النظام.

SOMMAIRE

I. Introduction générale	1
II. Problématique	2
III. Objectifs	3

Chapitre I : Les systèmes décisionnels et les entrepôts de données

Introduction	4
I. Le système d'information décisionnel	5
1. Introduction	5
2. Définitions	5
3. Buts d'un SIAD	7
4. Les apports du SIAD	7
4.1. Apport du SIAD à la gestion	7
4.2. Apport du SIAD à l'analyse	8
II. Les entrepôts de données	8
1. Introduction	8
2. Définitions	9
3. Objectifs du Data Warehouse	13
4. Les concepts de base d'un entrepôt de données	14
4.1. La composition du Data Warehouse	14
4.2. La structure d'un Data Warehouse	15
4.3. Les architectures d'implémentation de l'entrepôt de données	18
5. La conception d'un Data Warehouse	19
6. Le concept OLAP	21
6.1. À propos d'OLAP	21
6.2. Systèmes OLTP versus systèmes OLAP	21
6.3. Les différents outils OLAP	23
6.3.1. Les outils MOLAP	23
6.3.2. Les outils ROLAP	23
7. La modélisation des données	24
8. Construction d'un entrepôt de données	28
9. Le processus d'alimentation ETL	30
10. La fonction de pilotage et de contrôle de gestion	31
11. Méthodologie adoptée	32
Conclusion	35

Chapitre II : Présentation générale de Naftal Spa

Introduction	36
1. Présentation générale	36
1.1. Historique	36
1.2. Organisation	36
2. Les moyens humains & infrastructures de Naftal	44
3. L'environnement de l'entreprise	45

3.1. Les produits.....	45
3.2. La clientèle.....	49
3.3. Fournisseurs.....	50
3.4. La concurrence.....	50
4. La politique interne de l'entreprise.....	51
4.1. La stratégie adoptée par Naftal Spa.....	51
4.2. Les projets de Naftal Spa.....	52
Conclusion.....	53

Chapitre III : Identification des besoins

Introduction.....	54
1. La gestion des créances de Naftal Spa.....	54
2. Définition des besoins.....	55
2.1. Les documents.....	55
2.2. Les entretiens.....	55
2.3. Les besoins recensés.....	56
3. Le processus métier.....	57
3.1. Les activités clés.....	57
3.2. Les indicateurs d'analyse des activités métiers.....	57
3.3. La dimensionnalité métier.....	58
3.4. Les systèmes sources.....	59
Conclusion.....	60

Chapitre IV Conception et mise en œuvre du système

Introduction.....	61
1. La démarche de la conception du Data Warehouse.....	61
1.1. Choisir les processus d'activité à modéliser.....	61
1.2. Choisir le grain de chaque table de faits.....	61
1.3. Choisir les dimensions de chaque table de faits.....	61
1.4. Choisir les faits mesurés que contiendra chaque enregistrement de table de fait.....	62
1.5. Les attributs des dimensions.....	62
1.6. Comment suivre les dimensions à évolution lente.....	62
1.7. Les agrégats.....	63
1.8. L'étendue historique de la base de données.....	63
1.9. L'urgence avec laquelle les données doivent être extraites et chargées dans l'entrepôt de données.....	63
2. La modélisation dimensionnelle des activités métiers du pilotage des créances.....	63
2.1. Modélisation dimensionnelle de l'activité « Etat_Créance ».....	63
2.2. Modélisation dimensionnelle de l'activité « Chèques_Impayés ».....	71
2.3. Modélisation dimensionnelle de l'activité « Recouvrement ».....	74
3. Mise en œuvre de l'architecture.....	77
3.1. Construction de l'entrepôt de données.....	78
3.2. Construction des cubes OLAP.....	81
3.3. Construction de la zone d'alimentation.....	86

3.4. L'exploitation du datawarehouse « Le portail de restitution ».....	88
Conclusion.....	91

Chapitre V : Déploiement du système

Introduction.....	92
1. Architecture informatique du système.....	92
2. La disposition physique du système.....	92
3. Les utilisateurs du système.....	93
4. logiciels utilisés.....	93
5. Matériel utilisé.....	93
6. La sécurité du système.....	94
7. Présentation de l'application.....	95
Conclusion générale	100
Glossaire	
Bibliographie	

*Introduction et
problématique*

I. Introduction générale :

Face à la mondialisation et à la concurrence grandissante, la prise de décision est devenue cruciale pour les dirigeants d'entreprises. L'efficacité de cette prise de décision repose sur la mise à disposition d'informations pertinentes et outils adaptés.

Et avec la généralisation de l'informatique dans tous les secteurs d'activité, les entreprises produisent et manipulent de très importants volumes de données électroniques. Ces données sont stockées dans les systèmes opérationnels qui conviennent bien aux applications gérant l'activité quotidienne de l'entreprise, mais s'avèrent inadaptés au décisionnel.

Le problème des entreprises est d'exploiter efficacement d'importants volumes d'informations, provenant soit de leurs systèmes opérationnels, soit de leur environnement extérieur, pour supporter la prise de décision. Afin de pallier cet inconvénient, des systèmes décisionnels ont été développés. Ces systèmes reposent sur un espace de stockage centralisé, appelé entrepôt de données (*Data Warehouse*).

La réussite de l'entreprise est souvent mesurée par le nombre de ses clients et leurs satisfactions de ses services et produits. Au regard du contexte actuel, le suivi des entreprises est menacé par les changements de son environnement. Les systèmes décisionnels constituent une réponse appropriée pour faire face aux problèmes de pilotage des activités des entreprises.

Dans ce cadre, l'entreprise NAFTAL SPA nous a sollicité pour la conception et le déploiement d'un système décisionnel pour le pilotage des créances de ses clients.

II. Problématique :

NAFTAL SPA est une entreprise nationale chargée de la commercialisation et de la distribution des produits pétroliers à l'échelle nationale. Elle se trouve confrontée à des problèmes pour la gestion des créances de ses clients.

Par nature, la gestion des clients impose des structures de données complexes. Le dossier client regroupe un ensemble d'enregistrements variés et hétérogènes : il comporte des données administratives, des données sur le déploiement des paiements, ...

Enfin, l'importance des volumes de données mis en jeu dans les systèmes d'aide à la décision nécessite des mécanismes d'agrégation pour synthétiser l'information. Ces mécanismes sont d'autant plus importants qu'il est crucial d'offrir des moyens pour archiver les données temporelles à des niveaux de détail en adéquation avec les besoins des décideurs.

Alors, on peut dire que les principales difficultés auxquelles NAFTAL fait face sont :

- Complexité et diversité de la structure conceptuelle existante rendant l'intégration des données difficile pour une vue multiaxes.
- Manque de schéma global décisionnel.
- Difficulté de garder l'historique des mouvements, même sur une période de quelques mois.
- Absence de moyens d'analyse de l'activité créance (pas d'indicateurs de mesure).

III. Objectifs :

L'entreprise NAFTAL SPA souhaite simplifier le pilotage des créances de ses clients de différents types. Pour cela, elle a adopté une démarche de développement des processus de management des systèmes d'informations.

Notre objectif est de proposer un système décisionnel pour aider les gestionnaires et les décideurs de NAFTAL pour le pilotage de ses créances par :

- La spécification des besoins en matière d'analyse et de pilotage des créances, pour l'amélioration des performances décisionnelles de l'entreprise.
- Le déploiement d'un entrepôt de donnée « DATA WAREHOUSE », qui permettra aux utilisateurs une navigation libre des données.
- La proposition d'un modèle conceptuel d'analyse.
- Proposition d'une architecture de chargement des données.

Chapitre I :

*Les systèmes décisionnels
et
les entrepôts de données*

Introduction :

Jusqu'à la mise en place des solutions d'informatique décisionnelle, les bases de données étaient des solutions de sauvegardes des informations dont le principal atout était le gain de place : un serveur de base de données peut contenir l'équivalent en volume d'informations de très nombreux classeurs. L'accès aux données était possible, mais réservé aux informaticiens.

En effet, il s'agissait de base de données relationnelle, et il était indispensable de connaître la structure de la base pour s'y repérer, ce qui, lorsqu'il y a plusieurs bases contenant chacune plus d'une centaine de tables, peut réellement être difficile. Or, les marchés évoluant de plus en plus vite, il est devenu indispensable de trouver une solution permettant la prise de décisions rapide, c'est à dire une solution qui permet l'utilisation simplifiée des données, des informations.

Cette solution va consister à séparer le système opérationnel du décisionnel. Plus concrètement, le système d'information de l'entreprise sera constitué de deux pôles. :

- Un premier pôle, opérationnel, qui représentera la partie production de l'entreprise (ventes, commandes, gestion de stocks, données comptables, productions,...) et dont le rôle principal, au niveau informatique sera la saisie des données.
- Un second pôle, décisionnel celui-là, qui consistera à rapatrier les données saisies au niveau opérationnel, à les analyser et à les traiter afin de les rendre accessibles pour le personnel chargé de prendre les décisions. C'est dans le cadre de ce second pôle que la Business Intelligence intervient. L'accès aux données se fait grâce à des rapports, indicateurs, graphiques,... qui permettent d'avoir facilement et en temps réel la situation de l'entreprise, ainsi que des estimations sur l'avenir.

Dans ce chapitre nous allons définir qu'est ce qu'un système décisionnel, ses objectifs et ses apports dans le domaine de la prise de décision, nous allons voir aussi, les différences essentielles entre un système de production et de décision.

Et enfin, ce chapitre se termine par une présentation de la méthodologie adoptée pour la réalisation de notre système décisionnel.

I. Le système d'information décisionnel :

1. Introduction :

L'information de gestion a gagné sa place dans l'entreprise depuis les années 60 par une succession de progrès technologiques, logiciels et méthodologiques qui ont tous contribué à une réduction des coûts d'exploitation. Les bases de données relationnelles, les progiciels de gestion, ainsi que les premiers micro-ordinateurs des années 80 ont largement contribué à l'équipement des petites et moyennes entreprises, commerces, administrations. Jusque là, la plus grande partie des applications était dédiée au traitement des données directement liées à l'activité quotidienne des organisations : paie, comptabilité, commandes, facturation.

Avec l'apparition des ordinateurs personnels et des réseaux locaux, une autre activité a émergé, tout à fait distinct de l'informatique de production. Aux années 90, avec la montée en puissance des micro-ordinateurs et l'avènement de l'architecture client – serveur, on observe aujourd'hui un décloisonnement remarquable. Le mot d'ordre principal est : « Fournir à tout utilisateur reconnu et autorisé, les informations nécessaires à son travail ».

Ce slogan fait naître une nouvelle informatique, intégrante, orientée vers les utilisateurs et les centres de décision des organisations.

L'informatique décisionnelle s'est développée dans les années 70. Elle est alors essentiellement constituée d'outils d'édition de rapports, de statistiques, de simulation et d'optimisation.

Dans les années 90, deux phénomènes se produisent simultanément. Premièrement, il est possible de concevoir des environnements spécialisés pour l'aide à la décision. Deuxièmement, de nombreux algorithmes permettant d'extraire des informations à partir des données brutes sont arrivés à maturité. Certains de ces algorithmes sont issus des statistiques ; d'autres proviennent des recherches en Intelligence Artificielle. Ces algorithmes sont regroupés dans des logiciels de fouille de données ou Data Mining qui permettent la recherche d'informations nouvelles ou cachées à partir de données. Ainsi, dans le cas de systèmes à base de règles, plutôt que d'essayer d'extraire la connaissance d'experts et d'exprimer cette connaissance sous forme de règles, un logiciel va générer ces règles à partir de données.

2. Définitions :

2.1. L'aide à la décision :

La décision est souvent vue comme le fait d'un individu isolé (« le décideur ») exerçant un choix entre plusieurs possibilités d'actions à un moment donné.

Roy définit l'aide à la décision comme étant « l'activité de celui qui, prenant appui sur des modèles clairement explicités mais non nécessairement complètement formalisés, aide à obtenir des éléments de réponse aux questions que se pose un acteur dans un processus de décision, éléments concourant à éclairer la décision et normalement à recommander, ou simplement à favoriser, un comportement de

nature à accroître la cohérence entre l'évolution du processus d'une part, les objectifs et le système de valeurs au service desquels ces acteurs se trouvent placés d'autre part » [L13].

2.2. Système d'aide à la décision :

- Un SIAD (Système Interactif d'Aide à la Décision) est un outil d'analyse et de modélisation des données de l'entreprise qui permet de créer des représentations multidimensionnelles de l'information. Historiquement, il s'agit d'une terminologie et d'outils utilisés avant l'avènement et la maturité de Data Warehouse [L12].
- J. Johnson et J-C. Courbon définissent un SIAD comme un système homme – machine permettant à travers un dialogue une amplification du raisonnement humain lors de résolution d'un problème faiblement structuré [A1].
- P. Levine et J-C. Pomerole le définissent comme un système qui sélectionne, interprète et densifie l'information pour préparer la décision. Un SIAD correspond à un système d'information spécialement conçu pour traiter l'information impliquée dans un processus de décision [W04].
- Système Informatique Intégré, conçu spécialement pour la prise de décision, et qui est destiné particulièrement aux dirigeants d'entreprise. Le système d'aide à la décision est un des éléments du système d'information de gestion. Il se distingue du système d'information pour dirigeants, dans la mesure où sa fonction première est de fournir non seulement l'information, mais les outils d'analyse nécessaires à la prise de décision. Ainsi, il est habituellement constitué de programmes, d'une ou de plusieurs bases de données, internes ou externes, et d'une base de connaissances. Il fonctionne avec un langage et un programme de modélisation qui, permettent aux dirigeants d'étudier différentes hypothèses en matière de planification et d'en évaluer les conséquences [L02].
- La définition fonctionnelle proposée par Keen et Morton est la définition de référence pour le concept de SIAD : « Un SIAD est un système ouvert qui augmente la qualité des processus de décision plutôt que leur productivité » [L12]. Comme pour Roy et Bouyssou, « Les SIAD mettent l'accent sur le processus de recherche de solutions » [L13]. Ces systèmes n'ont pas pour objectif d'identifier une décision optimale qu'il 'suffirait' ensuite de mettre en application. Leur finalité consiste à « amplifier le raisonnement du décideur sur la base de ses propres représentations » [A1]. Ces systèmes ont une place privilégiée lorsque les problèmes sont peu, voire non structurés.

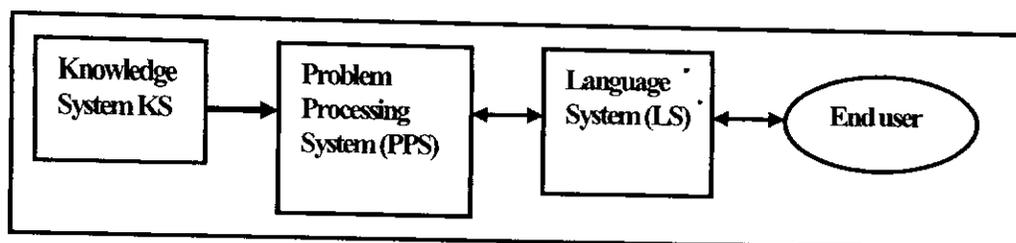


Fig-1- Structure d'un système d'aide à la décision (W02)

3. Buts d'un SIAD :

Le SIAD est un outil d'observation et de description qui vise, à partir de données de gestion et/ou de statistiques, à donner aux managers d'une entreprise les moyens d'identifier des alertes de gestion, de suivre l'évolution de l'activité et de disposer d'outils d'investigation de sujets ou phénomènes particuliers. Il ne fournit pas les explications ni les commentaires qui relèvent d'une phase de travail postérieure à l'observation [L05].

Le SIAD tire parti de l'ensemble des données produites ou acquises par l'entreprise, ensemble dont il fournit une présentation synthétique. Cela suppose (1) que le SIAD soit alimenté potentiellement par toutes les applications de l'entreprise, (2) qu'il résolve les problèmes de comparabilité et de redressement des données qui se posent inévitablement lorsque l'on utilise des sources diverses.

Le SIAD vise à présenter des informations utiles. Ceci implique qu'il soit construit selon des critères de sélectivité en choisissant, parmi toutes les statistiques qu'il est possible de produire, celles qui peuvent servir à telle ou telle catégorie d'utilisateurs. Sa construction suppose donc une analyse des besoins, elle-même fondée sur une segmentation des utilisateurs en sous – populations homogènes chacune en ce qui concerne les missions à remplir et les besoins correspondants [L11].

Le SIAD vise à fournir aux utilisateurs un outil de consultation commode, d'une ergonomie aisée, de façon à minimiser les tâches de recherche de l'information et de présentation des résultats.

Produire des statistiques en adressant au coup par coup des requêtes à une application opérationnelle est coûteux en traitement. Le SIAD protège donc les bases de données opérationnelles en s'intercalant comme un tampon entre elles et les utilisateurs et en préparant la plupart des statistiques dont ces derniers ont besoin.

4. Les apports du SIAD :

4.1. Apport du SIAD à la gestion :

Le SIAD a pour but de fournir des données observées alimentant, après recoupement avec d'autres sources (économiques, démographiques, marketing etc.) la compréhension du marché et permettant de réaliser le suivi de l'activité, l'analyse de son impact, l'optimisation des moyens, de façon à faciliter l'orientation de l'action.

Le SIAD a donc vocation à fournir les indicateurs de pilotage permettant à un responsable opérationnel d'évaluer la qualité et productivité du travail fourni par des établissements ou des équipes, indicateurs qui impliquent un recoupement avec des données que le SIAD ne comporte pas (volume et qualité des ressources employées, délai de traitement des affaires, etc.)

Le SIAD n'a, par contre, pas vocation à fournir des indicateurs pour un pilotage opérationnel au jour le jour ou pour le suivi de dossiers individuels. Il faut que chaque application soit munie des outils permettant aux responsables opérationnels de

piloter par domaine les travaux au plus près de leur réalisation. Cependant le SIAD peut contribuer à l'alimentation de ces outils : un responsable peut trouver, dans les hypercubes produits par le SIAD, telle série chronologique qu'il recoupera avec des données de gestion pour évaluer l'efficacité du travail de son unité [L03].

4.2. Apport du SIAD à l'analyse :

Certains représentent l'architecture du SIAD (on dit aussi " Data Warehouse") par une pyramide. Sa large base est constituée des diverses applications qui l'alimentent, le sommet par les hypercubes et autres outils d'observations synthétiques.

Un SIAD est bâti à partir des données d'observations, et il faut distinguer l'observation de l'explication. C'est en complétant le SIAD par des outils d'analyse des données et d'économétrie, et en le confrontant aux modèles explicatifs, que l'on pourra l'utiliser pour comprendre ce qui se passe sur le marché. Le SIAD alimente ces outils mais ne les comporte pas [L08].

Il est donc utile de représenter les opérations éditoriales s'appuyant sur le SIAD ; utilisant les données d'observation synthétiques, elles permettent de produire des résultats interprétés et commentés destinés à diverses populations d'utilisateurs (responsables régionaux, responsables de ligne de produit etc.)

S'il est souvent nécessaire pour l'interprétation d'utiliser les méthodes de l'analyse des données ou de l'économétrie, il est recommandé de rien laisser paraître de ces démarches techniques dans la publication qui ne doit en recueillir que les résultats (il ne convient pas en effet de laisser les échafaudages en place après la construction d'un immeuble).

Le SIAD est un puissant outil d'observation ; il appelle donc un dépassement nécessaire vers l'explication et le commentaire. Ce dépassement implique, pour pouvoir servir les diverses populations d'utilisateurs concernées, une diversification éditoriale.

II. Les entrepôts de données :

1. Introduction :

C'est dans les années 1990 que les entreprises comprennent que les données sont non seulement utiles dans le cadre d'une utilisation opérationnelle, mais qu'elles peuvent leur trouver une utilisation stratégique. Les systèmes opérationnels ont été créés pour répondre aux besoins de traitements de transaction en ligne (OLTP) et le traitement en batch. Au contraire, l'utilisation stratégique de données est caractérisée par le traitement de requêtes en ligne ou les fonctions de batch pour l'aide à la décision (OLAP).

Les entreprises recherchent toujours à obtenir un avantage concurrentiel, notamment au travers d'un SI informationnel performant. L'efficacité ou productivité n'est plus la seule clé de la réussite d'une entreprise ; cette dernière se doit d'être flexible, rapide et surtout réactive.

Le Data Warehouse est arrivé comme une reconnaissance de la valeur et du rôle de l'information, à savoir son utilisation à des fins stratégiques. Un Data Warehouse est une plateforme avec des données de très bonne qualité qui supportent des applications de type informationnel, comme les applications d'aide à la décision, ainsi que les processus dans l'entreprise. Le Data Warehousing augmente la productivité des décideurs d'entreprise en consolidant, convertissant, transformant et intégrant les données de différents systèmes opérationnels et en leur fournissant une vue consistante, globale et unifiée de leur entreprise.

En bref, le Data Warehousing est une alliance de technologies dont le but est d'intégrer les données des systèmes opérationnels dans un environnement qui permet l'utilisation stratégique de ces données. Ces technologies comprennent les systèmes de gestion de bases de données relationnelles et multidimensionnelles, l'architecture client – serveur, la modélisation de métas données, d'interfaces utilisateurs graphiques et bien d'autres encore.

2. Définitions :

Plusieurs auteurs ont proposé différentes définitions pour l'entrepôt de donnée ou le « Data Warehouse », mais toutes ces définitions se basent sur les mêmes principes.

- **Définition 1** : 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.
- **Définition 2** : « L'entrepôt de données est un lieu de stockage centralisé d'un extrait des bases de production. Cet extrait concerne les données pertinentes pour le support à la décision. Elles sont intégrées et historisées. L'organisation des données est faite selon un modèle qui facilite la gestion efficace des données et leur historisation » [L08].
- **Définition 3** : Un Data Warehouse est un entrepôt de données. Il s'agit d'un stockage intermédiaire des données issues des applications de productions, dans lesquelles les utilisateurs finaux puisent avec des outils de restitution et d'analyse.

« Un Data Warehouse est une collection de données thématiques, intégrées, non volatiles et historisées organisées pour la prise de décision » [L04].

- **Définition 4** : Une définition classique du Data Warehouse est donnée par William H. Inmon, dans son livre « Building the Data Warehouse » :

« Le Data Warehouse est une collection de données orientées sujet, intégrées, non volatiles et historisées, organisées pour le processus de décision » [L09].

* Orientées sujet

Le Data Warehouse est organisé autour des sujets majeurs de l'entreprise, contrairement aux données des systèmes de production. Ceux-ci sont généralement organisés par processus fonctionnels. Les données sont structurées par thème.

L'intérêt de cette organisation est de disposer de l'ensemble des informations utiles sur un sujet le plus souvent transversal aux structures fonctionnelles et organisationnelles de l'entreprise.

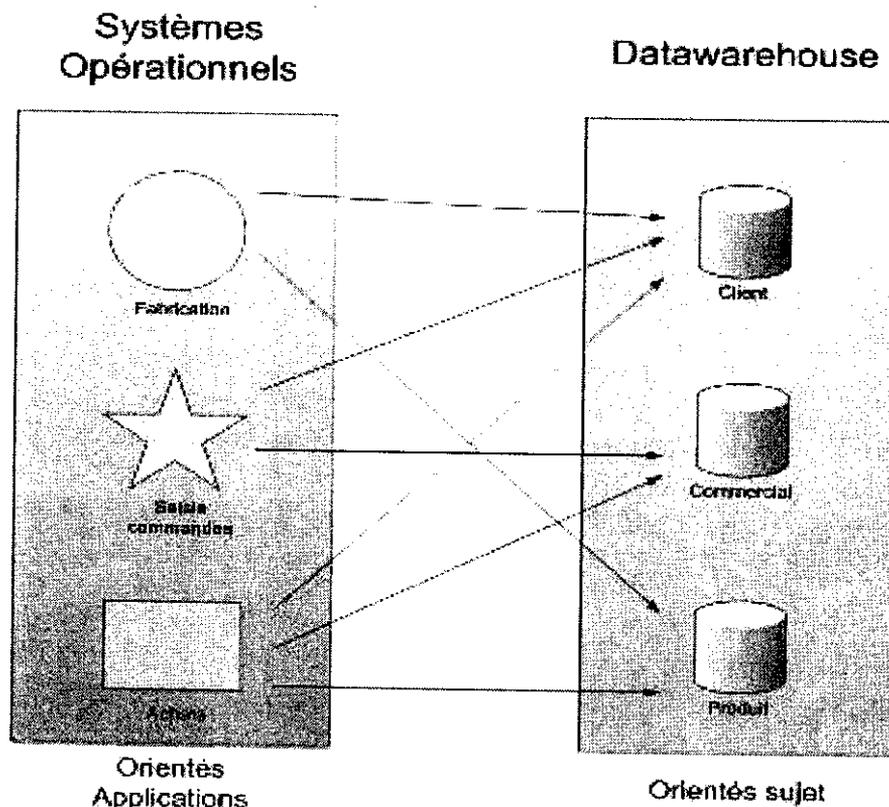


Fig-2- Organisation orienté sujet du data warehouse (W01)

* Données intégrées

Les données alimentant le Data Warehouse proviennent de multiples applications sources hétérogènes. Les données des systèmes de production doivent être converties, reformatées et nettoyées, de façon à avoir une seule vision globale dans le Data Warehouse.

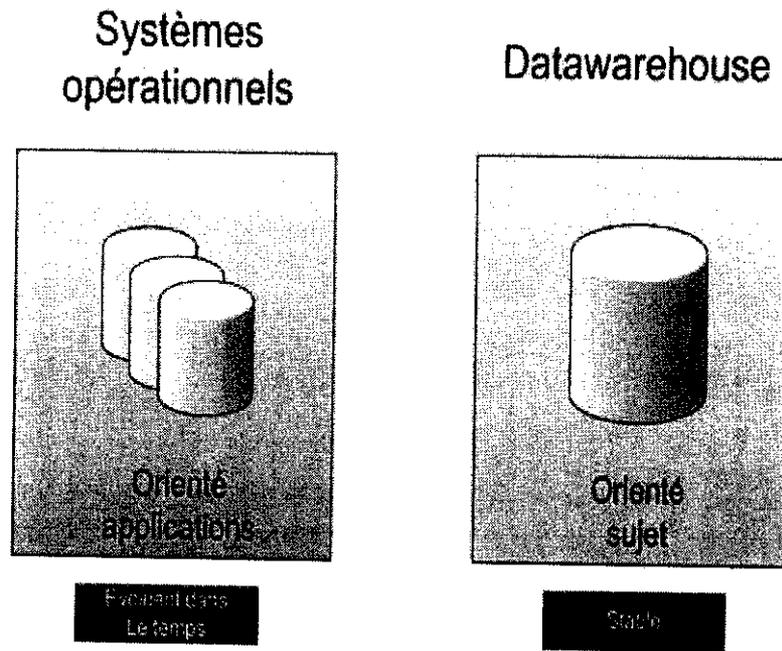


Fig-3- Intégrité des données du data warehouse (W01)

✿ **Données historisées**

Dans un système de production ; la donnée est mise à jour à chaque nouvelle transaction. Dans un Data Warehouse, la donnée ne doit jamais être mise à jour. Un référentiel temps doit être associé à la donnée afin d'être capable d'identifier une valeur particulière dans le temps.

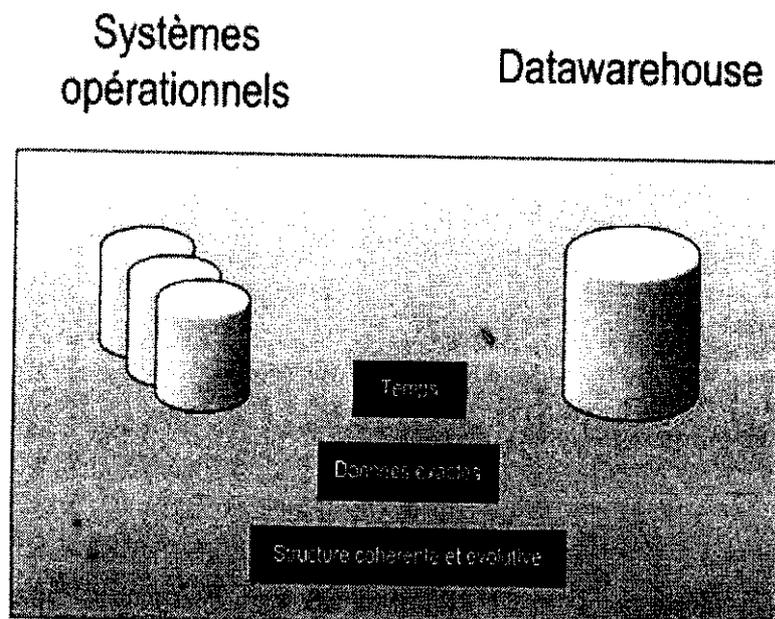


Fig-4- Historisation des données dans le data warehouse (W01)

✱ **Données non volatiles**

La non volatilité des données est en quelque sorte une conséquence de l'historisation. Une même requête effectuée à quelques mois d'intervalle en précisant la date de référence de l'information recherchée donnera le même résultat.

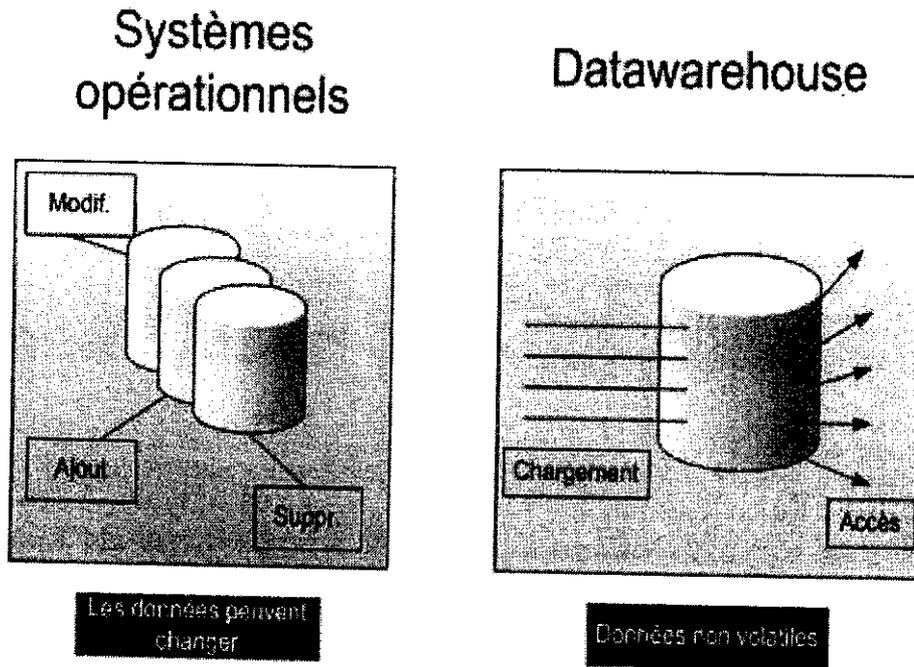


Fig-5- Non-Volatilité des données dans le data warehouse (W01)

Le tableau ci-après présente les principales différences entre le système de production et le Data Warehouse :

Critère	Système de production	Data Warehouse
Niveau de détail des informations utilisateurs	Très détaillé	Synthétique, parfois détaillé
Utilisateurs	Une ou quelques fonctions de l'entreprise	Plusieurs fonctions de l'entreprise
Données figées	Non – évolution en temps réel	Oui – archivage
Historique	Non	Oui
Opérations sur les données	Ajout/mise à jour/consultation	Consultation uniquement

Tab -1- : différences entre système de production et data warehouse

3. Objectifs du Data Warehouse :

L'entrepôt de données est lieu où les personnes peuvent accéder aux informations. Les objectifs fondamentaux d'un entrepôt de données peuvent être cernés en se promenant dans n'importe quelle grande organisation et en écoutant les conversations des dirigeants. Les thèmes récurrents sont à peu près les suivants [L01] :

- « Nous avons des montagnes d'informations dans cette société mais nous ne pouvons pas y accéder »
- « Rien ne met plus en fureur la direction générale que de voir deux personnes présenter les résultats de la même activité avec des nombres différents »
- « Nous voulons pouvoir découper les données en tranches et en dés de toutes les manières possibles »
- « Montrez moi seulement ce qui est important »

Ces préoccupations sont tellement universelles qu'elles représentent les exigences fondamentales d'un entrepôt de données. De ces problèmes, on peut tirer les principaux objectifs d'un entrepôt de données qui sont :

- ✓ **Accès aux informations de l'entreprise :** L'entrepôt de données assure l'accès aux informations de l'entreprise et de l'organisation
- ✓ **Les informations d'un entrepôt de données sont cohérentes :** Cela signifie que les requêtes faites à des moments différents doivent fournir les mêmes résultats. La cohérence implique que les données sont chargées dans leur totalité
- ✓ **Les outils de présentation d'informations font partie du Data Warehouse :** L'entrepôt de données ne comporte pas seulement des données, mais aussi un ensemble de requêtes, d'analyses et de présentations des informations.
- ✓ **Les données publiées sont stockées dans le Data Warehouse :** L'entrepôt de données est le lieu où sont publiées des données qui ont déjà servi
- ✓ **Elles sont nettoyées :** Leur qualité est vérifiée et elles ne sont diffusées que si elles sont prêtes à être utilisées
- ✓ **Qualité de l'information d'un Data Warehouse :** La qualité de l'information d'un entrepôt de données est très importante. En effet, comment obtenir des analyses fiables si les données brutes sont de mauvaise qualité

4. Les concepts de base d'un entrepôt de données :

4.1. La composition du Data Warehouse :

Dans ce paragraphe, on présente les composants de base du Data Warehouse, ainsi que les principales notions propres au data Warehousing.

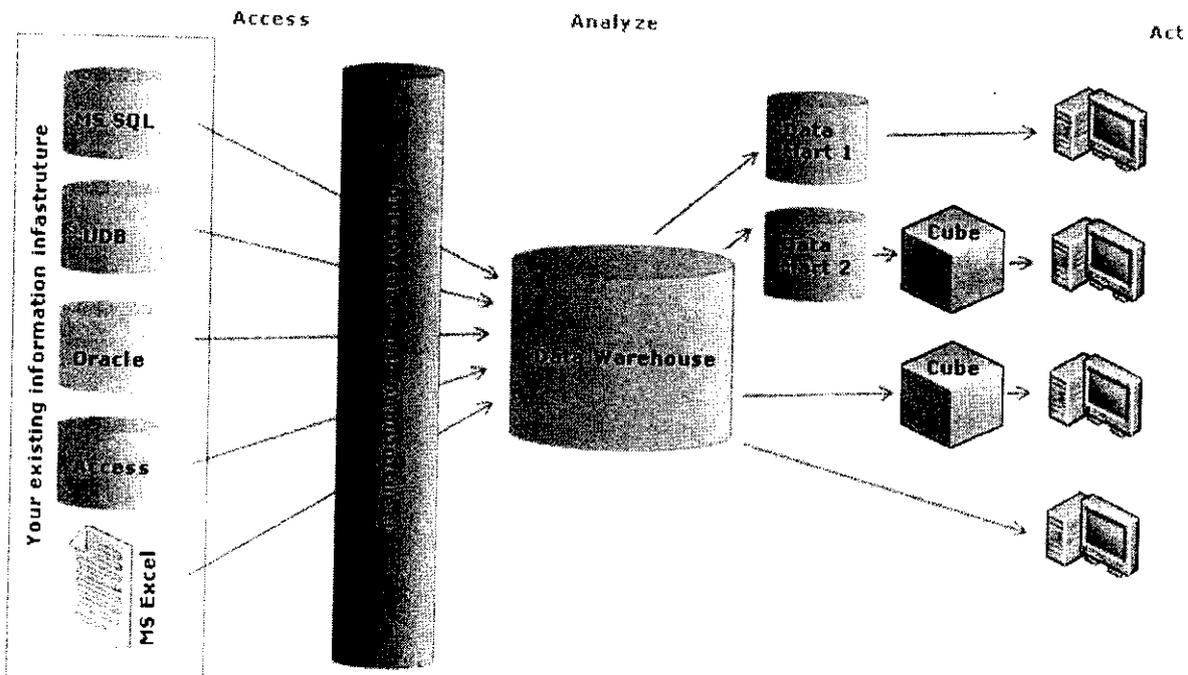


Fig-6- Les composants d'un Data Warehouse (W03)

4.1.1. Le système source :

Système opérationnel d'enregistrement, dont la fonction consiste à capturer les transactions liées à l'activité de l'entreprise. Il s'agit souvent de ce que l'on appelle les applications de gestion. La principale priorité du système source est le temps de disponibilité. Les systèmes sources, de par leur nature, « vivent leur vie ». Généralement, peu d'effort a été consenti pour homogénéiser les dimensions de base telles que le produit, le client, le lieu ou le calendrier par rapport aux autres SI opérationnels de l'entreprise.

4.1.2. La zone de préparation des données :

Cette zone comprend tout ce qui se trouve entre les systèmes sources et le serveur de présentation. Elle regroupe l'ensemble des processus qui nettoient, transforment, combinent, archivent, suppriment les doublons. Elle prépare les données sources en vue de leur intégration et de leur exploitation dans le Data Warehouse. La frontière qui détermine la zone de préparation des données est que cette dernière ne doit en aucun cas être accessible à l'utilisateur final par requêtes ou par un quelconque autre service de présentation.

4.1.3. Le serveur de présentation :

Ce composant correspond à la machine cible sur laquelle l'entrepôt de données est stocké et organisé pour répondre en accès direct aux requêtes provenant des utilisateurs, des générateurs d'états ou d'autres applications. Sur le serveur de présentation, les données seront stockées et présentées sous forme dimensionnelle, de façon à faciliter l'accès pour l'utilisateur final. Dans la majorité des cas, le serveur est basé sur une base de données relationnelle, de sorte que les tables y sont organisées sous forme de schémas en étoile.

4.1.4. Le data mart :

D'après R.Kimball, le data mart est défini comme un sous – ensemble logique d'un data warehouse. Il représente un projet réalisable. On considère souvent le data mart comme la réduction de l'entrepôt de données à un seul processus ou à un groupe de processus ciblant un groupe métier spécifique [L01].

Les data marts sont basés sur des données détaillées et peuvent contenir des agrégats de données destinées à optimiser les performances, contrairement à la définition de Bill Inmon, qui conçoit le data mart comme un ensemble de données agrégées [W11].

4.1.5. L'entrepôt de données :

Le Data Warehouse correspond à la source de données interrogeable de l'entreprise. Selon la définition du data mart, on peut voir le Data Warehouse comme l'union de tous les data mart qui le composent. L'entrepôt de données est alimenté par la zone de préparation des données.

4.1.6. Portail de restitution :

C'est la partie publique de l'entrepôt de données. Il représente ce que voient les utilisateurs, les outils avec lesquels ils travaillent. Sous ce terme sont regroupées toutes les applications qui s'appuient sur les données du Data Warehouse pour les restituer soit à l'utilisateur, soit à une autre application. Les services offerts par le portail de restitution sont les services d'accès aux données, les applications de modélisation et de data mining.

4.1.7. Méta données :

Ceux sont toutes les informations de l'environnement du Data Warehouse qui ne constituent pas les données proprement dites. Ce sont les « données sur les données »

4.2. La structure d'un Data Warehouse:

Un Data Warehouse se structure en quatre classes de données organisées selon un axe historique et un axe synthétique.

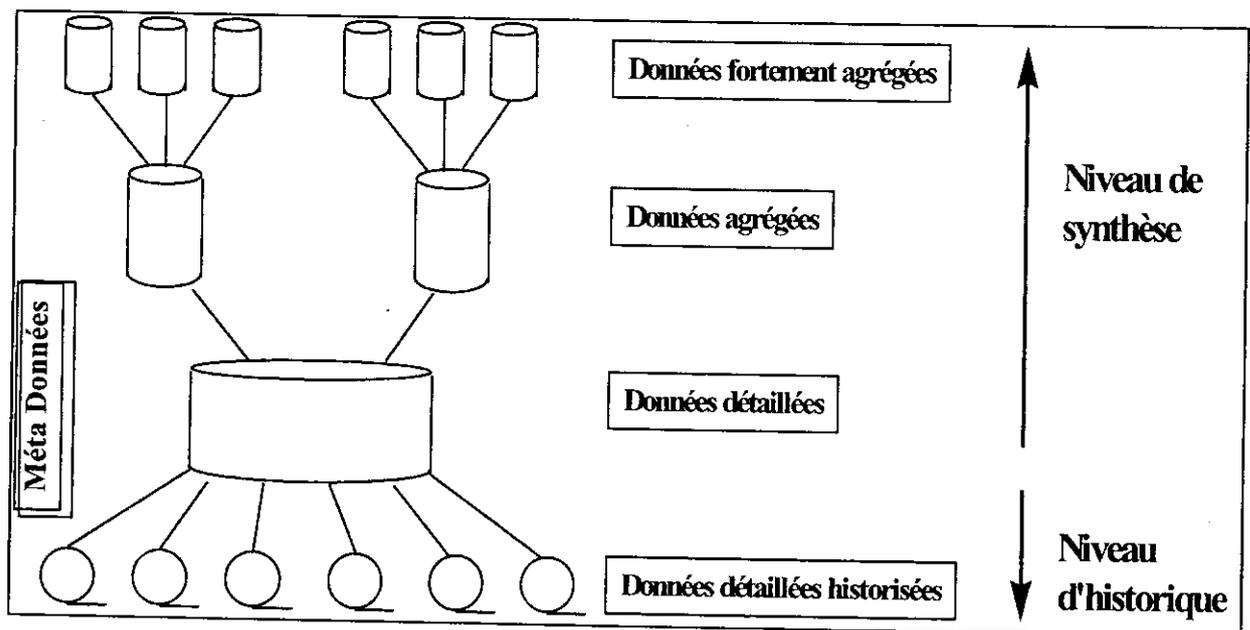


Fig -7 - Structure d'un data warehouse (W01)

✿ **Les données détaillées**

Elles reflètent les événements les plus récents. Les intégrations régulières des données issues des systèmes de production vont habituellement être réalisées à ce niveau.

Les volumes à traiter sont plus importants que ceux gérés en transactionnel et le niveau de détail géré dans le Data Warehouse n'est pas forcément identique au niveau de détail géré dans les systèmes opérationnels. La donnée insérée dans le Data Warehouse peut être déjà une agrégation ou une simplification d'informations tirées du système de production.

✿ **Les données agrégées**

Elles correspondent à des éléments d'analyse représentatifs des besoins utilisateurs. Elles constituent déjà un résultat d'analyse et une synthèse de l'information contenue dans le système décisionnel, et doivent être facilement accessibles et compréhensibles. La facilité d'accès est apportée par des structures multidimensionnelles qui permettent aux utilisateurs de naviguer dans les données suivant une logique intuitive, avec des performances optimales.

* Les méta – données

Elles regroupent l'ensemble des informations concernant le Data Warehouse et les processus associés. Elles constituent une véritable aide en ligne permettant de connaître l'information contenue dans le Data Warehouse. Elles sont idéalement intégrées dans un référentiel.

Les principales informations sont destinées :

- A l'utilisateur (sémantique, localisation).
- Aux équipes responsables des processus de transformation des données du système de production vers le Data Warehouse (localisation dans les systèmes de production, description des règles, processus de transformation).
- Aux équipes responsables des processus de création des données agrégées à partir des données détaillées.
- Aux équipes d'administration de la base de données (structure de la base implémentant le Data Warehouse).
- Aux équipes de production (procédures de changement, historique de mise à jour,...)

* Les données historisées

Un des objectifs du Data Warehouse est de conserver en ligne les données historisées. Chaque nouvelle insertion de données provenant du système de production ne détruit pas les anciennes valeurs, mais créer un nouvelle occurrence de la donnée. Le support de stockage dépend du volume des données, de la fréquence d'accès, du type d'accès. Les supports les plus couramment utilisés sont les disques, les disques optiques numériques, les cassettes.

La logique d'accès aux données la plus utilisée est la suivante : les utilisateurs commencent à attaquer les données par le niveau le plus agrégé, puis approfondissent leur recherche vers les données les plus détaillées (Drill Down).

L'accès des données se fait également directement par les données détaillées et historisées, ce qui conduit à des brassages de données lourds, demandant des machines très puissantes.

Le Data Warehouse est une réussite dans une entreprise lorsque le nombre d'utilisateur accédant aux données de détail augmente.

4.3. Les architectures d'implémentation de l'entrepôt de données :

Pour implémenter un Data Warehouse, trois types d'architectures sont possibles :

4.3.1. L'architecture réelle :

Elle est généralement l'architecture retenue pour les systèmes décisionnels. Le stockage des données est réalisé dans un SGBD séparé du système de production. Ce SGBD est alimenté par des extractions périodiques.

Avant le chargement, les données subissent d'important processus d'intégration, de nettoyage et de transformation. L'avantage de cette solution est de disposer de données préparées pour les besoins de la décision et répondant bien aux objectifs du Data Warehouse.

La principale raison justifiant l'architecture réelle est l'inadaptation des données de production aux besoins des systèmes décisionnels. De plus la signification des données dans un système de production peut être ambiguë ; elle peut dépendre de l'application utilisant la donnée.

4.3.2. L'architecture virtuelle :

Cette architecture n'est pratiquement pas utilisée pour le Data Warehouse. Les données résident dans le système de production. Elles sont rendues visibles par des produits middleware ou par des passerelles.

Il en résulte deux avantages : pas de coût de stockage supplémentaire et l'accès se fait en temps réel. L'inconvénient est que les données ne sont pas préparées.

4.3.3. L'architecture remote :

C'est une combinaison de l'architecture réelle et de l'architecture virtuelle. Elle est rarement utilisée.

L'objectif est d'implémenter physiquement les niveaux agrégés afin d'en faciliter l'accès et de garder le niveau de détail dans le système de production en y donnant l'accès par le biais de middleware ou de passerelle.

4.3.4. Synthèse :

Les différents éléments d'appréciation sont repris dans le tableau récapitulatif suivant :

	Architecture réelle	Architecture virtuelle	Architecture remote
Utilisation	Retenue pour les systèmes décisionnels	Rarement utilisée	Rarement utilisée
Stockage	SGBD séparé du système de production, alimenté par des extractions périodiques	Données résidant dans le système de production	Combinaison des architectures réelle et virtuelle
Avantages	Données préparées pour les besoins de la décision	Pas de coût de stockage supplémentaire, accès en temps réel	
Inconvénients	Coût de stockage supplémentaire, manque d'accès temps réel	Données non préparées	

Tab -2- : tableau de synthèse des architectures de data warehouse

5. La conception d'un Data Warehouse :

Quatre caractéristiques ont des effets déterminants sur la démarche de conception d'un Data Warehouse :

- **Les évolutions technologiques :**

Un système d'information peut se construire par intégration d'un certain nombre de composants, chacun pouvant être choisi par rapport à son contexte d'utilisation. L'entreprise définit son architecture en fonction de ses besoins.

- **La stratégie de l'entreprise :**

Le Data Warehouse est très proche de la stratégie de l'entreprise. L'objectif du Data Warehouse se définit en terme métier. Il faut donc impliquer les utilisateurs ayant le plus de connaissances dans leur entreprise ou dans leur métier.

- **L'amélioration continue :**

Un Data Warehouse doit évoluer en fonction des demandes utilisateurs ou des nouveaux objectifs de l'entreprise.

- **La maturité de l'entreprise :**

Certaines entreprises ont déjà un système décisionnel. D'autres n'ont aucun acquis.

Dans tous les cas, il n'existe pas de cadre figé pour la conception d'un Data Warehouse. Chaque entreprise doit adapter le projet à son contexte, en ne perdant pas les objectifs de vue. Cet objectif est de mettre en place un système d'information cohérent et intégré, le système devant être décomposé en applications, chacune s'intégrant dans le Data Warehouse.

Il est possible de proposer trois phases pour la conception :

- Définir le pourquoi du Data Warehouse et les objectifs à atteindre (impliquer les utilisateurs).
- Définir l'infrastructure technique et organisationnelle du Data Warehouse.
- Mettre en œuvre les applications.

5.1. Découvrir et définir les initiatives :

Cette phase consiste en l'étude stratégique du Data Warehouse et la définition du plan d'action.

5.1.1. L'étude stratégique :

L'étude stratégique permet d'identifier la stratégie de l'entreprise, son organisation, les processus qu'elle met en œuvre, la culture de l'entreprise.

Le but est de déterminer les domaines pour lesquels la mise en place d'un Data Warehouse peut être le plus bénéfique.

5.1.2. Le plan d'action :

Pour mettre en place le plan d'action, il faut :

- Vérifier la faisabilité de chaque projet (s'assurer de l'existence et de la qualité des données, des possibilités techniques, des possibilités organisationnelles).
- Estimer les ressources pour chaque projet, les besoins.
- Séquencer et planifier les projets.

Il est possible de commencer petit et voir grand : limiter l'objectif à un domaine précis (Data Mart) ou faire un prototype. Ceci permet d'avoir un retour sur investissement rapide et de prouver le bien fondé du concept.

Il faut déterminer la faisabilité de chaque projet envisagé et déterminer les moyens nécessaires (techniques, humains, organisationnels) à leur réalisation.

5.2. La formation :

Selon l'expérience de l'entreprise en matière de décisionnel et des technologies utilisées, un plan de formation sera nécessaire. Il sera impératif que les membres participants au projet soient favorables au changement.

5.3. La mise en œuvre des applications :

La mise en œuvre est réalisée pour chaque initiative. La démarche proposée est une démarche en cinq étapes :

- la spécification,
- la conception,
- la mise en œuvre et l'intégration,
- le déploiement et la mise en place des accompagnements,
- les mesures.

6- Le concept OLAP :

6 1- À propos d'OLAP :

Si les data warehouse ou les data mart sont les lieux de stockage des données analytiques, OLAP est la technologie qui permet aux applications clientes d'y accéder efficacement. OLAP présente de nombreux avantages pour les utilisateurs analytiques, notamment :

- Un modèle de données multidimensionnel intuitif qui permet de sélectionner, de parcourir et d'explorer plus facilement les données.
- Un langage de requête analytique puissant qui permet d'explorer les relations complexes entre les données de l'entreprise.
- Le pré-calcul des données fréquemment interrogées, ce qui permet d'accélérer le temps de réponse pour les requêtes concernées.
- Un outil puissant permettant de créer de nouvelles vues des données à l'aide d'un ensemble riche de fonctions de calcul appropriées.

6-2- Systèmes OLTP versus systèmes OLAP :

Les bases de données sont utilisées dans les entreprises pour gérer les importants volumes d'informations contenus dans leurs systèmes opérationnels. Ces données sont gérées selon des processus transactionnels en ligne (OLTP : «On-Line Transactional Processing») qui se caractérisent de la manière suivante [W12].

- Ils sont nombreux au sein d'une entreprise,
- Ils concernent essentiellement la mise à jour des données,
- Ils traitent un nombre d'enregistrements réduit,
- Ils sont définis et exécutés par de nombreux utilisateurs.
- Base de données partagées (lecture et mise à jour) par l'ensemble des utilisateurs.
- Flux de requête irrégulier : les requêtes individuelles des utilisateurs ne peuvent pas être planifiées.
- Travail répétitif : les utilisateurs demandent au système l'exécution de fonctions appartenant à un ensemble prédéfini.
- Fonctions simples : la plupart des fonctions sont peu complexes.
- Grand nombre de terminaux clients.
- Haute disponibilité : exigence typique des systèmes transactionnels.
- Cohérence et intégrité des données à assurer.
- Taille des bases de données : proportionnelle à l'activité de la société.
- Peu de données touchées par une transaction.
- Equilibrage de charge automatique : haut débit et temps de réponse garantis.

L'exploitation de l'information contenue dans ces systèmes opérationnels est devenue une préoccupation essentielle pour les dirigeants des entreprises qui désirent améliorer leur prise de décision par une meilleure connaissance de leur propre activité, de celle de la concurrence, des employés, des clients et des fournisseurs. Les entreprises sont donc à la recherche de systèmes supportant efficacement les applications d'aide à la décision. Ces applications décisionnelles utilisent des processus d'analyse en ligne de données (OLAP : ON-LINE Analytical Processing). Ces processus répondent aux besoins spécifiques des analyses d'information. Ces processus OLAP se caractérisent de la manière suivante [W12] :

- Ils sont peu nombreux, mais leurs données et traitements sont complexes,
- Il s'agit uniquement de traitements semi-automatiques visant à interroger, visualiser et synthétiser les données,
- Ils sont définis et mis en œuvre pour un nombre réduit d'utilisateurs qui sont les décideurs.
- Bases de données en lecture essentiellement, partagées par un petit nombre d'utilisateurs et souvent distinctes des bases de production.
- Flux de requête irrégulier : les requêtes individuelles des utilisateurs ne peuvent pas être planifiées.
- Travail répétitif : les utilisateurs demandent au système l'exécution de requêtes variées.
- Fonctions complexes : les requêtes sont souvent complexes et mettent en jeu un grand volume de données.
- Petit nombre de stations de travail clients.
- Disponibilité : ne représente pas une exigence première des systèmes d'aide à la décision.
- Taille des bases de données : proportionnelle à l'historique de l'activité de la société.
- Beaucoup de données touchées par une requête.
- Recherche de la performance au moyen du parallélisme intra-requête.

6-3- Les différents outils OLAP :

6-3-1- Les outils MOLAP :

MOLAP est conçue exclusivement pour l'analyse multidimensionnelle, avec un mode de stockage optimisé par rapport aux chemins d'accès prédéfinis. Ainsi, toute valeur d'indicateur associée à l'axe temps, sera pré calculée au chargement pour toutes ses valeurs hebdomadaires, mensuelles, etc.

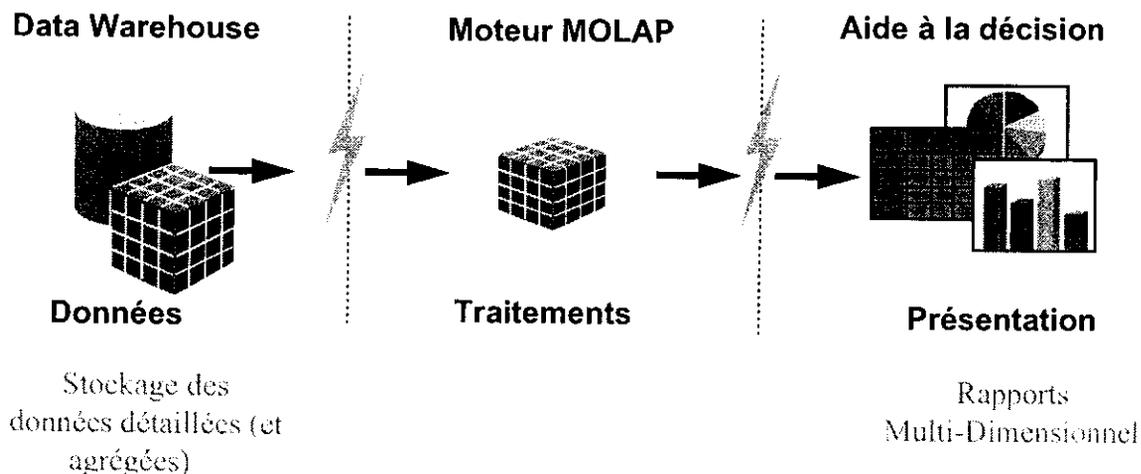


Fig -8- architecture d'un produit MOLAP (W01)

6-3-2- Les outils ROLAP :

Les outils ROLAP superposent au dessus des SGBD/R bidimensionnels un modèle qui représente les données dans un format multidimensionnel. Ces produits diminuent sensiblement le coût lié à la mise en œuvre d'un serveur de base de données multidimensionnelle supplémentaire.

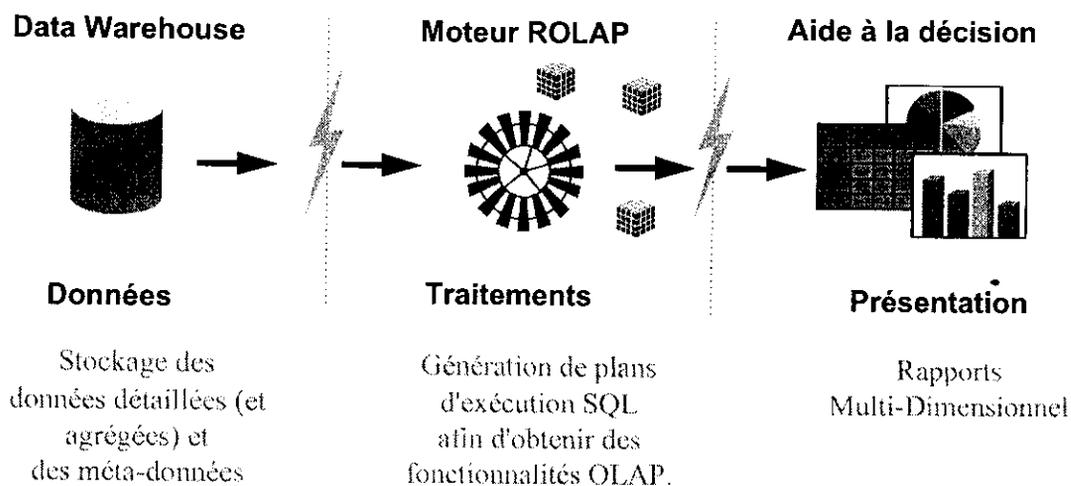


Fig -9- architecture d'un produit ROLAP (W01)

7- La modélisation des données :

7-1- Modélisation entité – relation :

La modélisation entité–relation vise à éliminer les données redondantes. Ceci apporte de nombreux avantages au niveau du traitement des transactions, qui deviennent simples et déterministes. Plus la transaction est «atomique», moins il existe d'entités concernées par cette transaction et plus le risque d'incohérence de la base de données est réduit.

Dans le cadre du data warehouse, la modélisation entité-relation n'est pas appropriée. L'utilisateur final peut difficilement comprendre le modèle entité relation. Il ne peut pas naviguer dans les données. C'est pourquoi d'autres techniques sont utilisées.

7-2- Modélisation dimensionnelle :

La modélisation dimensionnelle est une méthode de conception logique souvent associée aux entrepôts de données. Elle diffère de la modélisation entité-relation traditionnelle. Elle vise à présenter les données sous une forme standardisée intuitive et qui permet des accès performants.

Chaque modèle dimensionnel se compose d'une table contenant une clé multiple, la table de faits, et d'un ensemble de tables plus petites nommées tables «dimension». Chaque table «dimension» possède une clé primaire unique qui correspond exactement à l'un des composants de la clé multiple de la table de faits. La table de faits contient des faits numériques sur un sujet donné, par exemple les données statistiques d'une couverture d'assurance. Au contraire, les tables «dimension» contiennent le plus souvent des informations textuelles descriptives. Les attributs des dimensions sont les sources de contraintes les plus intéressantes dans les requêtes utilisées.

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

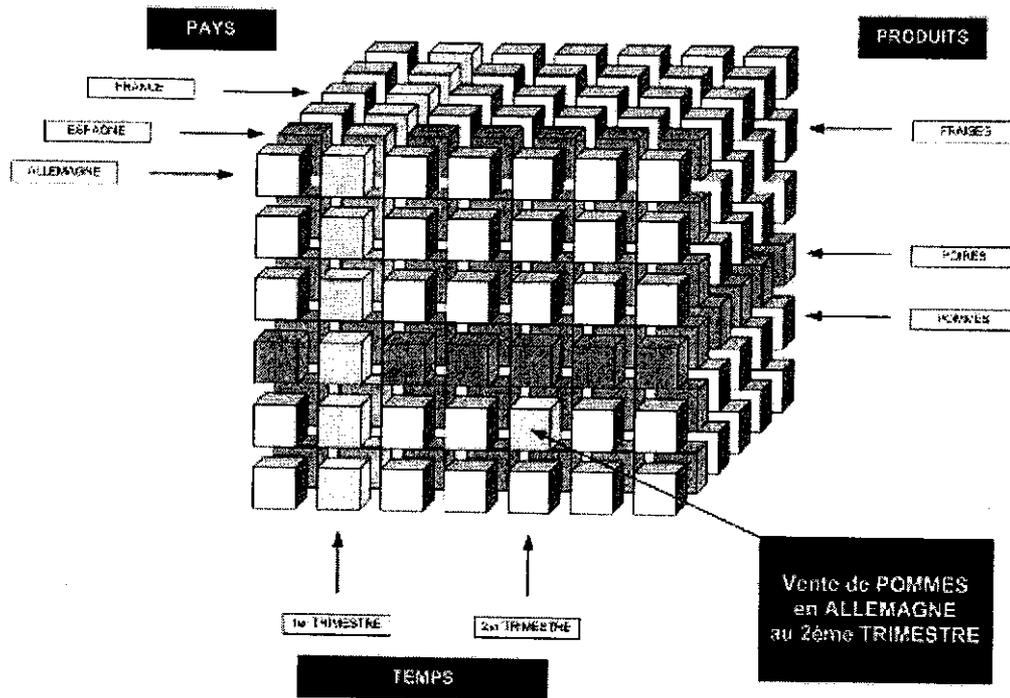


Fig -10- Exemple de représentation multidimensionnelle (W01)

7-2-1- Concept de fait :

Le sujet analysé est représenté par le concept de fait. Le fait modélise le sujet de l'analyse. Un fait est formé de mesures correspondant aux informations de l'activité analysée.

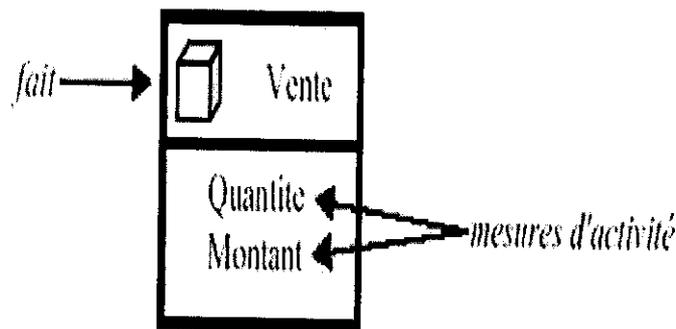


Fig -11- Exemple de fait (W04)

7-2-2- Concept de dimension :

Le sujet analysé, c'est-à-dire le fait, est analysé suivant différentes perspectives. Ces perspectives correspondent à une catégorie utilisée pour caractériser les mesures d'activité analysées ; on parle de dimensions.

Une dimension modélise une perspective de l'analyse. Une dimension se compose de paramètres correspondant aux informations faisant varier les mesures de l'activité.

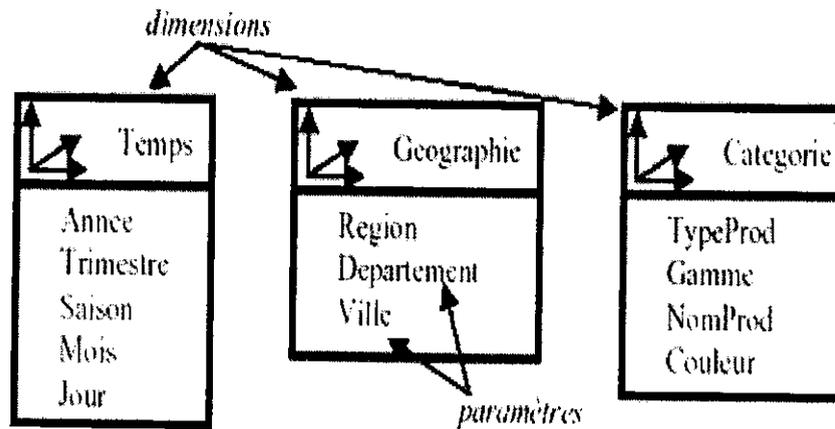


Fig -12- Exemple de dimension (W04)

7-2-3- Modèles en étoile, en flocon et en constellation :

A partir du fait et des dimensions, il est possible d'établir une structure de données simple qui correspond au besoin de la modélisation multidimensionnelle. Cette structure est constituée de fait central et des dimensions. Ce modèle représente visuellement une étoile, on parle de modèle en étoile.

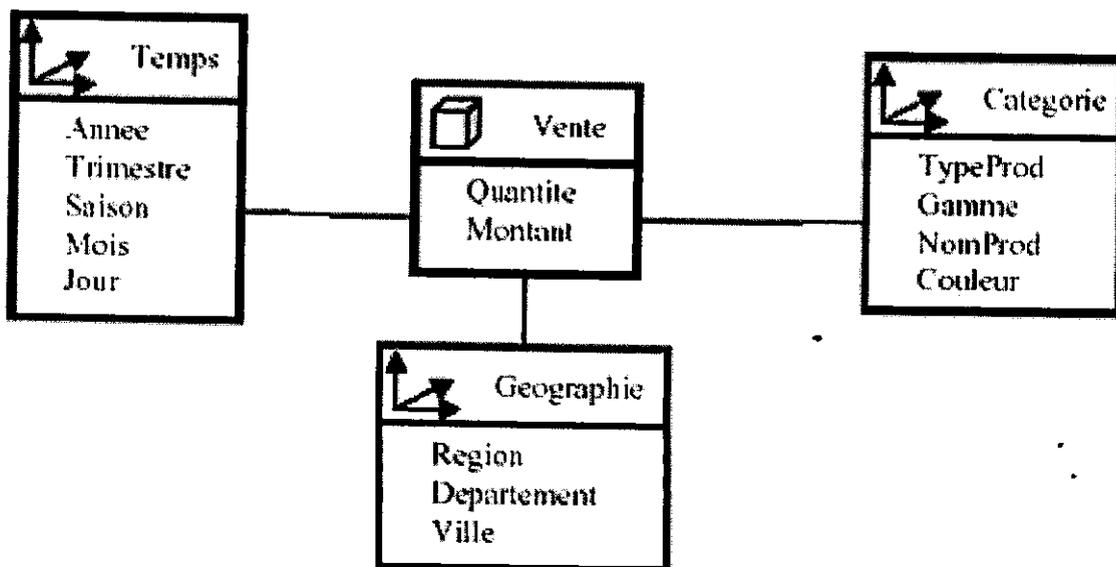


Fig -13- Exemple de modélisation en étoile (W04)

Il existe d'autres techniques de modélisation, notamment la modélisation en flocon. Une modélisation en flocon consiste à décomposer les dimensions du modèle en étoile en sous-hiérarchies. La modélisation en flocon est donc une émanation de la modélisation en étoile ; le fait est conservé et les dimensions sont éclatées conformément à l'hierarchie des paramètres.

L'avantage de cette modélisation est de formaliser une hiérarchie au sein d'une dimension. Par contre, la modélisation en flocon induit une dénormalisation des dimensions générant une plus grande complexité en termes de lisibilité et de gestion.

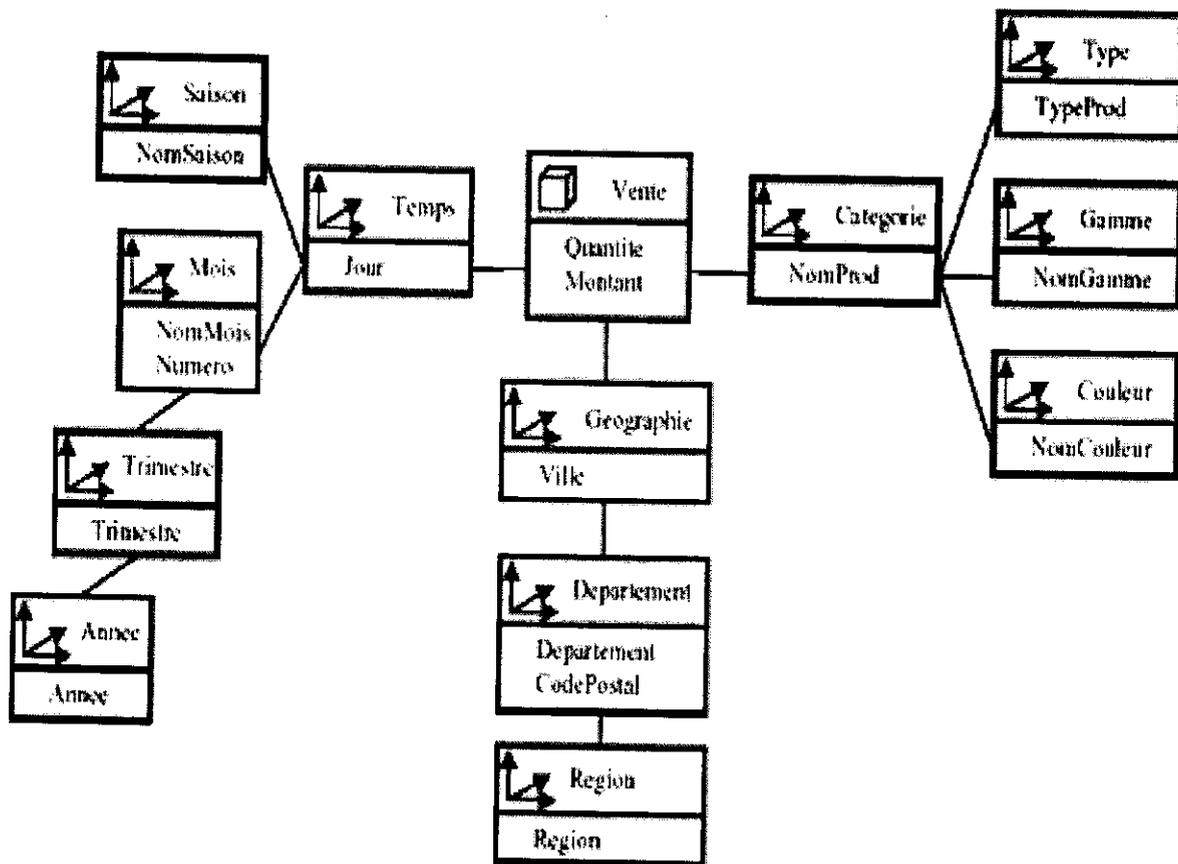


Fig -14- Exemple de modélisation en flocon (W04)

Une autre technique de modélisation, issue du modèle en étoile, est la modélisation en constellation. Il s'agit de fusionner plusieurs modèles en étoile qui utilisent des dimensions communes. Un modèle en constellation comprend donc, plusieurs faits et des dimensions communes ou non.

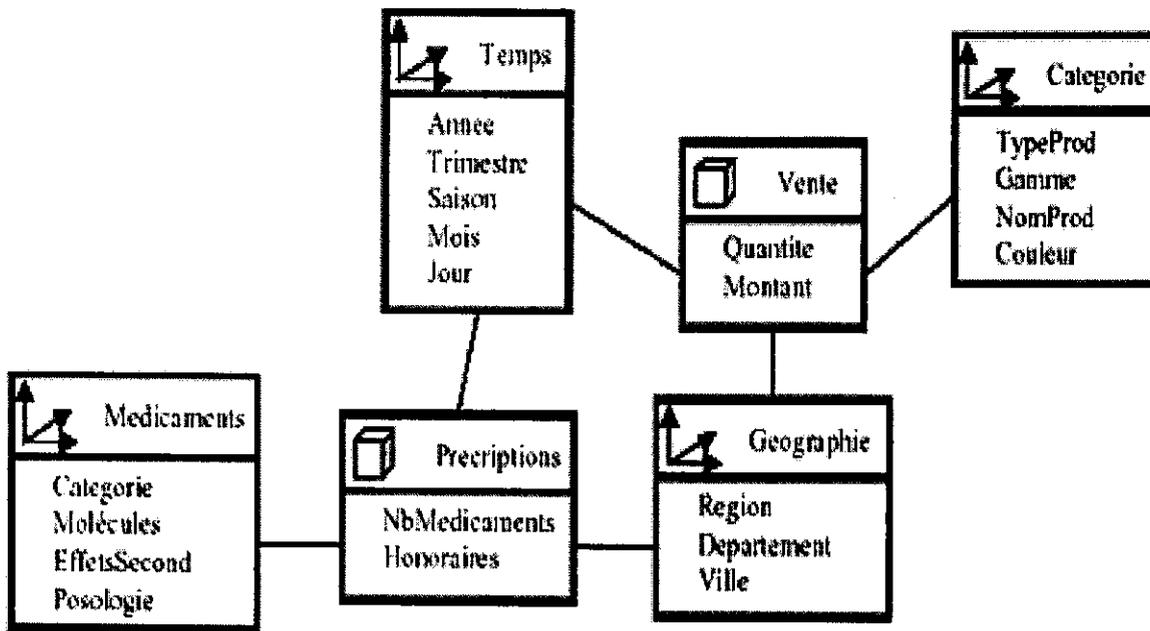


Fig -15- Exemple de modélisation en constellation (W04)

8- Construction d'un entrepôt de données :

L'objectif à atteindre est de recomposer les données disponibles pour obtenir une vision intégrée et transversale aux différentes fonctions de l'entreprises, une vision métier au travers des différents axes d'analyse et une vision agrégée ou détaillée, adaptée à son besoin.

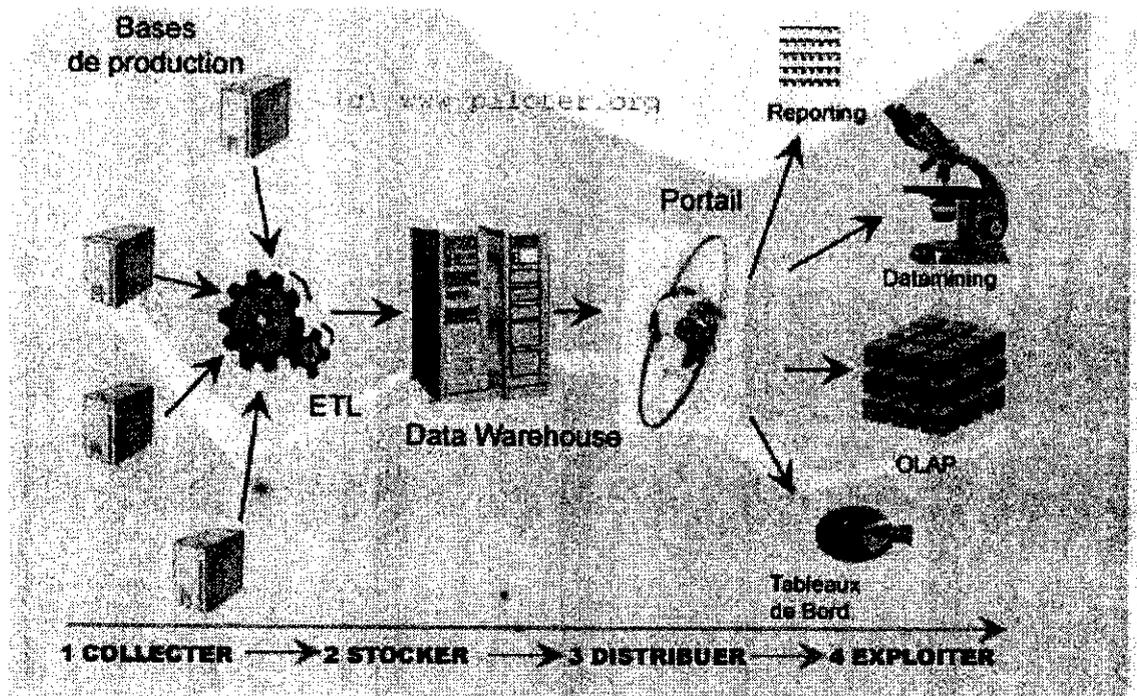


Fig -16- La construction d'un entrepôt de données (W05)

Le cadre général d'un Data Warehouse comporte trois domaines principaux :

- Les applications,
- Les composants fonctionnels du Data Warehouse (acquisition, stockage, accès),
- Les infrastructures (techniques et opérationnelles).

8-1- Les applications :

Un Data Warehouse se fait en plusieurs itérations. Chaque sujet traité est décomposé en un ensemble d'initiatives (projet décisionnel entrant en jeu dans la construction d'un Data Warehouse dans une démarche itérative).

Les applications doivent rester maîtrisables et fournir des résultats tangibles dans un délai de moins de six mois, qui correspond au délai moyen de réalisation d'applications.

8-2- Les composants fonctionnels :

Trois composants caractérisent un Data Warehouse :

8-2-1- L'acquisition (ETL) :

Elle se compose de trois phases : l'extraction, la transformation et le chargement (Extraction, Transformation, Loading). On va détailler le processus ETL dans la suite.

8-2-2- Le stockage :

Le composant de base est le SGBD. Il doit être spécifiquement adapté aux caractéristiques de l'accès décisionnel. Du fait de l'importance de l'historique, la structuration physique des données est également très importante.

8-2-3- L'accès :

Définir une architecture globale servant de support aux accès décisionnels impose des choix technologiques non structurants. Il faudra mettre en place une infrastructure commune à toutes les applications décisionnelles, tout en laissant aux utilisateurs l'opportunité d'utiliser les solutions d'accès les mieux adaptées à leur problématique.

8-3- Les infrastructures :

Il y a deux niveaux d'infrastructures :

a- L'infrastructure technique :

L'ensemble des composants matériels et logiciels (à associer aux composants fonctionnels : alimentation, stockage et accès).

b- L'infrastructure opérationnelle :

L'ensemble des procédures et des services pour administrer les données, gérer les utilisateurs et exploiter le système.

9. Le processus d'alimentation ETL :

La démarche utilisée pour construire une couche d'alimentation se décompose en trois étapes essentielles : Extraction, Transformation et Chargement.

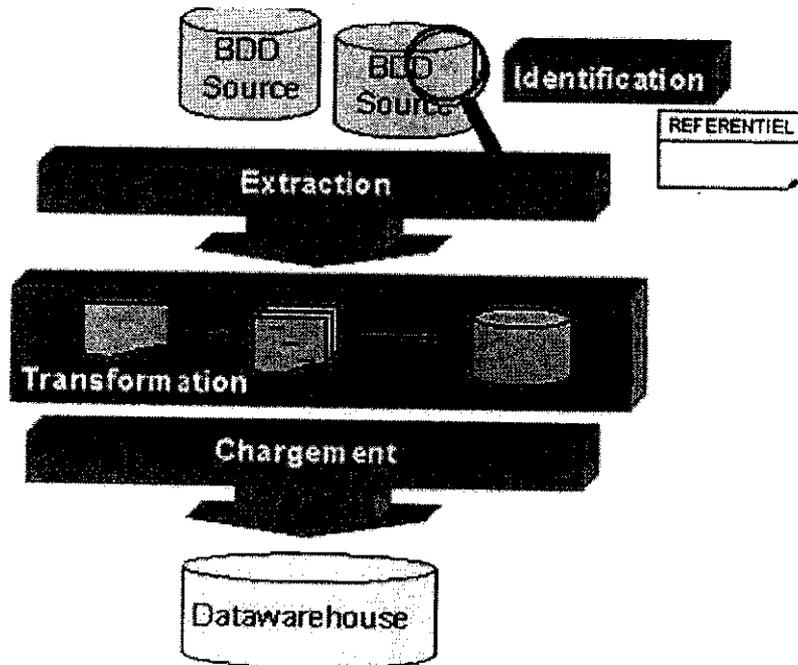


Fig-17- Le processus d'alimentation ETL (W01)

a- L'extraction des données :

Consiste à collecter des données utiles dans le système de production. Il faut donc d'abord identifier les données ayant évoluées afin d'extraire le minimum de données, puis planifier ces extractions afin d'éviter les saturations du système de production.

b- La transformation des données :

Correspond à la transformation des caractéristiques des données du système opérationnel dans la forme définie du Data Warehouse. Cette préparation inclut la mise en correspondance des formats de données, le nettoyage, la transformation et l'agrégation. Ce processus de transformation va accéder naturellement à toutes les informations stockées dans le dictionnaire, notamment la localisation des sources de données et leurs structures dans le système de production.

c- Le chargement :

C'est une phase délicate notamment lorsque les volumes sont importants. Pour obtenir de bonnes performances en chargement, il est impératif de maîtriser les structures du SGBD associées aux données chargées afin d'optimiser au mieux ces processus.

10. La fonction de pilotage et de contrôle de gestion :

A côté de la fonction informatique devenue centrale, la fonction de pilotage et de contrôle de gestion constitue une autre fonction support particulièrement importante dans la nouvelle configuration des organisations.

10.1. Les indicateurs comme outil de pilotage :

L'évaluation de la performance à l'aide d'indicateurs appropriés est un exercice crucial afin d'assurer le succès des organisations quels que soit leurs secteurs d'activités. Les clients sont de plus en plus exigeants et leurs attentes sont de plus en plus grandes [W10].

a. Définition :

Un indicateur est une information ou regroupement d'informations contribuant à l'appréciation d'une situation par le décideur. Il n'est jamais muet et entraîne toujours une action ou réaction du décideur. Cette action peut être de ne rien faire. Mais il s'agit là d'une démarche active et non passive.

b. Catégories d'indicateurs :

Les indicateurs peuvent être classés en trois catégories selon le rôle qu'ils jouent pour le décideur :

- ✓ **Les indicateurs d'alerte** signalent un dysfonctionnement, donc un état anormal impliquant des actions réparatrices. Au contraire, leur silence indique un état normal. Par exemple, un indicateur d'alerte pourra signaler une augmentation ou une diminution de vente d'un produit donné.
- ✓ **Les indicateurs d'équilibration** signalent l'avancement par rapport aux objectifs. Ils peuvent induire des actions correctives. Par exemple, un indicateur d'équilibration pourrait mesurer les ventes d'un courtier en assurance donné et comparer avec l'objectif préalablement fixé.
- ✓ **Les indicateurs d'anticipation** informent sur le système dans son environnement et permettent de reconsidérer la stratégie choisie. Par exemple, l'introduction par un concurrent d'un nouveau produit d'assurance qui se vend très facilement devrait inciter les décideurs d'une compagnie d'assurance à revoir leur offre de façon à ne pas perdre de parts de marché.

11. Méthodologie adoptée :

Deux approches sont disponibles concernant le développement des systèmes décisionnels. Une approche basée sur le génie logiciel et une approche basée sur la gestion de projet par prototypage qui est proposée et privilégiée par Ralph Kimball qui est l'un des pionniers des systèmes décisionnels, cette approche repose sur la modélisation multidimensionnelle des données et les données décisionnelles sont organisées en fonction des analyses multidimensionnelles effectuées (analyse des données suivant plusieurs axes).

Pour l'implémentation de notre système, on va adopter l'approche proposée par R.Kimball.

Le schéma de cycle de vie dimensionnel proposé par R.Kimball matérialise l'enchaînement des grandes étapes de l'implémentation d'un Data Warehouse.

Cette démarche est illustrée dans le schéma ci-dessus. Ce schéma représente la succession des tâches de haut niveau nécessaires à la conception et au déploiement d'un entrepôt de données efficace [L01].

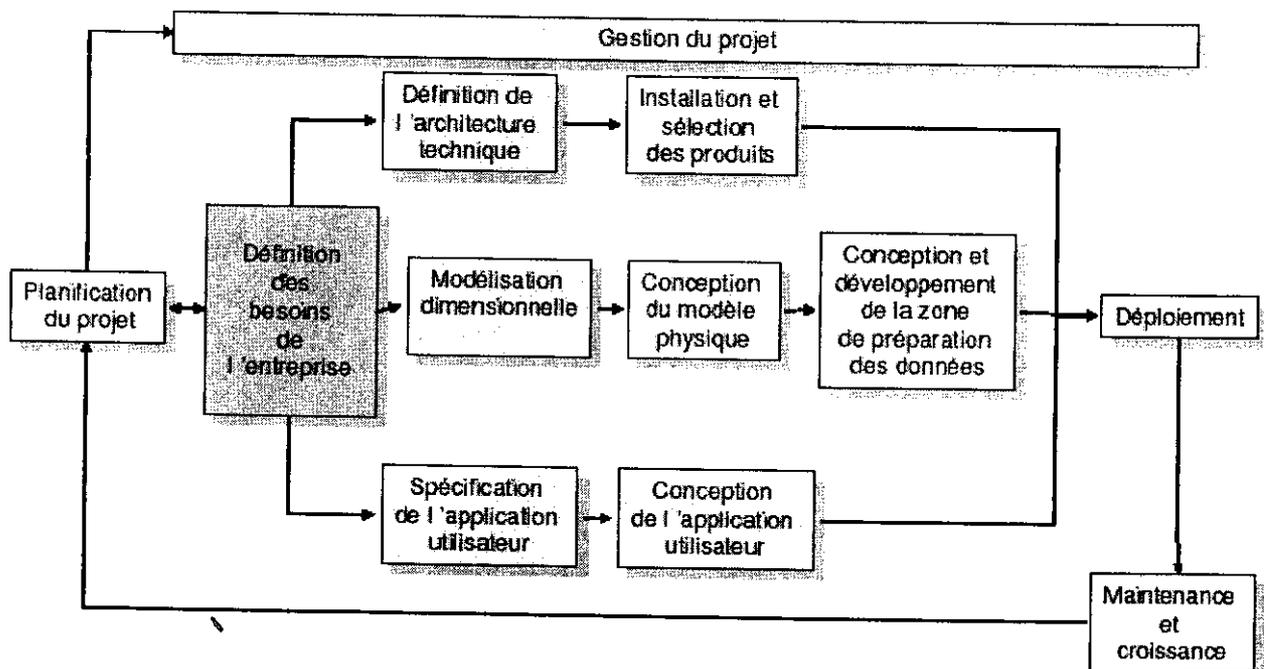


Fig-18- Schéma du cycle de vie dimensionnel (L01)

Description des phases :

1. **Planification du projet :** Aborde la définition et l'étendu du projet de Data Warehouse, y compris l'appréciation du niveau de maturité de l'organisation face à ce type d'approche et sa justification fonctionnelle.
2. **Définition des besoins :** Les chances de succès d'un entrepôt de données se trouvent considérablement accrues par la bonne compréhension des utilisateurs et de leurs besoins. Les constructeurs d'entrepôts de données doivent appréhender les facteurs clés qui conduisent l'entreprise à vouloir définir de manière efficace ses besoins et les traduire pour les intégrer lors de la conception.
3. **Modélisation dimensionnelle des données :** La conception de modèles de données destinés à corroborer les analyses des utilisateurs nécessite une approche différente de celle utilisée lors de la conception de systèmes opérationnels. Nous commencerons par construire une matrice qui représentera les processus métier clés et leur dimensionnalité. De là nous effectuerons une analyse plus détaillée des données des systèmes sources opérationnels. En couplant cette analyse à la compréhension des besoins précédemment établie, nous développerons un modèle dimensionnel.
4. **Conception du modèle physique de données :** La conception physique d'une base de données définit les structures physiques nécessaires pour l'implémentation de la base de données logique.
5. **Conception et déploiement de la zone de préparation des données :** Le processus de conception et de déploiement des éléments de la zone de préparation des données constitue souvent la tâche la plus sous-estimée du projet de Data Warehouse. Le processus de préparation se déroule en trois étapes majeures : l'extraction, la transformation et le chargement des données.
6. **Définition de l'architecture technique :** Les environnements de Data Warehouse nécessitent l'intégration de nombreuses technologies. Cette étape de définition donne une vision globale de la structure de l'architecture technique à mettre en œuvre.
7. **Choix technologiques et mise en œuvre :** A partir de l'étude de l'architecture technique, on sélectionne les composants spécifiques, tels que la plate forme matérielle, la système de gestion de base de données et les outils de préparation et d'accès aux données.
8. **Développement de l'application utilisateur :** On définit une série d'applications standard destinées à l'utilisateur final, car tous les utilisateurs n'ont pas besoin d'un accès *ad hoc* à l'entrepôt de données.
9. **Déploiement :** Le déploiement est le point de convergence de la technologie, des données et des applications utilisateurs. Durant cette phase, on va définir le « comment faire ? ».
10. **Maintenance et croissance :** Après le déploiement initial de l'entrepôt de données, il faut s'occuper des utilisateurs en leur procurant un service de support et une formation continue.

L'approche de cycle de vie dimensionnelle décrit le cheminement du projet dans son ensemble. Pour notre cas on va :

- Abandonner la phase de l'installation et la sélection des produits à cause de l'existence d'une infrastructure technique (matériel et logiciel).
- Regrouper les deux phases « spécification et développement des applications utilisateurs » en une seule phase.
- La phase maintenance qui traite de continuation d'assistance et de formation des utilisateurs ainsi l'optimisation des performances de la base de données, se déroule après le déploiement de l'entrepôt, donc elle est elle aussi abandonnée.

Conclusion :

Les systèmes d'information décisionnels constituent un domaine d'expertise de plus en plus répandue dans les entreprises soucieuses d'extraire l'information pertinente cachée dans leurs bases de données, en vue notamment d'améliorer leur gestion de la relation client et leur maîtrise des risques.

Après avoir défini ce qu'est un système d'information décisionnel et dresser un panorama de ses buts et ses apports à la gestion et à l'analyse, on a enchaîné vers le DataWarehouse.

L'entrepôt de données est la base dans laquelle est stocké l'ensemble des informations d'un système décisionnel. Ces informations permettent aux décideurs de travailler dans un environnement informationnel, référence, homogène et historisé.

L'objectif de l'entrepôt est de centraliser l'information décisionnelle en assurant l'intégration des données extraites, la pérennité de ces données dans le temps et la conservation de leurs évolutions. Ainsi, les données stockées dans un magasin doivent correspondre à une structuration adaptée des données (selon plusieurs axes d'analyses) reflétant la vision des analystes. Cette représentation des données est basée sur une modélisation multidimensionnelle.

Les progrès des technologies de l'information, et en particulier les systèmes d'aide à la décision, ont apportés des améliorations sur la fonction du pilotage et la comptabilité des entreprises. Cela par les différents indicateurs de gestion qui présentent un outil d'aide à la décision pour les gestionnaires et les décideurs des organisations.

Chapitre II :

*Présentation générale
de Naftal Spa*

Introduction :

Afin de bien comprendre le problème de notre étude, et pour mieux définir les objectifs et les besoins attendus, nous allons présenter, dans ce chapitre, l'entreprise Naftal Spa, avec ses différentes directions qui la composent et ses principales missions, ainsi que la structure d'accueil qui n'est d'autre que la Direction Centrale des Systèmes d'Informations (DCSI).

1. Présentation générale :**1.1. Historique :**

Issue de SONATRACH, l'entreprise ERDP a été créée par le décret N° 80/101 du 06 avril 1981.

Entrée en activité le 1er janvier 1982, elle est chargée de l'industrie du raffinage et de la distribution des produits pétroliers sous le sigle NAFTAL.

NAFTAL a pour mission principale, la distribution et la commercialisation des produits pétroliers sur le marché national. Elle intervient dans les domaines :

- ◆ De l'enfûtage GPL
- ◆ De la formulation de bitumes
- ◆ De la distribution, stockage et commercialisation des carburants, GPL, lubrifiants, bitumes, pneumatiques, GPL/carburant, produits spéciaux.
- ◆ Du transport des produits pétroliers.

1.2. Organisation :

La gestion a pour but d'atteindre, de la façon la plus efficace possible, les objectifs de l'entreprise.

Le but de la fonction organisation consiste à regrouper les ressources de l'entreprise d'une manière ordonnée et à répartir les individus selon un model acceptable leur permettant de remplir les taches qui leur sont assignées.

Ce model doit établir aussi la relation d'autorité entre les diverses structures, il s'agit de diviser l'objectif global de l'entreprise en diverses taches individuelles et de les regrouper en services, département, direction, sous la responsabilité d'un manager à qui l'autorité nécessaire sera déléguée.

En 1987, l'activité raffinage est séparée de l'activité distribution.

La raison sociale de la société change suite à cette séparation des activités et NAFTAL est désormais chargée de la commercialisation et de la distribution des produits pétroliers et dérivés.

A partir de 1998, elle change de statut et devient Société par actions filiale à 100% de SONATRACH.

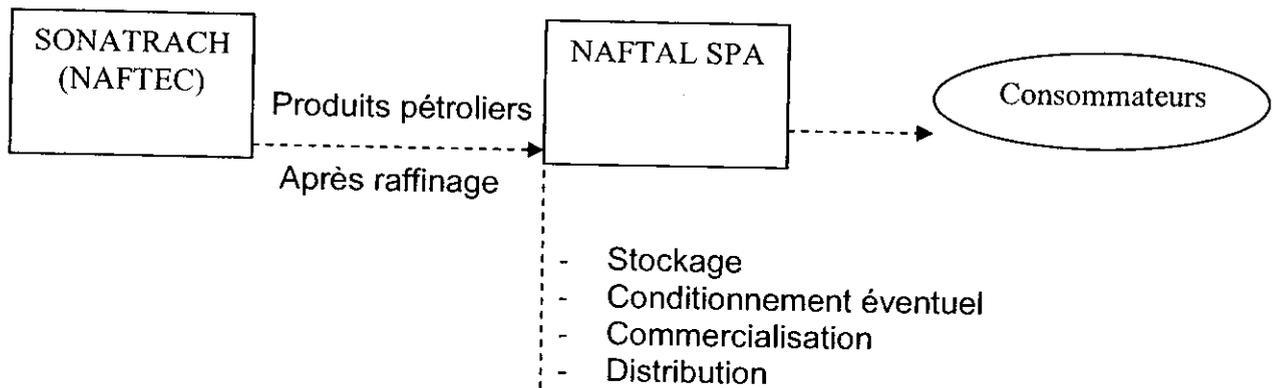


Fig-19- Création de l'entreprise

Avant cette nouvelle organisation, l'entreprise NAFTAL SPA se chargeait de commercialiser et de distribuer tous ces produits (un grand nombre). L'entreprise était structurée en 48 UNDS (Unité Naftal Spa de Distribution, une UND par Wilaya) et chacune reprenait toute l'activité de Naftal Spa dans sa Wilaya.

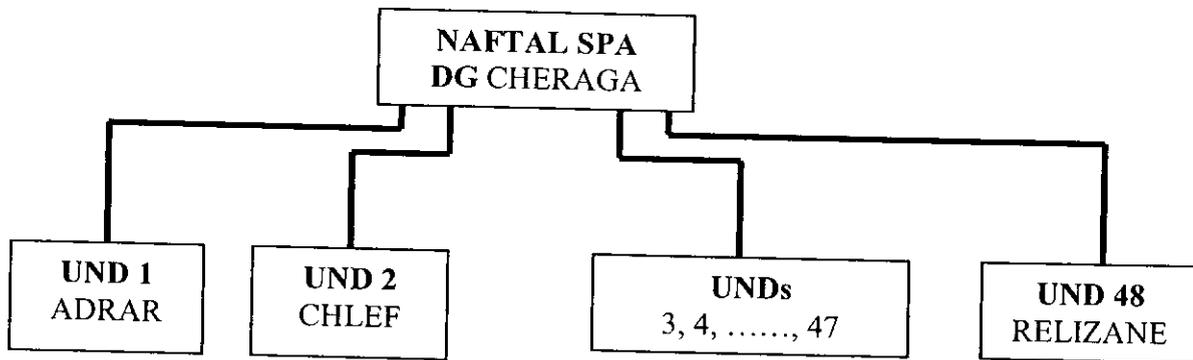


Fig -20- L'ancienne structure en UNDS de l'entreprise Naftal Spa

Les insuffisances et les dysfonctionnements divers élevés lors de l'accomplissement de la mission de Naftal Spa, ont poussé les responsables à penser à une restructuration de l'entreprise.

La réorganisation de la société engagée par la Direction Générale a abouti à la création des Divisions (structure par produit) CLPM, AVM et GPL.

La deuxième phase a été marquée par la séparation de BITUME du CLPB et devient une Division BITUME et la Division CLPM change de nom et devient une division CLP.

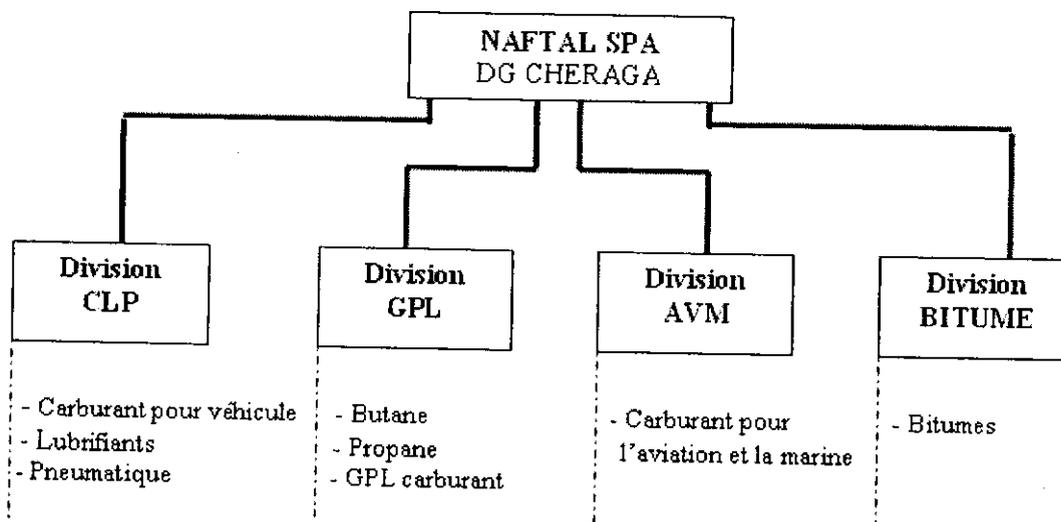


Fig -21- La nouvelle structure divisionnaire de Naftal Spa

L'organisation de NAFTAL SPA est articulée autour de structures centrales chargées de la définition de la politique du suivi et du contrôle des activités de l'entreprise, et des structures opérationnelles qui bénéficient de produits pétroliers.

Ces structures opérationnelles bénéficient d'un soutien logistique et la maintenance assurée par des directions régionales spécialisées.

Ces centres opérationnels ont pour mission la distribution de l'ensemble des produits commercialisés par Naftal Spa dans les champs d'influence d'une ou de plusieurs wilayas.

Les divisions ont pour mission de définir avec la direction générale, la stratégie de distribution et de commercialisation des produits pétroliers en veillant à rassembler toutes les conditions de son application dans les centres opérationnels de la société.

Le schéma d'organisation de la macrostructure de la société s'articule autour des principales structures ci-après énumérées :

- ✦ La direction générale.
- ✦ Les structures fonctionnelles.
- ✦ Les structures opérationnelles (division par produit).

Schéma actuel d'organisation de la société NAFTAL

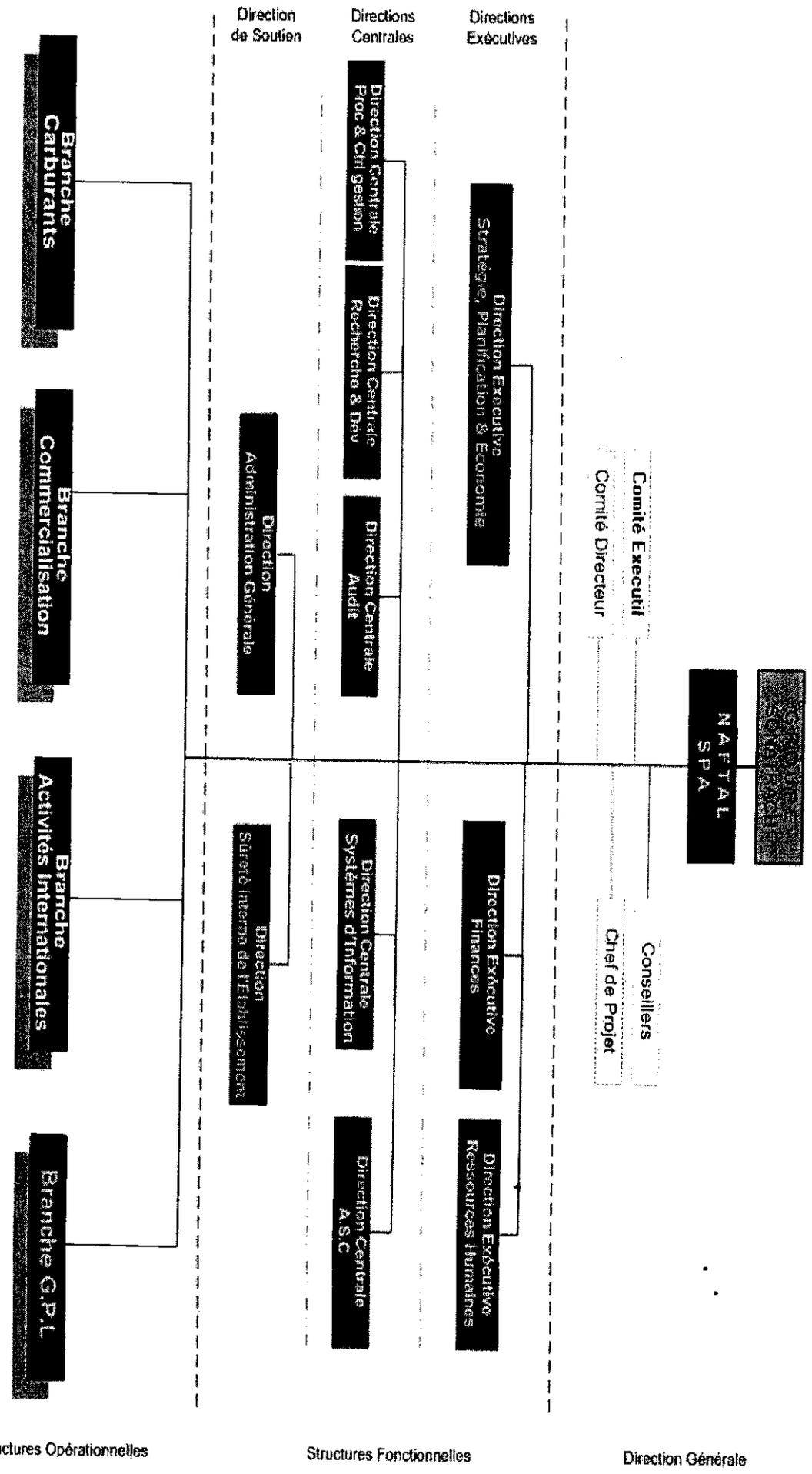


Fig -22- L'organigramme général de l'entreprise Naftal Spa (W06)

Structures Opérationnelles

Structures Fonctionnelles

Direction Générale

La direction générale de la société est assurée par le président du conseil d'administration conformément à l'article 17 des statuts de la société Naftal Spa

A ce titre, le président directeur général est secondé par des conseillers et assistants, et deux organes permanents (comité exécutif et comité directeur).

Les conseillers et assistants sont chargés d'assister le président directeur général dans l'exercice de ses fonctions et dans la prise de décision stratégique ainsi qu'en matière de :

- activités internationales.
- relations presse et médias.

1.2.1. Le comité exécutif :

Est une instance présidée par le président directeur général, composée des directeurs de divisions et directeurs exécutifs ; ce comité est chargé de définir et d'arrêter les grands axes stratégiques et développement, d'engineering financier, de sécurité.

1.2.2. Le comité directeur :

Est une instance présidée par le président directeur général, composée des directeurs de divisions et directeurs exécutifs ; directeurs centraux, directeur de l'administration générale et un représentant du partenaire social.

Ce comité assiste le président directeur général dans la gestion et le contrôle des activités de la société.

1.2.3. La direction de l'administration générale :

La direction de l'administration générale (DAG) assure la gestion des moyens humains et matériels de siège, et les prestations de services... elle a pour missions principales :

- Gérer le personnel de l'unité siège et mettre à la disposition de toutes les directions du siège de l'entreprise, les moyens nécessaires à la réalisation des objectifs assignés par la direction générale.
- Coordonner l'élaboration et la consolidation des budgets de fonctionnements.
- Faire assurer la gestion et la paie du personnel de l'unité siège.
- Mettre en œuvre les moyens nécessaires au bon fonctionnement des directions du siège en assurant la gestion du parc auto, des infrastructures (bâtiment, cantine, des moyens de communication et de l'archivage).
- Veiller à la sécurité au niveau du siège de la direction générale.

- Veiller à l'application effective des directives, règles et procédures de l'entreprise dans le domaine de l'administration et de la gestion des personnels et des moyens.
- Veiller à la disponibilité et la circulation des informations nécessaires au bon fonctionnement des structures.
- Veiller en permanence, au maintien d'une image de marque de l'entreprise.

Elle se compose de quatre départements, un service et un centre :

- Département gestion des moyens.
- Département immeuble.
- Département personnel.
- Département finances et comptabilité.
- Service documentation.
- Un centre d'archive.

1.2.4. Les structures fonctionnelles :

Les structures fonctionnelles de la société sont organisées en trois directions exécutives, six directions centrales et deux directions de soutien.

La direction exécutive, chacune dans son domaine d'activités, est chargée de :

- Définir la politique et la stratégie de la société.
- Anticiper les tendances.
- Concevoir et mettre en place les instruments de pilotage et les outils de contrôle.
- Assurer le management stratégique.
- Assister les structures opérationnelles.
- Assurer la coordination et la cohérence d'ensemble.

La direction centrale est un centre d'expertise pour les activités de recherche et de développement, d'audit de la protection du patrimoine et de la sécurité industrielle et de l'environnement.

Les directions de soutien assurent la gestion administrative du siège social de la société.

Notre travail se déroule à la Direction Centrale des Systèmes d'Information (DCSI).

1.2.4.1. La direction centrale des systèmes d'information :

Les missions principales de la direction centrale des systèmes d'information sont:

- intègre au sein de l'entreprise les différentes technologies de l'Internet adaptées à la demande des différentes directions.
- Définit une politique de soutien WEB en cohésion avec la stratégie générale de la société.
- Réalise la conception et le développement des systèmes informatiques et met en place un système décisionnel.
- Réalise la consolidation centrale des données issues de l'exploitation de tous les systèmes.
- Réalise les études d'installation des réseaux informatiques.
- Identifie auprès de toutes les structures de la société, les axes ou les domaines sur lesquels sont basées leur stratégie et leurs analyses.
- Collabore avec les différentes structures de la société pour collecter et analyser les besoins en termes de solution Web.

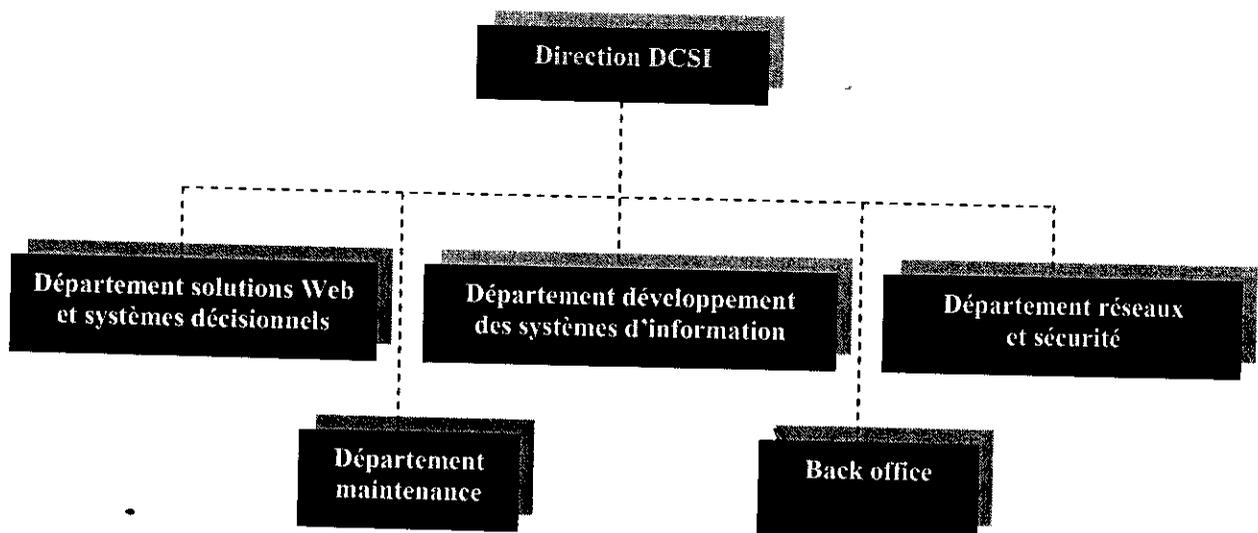


Fig -23- Organigramme de la DCSI

1.2.5. Les structures opérationnelles :

Les structures opérationnelles de la société sont organisées en quatre branches. Chaque branche exerce ses métiers de base et développe ses activités conformément à la politique générale de la société, dans la limite de son objet social.

1.2.5.1. Branche Carburants :

Les missions principales de cette branche sont :

- Approvisionner, stocker, distribuer et commercialiser les produits pétroliers, carburants terre, aviation et marine.
- Définir la stratégie de commercialisation qui permettra de consolider et développer les parts du marché de la société.
- Faire les études et les programmations de développement, rénovation et maintenance des centres carburants.
- Veiller, avec l'assistance de la direction des finances & comptabilité, à l'équilibre financier de l'activité carburants, le cas échéant, recommander et mettre en œuvre les mesures correctives.

1.2.5.2. Branche Commercialisation :

Les missions principales de cette branche sont :

- Commercialise les produits pétroliers ainsi que les prestations de services sur l'ensemble de points de vente.
- Elabore et met en œuvre la stratégie de développement, contrôle et la gestion du réseau (GD, GL, PVA).
- Trace une politique d'amélioration de développement de la qualité de service et d'accueil de la clientèle.
- Arrête un plan de production par station de service, centre de conversion et autres points de vente.
- Veille à l'équilibre financier des activités de commercialisation, le cas échéant, recommande et met en œuvre les mesures correctives.
- Analyse, consolide et recommande à l'approbation de la hiérarchie, les plans & budgets de fonctionnement et d'investissement de la branche.

1.2.5.3. Branche GPL :

Les missions principales de cette branche sont :

- L'approvisionnement en gaz du pétrole liquéfié provenant des raffineries.
- Le stockage, le conditionnement, la distribution et la commercialisation des produits GPL, à travers le territoire national.
- Gestion, organisation, promotion et développement de l'activité enfûtage.
- Commercialisation des GPL vrac et conditionnés, leurs emballages et accessoires.
- Contrôle et respect des normes et consignes de sécurité sur toute la chaîne GPL (transport, installations d'enfûtage et de stockage, bouteilles, citernes, accessoires, etc).

1.2.5.4. Branche activités internationales et partenariats :

Cette branche n'est pas encore en activité, ses missions ne sont pas précisées.

2. Les moyens humains & infrastructures de Naftal :

2.1. Les moyens humains :

Naftal axe ses efforts sur la stabilisation des effectifs, la promotion de l'emploi féminin, le recrutement qualitatif et la formation pour le développement permanent des ressources humains.

Naftal compte 30.000 agents dont 16.000 ont bénéficié de formation. En 2005 elle a concerné un effectif de 6.100 agents dans différents domaines dont : La commercialisation des produits pétroliers, la sécurité industrielle, le management qualité et les normes ISO, les finances et la comptabilité ainsi que la langue anglais.

2.2. Les infrastructures de Naftal :

Naftal dispose de :

- ❖ 67 centres de stockage de carburants, lubrifiants et pneumatiques, dont :
7 entrepôts ,42 dépôts carburants terre.
- ❖ 31 centres aviation, dont : 5 entrepôts, 26 dépôts.
- ❖ 6 centres carburants marine.
- ❖ 41 centres d'emplissage GPL et 7 centres GPL vrac.
- ❖ 15 unités de formulation bitumes.
- ❖ 49 dépôts relais de GPL conditionnées.
- ❖ 6.000 véhicules et engins.
- ❖ 750 Km de pipes (carburants « terre, aviation-marine » et GPL).
- ❖ 1.860 stations de services, dont 1.176 appartenant aux privés.

3. L'environnement de l'entreprise :

Notre objectif est de connaître le positionnement de l'entreprise en terme de marché, non pas en qualité de juge et de conseil, et non pas en détails, mais pour en identifier la complexité et faciliter la compréhension de la stratégie de l'entreprise.

Les principaux critères définissant la complexité du marché de NAFTAL SPA sont résumés dans le schéma suivant :

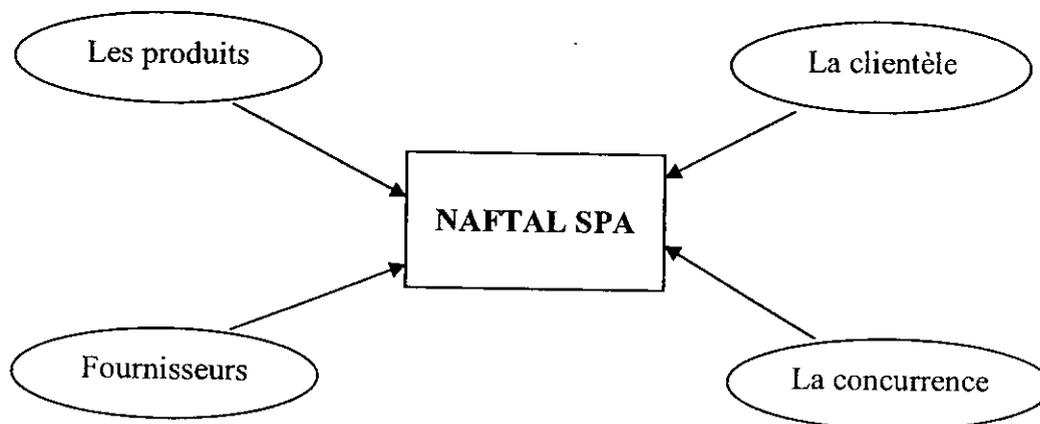


Fig -24- Schéma de l'environnement de Naftal Spa

3.1. Les produits :

Naftal spa est chargée dans le cadre du plan national de la commercialisation et de distribution des différents produits pétroliers qui sont classés selon les branches :

3.1.1. Carburants :

La branche carburants dans ses activités aviations et marine est certifiée ISO 9001/2000 par l'organisme belge AIB VIN90TTE international SA depuis le 22 août 2005.

A travers cette certification et dans le cadre de l'amélioration continue, les activités aviations et marine assureront des prestations de meilleure qualité en vue d'accroître encore plus la satisfaction de ses clients.

La qualité des produits commercialisés et leurs conformités par rapport aux spécifications internationales est assurée par des contrôles permanents.

Naftal spa commercialise trois types de carburants à savoir :

a/ **Carburant terre** : avec une capacité de stockage de 708.130 m³, il y a :

- Essence normale
- Essence super
- Essence sans plomb, et
- Gasoil

b/ Carburant aviation : avec une capacité de stockage de 50.500 m³, il y a :

- Le carburateur JET-A1 (kérosène)
- L'essence avion AVGAS-100LL
- METHMIX

c/ Carburant marine : Avec une capacité de stockage de 91.000 m³, il y a :

- Le fuel-oil Bunker C
- Le fuel-oil BTS
- Les inters fuel-oils de différentes viscosités
- Le gas-oil (DMA)

3.1.2. Lubrifiants :

Naftal Spa à travers son réseau de distribution déployé sur tout le territoire national, commercialise une gamme complète de lubrifiants qui couvre toutes les applications du secteur automobile et industriel, répondant aux exigences d'une clientèle professionnelle et rigoureuse.

Les produits lubrifiants commercialisés par Naftal spa sont :

- Les huiles moteurs (moteurs a essence et moteurs diesel)
- Les huiles de transmission automobile
- Les graisses (pour automobiles, graisse vaseline blanche, graisse spéciales).
- Les huiles industrielles (huiles de mouvements, huiles hydrauliques, huiles pour engrenage sous carter, huiles pour engrenages découverts, huiles blanches, huiles blanches, huiles solubles, huiles de protection)

3.1.3. Pneumatiques :

Naftal, en s'alliant aux plus grandes marques de fabricants de pneumatiques telles que : Good-year, semperit, uniroyal, continental..., assure des pneumatiques de premier choix garantis contre tous vices de fabrication.

Naftal propose des pneumatiques dans les catégories :

- Tourisme
- Camionnette
- Poids lourds
- Génie civil
- Agraire
- Manutention
- Industriel.

3.1.4. Bitumes :

Naftal spa commercialise à partir de ses centres trois formes de bitumes :

➤ **les bitumes purs :**

Les bitumes purs sont obtenus par la distillation sous vide du pétrole brut suivi d'un soufflage à l'aire. Utilisés dans les domaines routiers pour la construction et l'entretien des chaussées.

Les bitumes purs sont également utilisés comme matière premier de base pour la formulation des bitumes fluidifiés (cut-backs), des bitumes oxydés et des émulsions de bitumes

➤ **les bitumes oxydés :**

Les bitumes oxydés sont obtenus par soufflage d'air d'une charge bitumineuse. Naftal commercialise le bitume oxydé 85/25 (le 85 représente le point de ramollissement et le 25 la pénétrabilité a l'aiguille).

Les bitumes oxydés sont utilisés pour l'étanchéité multicouche, pour l'isolation thermique et phonique et pour la protection anticorrosion des ouvrages d'art.

➤ **Les bitumes fluidifiés ou CUT-BCKS :**

Les bitumes fluidifiés ou cut-backs sont formulés dans les centres bitumes de Naftal Spa, ils sont obtenus en fluidifiant les bitumes purs avec du kérosène.

- ❖ 0/1 (cut-back très fluide)
- ❖ 150/250 (cut-back visqueux)
- ❖ 400/600 (cut-back tiré visqueux)

3.1.5. GPL :

Comme son nom l'indique, Naftal spa distribue et commercialise les Gaz du pétrole liquéfiés, qui sont des gaz extraits à partir du pétrole pendant l'opération du raffinage, et stockés dans l'état liquide. Cela au contraire du gaz naturel qui est un gaz propre, non extrait du pétrole (se trouvant dans des gisements de gaz).

Les produits GPL distribués et commercialisés par Naftal sont :

- propane (vrac ou conditionné)
- Butane (vrac ou conditionné)
- GPL/C : un mélange de propane commercial et de butane commercial.

Notons que les produits GPL sont des produits standardisés et universels qui ne nécessitent aucune personnalisation. Ils sont utilisés essentiellement comme source d'énergie et entrent dans la composition chimique de certains produits industriels.

Vente et commercialisation des produits de Naftal Spa :

Naftal a commercialisé en 2005 :

- ❖ 7.300.000 tonnes de Carburants.
- ❖ 1.800.000 tonnes de GPL.
- ❖ 80.000 tonnes de lubrifiants.
- ❖ 340.000 tonnes de bitumes.
- ❖ 23.000 tonnes de produits spéciaux.
- ❖ 200.000 unités pneumatiques.

Gamme de produits :

Naftal Spa offre une gamme de produits diversifiée et complète qui permet de répondre aux besoins de tous les secteurs d'activité économique ; elle comprend notamment les huiles automobiles (moteur essences et diesel et huiles de transmission), les huiles industrielles, les huiles aviation, les huiles marines et les graisses et paraffines.

Dans le but d'enrichir davantage sa gamme de produits en y intégrant des lubrifiants de haute technologie, Naftal Spa a programmé la mise sur le marché de lubrifiants à base synthétique de dernière génération dès le premier trimestre 2008 ; ce lubrifiant sera commercialisé sous l'appellation NAFTILIA SYNT.

Spécification technique des produits :

Les lubrifiants automobiles commercialisés par Naftal spa ont toujours été d'un niveau de qualité adapté au parc automobile Algérien. Cependant et dans le souci d'amélioration continue de la qualité de ses produits, Naftal Spa a augmenté, depuis le début de l'année 2003, le niveau des spécifications de ces lubrifiants pour satisfaire aux exigences internationales les plus sévères.

Pour les autres catégories de lubrifiants, les spécifications répondent aux normes internationales les plus récentes et peuvent, en cas de besoin, être déterminées en collaboration avec le client pour satisfaire ses besoins spécifiques les plus stricts. A ce sujet, il est noter que les spécifications des lubrifiants aviation, marine et ceux utilisés dans le transport ferroviaire sont fixés conjointement avec les sociétés (SNTF, Air Algérie, etc.....) opérant dans ces secteurs.

Notre première étude nous a permis de constater que Naftal Spa répond aux normes de sécurité internationales. L'entrepris se soucie de la sécurité de ses clients, bien que Naftal Spa soit sur de la meilleure qualité de ses produits et de la confiance que lui témoigne sa clientèle dans ses domaines.

3.2. La clientèle : la clientèle de Naftal spa est selon :

a/ Le produit consommé :

Pour les produits GPL, on trouve les clientèles butane ainsi les clientèles propane ; plus de 90% du butane est consommé par les ménages, le reste est destiné pour l'artisanat et l'industrie chimiste. Or, le client par rapport à Naftal Spa n'est pas le consommateur final du produit, mais le revendeur chez qui ce consommateur s'approvisionne.

Année	1962	1972	1982	1992	2002	2004
Butane TM	42.000	211.000	824.700	1.400.000	1.362.800	1.420.000
Propane TM	9.500	14.000	31.600	50.300	109.400	140.000
Total GPL TM	51.500	225.000	856.300	1.482.800	1.726.000	1.860.000

Tab-3- Evolution des consommations GPL

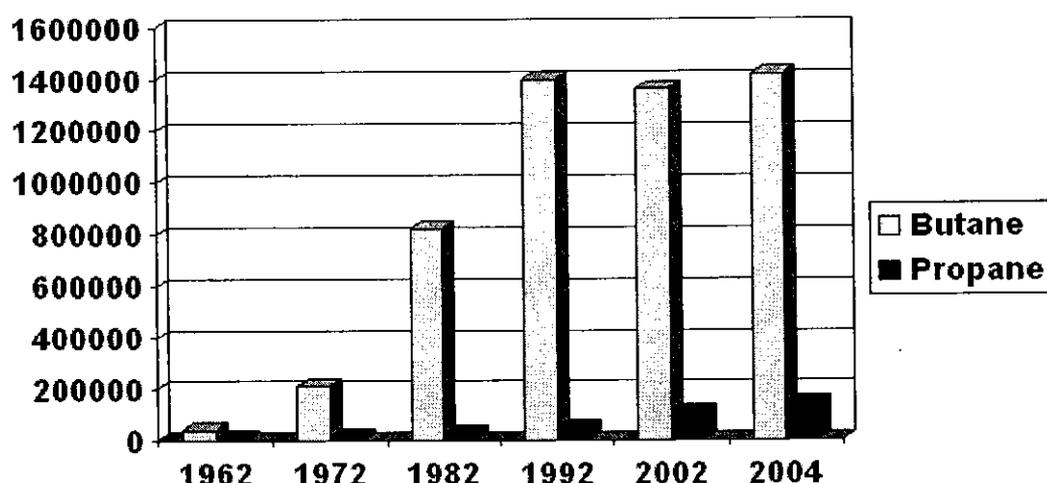


Fig -25- L'évolution des consommations GPL

Pour les autres produits, les clients sont les points de vente suivants :

Le client	signification	Nombre total
PVS	Point de Vente Spécialisé (centre de Naftal)	4060
PVO	Point de Vente Ordinaire (station service privée)	8123
PVA	Point de Vente Agréé (revendeur privé)	
GD	Station service Naftal à Gestion Direct	904
GL	Station service Naftal à Gestion Libre	1176
ASR	Aire de Stockage Régulateur	345

Tab-4- Les points de ventes

b/ Les segments de marché :

Les clients sont classés selon les 5 segments de marché suivants :

- 1- Domestique : cuisine, chauffage, avec une consommation de 7.100 tm
- 2- Industrie : fours, soudure, composants chimique, avec une consommation de 111.190 tm
- 3- Collectivités : hôtels, restaurants, bureaux, hôpitaux, casernes, avec une consommation de 18.420 tm.
- 4- Agriculture: chauffage et séchage (élevage), avec une consommation de 3.300 tm.
- 5- Transport : pour l'utilisation du GPL/C, avec une consommation de 300.000 tm.

c/ Le secteur d'activité :

Une troisième classification des clients est selon le secteur d'activité, on trouve six secteurs :

- A. administration centrales : présidence, ministères, forces terrestres, forces aériennes, forces navales, autres corps du ministre de la défense nationale MDN.
- B. Administrations locales : wilaya, daïra, APC, santé, Enseignement, justice,.....
- C. Intra group SONETRACH: Sonatrach, Sonatrach EGZ, NAFTEC,
- D. Entreprises publics : air Algérie, CNAN, SNTF, SNTR, SONALGAZ,...
- E. Entreprises privées : KHALIFA airways, ANTINEA airlines, SAHARA airlines, entreprises privées aviation et marine, clients étrangers aviation et marine,....
- F. Autre secteurs inconnus.

3.3. Fournisseurs :

Naftal Spa a comme fournisseur principal : SONATRACH et filiale NAFTEC qui assure la production des produits pétroliers au niveau des raffineries du pétrole (arzew. Skikda).

L'acheminement vers les centres de Naftal Spa se fait par :

- pipe-line.
- Cabotage (bateau).
- Camions citernes.

Les décideurs de Naftal Spa se déclarent totalement satisfaits de leurs fournisseurs, qui n'exercent aucune pression sur la division.

3.4. La concurrence :

A nos jours, il n'existe pas une politique algérienne sur l'ouverture des portes de la concurrence dans le domaine pétrolier. En effet, l'entreprise n'a que des petits concurrents locaux sur quelques produits (pneumatiques).

Les produits de remplacement sont les autres sources d'énergie, à savoir : l'électricité, le gaz naturel, les carburants (essence, gasoil,). Cela sous entend que les concurrents de Naftal sont les entreprises du secteur d'énergie. Or, les dirigeants de l'entreprise considèrent ces produits comme des compléments et non pas des concurrents de leurs produits. D'ailleurs, l'utilisation des produits se fait, souvent, en parallèle avec d'autres sources d'énergie comme un outil de secours ou dans le cas où l'utilisation des autres produits pose des problèmes.

Selon ses dirigeants, une entreprise concurrente à Naftal est celle qui répondra le processus d'activité de Naftal dans sa totalité : elle doit acheter directement les produits de chez NAFTEC (ne pas être livrée par Naftal) assurer le stockage, le conditionnement et la distribution avec ses propres moyens et s'implanter sur une grande surface du territoire national.

4. La politique interne de l'entreprise :

Les nouvelles mutations qui s'opèrent tant sur le plan national qu'international ont imposé à Naftal spa de définir une politique et des stratégies à l'effet d'adapter progressivement ses activités aux conditions de fonctionnement d'un marché libre et ouvert à la concurrence.

Pour ce faire, elle s'est inscrite dans cette dynamique et se prépare à affronter les règles de la concurrence et se rapprocher progressivement aux niveaux des standards des grandes entreprises de distribution des produits pétroliers.

L'ouverture du marché national de la distribution des lubrifiants depuis 1995 a amené Naftal Spa à revoir profondément son comportement sur le marché et à élaborer une nouvelle stratégie qui lui permettra de s'adapter à ce nouvel environnement économique fortement concurrentiel.

4.1. La stratégie adoptée par Naftal Spa :

Naftal Spa a défini et mis en place une stratégie dont les grandes axes portent sur :

- La mise en place d'une nouvelle organisation alignée sur les standards internationaux s'inscrivant pleinement dans les réformes du secteur de l'énergie et des mines.
- La mise en œuvre d'un programme ambitieux de développement des ressources humaines et ce, par une valorisation et une présentation des compétences actuelles, ainsi que la préparation d'une relève de haut niveau dans tous les domaines d'activités de la société.
- La mise en place d'une stratégie de marketing et de communication afin de répondre aux attentes de la clientèle et pour la fidéliser.
- Une exploitation de toutes les opportunités offerts par la mondialisation tant au niveau national qu'international.
- Le renforcement de sa position de leader de la distribution des produits pétrolier, au plant national
- L'intensification des actions visant la maîtrise, la réduction des coûts et l'amélioration de la qualité des produits et des services.

- La pénétration des marchés extérieurs tout en axant ses efforts sur les marchés maghrébins et africains.
- La rationalisation de son portefeuille en donnant la priorité au développement et à la croissance sélective des produits forte valeur ajoutée.
- Le développement d'une politique HSEQ pour améliorer les performances dans les domaines de l'hygiène, la sécurité, de l'Environnement et la qualité.
- Le développement des énergies propres renouvelables et de nouveaux produits et services.

4.2. Les projets de Naftal Spa :

Naftal, en particulier la branche GPL, développe, rénove et modernise ses infrastructures de stockage, d'emplissage et de transport au profit des utilisateurs.

A/ Projets en cours de réalisation :

➤ Pipes GPL :

- Skikda/khroub : avec une capacité de 923000 tm/an, longueur de 105 km
- Arzew/Sidi Bel abbes/Telemcen : avec une capacité de 577000 tm/an et longueur de 165 Km

➤ stockage :

Extension du centre GPL vrac de khroub en 2 sphères de 2000 tm.

➤ Enfûtage :

- Centre enfûteur GPL d'Arzew, capacité : 50000 tm/an.
- Centre enfûteur GPL de Berrahal (Annaba), capacité : 50000 tm/an.
- Centre enfûteur GPL de M'sila, capacité : 25000 tm/an.

B / Projets en perspective :

➤ Pipes GPL :

- Arzew/Blida : capacité : 1385000 tm/an, longueur : 600 km.

➤ Modernisation des centres GPL vrac :

Blida, Sidi Arcine (Alger), Skikda et Tlemcen (télé jaugeage et chargement automatisé).

Conclusion :

Nous avons essayé à travers cette étape d'étudier l'environnement le l'organisme d'accueil. Notre constat est que Naftal Spa, tout comme la plupart des entreprises algériennes, est instable, vu que l'organigramme général de l'entreprise change constamment.

Un entrepôt de données est nécessaire aux décideurs pour bien piloter leurs activités de base. Car avec ce qui se fait actuellement, l'avenir de celle-ci reste très risqué en cas d'arriver de nouveaux concurrents. Donc, elle doit décider et anticiper en fonction de l'information disponible et capitaliser sur ses expériences.

Pourtant, les données représentent une mine d'information. Il devient fondamental de rassembler et d'homogénéiser les données de Naftal afin de permettre d'analyser les indicateurs pertinents pour faciliter les prises de décisions.

Notre entrepôt de données se limite à la gestion des créances des clients de différents types de Naftal Spa.

Au cours de l'étape prochaine, nous allons identifier l'ensemble des besoins de décideurs de Naftal, comme point important placé au centre de l'univers des entrepôts de données.

Chapitre III :

Identification des besoins

Introduction :

Après avoir présenté l'entreprise Naftal Spa, dans sa globalité, nous pouvons commencer la construction de l'entrepôt de données. Comme nous l'avons déjà précisé, la première étape dans l'approche globale de l'implémentation d'entrepôt de données par le cycle de vie, est l'identification des besoins.

Les constructeurs d'entrepôts de données doivent appréhender les facteurs clés qui conduisent l'entreprise à vouloir définir de manière efficace ses besoins et les traduire pour les intégrer lors de la conception. Ces besoins constituent le point de départ des trois trajectoires parallèles que sont la technologie, les données et les interfaces utilisateur.

La première phase, pour notre projet de construction d'un entrepôt de données, consiste à étudier la stratégie de l'entreprise adaptée pour le pilotage de créances de ses différents clients. Cette étude nous permet de dresser la liste des indicateurs et données nécessaires pour la réalisation de cette fonction.

Avant d'entrer dans le vif du sujet, à savoir identifier les besoins en informations de Naftal Spa pour le pilotage de ses créances à travers une démarche data warehouse, il convient de donner une idée générale sur la procédure de gestion des créances suivies par les agents de Naftal Spa.

1. La gestion des créances de Naftal Spa :

La gestion des créances des clients de Naftal Spa se fait par niveaux. On trouve trois branches (GPL, Commerciale, Carburant) chacune est divisée en districts (presque un district par wilaya) et chaque district est responsable de la gestion de centres de services, chaque centre a ses propres clients.

Au niveau de chaque district ; la gestion des créances se fait par un SGC (Système de gestion de créances). C'est un système de gestion informatisé qui permet le suivi des enlèvements et les règlements des clients à terme.

Si un client, pour une raison ou une autre, ne paye pas, son compte est bloqué automatiquement et son dossier est transmis au service juridique. La créance que présente ce client est dite contentieuse. Par suite, ce client peut négocier pour avoir une échéance sur sa créance qui est dite alors une créance échéanciée.

Le client se met en accord avec le représentant de Naftal sur les tranches à payer et la durée de chaque tranche (soit un échéancier de 15, 21, 30, 45, 60 ou bien 90 jours comme le plus grand échéancier à avoir).

Afin de faire face à ces situations et pour mieux répondre à ces exigences complexes, la mise en œuvre d'un data warehouse métier pour la fonction des créances semble nécessaire.

2. Définition des besoins :

L'étude des besoins doit déterminer le contenu de l'entrepôt et son organisation, d'après les résultats attendus par les utilisateurs, les requêtes qu'ils formuleront, les projets qui ont été définis. Le besoin d'informatisation peut provenir du système de pilotage ou d'un service particulier de l'entreprise. L'information est dupliquée, les traitements répétés et aucune stratégie d'ensemble n'est définie. Le projet va donc s'orienter vers la recherche d'une solution intégrante résolument tournée vers l'utilisateur.

Pour mieux cerner les besoins des décideurs de Naftal Spa (pour l'activité créances), nous avons organisé des entretiens avec ceux concernés par notre nouveau système. L'objectif de nos interviews est de permettre de préciser les faits à suivre et dans quelles dimensions. Il faut alors recenser les données disponibles dans les bases de production, les données de production ne sont pas toutes utiles dans l'entrepôt. Il faut aussi identifier les données supplémentaires requises.

Avant de commencer à fixer les questions que nous allons poser aux responsables de l'entreprise et pour mieux conduire les entretiens, il nous semble nécessaire d'examiner les documents et les rapports d'analyse utilisés par les différents services, mis à notre disposition et concernant les créances clients.

2.1. Les documents :

- Rapport d'activité sur la situation des créances : ce rapport est établie chaque mois par la direction exécutive Stratégie / Planification & Economie (SPE)
- Analyse des créances : ce rapport est établi chaque mois par la Direction Exécutive Finances (DEF)

2.2. Les entretiens :

Les entretiens permettent de recueillir des informations à l'aide de questions posées verbalement aux analystes et aux décideurs. Ils sont la meilleure source d'informations qualitatives, ils permettent de connaître les opinions, les commentaires, les idées et les suggestions.

Les buts recherchés de ces questions sont :

- ✓ D'en savoir plus sur les principales activités pilotées par les dirigeants interrogés ;
- ✓ De bien comprendre les métriques d'évaluations de la réussite utilisées qui mettront en évidence les données qui devront être stockés dans l'entrepôt ;
- ✓ Et bien sur pour bien préciser les attentes des dirigeants.
- ✓ D'avoir une idée sur le type d'analyse effectué dans l'entreprise concernant les créances, ainsi que le genre d'analyse souhaité par les décideurs ;

- ✓ La compréhension des différentes relations entre les attributs essentiels de l'entreprise ;
- ✓ La détermination des dimensions et des faits de mesure qui nous permettront d'établir la modélisation dimensionnelle.
- ✓ Avoir une idée sur la définition des champs, le niveau de détail des données, leurs volumes et leurs localisations.
- ✓ Rassurer de la disponibilité des données en réponse aux besoins des utilisateurs.

2.3. Les besoins recensés :

Après avoir fait les entretiens auprès des responsables, des analystes et des audit de données, on a pu rassembler les différentes informations et données existantes pour la définition des besoins des décideurs, savoir le type d'analyse demandé et souhaité par les analyses pour répondre aux besoins déclarés par les responsables, et bien sur on s'est rassuré de l'existence des données nécessaires pour effectuer ces analyses et la possibilité d'ajouter les données manquantes.

Comme résultat des entretiens faits, on est arrivé à bien préciser les objectifs des responsables, les activités et processus clés des analystes et les indicateurs d'analyse des performances concernant le pilotage des créances de Naftal Spa.

Le grand souci des dirigeants et décideurs centraux de Naftal Spa est d'arriver à réduire la créance au minimum, cet objectif est d'une importance capitale pour l'entreprise et représente une nécessité primordiale, et afin d'atteindre ce but il est indispensable de :

- ✓ Avoir un suivi périodique de la créance, en d'autre terme, les dirigeants veulent savoir à chaque mois, la situation des créances (échue, non échue, antérieure) pour prendre la décision qu'il faut au temps qu'il faut.
- ✓ Optimiser l'activité de recouvrement, c'est-à-dire améliorer au maximum le rendement des agents de Naftal Spa qui sont responsables de recouvrement en leur fournissant un état de classement des clients par :
 - Ancienneté de la créance : cibler les clients qui ont une ancienne créance qui n'est pas encore réglée
 - Montant de créance : cibler les clients qui ont une créance très élevée
- ✓ Avoir un suivi périodique des chèques impayés

Après avoir recensé les objectifs de l'entreprise nous allons maintenant déterminer les différentes activités clés ainsi que les indicateurs et axes d'analyse qui caractérisent le processus métier des créances clients, mais avant cela il faut d'abord définir qu'est ce qu'un processus métier.

3. Le processus métier :

« Un processus est un ensemble cohérent d'activités organisées dans le temps produisant un résultat précis et mesurable »

Un processus reçoit des objets en entrée et leur rajoute de la valeur, par le moyen de ressources, tout en fournissant des objets de sortie remplissant les besoins et les exigences (atteindre les objectifs) dans le temps.

Dans notre cas le processus métier est le pilotage des créances clients de l'entreprise Naftal Spa.

3.1. Les activités clés :

Le pilotage des créances de Naftal Spa réalisé selon les différentes activités métiers suivantes : Etat_créance, Recouvrement, Chèques_Impayés.

Etat_créance : c'est la principale activité de ce processus, elle concerne les différents états des créances de Naftal Spa pour ses différents clients.

Recouvrement : qui présente un souci majeur pour les responsables de Naftal concernant le pilotage des créances, cette activité consiste à utiliser tous les moyens de droit pour obtenir le paiement en comptant.

Chèques impayés : vu que la majorité des paiements sont effectués par les chèques, il s'est avéré important d'avoir une image sur la situation des chèques impayés, et en plus, un chèque impayé représente une créance qui risque d'être perdue.

3.2. Les indicateurs d'analyse des activités métiers :

Pour mieux comprendre leur environnement (clients, produits,.....), les utilisateurs du service créance souhaitent l'emploi efficace des informations produits par les activités précisées précédemment. Pour cela, voici les principaux indicateurs que nous avons extraits :

- Le CA réalisé par : client, catégorie, sous-secteur, secteur, produit, famille produit, date, centre (district) et par branche.
- Le montant total des créances par : client, catégorie, sous-secteur, secteur, produit, famille produit, date, centre (district), branche et par type_échéance.
- Le montant des créances échues par : client, catégorie, sous-secteur, secteur, produit, famille produit, date, centre (district), branche et par type_échéance.

- Le montant des créances antérieures par : client, catégorie, sous-secteur, secteur, produit, famille produit, date, centre (district), branche et par type_échéance.
- Le montant des créances non échues par : client, catégorie, sous-secteur, secteur, produit, famille produit, date, centre (district), branche et par type_échéance.
- Le montant des créances recouvertes par : client, catégorie, sous-secteur, secteur, produit, famille produit, date, centre (district) et par branche.
- Le montant et la situation des chèques impayés par : client, catégorie, sous-secteur, secteur, produit, famille produit, date, centre (district) et par branche.

Les ratios :

Un ratio est un rapport entre deux valeurs caractéristiques de l'activité et de la situation économique de l'entreprise. Un ratio ne doit pas s'analyser seul, il doit, si possible, en fonction des données fournies par l'entreprise, se calculer sur plusieurs années afin de permettre le suivi de l'évolution de l'entreprise, voici les principaux ratios que nous avons extraits :

- Ratio (Taux de recouvrement) par : client, catégorie, sous-secteur, secteur, produit, famille produit, date, centre (district) et par branche.
- Ratio (Le délai moyen des encaissements clients) par : client, catégorie, sous-secteur, secteur, produit, famille produit, date, centre (district), branche et par type_échéance.
- Ratio (Montant des créances / Chiffre d'affaire) par : client, catégorie, sous-secteur, secteur, produit, famille produit, date, centre (district), branche et par type_échéance.
- Ratio (Montant des chèques impayés / Montant total des créances) par : client, catégorie, sous-secteur, secteur, produit, famille produit, date, centre (district) et par branche.

3.3. La dimensionnalité métier :

Les dimensions sont les axes sur lesquels l'utilisateur portera son analyse, pour chaque processus nous avons déterminé l'ensemble des dimensions qui lui sont associées.

Comme tout entrepôt de données, la première dimension présente est le « temps ». Pour notre entrepôt de données, les autres dimensions sont : client, type_échéance, produit, centre.

Pour illustrer le principe de dimensionnalité, on prend l'exemple suivant :

Un décideur se demande : « Quel est le montant des créances pour tel client de tel CDS pour le mois dernier ? ».

Ce besoin peut être exprimé en terme de cube comme le montre la figure suivante :

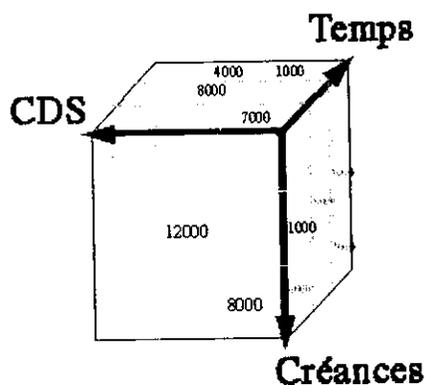


Fig -25- La représentation sous forme de cube

3.4. Les systèmes sources :

Notre entrepôt sera alimenté à partir de deux systèmes opérationnels différents qui sont :

- SGC (Système de Gestion des Créances)
- NAFTCOM (Naftal Commercialisation).

Conclusion :

Dans ce chapitre, nous avons listé les objectifs que les décideurs de Naftal Spa cherchent à atteindre, en se basant sur une analyse de leurs besoins concernant le pilotage des créances des différents clients de l'entreprise.

Nous avons utilisé une approche fondée sur l'organisation d'entretiens pour le recensement des besoins et données nécessaires pour la construction de notre entrepôt de données.

Et c'est à partir de ces besoins que nous avons extrait les processus métiers, les indicateurs et les axes d'analyse de notre modélisation dimensionnelle qui sera traitée dans le prochain chapitre.

Chapitre IV :

Conception et mise en oeuvre du système

Introduction :

Ce chapitre traite la conception logique et physique de notre système décisionnel. Nous allons commencer par expliquer la démarche utilisée pour la conception du notre entrepôt de données, ensuite nous allons passer à la modélisation dimensionnelle des sujets précédemment sélectionnés et enfin la mise en œuvre de l'architecture du système

1. La démarche de la conception du Data Warehouse :

Consiste à présenter un modèle idéal, en tenant compte des besoins recensés et des données disponibles. Pour présenter un modèle idéal, il faut d'abord avoir une description complète de l'environnement de l'organisation, permettant la conception adéquate de l'entrepôt de données dimensionnel.

Une fois dans la conception nous devons toujours nous demander si nous répondons avec la meilleure efficacité possible aux besoins les plus importants de l'organisation.

Ralph Kimball propose une démarche qui se résume en neuf décisions dans un ordre bien défini, qui sont à la base pour la conception complète d'un entrepôt de données dimensionnel, ces décisions portent sur les besoins suivants :

1.1. Choisir les processus d'activité à modéliser :

On commence par identifier les processus majeurs de l'entreprise dans lesquels les informations sont collectées au profit de l'entrepôt de données à partir d'une application existante (ou des applications). Des exemples de processus d'activité sont les stocks, les comptes clients, les ventes et la comptabilité. Une fois les processus identifiés, une ou plusieurs tables de faits sont construites à partir de chacun des processus choisis.

1.2. Choisir le grain de chaque table de faits :

Le grain est le niveau de détail fondamental, atomique, des figurantes dans la table de faits, autrement dit c'est la signification précise d'un enregistrement du plus bas niveau dans la table de faits. Des grains typiques sont la transaction individuelle, la ligne d'articles individuels, des récapitulatifs individuelles quotidiennes ou mensuelles, un instantané journalier, ou instantané mensuel.

1.3 Choisir les dimensions de chaque table de faits :

Lorsque le grain de la table de faits est connu, les dimensions et leurs grains respectifs peuvent être identifiés. Des dimensions typiques sont le temps, le produit, le client. Le choix des dimensions s'accompagne de la définition de tous les attributs textuels (les champs) qui garniront la table de dimension.

Les tables de dimensions servent à enregistrer les descriptions textuelles des dimensions de l'activité, c'est les attributs, ils sont utilisés pour décrire des aspects d'une dimension qui peuvent être utilisés comme source de contraintes par les utilisateurs.

1.4 Choisir les faits mesurés que contiendra chaque enregistrement de table de fait :

Des faits mesurés typiques sont des quantités additives telles que la quantité vendue ou les montants de vente, le chiffre d'affaire.

1.5 Les attributs des dimensions :

Avec des descriptions complètes et la terminologie adéquate.

A ce stade, la conception de la structure logique principale est terminée, les autres étapes concernent plus généralement la structure physique.

1.6 Comment suivre les dimensions à évolution lente :

Dans la réalité, les dimensions ne sont pas indépendantes les unes par rapport aux autres. En particulier, comme exemple les dimensions produit et client, ne sont pas indépendantes de la dimension temps, sur la durée, les descriptions et la formulation des produits réels, subissent de lentes évolutions. Le client, notamment, change constamment. Les humains changent de nom, se marient et divorcent, ont davantage d'enfants et changent d'adresse. Les équipes de vente changent périodiquement... etc.

Nous appelons ces dimensions quasi constantes, des dimensions à évolution lente. Nous devons décider de quelle manière nous allons traiter ces changements, ce choix se résume entre trois solutions, chacune assurent un suivi plus ou moins strict des modifications dans le temps :

Premier type : recouvrir et perdre les valeurs anciennes dans l'enregistrement de la dimension, renonçant ainsi à la possibilité de suivre les événements ou situations passées.

Deuxième type : Créer un enregistrement de dimension supplémentaire lors du changement, comptant les nouvelles valeurs de l'attribut, ce qui revient à segmenter l'historique très exactement selon l'ancienne et la nouvelle description.

Troisième type : Créer de nouveaux champs « actuel » à l'intérieur de l'enregistrement d'origine de la dimension, tout en conservant en même temps les valeurs enregistrées en premier lieu. Ceci permet de décrire l'historique en amont du changement en utilisant les valeurs d'origine, et de décrire l'historique aval selon les nouvelles valeurs des attributs.

1.7. Les agrégats :

Les dimensions hétérogènes, les minidimensions, les modes de requêtes et autres décisions sur le stockage physique.

1.8. L'étendue historique de la base de données :

On peut permettre à l'historique de s'accumuler au-delà des trois ans envisagés pour n'importe quel entrepôt de données.

1.9. L'urgence avec laquelle les données doivent être extraites et chargées dans l'entrepôt de données :

Plus généralement le chargement se fait chaque jour les premières heures du matin, on peut décider de le faire à la date de la veille, ou attendre jusqu'à la fin de la semaine.

Enfin, nous devons garder à l'esprit qu'une bonne conception ne peut que résulter d'un équilibre entre les besoins des utilisateurs finaux (déterminés par leur interview ou autre technique de collectes de données) et les réalités des données disponibles à partir des applications existantes.

2. La modélisation dimensionnelle des activités métiers du pilotage des créances :**2.1. Modélisation dimensionnelle de l'activité « Etat_Créance » :****2.1.1. Le processus d'activité :**

La première activité importante choisie pour la création de l'entrepôt de données est l'activité « Etat_Créance » qui est l'activité clé des créanciers qui leur permet d'avoir une vue générale sur l'état des créances.

Les décideurs auront besoin de savoir le montant total des créances, et pour s'assurer qu'ils continuent d'avoir les capacités nécessaires pour gérer le crédit accordé aux clients, ils doivent savoir :

- a) Le taux de croissance des créances par rapport au chiffre d'affaire mensuel qui est égal à :

$$\text{RATIO1} = \frac{\text{Montant Total des Créances}}{\text{Chiffre d'Affaire mensuel}}$$

- b) Le délai moyen de paiement des clients qui est exprimé en nombre de jours et qui est égal à :

$$\text{RATIO2} = \frac{\text{Montant Total des Créances} \times 30}{\text{Chiffre d'Affaire mensuel}}$$

Il est à noter que ce ratio doit être toujours inférieur au délai moyen de paiement accordé à l'entreprise Naftal Spa, par ses fournisseurs, qui est égal à 45 jours.

2.1.2. Le grain du processus d'activité :

Le grain est important car il détermine les différentes dimensions de la base de données. La granularité de la table des faits choisie est, la plus fine possible. Plus le niveau de détail est fin, plus la conception est robuste.

Dans l'analyse des créances, il est important de connaître l'état des créances. Ainsi, chaque enregistrement de la table de faits est le montant de la créance par type_échéance, par client, par centre, par produit et par période. Donc les dimensions : type_échéance, centre, client, produit et temps (mois) s'imposent instantanément.

La figure suivante illustre le modèle dimensionnel de base de cette activité :

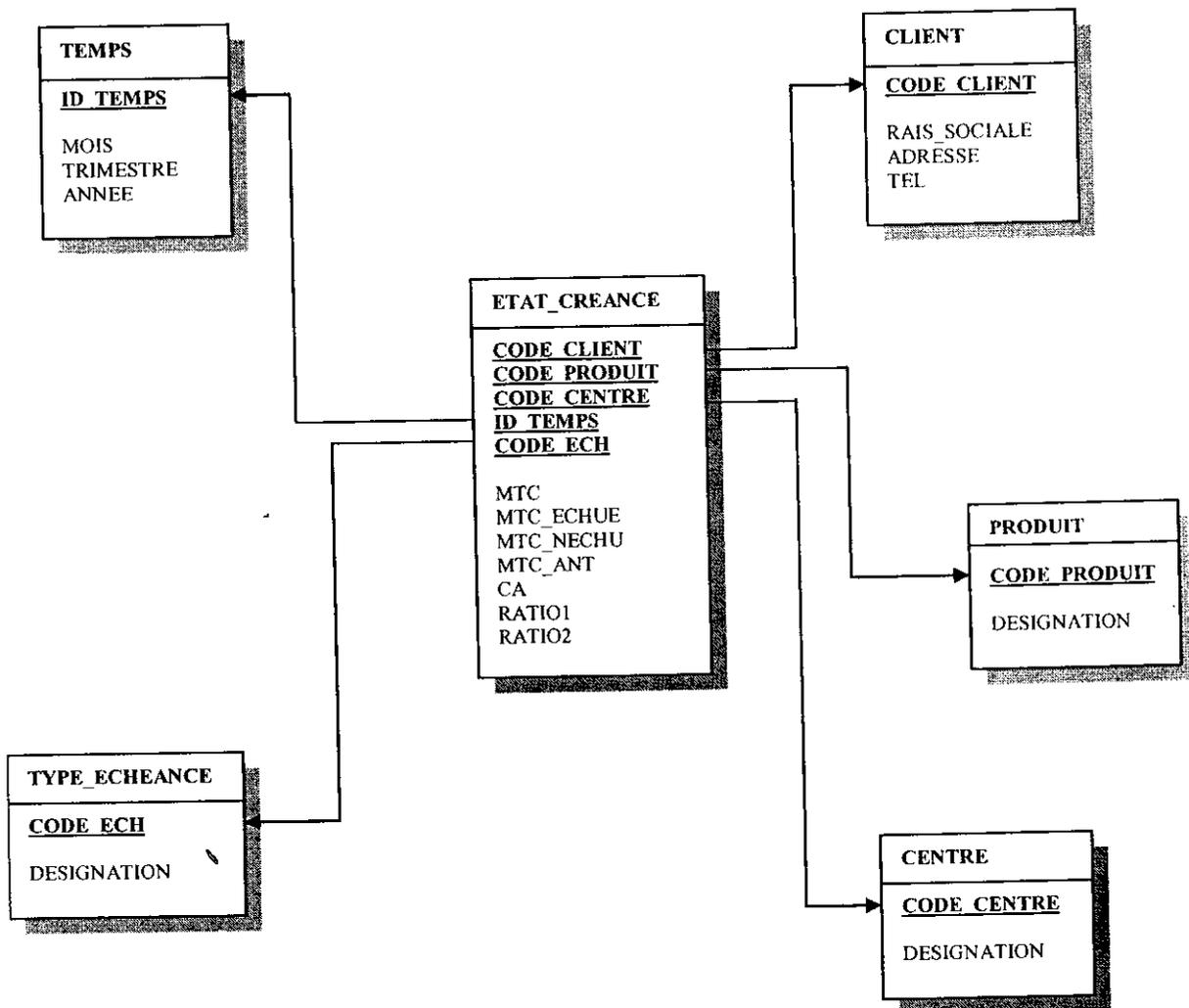


Fig-27- Schéma du modèle dimensionnel de base de l'activité « Etat_Créance »

2.1.3. Les dimensions :

D'après la granularité de la table de faits, nous avons déterminé une série de dimensions principales, qui sont : Client, Produit, Centre (ou Unité Opérationnelle : Zone, District), Type_Echéance et Temps.

2.1.3.1. Dimension « Temps » :

La définition de la dimension Temps est selon Ralph Kimball : « La seule dimension qui figure systématiquement dans tout entrepôt de données, car en pratique tout entrepôt de données est une série temporelle. Le temps est le plus souvent la première dimension dans le classement sous-jacent de la base de données ».

Pour cette activité et même les autres activités, on a opté pour un grain mensuel des données. Ainsi chacun des enregistrements de notre table de dimension Temps représente un mois.

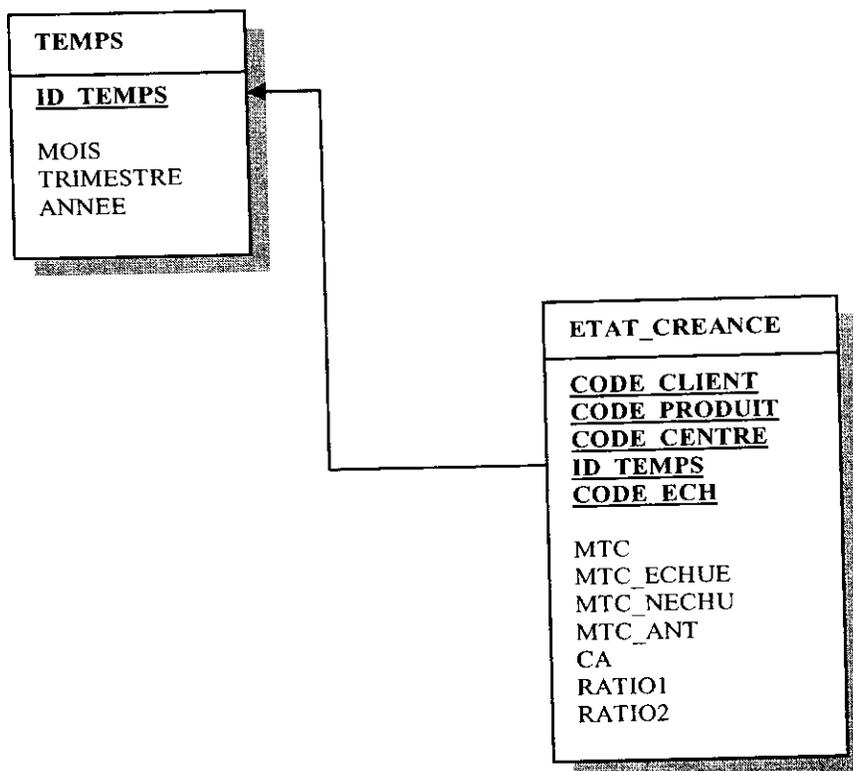


Fig-28- Dimension Temps du schéma activité « Etat_Créance »

Le champ « Id_Temps » contient la date de la créance. Le champs « mois » sert à connaître le mois, 02 pour Février. Le champ trimestre sert à avoir un suivi des créances par tranche de trois mois. Le champ « année » est utilisé pour stocker l'historique des créances annuelles et aussi pour préparer les rapports de comparaison.

2.1.3.2. Dimension « Produit » :

La dimension « produit » décrit tous les produits vendus par l'entreprise. Vu la diversité des produits pétroliers commercialisés par Naftal Spa, l'entreprise les a classés en familles.

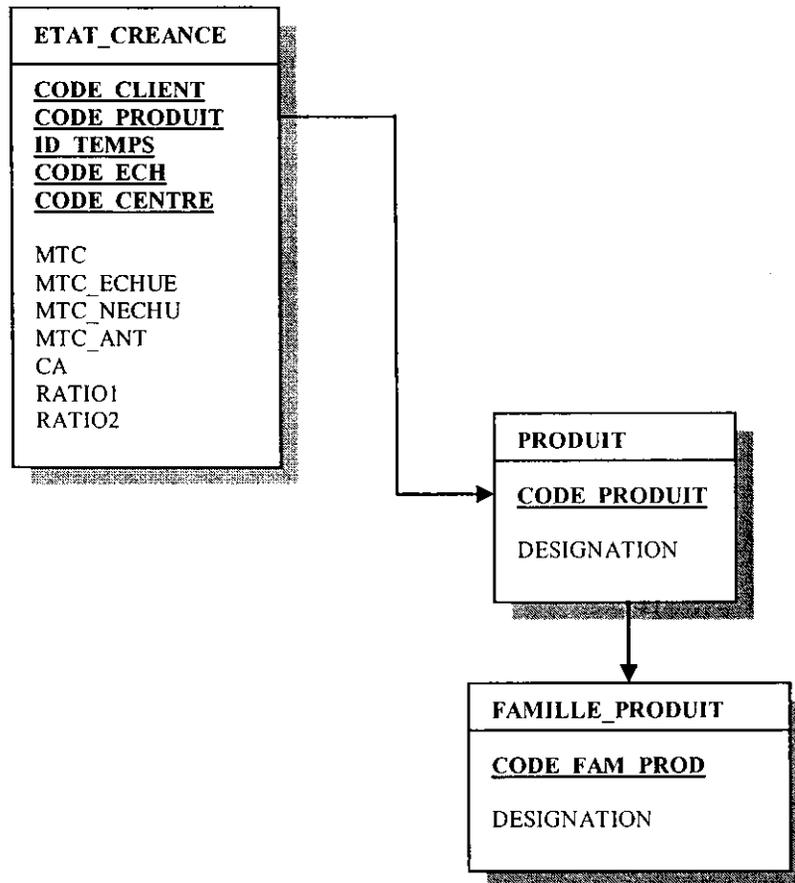


Fig-29- Dimension Produit du schéma activité « Etat Créance »

2.1.3.3. Dimension « client » :

L'entreprise a comme clients des êtres humains individuels et des organisations, des institutions, des compagnies ... etc. Ces clients sont groupés par catégorie et chaque catégorie appartient à un sous-secteur d'activité (Présidence, Ministères, Wilaya,...etc.), lui aussi appartient à un secteur d'activité (Administrations centrales, Administrations locales, Entreprise privée,...).

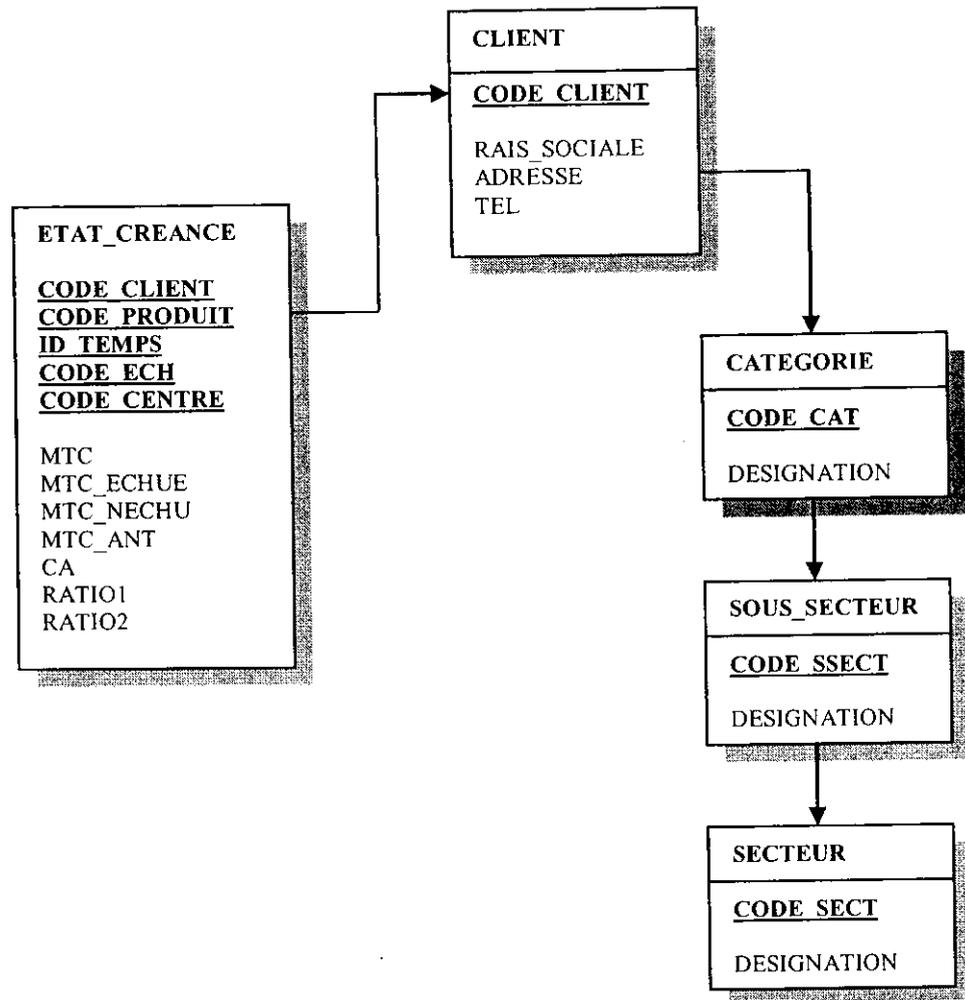


Fig-30- Dimension Client du schéma de l'activité « Etat_Créance »

2.1.3.4. Dimension « Centre » :

L'entreprise est organisée en trois branches (Carburants, Commercialisation et la branche GPL). Chaque branche est divisée en zone, et chacune est elle-même divisée en un ensemble de centre de Dépôt et Stock (CDS). C'est au niveau des CDS que les clients font leurs demandes, seront livrés et payent leurs factures.

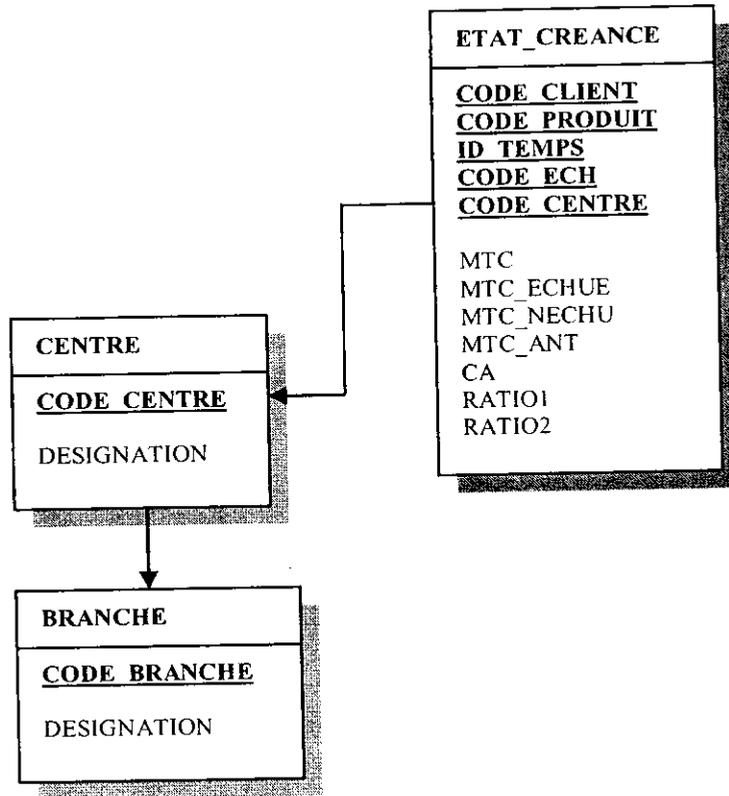


Fig-31- Dimension Centre du schéma activité « Etat_Créance »

2.1.3.5. Dimension « Type_Echéance » :

La dimension « Type_Echéance » désigne le délai de paiement accordé aux clients. Au niveau de l'entreprise, il existe plusieurs types d'échéances : 15, 21, 30,.... Jours.

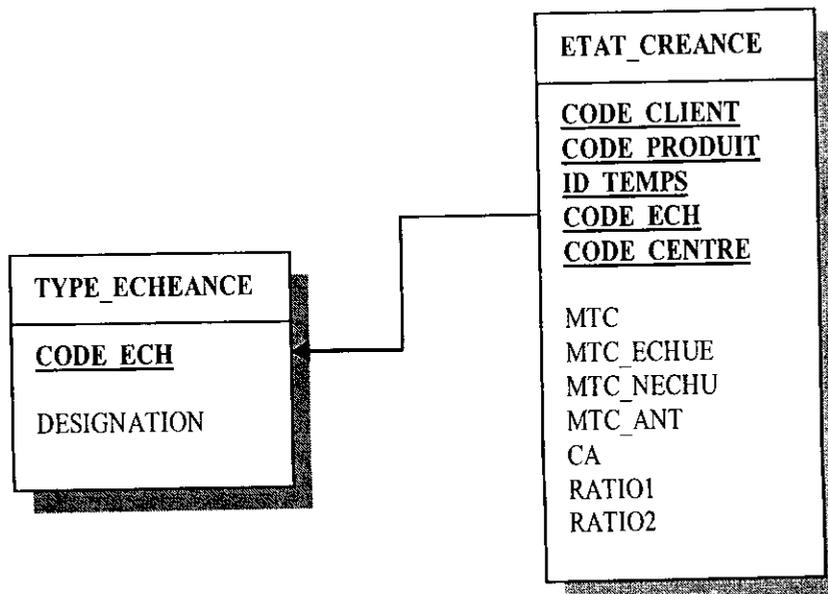


Fig-32- Dimension Type_Echéance du schéma activité « Etat_Créance »

2.1.4. Les faits mesurés :

Notre activité est mesurée par des faits mesurés que chacun d'eux doit être pris à l'intersection de toutes les dimensions de l'activité concernée. Ces faits sont définis tels qu'on peut couper la table en tranches et en dès sans risque.

Dans notre cas (pilotage des créances), il est important de connaître le montant total des créances mensuelles et annuelles, le taux de leur croissance mensuelle et annuelle, c'est pourquoi la table de faits doit contenir des indicateurs qui permettent le calcul de ces besoins.

Les faits mesurés que nous avons enregistrés sont :

- Montant total des créances « MTC ».
- Montant total des créances échues « MTC_ECHUE ».
- Montant total des créances non – échues « MTC_NECHUE ».
- Montant total des créances antérieures « MTC_ANT ».
- CA (chiffre d'affaire) réalisé (mensuel).
- Ratio1 qui sert à calculer le taux de croissance de la créance par rapport au chiffre d'affaire mensuel.
- Ratio2 : ce ratio permet de répondre à la question : en moyenne, à combien de jours les clients paient-ils ?

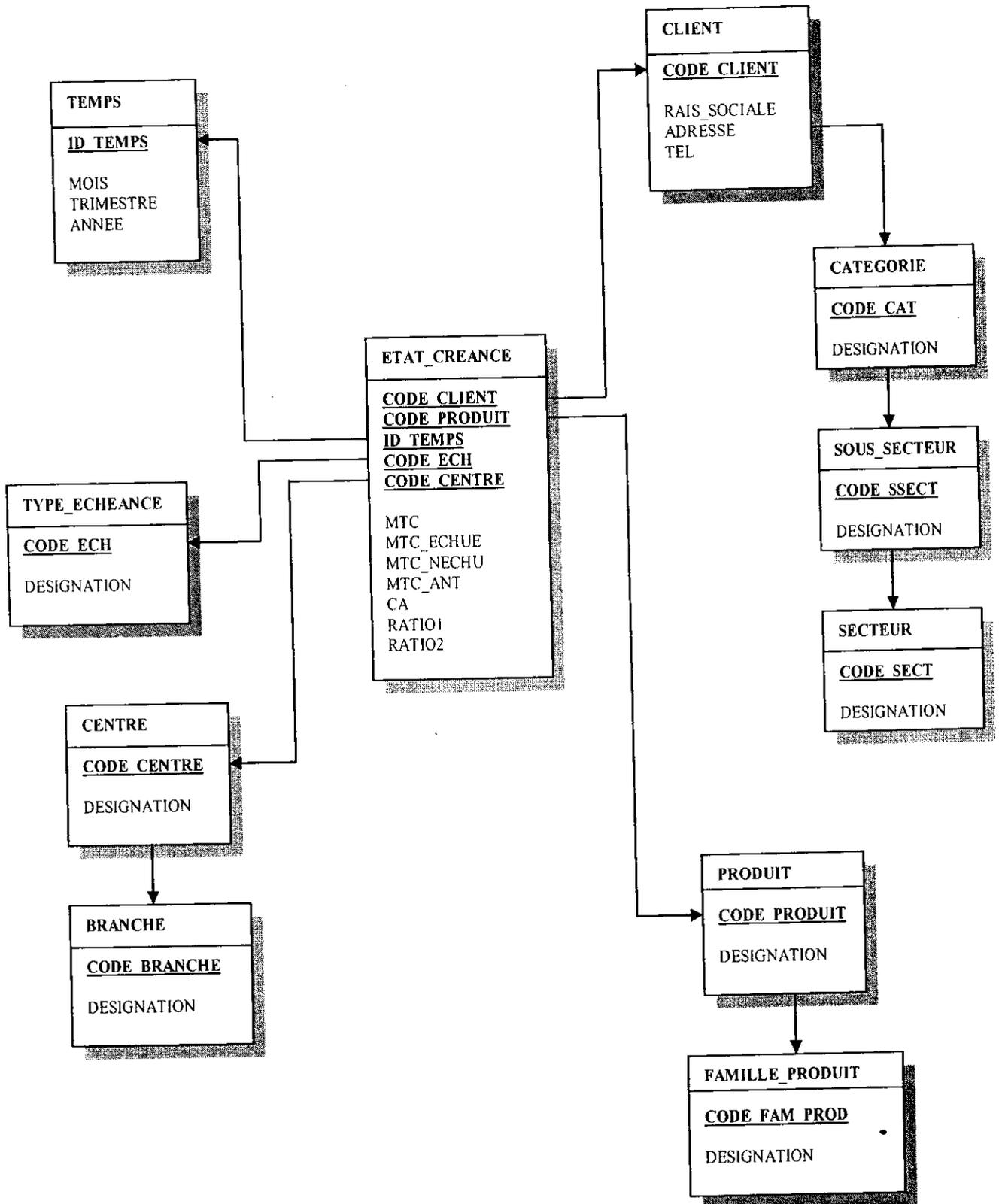


Fig-33- Modèle dimensionnel en flocon de l'activité « Etat_Créance »

2.2. Modélisation dimensionnelle de l'activité « Chèques_Impayés » :

2.2.1. Le processus d'activité :

La deuxième activité choisie pour la construction de notre entrepôt de données est l'analyse des chèques impayés. Pour avoir une vue sur la situation des chèques impayés les décideurs auront besoin de savoir le montant des chèques impayés. Mais cet indicateur peut ne pas donner l'image complète des créances non payées. D'où le calcul du portefeuille classé à risque qui est le rapport entre le montant des chèques impayés et le montant total de la créance :

$$\text{Ratio3} = \frac{\text{Montant total des chèques impayés}}{\text{Montant total des créances}}$$

2.2.2. Le grain du processus d'activité :

Dans ce processus, il est important de connaître le montant total des chèques impayés ainsi que leurs taux de croissance par rapport au montant total des créances, par centre, client, produit, pour chaque mois.

La figure suivante illustre le modèle dimensionnel de base de cette activité :

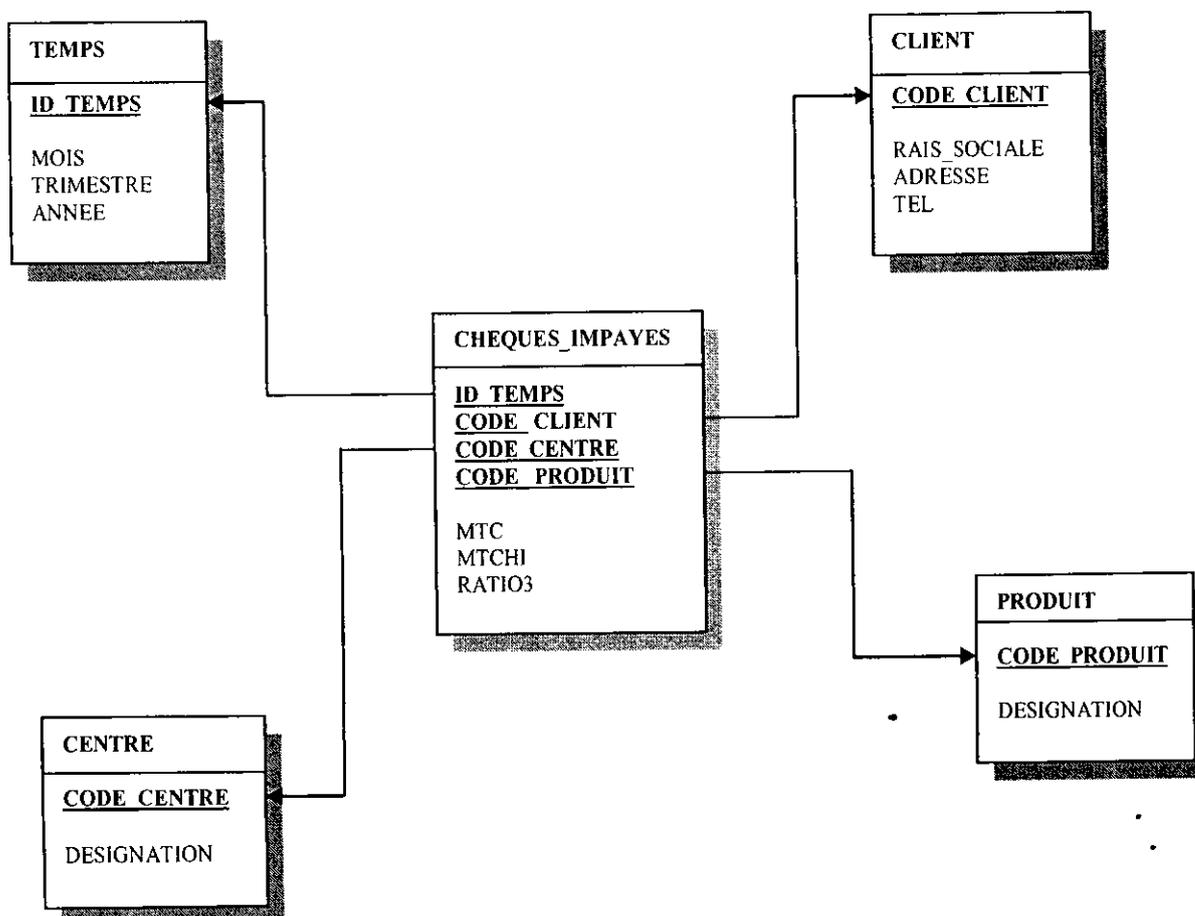


Fig-34- Schéma du modèle dimensionnel de base de l'activité « Chèques_Impayés »

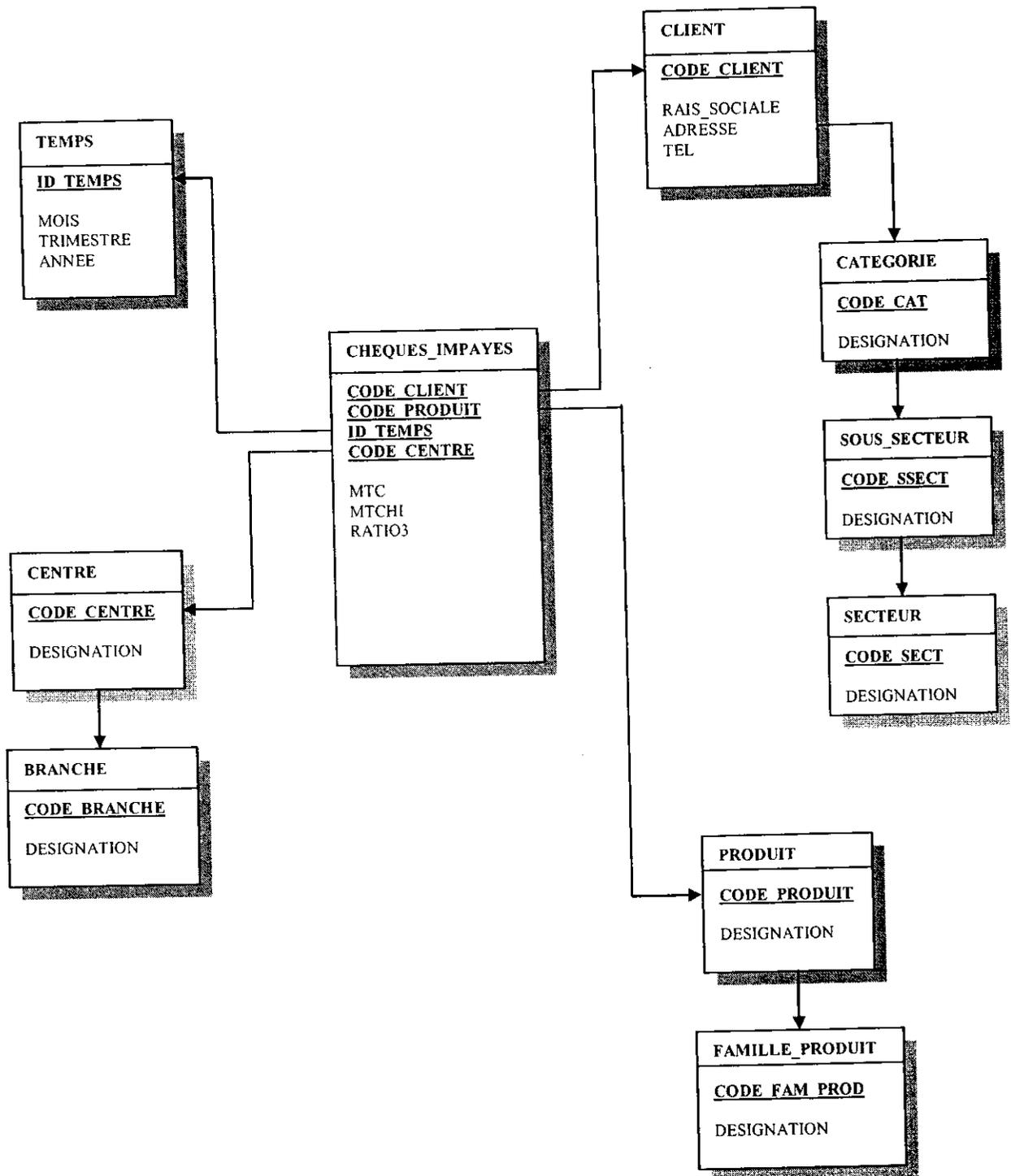
2.2.3. Les dimensions :

D'après la granularité de la table de faits, on peut extraire les dimensions de cette activité. A savoir, client, produit, centre et temps, et qui sont toutes décrites dans l'activité « Etat_Créance ».

2.2.4. Les faits mesurés :

Les faits mesurés de cette activité sont :

- Montant Total des Chèques Impayés « MTCI ».
- Montant total des créances « MTC ».
- Ratio3 qui permet de calculer le taux de croissance du montant des chèques impayés par rapport au montant total des créances. Ce ratio va nous permettre de déterminer le portefeuille classé à risque.



2.3. Modélisation dimensionnelle de l'activité « Recouvrement » :

2.3.1. Le processus d'activité :

La troisième activité jugée importante pour la création de notre entrepôt de données correspond au recouvrement des créances. Dans cette activité, il faut plutôt savoir le montant des créances qui sont recouvertes et le taux de recouvrement qui est le rapport entre les créances recouvertes et le montant des créances échues :

$$\text{Ratio4} = \frac{\text{Montant des créances recouvertes}}{\text{Montant des créances échues}}$$

2.3.2. Le grain du processus d'activité :

Dans ce processus, il est important de connaître le montant total des créances recouvertes ainsi que le taux de recouvrement par centres, client, produit et par période.

La figure suivante illustre le modèle dimensionnel de base de cette activité :

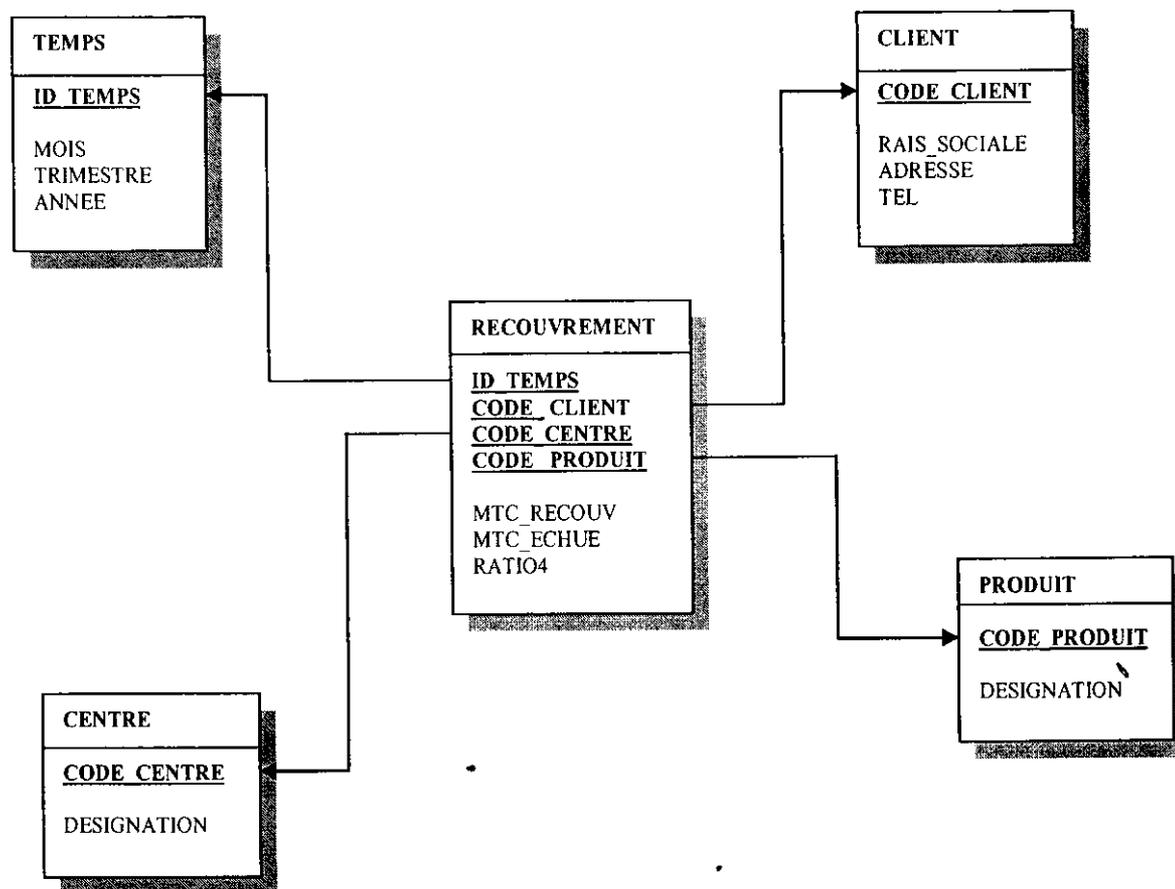


Fig-36- Schéma du modèle dimensionnel de base de l'activité « Recouvrement »

2.3.3. Les dimensions :

D'après la granularité de la table de fait, on peut extraire les dimensions de cette activité. A savoir, client, produit, centre et temps, et qui sont toutes décrites dans l'activité « Etat_Créance ».

2.3.4. Les faits mesurés :

Les faits mesurables pour cette activité sont :

- Le montant total des créances recouvertes « MTC_Recouv ».
- Le montant des créances échues « MTC_ECHUE ».

Et pour le calcul du taux de recouvrement des créances, on a :

- Ratio4 : cet indicateur est très important parce qu'il reflète l'état de l'activité de recouvrement, plus il est grand, plus le recouvrement est important.

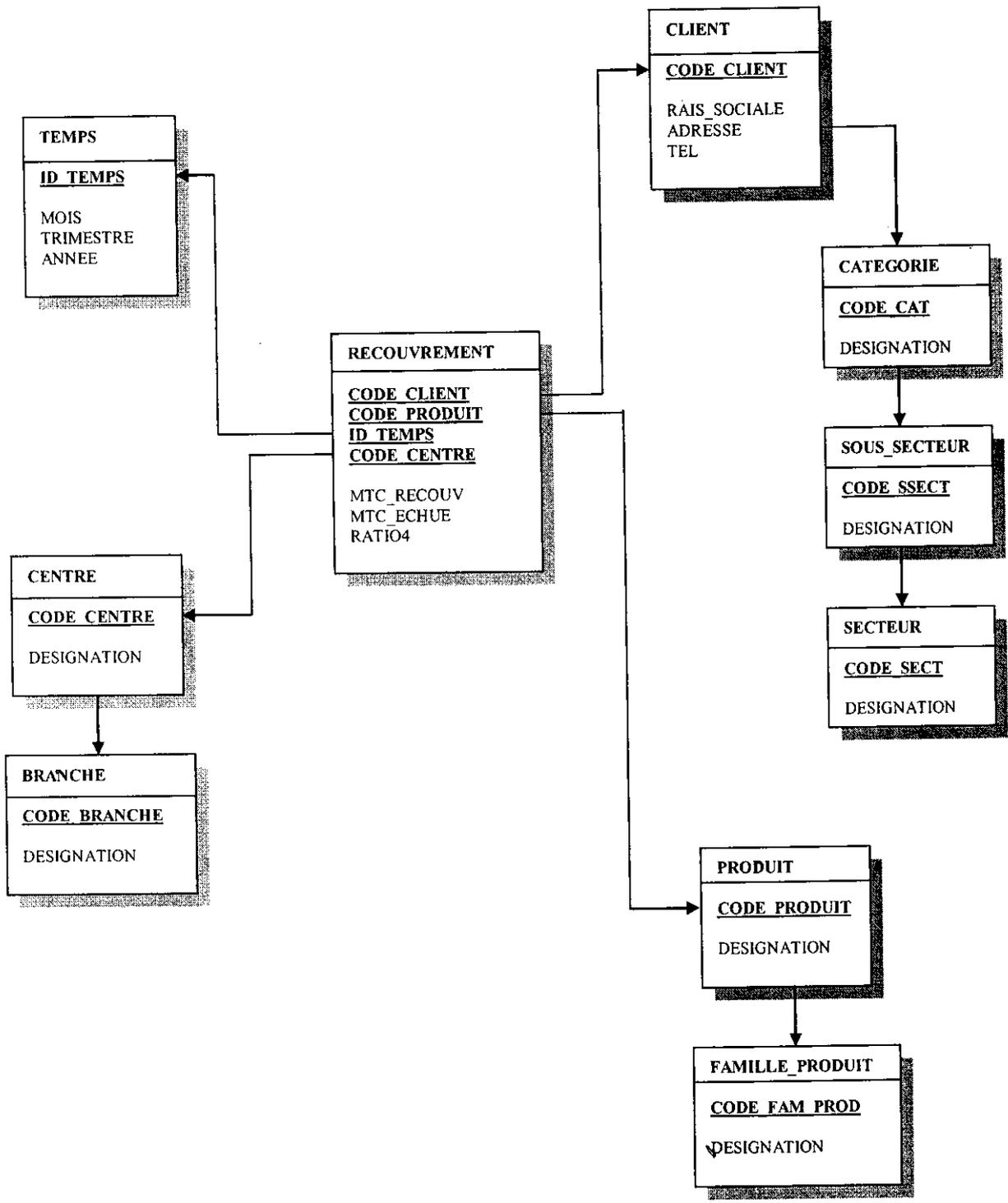


Fig-37- Modèle dimensionnel en flocon de l'activité « Recouvrement »

3. Mise en œuvre de l'architecture :

Le processus de mise en œuvre de notre système décisionnel est constitué de plusieurs composants visibles dans l'architecture des systèmes décisionnels basés sur le concept « Entrepôt de données ».

La figure suivante illustre la conception architecturale de notre système décisionnel :

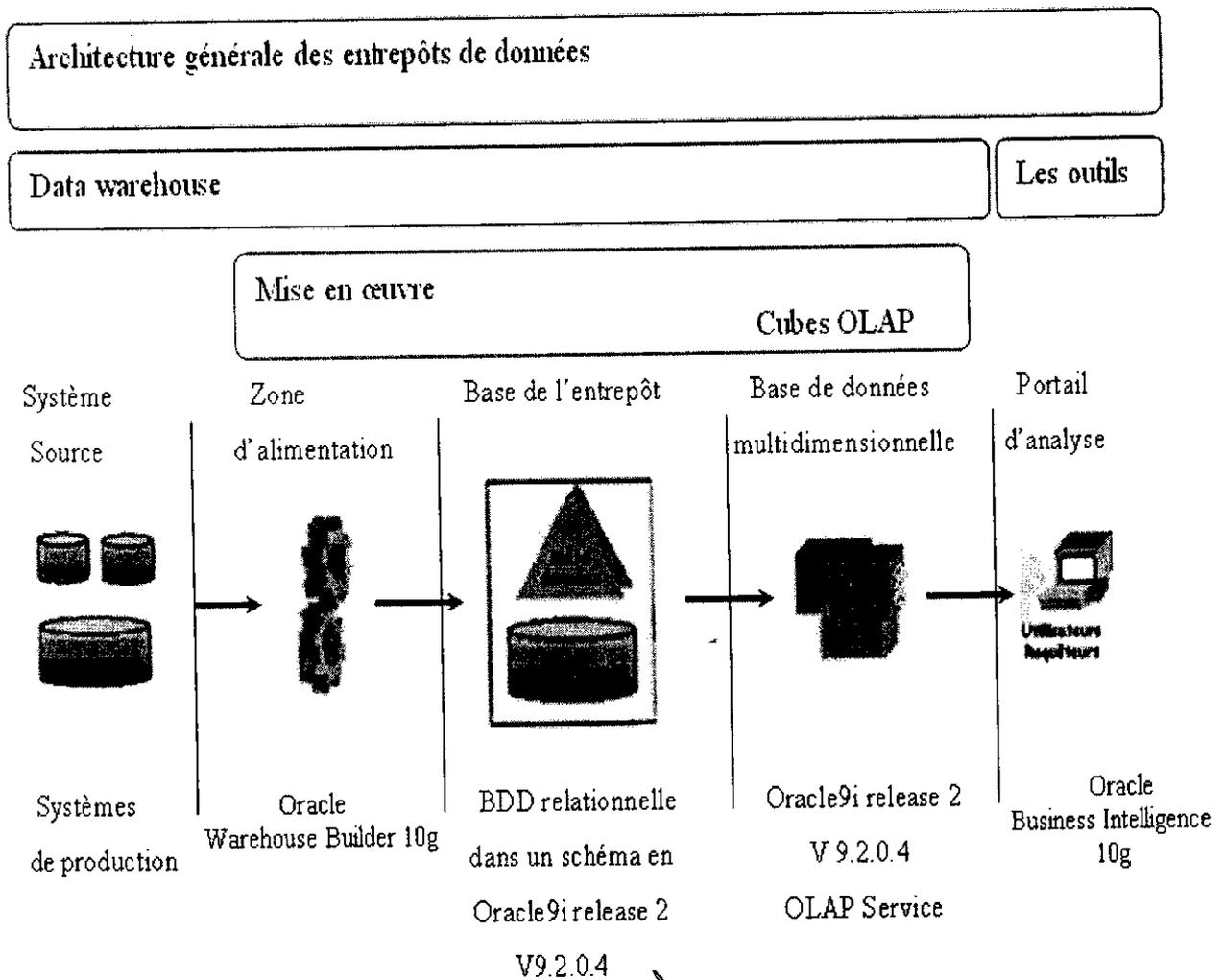


Fig-38- Architecture détaillée du système

- ❖ **ETL « Couche d'alimentation »** : les données stockées dans l'entrepôt de données sont différentes de celles du système source en terme d'intégration, d'agrégation et de consolidation, pour cela le rôle de cette couche est de préparer les données du système source pour qu'elles puissent être stockées dans l'entrepôt de données. L'outil que nous avons choisi est « Oracle Warehouse Builder 10g ».
- ❖ **Entrepôt de données** : c'est une base de données qui contient les données déjà alimentées à partir du système source par l'intermédiaire de la couche ETL. Ces données sont utilisées par la communauté des utilisateurs pour l'analyse. Cette composante regroupe une deuxième base de données qui est la base de méta données, cette base contient des données sur les données de l'entrepôt, comme le détail de chargement (la relation entre les objets sources et les objets d'entrepôt), la fréquence d'alimentation, etc.
- ❖ **Base OLAP** : Ou la base de données multidimensionnelle. L'objectif de cette couche est de stocker les agrégations des données de l'entrepôt dans des cubes multidimensionnelles, ces cubes seront interrogés par les utilisateurs finaux. Notre serveur OLAP utilisé est Oracle 9i Olap Service.
- ❖ **Portail de restitution** : c'est la couche visible aux utilisateurs. Elle contient tous les outils d'analyse et de reporting de données. Notre portail de restitution comporte une application web sur la base d'oracle discoverer.

Le processus de mise en œuvre se déroule en trois étapes :

3.1. Construction de l'entrepôt de données :

C'est une base de données dont l'objectif est de centraliser l'information décisionnelle, elle représente la base de données physique de notre entrepôt.

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

3.1.1. Liste des tables :

Chaque entité dans notre modélisation, table de faits ou dimension, est transformée en une table de données relationnelle. La liste des tables de notre entrepôt de données relationnel et leurs identifiants est illustrée dans le tableau suivant :

Table	Identifiant
Etat_Créance	Code_Client, Code_Produit, Code_Centre, Code_Ech, Id_Temps
Recouvrement	Code_Client, Code_Produit, Code_Centre, Id_Temps
Chèques_Impayés	Code_Client, Code_Produit, Code_Centre, Id_Temps
Temps	Id_Temps
Client	Code_Client
Catégorie	Code_Cat
Sous_Secteur	Code_SSect
Secteur	Code_Sect
Produit	Code_Produit
Famille_Produit	Code_Fam_Prod
Centre	Code_Centre
Branche	Code_Branche
Type_Echéance	Code_Ech

Tab-5- Liste des tables de l'entrepôt de données

3.1.2. Description des attributs :

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

Désignation	Code	Table	Type de données
Code client	CODE_CLIENT	CLIENT	Varchar2 (5)
Raison sociale	RAISON_SOCIALE	CLIENT	Varchar2 (40)
Adresse	ADRESSE	CLIENT	Varchar2 (40)
Code catégorie	CODE_CAT	CATEGORIE	Varchar2 (5)
Désignation catégorie	DESIGNATION_CAT	CATEGORIE	Varchar2 (80)
Code sous secteur	CODE_SSECT	SOUS_SECT	Varchar2 (3)
Désignation sous secteur	DESIGNATION_SSECT	SOUS_SECT	Varchar2 (30)
Code secteur	CODE_SECT	SECTEUR	Char (1)
Désignation secteur	DESIGNATION_SECT	SECTEUR	Varchar2 (30)
Code produit	CODE_PROD	PRODUIT	Varchar2 (5)
Désignation produit	DESIGNATION_PROD	PRODUIT	Varchar2 (30)
Code famille produit	CODE_FAM_PROD	FAMILLE_PROD	Varchar2 (3)

Désignation famille produit	DESIGNATION_FAM_PROD	FAMILLE_PROD	Varchar2 (30)
Code échéance	CODE_ECH	TYPE_ECHEANCE	Number (2)
Désignation échéance	DESIGNATION_ECH	TYPE_ECHEANCE	Varchar2 (15)
Code centre	CODE_CENTRE	CENTRE	Varchar2 (5)
Désignation centre	DESIGNATION_CENTRE	CENTRE	Varchar2 (30)
Code branche	CODE_BRANCHE	BRANCHE	Varchar2 (3)
Désignation branche	DESIGNATION_BRANCHE	BRANCHE	Varchar2 (20)
Identifiant temps	ID_TEMPS	TEMPS	Date
Année	ANNEE	TEMPS	Number (4)
Trimestre	TRIMESTRE	TEMPS	Number (1)
Mois	MOIS	TEMPS	Number (2)
Montant total des créances	MTC	ETAT_CREANCE	Number (15)
Montant des créances échues	MTC_ECHUE	ETAT_CREANCE	Number (15)
Montant des créances non échues	MTC_NECHUE	ETAT_CREANCE	Number (15)
Montant des créances antérieures	MTC_ANT	ETAT_CREANCE	Number (15)
Chiffre d'affaire	CA	ETAT_CREANCE	Number (15)
Taux de croissance de la créance / CA	RATIO1	ETAT_CREANCE	Number (2)
Délai moyen de paiement	RATIO2	ETAT_CREANCE	Number (3)
Montant des créances recouvertes	MTC_RECOUV	RECOUVREMENT	Number (15)
Montant des créances échues	MTC_ECHUE	RECOUVREMENT	Number (15)
Taux de recouvrement	RATIO4	RECOUVREMENT	Number (2)
Montant des chèques impayés	MTCHI	CHEQUES_IMPAYES	Number (15)
Montant total des créances	MTC	CHEQUES_IMPAYES	Number (15)
Taux de croissance des chèques impayés / montant total des créances	RATIO3	CHEQUES_IMPAYES	Number (2)

Tab-6- Liste des attributs de l'entrepôt de données

3.2. Construction des cubes OLAP :

Le gestionnaire OLAP nous permet de transformer les données stockées dans des bases de données relationnelles en informations commerciales pertinentes et faciles à exploiter, grâce à la création d'un cube de données. La modélisation multidimensionnelle des données constitue une façon de faciliter l'analyse commerciale en ligne ainsi que les performances des requêtes. Cette étape consiste à concevoir notre base de données multidimensionnelle (cube OLAP).

Pour la mise en œuvre des cubes OLAP, on a utilisé comme technique de modélisation, la modélisation MOLAP « Multidimensional OLAP » car :

- Cette méthode est facile à mettre en place,
- En plus, la facilité de l'analyse multiaxes sur les données agrégées est garantie par cette méthode.

Le serveur OLAP utilisé pour notre travail est Oracle 9i-OLAP Service.

On a résumé dans le tableau suivant les tables de faits choisies avec leurs mesures et par leurs dimensions d'analyse sollicitées.

Mesures	Dimensions sollicitées
Etat_Créance	
Montant total des créances Montant des créances échues Montant des créances non échues Montant des créances antérieures Chiffre d'affaire Ratio1 Ratio2	client → centre → produit → temps → type_échéance
Chèques_Impayés	
Montant total des créances Montant des chèques impayés Ratio3	client → centre → produit → temps
Recouvrement	
Montant des créances recouvertes Montant des créances échues Ratio4	client → centre → produit → temps

Tab-7- Liste des tables de faits choisies avec leurs mesures et dimensions d'analyse sollicitées

3.2.1. Formalisme utilisé :

✓ **Concept de fait** : Le fait modélise le sujet de l'analyse. Un fait est formé de mesures correspondantes aux informations de l'activité analysée. Nous représentons le fait par un cube englobant les différentes mesures d'activités qu'il contient. En outre le symbole d'un cube estampille

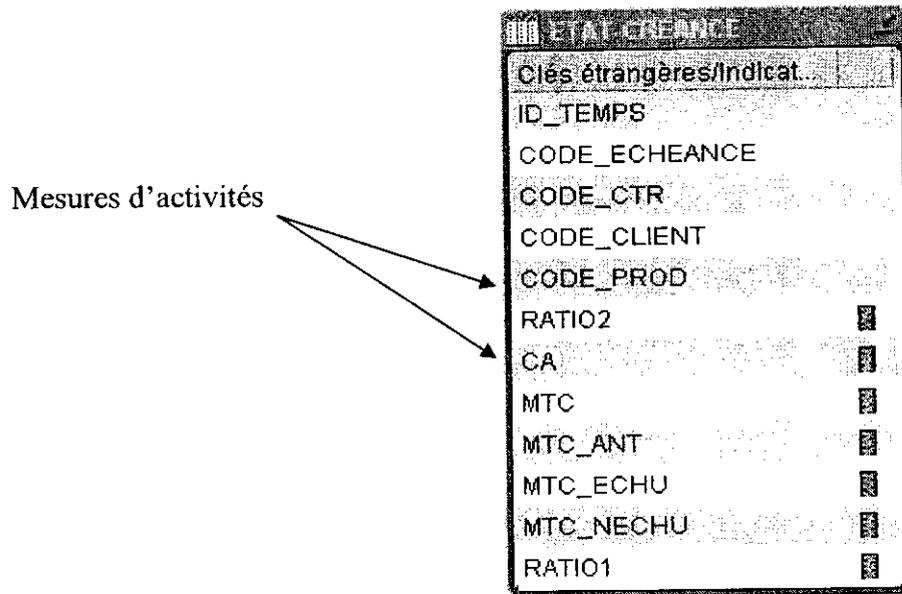


Fig -39- Concept de fait

✓ **Concept de dimension** : Une dimension modélise une perspective de l'analyse. Une dimension se compose de paramètres correspondant aux informations faisant varier les mesures de l'activité. Nous représentons la dimension par un rectangle englobant les différents paramètres d'analyse qu'il contient. En outre le symbole représentant trois axes estampille les dimensions pour le distinguer du fait.

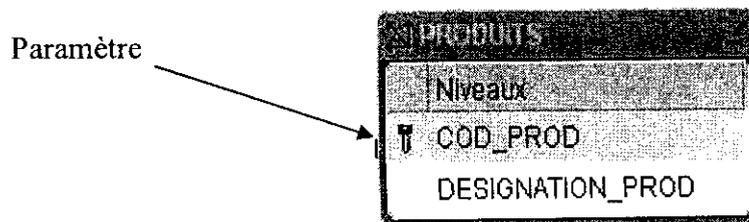


Fig -40- Concept de dimension

3.2.2. Schéma de modèle MOLAP des cubes :

Les figures suivantes illustrent le modèle physique de chaque cube MOLAP selon la plate forme Oracle 9i OLAP Service.

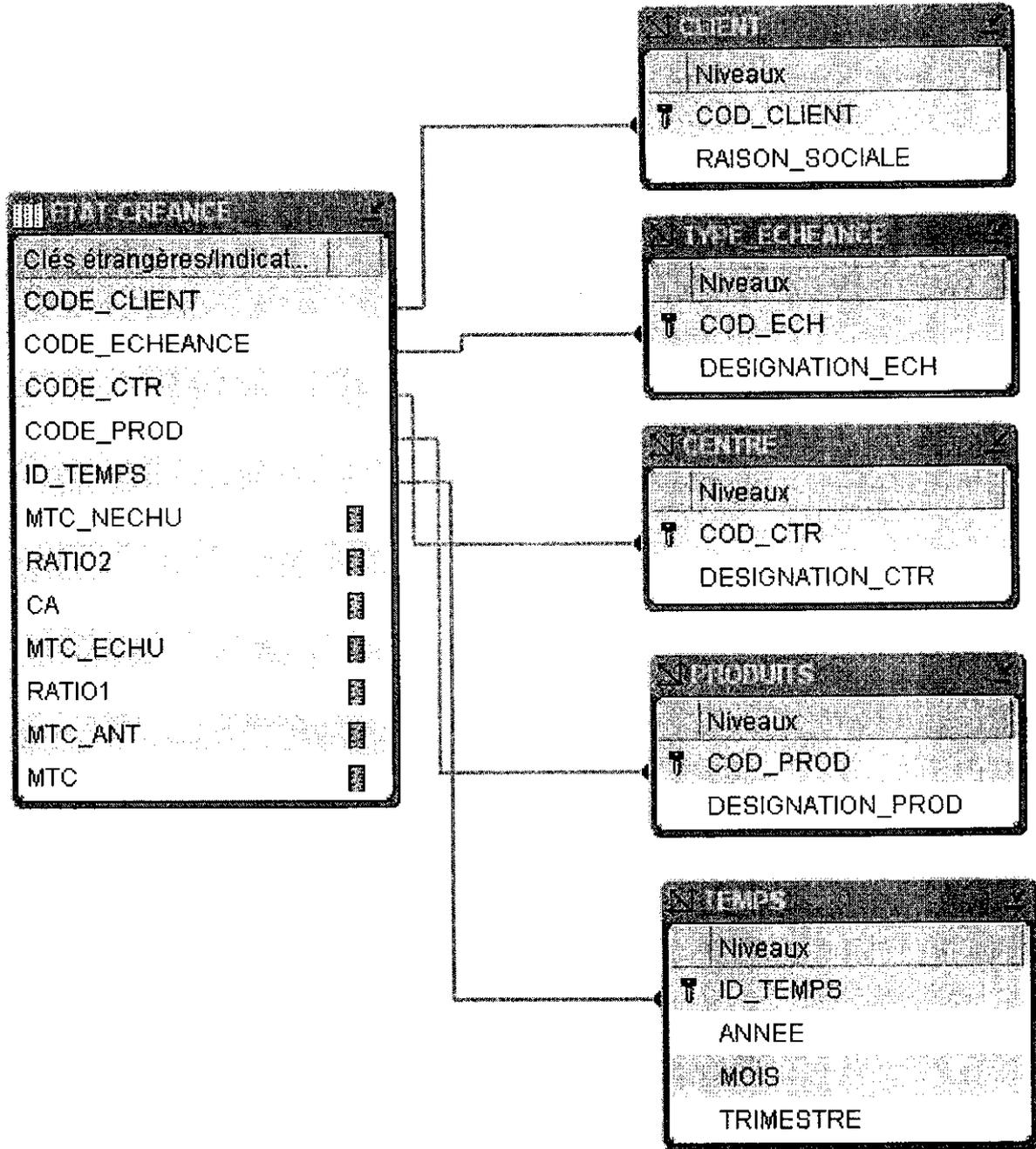


Fig-41- Cube MOLAP de l'activité « Etat Créance »

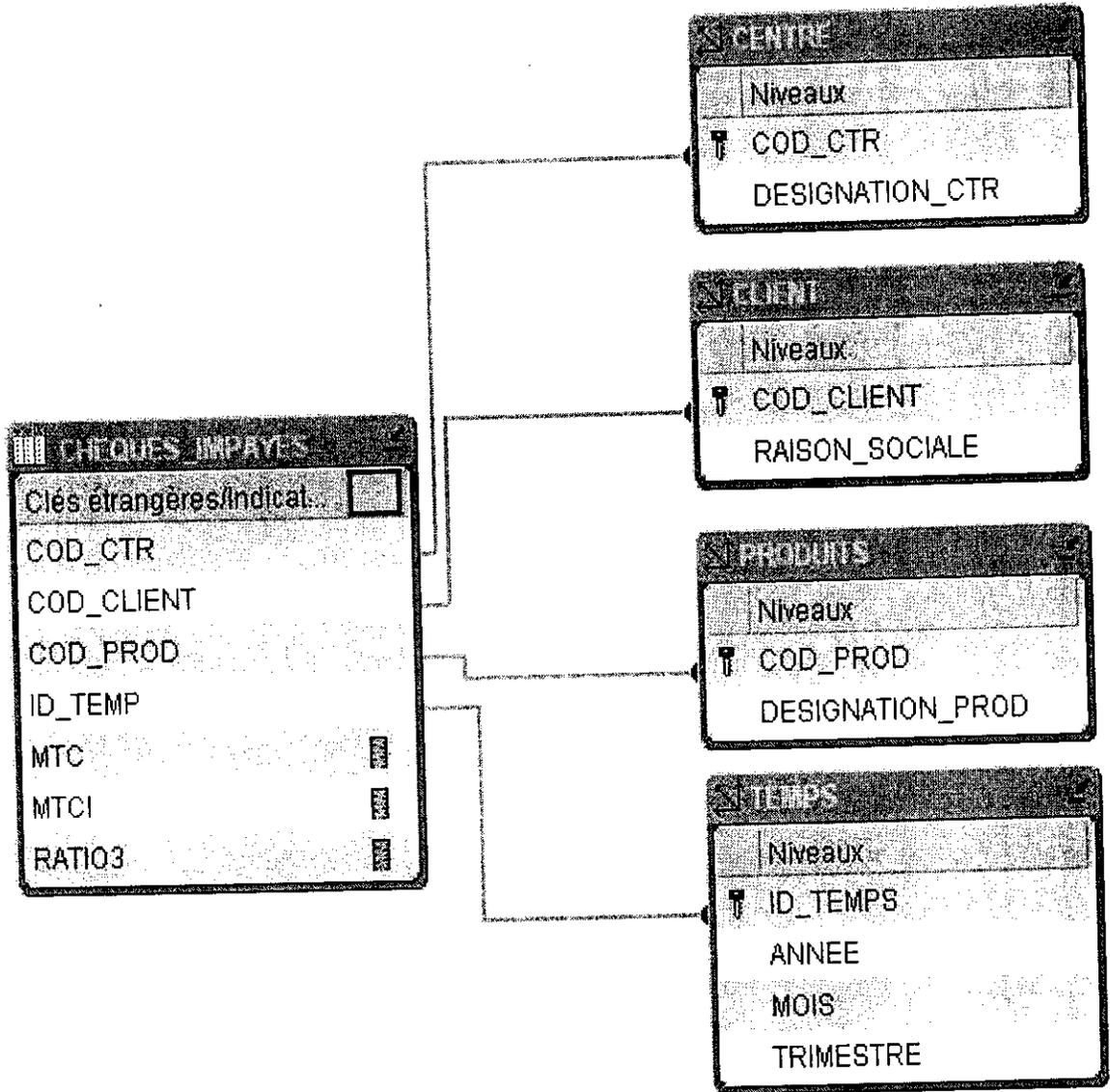


Fig-42- Cube MOLAP de l'activité « Chèques Impayés »

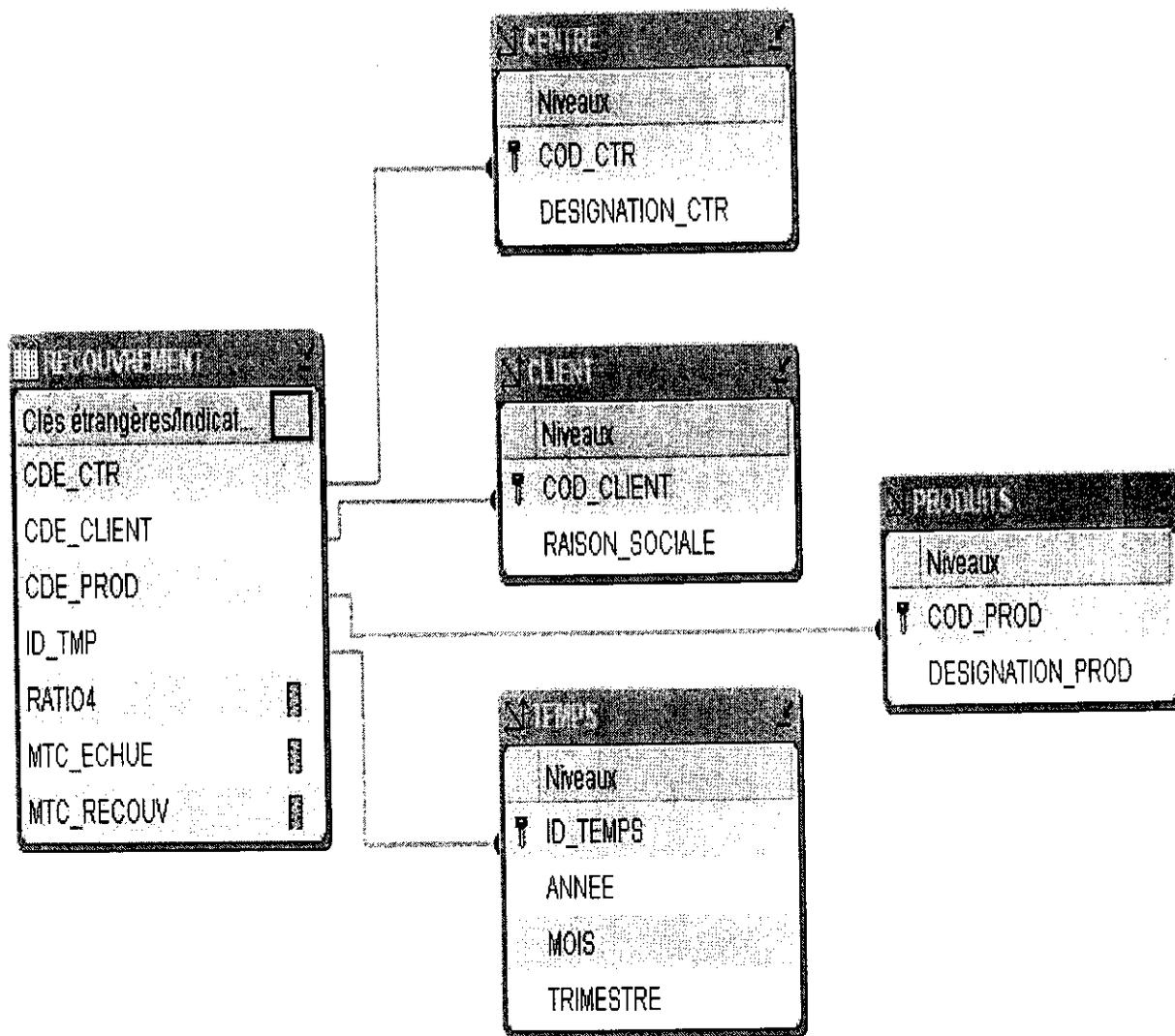


Fig-43- Cube MOLAP de l'activité « Recouvrement »

3.3. Construction de la zone d'alimentation :

L'alimentation est la procédure qui permet de transférer des données du système opérationnel vers l'entrepôt de données en les adaptant. La conception de cette opération est une tâche assez complexe, elle doit être faite en collaboration avec les administrateurs des bases de production qui connaissent les données disponibles et vérifient la faisabilité des demandes. Il est nécessaire de déterminer quelles données seront chargées, quelles informations et vérifications seront nécessaires, la périodicité et le moment auxquels les transferts auront lieu.

Cependant il existe des outils d'alimentation ; l'outil ETL (Extract, Transform and Load) récupère toutes les données et les centralise dans le Data Warehouse, quels que soit leurs sources et les systèmes qui les supportent (système d'exploitation, SGBD, fichiers plats, bases hiérarchiques...), d'automatiser et d'industrialiser le processus d'alimentation, de faciliter la maintenance des données et de limiter les développements spécifiques. Ces outils permettent de construire et de maîtriser à jour et de maintenir le dictionnaire de méta données.

Un processus ETL se décompose en trois phases : l'extraction, la préparation/transformation et le chargement.

3.3.1. L'extraction :

Il s'agit en premier lieu d'aller chercher les données là où elles se trouvent. L'outil ETL a la capacité de se connecter aux différentes applications, bases de données ou fichiers.

L'outil doit être capable de lire sélectivement les données sources, et donc de filtrer les données en lecture afin de n'extraire que l'information pertinente.

3.3.2. La préparation / transformation :

Les bases sources sont organisées dans un format répondant aux besoins d'une application. Ce format n'a que peu de chance d'être compatible avec celui d'une autre application. Les données rendues accessibles puis acquises dans la phase extraction doivent donc être transformées.

L'outil utilisé (Oracle Warehouse Builder 10g) permet de modéliser graphiquement le processus de transformation de données, quelque soit la source de données. Il possède aussi des fonctions qui permettent de transformer les données issues des systèmes sources avant chargement dans l'entrepôt de données (to_char, to_number, concat, substr, etc...) vu que les données sources sont hétérogènes.

Il s'agit d'un prétraitement incluant la mise en correspondance des formats, de données, le nettoyage, la transformation et l'agrégation.

L'étape de contrôle s'effectue par application de règles adaptées sur les flux de données entrants. Bien souvent les incohérences ne sont découvertes que dans cette phase.

3.3.3. Le chargement :

Le chargement prend en compte la gestion du format des données. Les données de l'environnement Data Warehouse sont chargées à partir des données trouvées dans l'environnement transactionnel de l'entreprise. Le processus de chargement des données transactionnelles nécessite un regard périodique sur les données sources et une sélection des données appropriées qui ont besoin d'être entrées dans le système décisionnel. Les données sélectionnées de l'environnement transactionnel entrent dans le processus d'intégration et de transformation. Les données sont intégrées et envoyées au système décisionnel.

Notre processus d'alimentation se déroule en trois étapes nécessaires. L'extraction depuis le système source, le regroupement des données en un seul ensemble et en fin le chargement dans l'entrepôt de données.

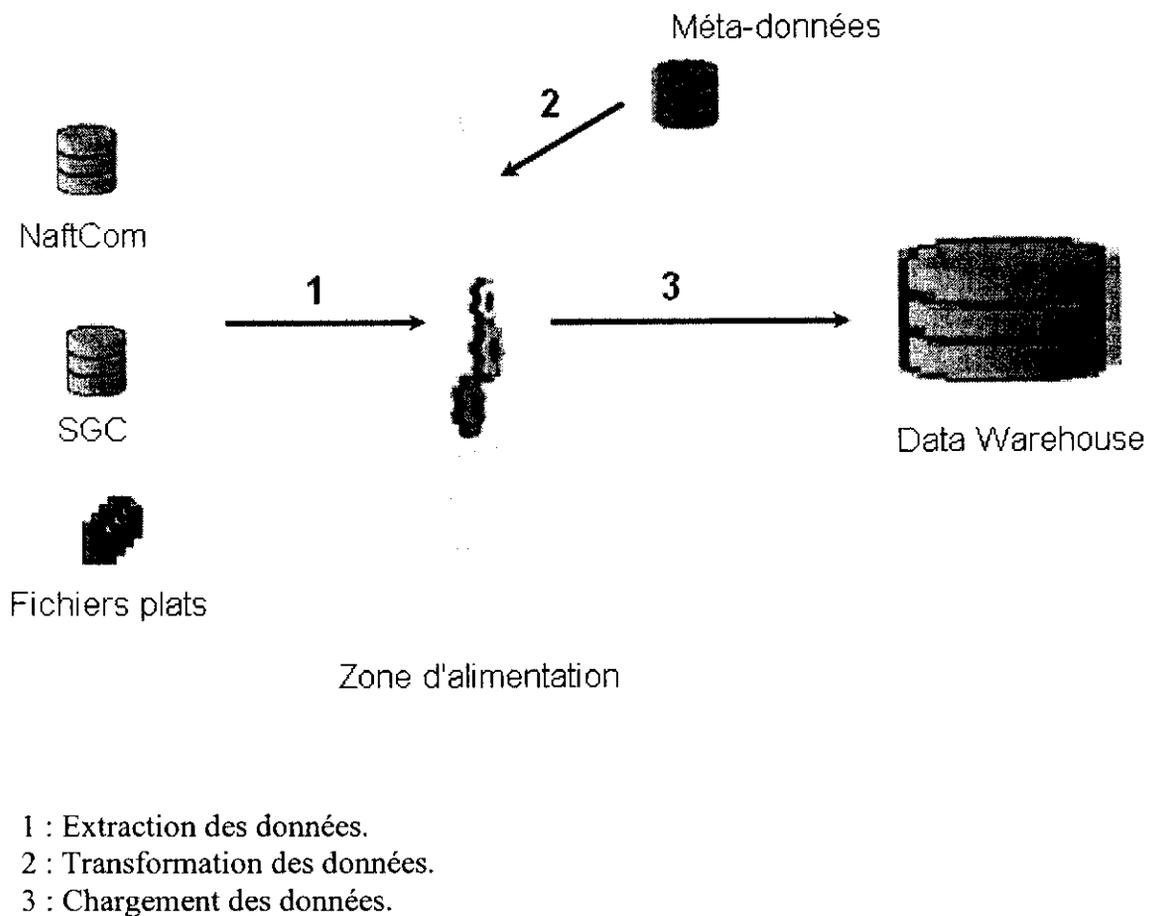


Fig-44- Le processus d'alimentation de notre entrepôt de données

Pour réaliser ces étapes, nous avons besoin d'une base de données qui stocke les informations sur les données. Il s'agit d'une métabase qui nous décrit les données que doit contenir l'entrepôt de données, en donnant où et comment extraire les données nécessaires.

Le processus ETL doit aussi s'effectuer périodiquement. La cible doit être donc optimisée en terme de temps de réponse et cela détermine la façon dont le processus de transformation doit être effectué.

Pour notre cas, l'opération de transformation de la source vers l'entrepôt de données sera exécutée à chaque fin du mois. D'où notre analyse sera suivie par l'hierarchie suivante : Année → Trimestre → Moi.

La figure suivante illustre la fréquence d'alimentation :

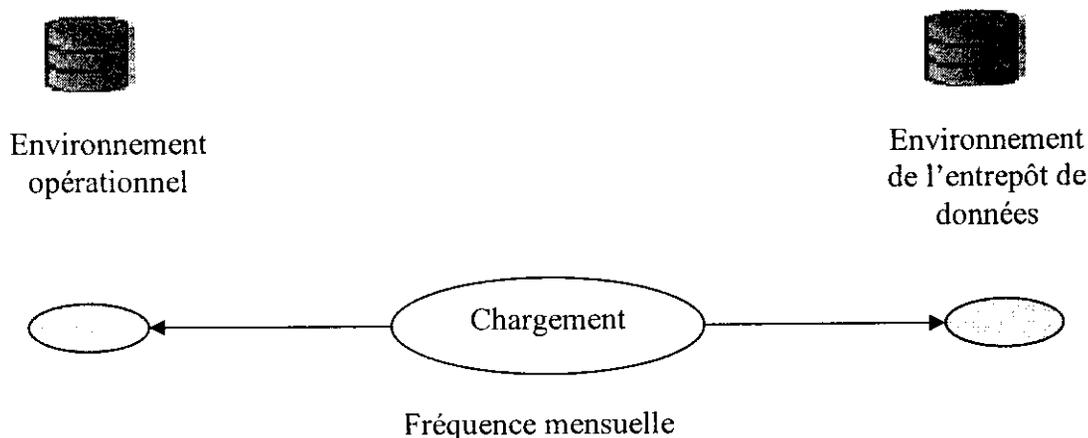


Fig-45- La fréquence de chargement des systèmes sources vers l'entrepôt de données

3.4. L'exploitation du Data Warehouse « Le portail de restitution » :

C'est la part publique de l'entrepôt de données. Il représente ce que voient les utilisateurs, les outils avec lesquels ils travaillent. Sous ce terme sont regroupées toutes les applications qui s'appuient sur les données de l'entrepôt de données pour les restituer aux utilisateurs.

La valeur d'un entrepôt de données est largement déterminée par l'aptitude de l'utilisateur à extraire des informations pertinentes sur lesquelles il peut baser son processus de prise de décision.

Il s'agit de mettre en œuvre une interface utilisateur efficace minimisant, pour l'utilisateur, le nombre d'actions pour obtenir le résultat désiré.

Il faut signaler que les utilisateurs ont des besoins différents tels que :

- Des besoins d'interrogation classique afin de retrouver des rapports préalablement établis sur un thème défini.
- Des besoins de consultation évolués (Analyse OLAP multidimensionnelle dans le cadre d'application d'aide à la décision). En mettant en réalisation les indicateurs de performances avec les dimensions, permettant de voir l'impact d'une décision.

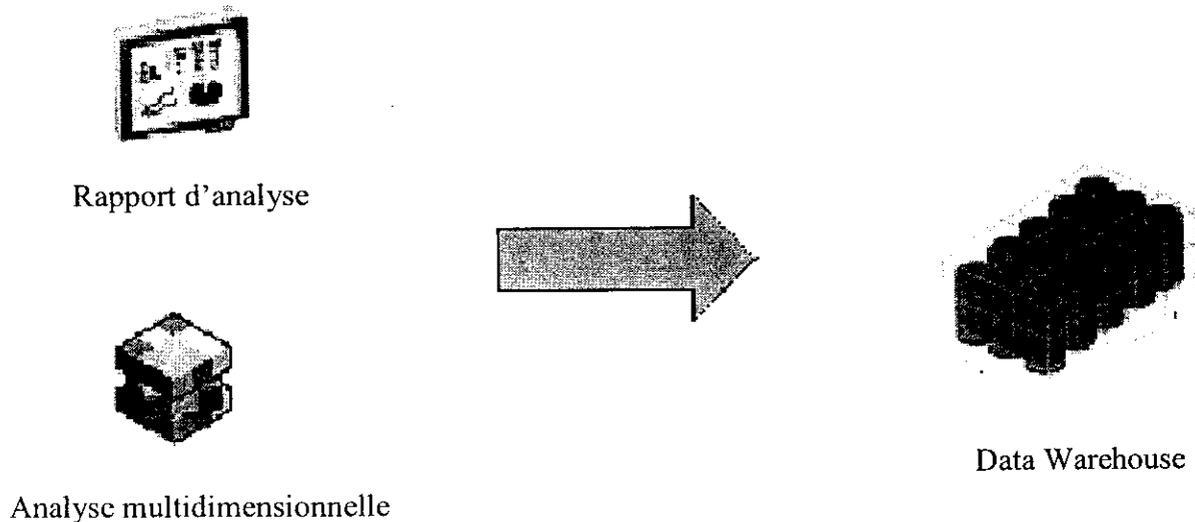


Fig-46- Les outils d'accès à notre entrepôt de données

3.4.1. Le reporting :

Il s'agit de développer des rapports d'activité prédéfinis dont la diffusion peut être soit planifiée soit laissée à l'initiative de l'utilisateur.

Le reporting est dynamique puisque l'utilisateur a la possibilité de rafraîchir les données se trouvant sur son rapport ou d'en modifier la mise en page.

3.4.2. L'analyse multidimensionnelle :

Elle nous permet de construire dynamiquement nos analyses à partir des données qui sont à sa disposition. Pour une première approche de notre solution, il s'agit de mettre en œuvre des analyses multidimensionnelles s'appuyant sur des cubes OLAP prédéfinis.

Notre outil multidimensionnel permettra de naviguer à l'intérieur d'un cube de données par l'intermédiaire de « Drill ».

- **Drill Down, Drill Up** : Les drill-down/up désigne la faculté d'aller du niveau global vers le niveau détaillé, et inversement. Ce mécanisme est totalement basé sur la notion de hiérarchie.

Chacun des axes d'analyse se décompose en attributs reliés entre eux par des relations père/fils. Une dimension doit normalement pouvoir comporter plusieurs hiérarchies.

- **Slice and Dice** : Le Slice and Dice désigne la possibilité de faire pivoter dynamiquement les axes du tableau d'analyse. Il est possible ainsi de passer d'un tableau présentant les ventes par magasin en lignes et jour en colonnes par un tableau similaire présentant les ventes par magasin en colonnes et jour en lignes.
- **Data Surfing** : Le Data surfing est la possibilité laissée à l'utilisateur de circuler librement, de manière intuitive et ergonomique dans un modèle dimensionnel, au-delà d'un simple drill-down ou slice and dice. L'utilisateur peut alors modifier dynamiquement ses axes d'analyse ou appliquer un nouveau filtre à ses données.

Conclusion :

Au cours de ce chapitre, nous avons présenté la partie qui constitue l'arrière plan de notre système décisionnel, la modélisation et la mise en œuvre de notre Data Warehouse.

Nous avons commencé par la conception logique sous forme de schéma multidimensionnel qui intègre les concepts de faits et de dimensions. Cette modélisation multidimensionnelle est utilisée car sa manière de représentation des informations est adaptée aux analyses et technologie du système OLAP.

Le modèle de données de notre entrepôt est basé sur un schéma en flocon comportant un unique fait et une hiérarchie ainsi que des dimensions, ces dimensions peuvent être partagées par plusieurs tables de faits. Le partage des dimensions entre plusieurs faits limite les redondances et la complexité.

Nous avons aussi présenté l'architecture type d'un système décisionnel basé sur l'approche entrepôt de données qui contient trois couches essentielles (ETL, entrepôt de données et le portail de restitution).

Après avoir construit une base de données multidimensionnelle supportant les requêtes d'analyse, il faut spécifier la disposition physique de notre système décisionnel. En d'autre terme, il faut sélectionner les offres logicielles et la conception architecturale pour le déploiement de notre système décisionnel. Et c'est ce que nous allons voir dans le chapitre suivant.

Chapitre V :

Déploiement du système

Introduction :

Après la conception de notre entrepôt de données, et afin de valider notre modèle, nous passerons au déploiement du prototype d'un système d'aide à la décision en se basant sur les fonctionnalités des cubes OLAP et les avantages de l'architecture multidimensionnelle.

Notre solution informatique pour la mise en place du système décisionnel proposé est présentée dans ce chapitre.

1. Architecture informatique du système :

En terme de contraintes techniques, l'entreprise Naftal Spa implémente au niveau de son siège général un réseau intranet basé sur la plate-forme Microsoft (Windows 2003 Server comme serveur d'annuaire, Oracle 9i comme serveur de base de données et Exchange Server 2000 comme serveur de messagerie électronique). Ce réseau intranet sera étendu vers un réseau extranet pour l'objectif de regrouper les autres structures en un seul réseau ouvert.

Notre système décisionnel est destiné comme premier lieu à un groupe d'analystes. La mise en place du système dépend donc de la plate-forme existante. L'architecture informatique est de type « Architecture trois-tiers ».

2. La disposition physique du système :

Notre système décisionnel est basé sur l'architecture illustrée dans la figure suivante.

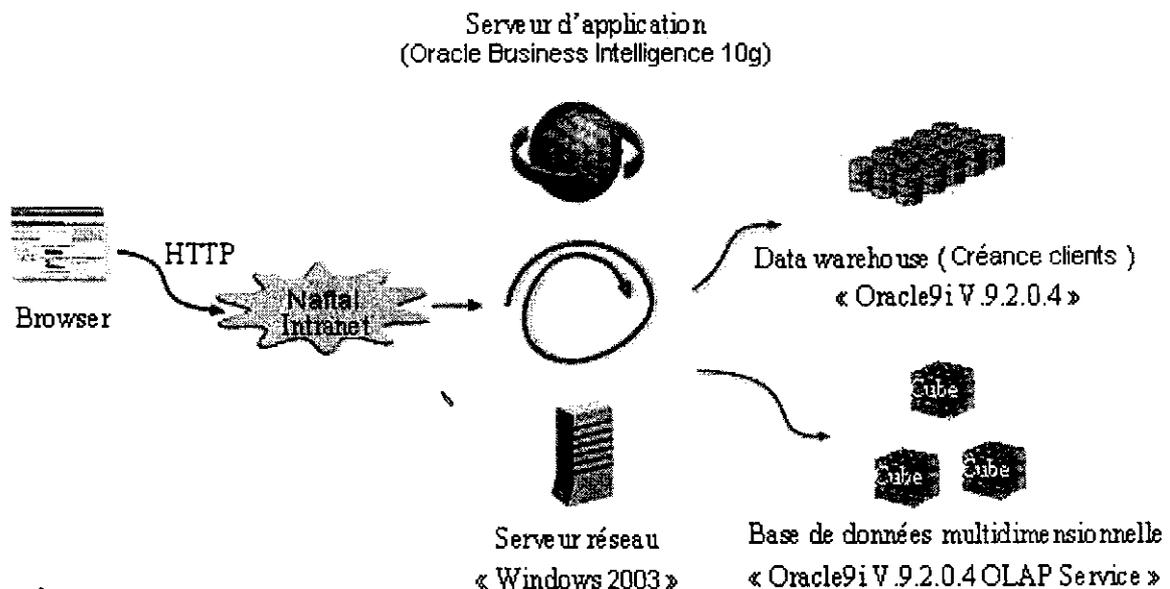


Fig-46- Disposition physique du système

Cette architecture est divisée en trois couches essentielles : Présentation, Métiers (traitements) et stockage (données), le détail de ces couches est présenté dans ce qui suit :

- **Présentation** : L'objectif final de cette couche est de présenter les données stockées aux serveurs de données en informations utiles et simples à analyser. L'accès à notre application se fait via un navigateur web, l'utilisateur peut accéder à l'information à travers le portail décisionnel développé.
- **Le traitement** : cette couche regroupe en général tous les processus (Objets) métiers des applications de l'entreprise (la logique métier). La connexion aux serveurs de données, aux serveurs de gestion des requêtes utilisateurs est l'un des processus de base dans le système informatique de l'entreprise.

Elle contient dans notre cas le serveur d'annuaire.

- **Les données** : ou stockage. Cette couche englobe tous types des serveurs de données, soit un SGBD, un serveur de fichiers ou un serveur de messagerie, etc. dans notre cas, cette couche contient le serveur de base de données Oracle 9i Server pour l'entrepôt de données et le serveur de base de données multidimensionnelle Oracle 9i OLAP service.

3. Les utilisateurs du système :

Notre système est utilisé par l'ensemble des analystes et décideurs intéressés par le déploiement d'un système d'aide à la décision basé sur les dimensions, leur objectif est d'avoir un aperçu sur les états d'analyse des activités de base de l'entreprise.

Il est à noter que chaque utilisateur du système est identifié par son « user » et son « password » et a ses propres « privilèges ».

4. logiciels utilisés :

- Microsoft Windows 2003 Server comme serveur de sécurité et serveur réseau.
- Oracle9i Entreprise Edition version 9.2.0.4 comme serveur de données qui implémente le data Warehouse.
- Oracle9i Entreprise Edition version 9.2.0.4 OLAP Service comme serveur de base de données multidimensionnelle qui implémente les cubes OLAP.
- Oracle Business Intelligence 10g comme serveur d'application.

5. Matériel utilisé :

Le matériel utilisé pour le déploiement de notre système est basé sur la plateforme Intel biprocesseur. Trois serveurs sont installés dans une salle bien spécifique à héberger les serveurs de l'entreprise. Deux serveurs de 80 Gëga Octets chacun pour la sauvegarde des données, et un troisième de 20 Gëga Octets pour héberger la base de données multidimensionnelle.

6. La sécurité du système :

La sécurité d'un système ne se décide pas à la fin, il faut prendre son temps à évaluer les risques, lors de la conception de l'application, les risques soulevés par chacun des choix d'implémentation ou de déploiement.

L'information tend à devenir l'un des biens les plus précieux de l'entreprise, donc il faut savoir la protéger, sa destruction ou bien sa perte d'intégrité entraîne des conséquences graves au fonctionnement du système surtout quand il s'agit d'information sensible tel que des données décisionnelles.

La qualité de fonctionnement d'un système d'information suppose au préalable une prévention et une protection suffisante et cohérente vis-à-vis à des risques encourus, il est impératif que ce système projeté doit être en mesure de se protéger contre les indiscretions et les destructions. Ceci n'est possible que par la mise en place d'un plan de sécurité rigoureux afin de minimiser les risques.

En général, deux aspects de sécurité possible : la sécurité physique et la sécurité logique. La première contre les dégâts matériels (Machine tombe en panne) et la deuxième contre les dégâts logiciels (Virus, Perte de mot de passe, etc.).

Notre approche de sécurité pour les deux types de pannes est basée sur les normes standard.

Sécurité physique :

Les données de l'entrepôt et de la base multidimensionnelle sont sauvegardées périodiquement dans des bandes magnétiques de grande capacité.

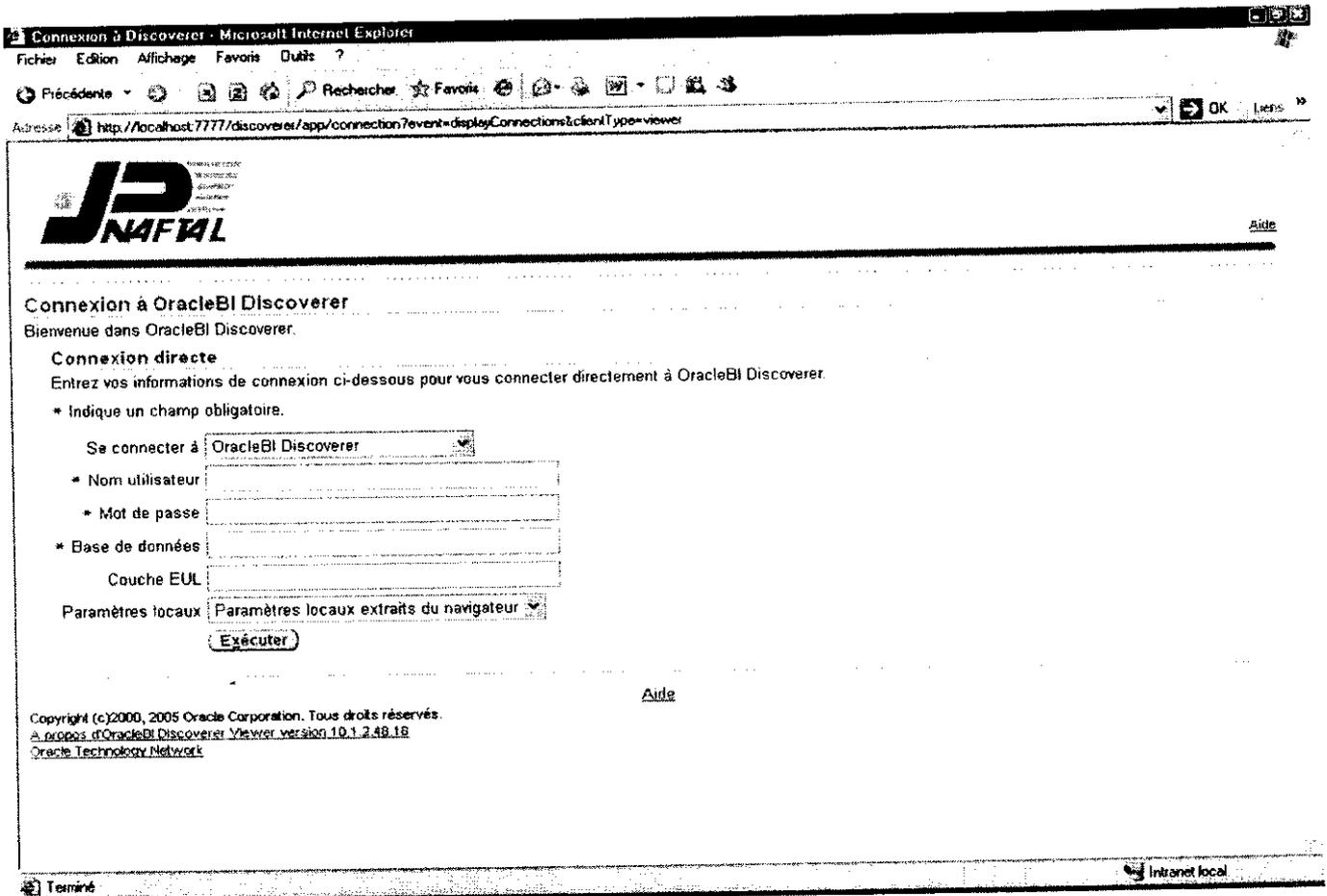
Sécurité logique :

Consiste à protéger les données logiquement contre les dégâts provoqués par des logiciels tiers (par exemple les virus, accès non autorisé...). Notre solution de sécurité est basée sur la sécurité du système d'exploitation et la sécurité du SGBD pour les accès à notre système décisionnel où chaque utilisateur se voit octroyé des droits d'accès et privilèges spécifiques.

7. Présentation de l'application :

Cette page représente la page d'accueil de notre application, à partir de laquelle l'utilisateur peut accéder au système après avoir saisi les champs qui lui sont appropriés :

- Nom d'utilisateur
- Mot de passe
- Base de données



Après avoir saisi les champs d'identification, l'utilisateur du système voit apparaître cette page qui contient les classeurs de notre base de données multidimensionnelle, chaque classeur regroupe un certain nombre de feuilles qui englobent les rapports d'analyse pré-établis, il est à noter que chaque utilisateur a le droit d'accès à un nombre limité de classeurs selon ses privilèges et ses besoins d'analyse.

Liste des feuilles de calcul

Rechercher
Entrez une chaîne de recherche pour trouver un classeur ou un jeu de résultats (ResultSet). Utilisez le menu déroulant pour sélectionner des classeurs de base de données ou des classeurs programmés.

Rechercher:

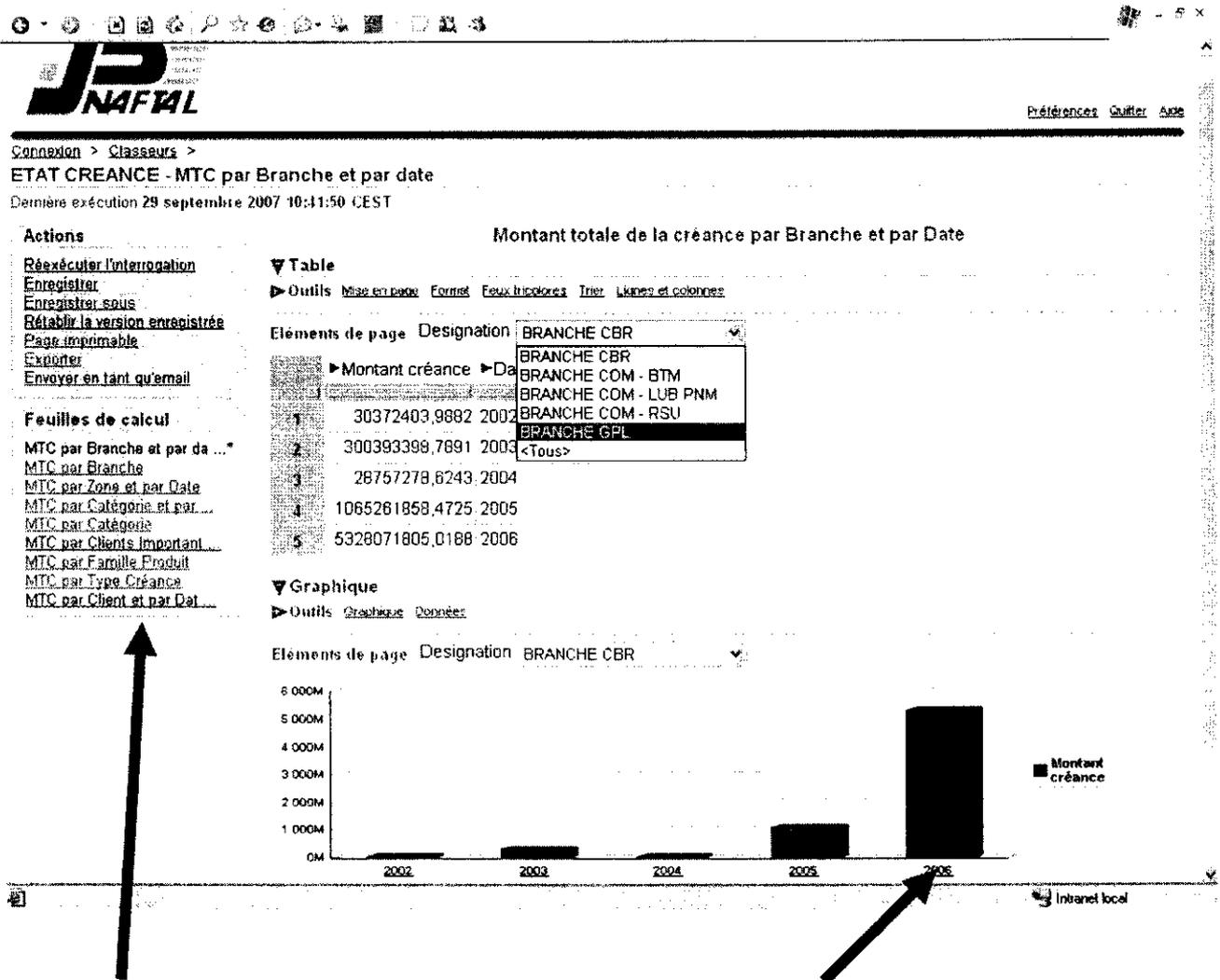
Liste des résultats

 |

Sélection	Nom	Description	Propriétaire	Dernière modification
▼	Discoverer Workbooks			
+	<ul style="list-style-type: none"> CHEQUES IMPAYES <ul style="list-style-type: none"> MTCI par Branche MTCI par Clients ayant des CI Importants MTCI MTC par Branche MTCI par U.O MTCI MTC par U.O MTCI par Branche et par Date MTCI par U.O et par Date MTCI par Année et par Branche 	Situation des chèques impayés CREANCE	CREANCE	29 août 2007 02:05:50 CEST
+	▶ ECHEANCE		CREANCE	12 septembre 2007 22:05:26 CEST
+	▶ ETAT CREANCE		CREANCE	12 septembre 2007 21:53:59 CEST
+	▶ RECOUVREMENT		CREANCE	22 septembre 2007 23:09:59 CEST

Préférences | Quitter | Aide

Cette page illustre un exemple de ce que peut contenir une feuille d'analyse. L'utilisateur peut naviguer librement dans un classeur avec la possibilité d'accéder à n'importe quelle feuille du classeur, en choisissant l'une des feuilles sur le menu qui se trouve à gauche de la page.



Liste des feuilles d'un classeur

L'utilisateur peut choisir le niveau de détail de son analyse par un simple clique, ici en cliquant sur l'année, une autre page s'ouvre affichant les résultats par trimestre et ainsi de suite, comme le montre la figure suivante.

Ici le niveau de détail des résultats est par trimestre.

Connexion > Classeurs >
ETAT CREANCE - MTC par Branche et par date
 Dernière exécution 27 septembre 2007 14:31:40 CEST

Actions
 Réexécuter l'interrogation
 Enregistrer
 Enregistrer sous
 Rétablir la version enregistrée
 Page imprimable
 Exporter
 Envoyer en tant qu'email

Feuilles de calcul
 MTC par Branche et par da...
 MTC par Branche
 MTC par Zone et par Date
 MTC par Catégorie et par...
 MTC par Catégorie
 MTC par Clients Important...
 MTC par Famille Produit
 MTC par Type Créance
 MTC par Client et par Dat...

Table
 ▶ Outils Mise en page Format Faux-tricolores Trier Lignes et colonnes

Eléments de page Designation BRANCHE GPL

▶ Montant créance ▶ Datedoc: Quarter

Id	Montant	Trimestre
1	43406078,87	Q1
2	31549301,58	Q2
3	56541366,31	Q3
4	33431005,37	Q4

Graphique
 ▶ Outils Graphique Données

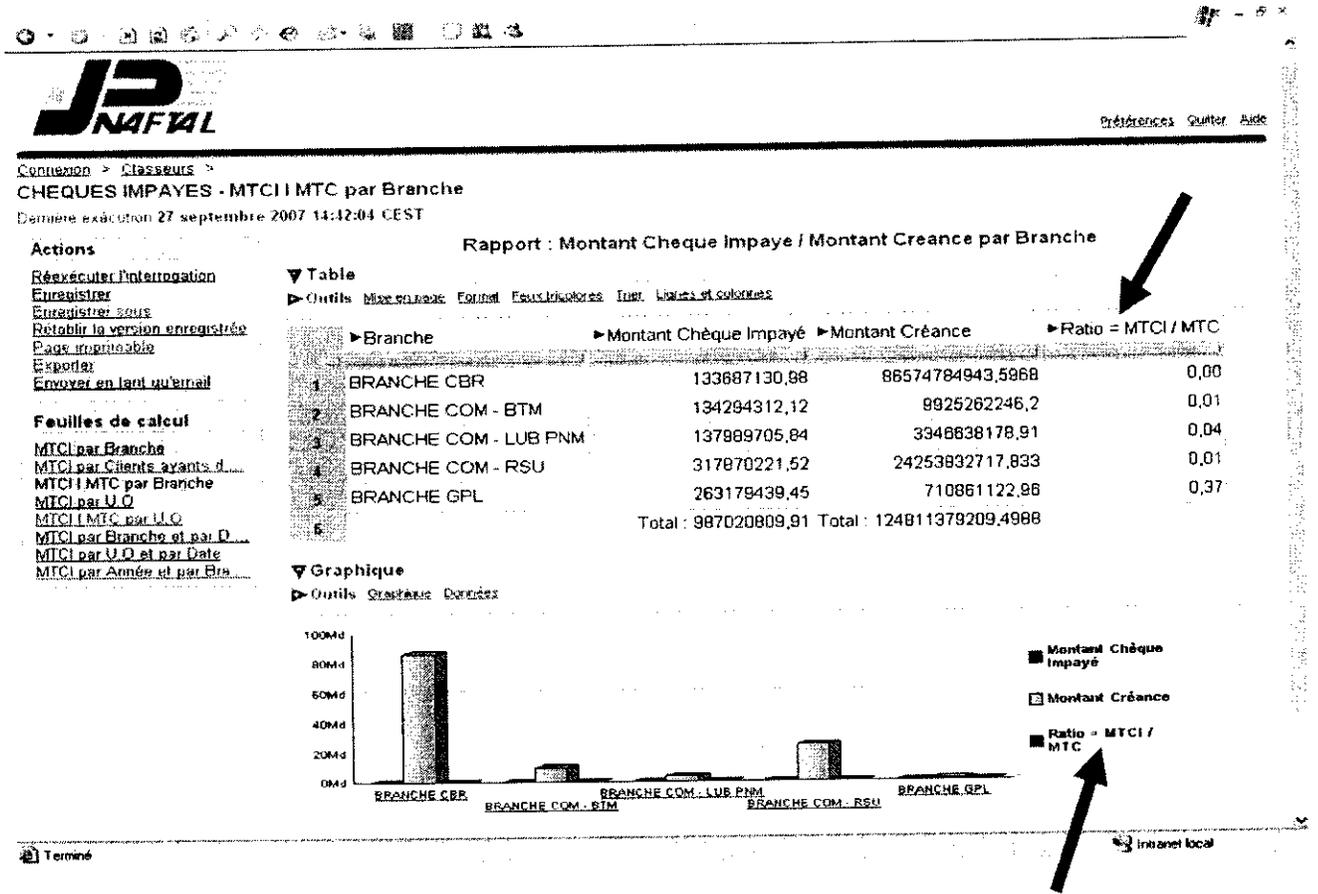
Eléments de page Designation BRANCHE GPL

Legend: Montant créance

X-axis: Datedoc: Year (Q1, Q2, Q3, Q4)

Trimestre

Cette figure représente un exemple de calcul de ratio à partir d'autres indicateurs, avec affichage des résultats sous forme de graphe comparatif.



Conclusion générale :

Le travail qui nous a été confié, était de concevoir et réaliser un système décisionnel basé sur l'architecture des entrepôts de données « Data Warehouse » pour le pilotage des créances des clients de différents types de l'entreprise NAFTAL Spa.

Lors de la conception et le déploiement de notre système, nous avons utilisé une méthodologie basée sur l'approche du cycle de vie dimensionnel initié par Ralph Kimball, l'un des pionniers dans le domaine des entrepôts de données.

Notre démarche est principalement structurée en trois grandes phases : Identification des besoins, Conception du système et en fin le Déploiement du système.

Dans la première phase, nous avons organisé des entretiens avec les dirigeants, les responsables, les analystes et les audits de données en vue d'identifier les différents besoins des décideurs et des utilisateurs. Comme conclusion de cette phase, nous avons extraire l'ensemble des axes (dimension) et indicateurs sur lesquels est basée la conception de notre modèle dimensionnel.

La deuxième phase qui est la conception du système peut être elle-même divisée en trois étapes :

- ❖ La conception du modèle dimensionnel qui consiste à faire correspondre les besoins de l'ensemble des utilisateurs avec la réalité des informations disponibles. Pour chaque activité est associé un groupe de dimension et indicateurs. L'avantage majeur de notre modèle proposé est d'être simple et facilement compréhensible par toute personne. La mise en œuvre est basée sur la modélisation dans des cubes en utilisant Oracle 9i OLAP Service.
- ❖ La construction de la zone d'alimentation qui est à déclenchement automatique sur une durée prédéterminée (un mois dans notre cas). Ce processus récupère les données de la source vers le data Warehouse.
- ❖ L'exploitation de l'entrepôt de données par un portail de restitution qui permet d'analyser les données stockées dans le data Warehouse.

La dernière phase concerne le déploiement du système. La base de données de l'entrepôt est créée et réalisée avec le système opérationnel par l'intermédiaire de la zone d'alimentation que nous avons programmé. Nous avons aussi créé la base de données multidimensionnelle et installé l'application de l'interface utilisateur. Notre système décisionnel est implémenté sur :

- Oracle 9i comme serveur de données de l'entrepôt et Oracle 9i OLAP Service comme serveur de la base de données multidimensionnelle, Intranet comme infrastructure réseau.

Ce projet nous a permis d'acquérir une connaissance pointue dans les systèmes d'informations décisionnels en particulier le data warehouse. Il présente une année de travail dans la quelle on a appris comment chercher et adopter les méthodes, la meilleure manière d'interroger les personnes afin d'obtenir les informations pertinentes et mettre la technologie à disposition de nos besoins.

Aussi, c'est notre première mise en œuvre de nouveaux systèmes décisionnels. Ces

systemes sont d'actualité pour un pilotage et une gestion efficace de l'entreprise. Nous avons tout de même éprouvé des difficultés qui nous ont empêché, dans une sorte, d'atteindre des résultats parfaits à savoir : le manque de la documentation et le manque des experts dans ce domaine.....

Perspectives :

En vue d'améliorer ses performances dans le domaine du décisionnel et afin d'atteindre une maturité en la matière, l'entreprise NAFTAL Spa, doit faire face à plusieurs défis :

- Les transferts de données des systèmes opérationnels vers le système décisionnel seront réguliers avec une périodicité bien choisie dépendante de l'activité de l'entreprise. Chaque transfert sera contrôlé avant d'être diffusé.
- Les entrepôts de données assureront une cohérence globale des données. Pour cette raison, leur alimentation sera un acte réfléchi et planifié dans le temps. Un grand nombre d'informations sera importé des systèmes transactionnels lorsqu'on aura la garantie que toutes les données nécessaires auront été produites et mémorisées.
- Une caractéristique importante des entrepôts, qui est aussi une différence fondamentale avec les bases de production, est qu'aucune information n'y est jamais modifiée. En effet, on mémorise toutes les données sur une période donnée et terminée, il n'y aura donc jamais à remettre en cause ces données car toutes les vérifications utiles auront été faites lors de l'alimentation. L'utilisation se résume donc à un chargement périodique, puis à des interrogations non régulières, non prévisibles, parfois longues à exécuter.
- Un entrepôt de données doit évoluer en fonction des demandes utilisateurs ou des nouveaux objectifs de l'entreprise.
- Mettre en œuvre de nouveaux systèmes décisionnels pour les métiers : commercialisation, ressources humaines dans le futur proche.

Glossaire

Agrégation	Assemblage des données hétérogènes formant un ensemble homogène.
Alimentation	Phase initiale de constitution de l'entrepôt de données, qui consiste à extraire les données de production, les rendre homogènes, et les charger dans la base de données décisionnelle. Les outils d'alimentation sont également appelés ETL (voir ETL).
Business Intelligence	Regroupement des technologies et principes d'aide à la décision. La Business Intelligence regroupe les fonctions de reporting et d'analyse des données.
Cube	Un nom pour une base de données dimensionnelle, utilisé habituellement lorsqu'on se réfère uniquement au cas simple du temps, du produit ou du marché.
Data mart	Data warehouse lié à un métier ou une fonction de l'entreprise (magasin de données).
Data Mining	Techniques statistiques de découverte des données. Ces outils utilisent des algorithmes pointus pour découvrir dans les données des informations qui n'apparaîtraient pas instinctivement à l'utilisateur.
Data Warehouse	Entrepôt de données. Lieu de stockage des gros volumes d'informations, destinées à être analysées.
Décision	Processus qui, en s'appuyant sur des données, conduit à initier une action.
Dimension	Principal critère d'analyse des informations stockées dans un cube multidimensionnel.
Drill down	Forage vers le bas.
ETL	Extract, Transform and Load.
Indicateur	C'est l'élément de base d'un cube de données multidimensionnel. Il est analysé par dimension, et le cube conserve les croisements de chaque dimension pour chaque indicateur.
Infocentre	Ancêtre des systèmes décisionnels. Il consiste en la copie à l'identique des bases de production dans un nouvel environnement, dédié aux requêtes décisionnelles.
Métadonnées	« les données sur les données » : L'ensemble des informations liées à une donnée de base : son origine, sa date de mise à jour, sa date d'import dans la base de données, la personne qui l'a modifié en dernier, sa date de validité, son format... Les métadonnées sont indispensables à l'échange d'informations entre les briques du système décisionnel. Elles sont stockées dans le référentiel (voir « référentiel »).

Multidimensionnel	Se dit d'une requête ou d'une base de données dans laquelle chaque indicateur est analysé en fonction de plusieurs critères, les dimensions.
OLAP	On-Line Transactional Processing – Modèle conceptuel des bases de données multidimensionnelles, inventé par le Dr Codd. Ce modèle a connu quelques variations (MOLAP, ROLAP, ...) qui ne sont que des variantes et qui restent rattachées au même modèle initial.
OLTP	On-Line Transactional Processing – Modèle conceptuel des bases de données relationnelles inventé par le Dr Codd.
Référentiel	Base de données dans laquelle seront stockées les métadonnées (voir « métadonnées ») et qui permet à toutes les applications du système décisionnel de s'échanger des données homogènes.
Relationnel	Se dit d'une base de données dans laquelle chaque information est stockée sous forme de table à deux dimensions. Les relations entre ces tables sont créées sous forme de « jointures ».
Reporting	Outil permettant de réaliser des tableaux, des rapports, sous forme de colonnes de chiffres ou de graphiques, qui seront ensuite diffusés dans l'entreprise ou à l'extérieur.
Requête	Question posée à la base de données sous forme de champs assortis de conditions. Les requêtes utilisent fréquemment le langage SQL (voir « SQL »).
Restitution	Phase ultime du système décisionnel, qui permet de visualiser sous forme de rapports, de graphiques, d'analyses, les données stockées dans les bases décisionnelles.
SQL	Structured Query Language, langage de requêtes qui permet d'interroger toute base compatible en utilisant les mêmes ordres.
Stockage	Phase intermédiaire de conservation des données dédiées à l'analyse dans un entrepôt ad hoc, qui peut être relationnel ou multidimensionnel.
Table	Les données sont stockées dans les bases de données relationnelles sous forme de tables, c'est-à-dire de feuilles à deux dimensions (lignes et colonnes).
Tableau de bord	Outil de synthèse qui affiche les principaux indicateurs de l'entreprise et permet de la piloter.
Table de faits	La table centrale du schéma de jointure en étoile.
Transaction	Dialogue informatique entre une application et un utilisateur qui aboutit à la prise d'une commande, l'édition d'un document.

Bibliographie :

Les ouvrages :

- L01 : « Concevoir et déployer un Data Warehouse – Guide de conduite de projet »
Ralph Kimball, Laura Reeves, Margy Ross, Warren Thornwaite
(Eyrolles 2000)
- L02 : « Entrepôts de données – Guide pratique de modélisation dimensionnelle »
Ralph Kimball, Margy Ross
(Vuibert 2003 – 2^{ème} édition)
- L03 : « Piloter l'entreprise grâce au Data Warehouse »
Jean-Michel Fronco, Sandrine de Lignerolles
(Eyrolles 2001 - 2^{ème} tirage)
- L04 : « La construction du Data Warehouse – du Data Mart au Data Web »
Jean-Fraçois Goglin
(Hemes 2001 – 2^{ème} édition)
- L05 : « Le Data Webhouse »
Ralph Kimball, Richard Merz
(Eyrolles 2000)
- L08 : « Le projet décisionnel, Enjeux, Modèles, architecture du Data Warehouse »
Marie Gouarne
(Eyrolles 1998)
- L09 : « Building the Data Warehouse »
William H. Inmon
(WILEY 2002 3^{ème} édition)
- L11 : « Introduction au Data Warehousing »
Kevin Viers
(Le Compus 2002)
- L12 : « Les systèmes interactifs d'aide à la décision »
Peter G.W. Keen & Michael S. Scott Morton
(Addison-Wesley 1978)
- L13 : « Aide multicritère à la décision : Méthodes & cas »
Bernard Roy & Denis Bouyssou
(Economica 1993)

Les sites :

W01 : <http://www.supinfo-projects.com>

W02 : <http://www.datawarehousingonline.com>

W03 : <http://www.businessobjects.com>

W04: <http://www.decisionnel.net>

W05: <http://www.piloter.org>

W06: <http://www.naftal.dz>

W10 : <http://univ-valenciennes.fr>

W11 : <http://www.cait.wustl.edu>

W12 : <http://www.olapcouncil.org>

Mémoire :

M01 : « Conception et déploiement d'un Data Warehouse métier pour la fonction commerciale de Naftal Spa »
Haceini H.L & Sebihi G. (USDB 2005-2006)

Article :

A1 : « Système interactif d'aide à la décision, in dictionnaire critique de la Communication »
(L.Sfez Edr, PUF 1993)