

MA-004-217-1

République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Saâd Dahlab, Blida



Faculté des sciences

Département Informatique

Mémoire de fin d'études

Pour l'obtention du diplôme de Master en informatique

Option : Systèmes d'Information

Thème :

**Conception et réalisation d'un système décisionnel pour le suivi
des coûts de production de la SONATRACH en association
(Activité AMONT).**



Réalisé par

EL ARBI RABAH Mohamed

OUZZANI Fatima Zohra

Devant le jury composé de :

Président :

Examineur :

Examineur :

Encadré par :

Mme. N.Rezoug

M. A.Felfoul

Promotion : 2013/2014

Remerciements

Avant toute chose, Nous remercierons ALLAH le tout puissant, de nous avoir donnée la force et la patience pour mener à terme ce travail

Nous présentons nos sincères remerciements à notre promotrice Mme REZOUG Nachida et notre encadreur M. FELFOUL Allaoua pour avoir assuré le suivi de ce projet mais aussi pour leur soutien, leurs conseils et leur disponibilité.

Nous tenons également à remercier, M. HASSAINE Yazid, chef département de gestion Activité AMONT, pour nous avoir acceptés et suivis tout au long de ce projet de fin d'études.

Nous tenons à remercier aussi tous nos enseignants qui ont participé à notre formation tout au long de notre cursus.

Nous tenons à remercier nos parents, nos familles et amis pour leur soutien tout au long de ce mémoire.

Nous remercions également toutes ces personnes sur internet qui ont partagé leurs connaissances et expériences et répondu à nos questions, ils nous ont été d'une grande aide.

Enfin, nous remercions tous ceux qui ont aidé de près ou de loin à l'élaboration de ce mémoire et à la réussite de ce projet.

Dédicaces

Je dédie ce travail :

À mes parents,

À ma sœur, À mes frères,

À tous les membres de ma famille,

À mon binôme,

À tous mes amis,

*À la personne qui m'a supporté et soutenu tout long de ce travail, à
NISSA*

À tous ceux qui me sont chers

Mohamed

Dédicaces

Je dédie ce travail :

*À mon adorable mère **NOUR EL HOUDA***

Qui j'ai aimé bien sa présence dans ce jour ; Que Dieu ait pitié et entrez paradis sans détriment.

*À mon père **ABD EL KADER***

Grâce à sa tendre, encouragement et son grand sacrifice, il a pu créer le climat affectueux et propice à la poursuite de mes études. Aucune dédicace ne pourrait exprimer mon respect, ma considération et mes profonds sentiments envers lui. Je prie le bon Dieu de le bénir, de veiller sur lui, en espérant qu'il se sent fière de moi.

*À ma jeune sœur et mon amour **AICHA** et à mes chers frères **ALI** et **ACHREF**, Ma fierté est leur présence dans ma vie*

*À l'épouse de mon père **RBIHA** pour son soutien moral et ses sacrifices le long de ma formation.*

*À toute la famille **OUZZANI**. Je le dédie à tous mes oncles et tantes, cousins et cousines Ils vont trouver ici l'expression de mes sentiments de respect et de reconnaissance pour le soutien qu'ils n'ont cessé de me le porter.*

*Sans oublier la précieuse aide de mon très cher fiançais **ALI** qui a été toujours présent pour m'aider et m'encourager et sa mère **Amina**, son père **Ahmed**, ses frères.*

*À la personne qui m'a toujours aidé et encouragé, mon binôme **Mohamed***

*À mes chères amies **yasmina**, **maha-emily** elles sont toujours à mes côtés.*

Fatima Zohra

Résumé

L'accès à la bonne information au bon moment telle est la principale préoccupation des entreprises qui veulent assurer leur croissance et leur pérennité. Evoluant dans des environnements à la fois complexes et compétitifs, les dirigeants s'orientent de plus en plus vers la mise en place de systèmes d'aide à la décision, aux seins de leurs entreprises, leur fournissant l'information nécessaire au bon pilotage opérationnel et stratégique.

SONATRACH est l'opérateur historique et leader du domaine de production des hydrocarbures et hydrocarbures gazeux ainsi que leur commercialisation en Algérie, a pour objectifs de préserver la place qu'elle occupe et de rester en constante évolution. Dans ce cadre et afin de répondre aux besoins des dirigeants en termes d'aide à la décision, ce qui a conduit à la mise en place de projets de systèmes d'aide à la décision.

Il nous a été confié, dans le cadre de ce projet de fin d'étude, le développement d'un système décisionnel pour le suivi des coûts de production de la Sonatrach en association concernant l'activité AMONT. Ce système a pour objectif de donner à tout moment une vision claire sur la situation actuelle de l'entreprise et ses partenaires (dans le domaine de production), et de permettre l'accès aux informations nécessaire pour le pilotage et la prise de décision. Notre travail consiste en la mise en place d'un entrepôt de données accessible de manière simple et personnalisable (à travers des tableaux croisés, de graphes, de rapport, tableau de bord... etc.,)

Mots-clés : Système décisionnel, entrepôt de données, Modélisation multidimensionnelle, ETL, Tableau de bord, Reporting, OLAP.

Abstract

Accessing to the right information at the right time that is the main concern for companies who want to ensure their growth and sustainability. Operating in complex and competitive environments, managers are moving more and more towards the use of decision support systems in their businesses, to provide them with the information necessary for the proper operational and strategic management.

Sonatrach is the incumbent and leader in the field of oil and natural gas production and marketing in Algeria hydrocarbons. It aims to preserve its place and stay in constant evolution. In this context and in order to meet the needs of leaders in terms of decision support, which has led to the implementation of projects of decision support systems.

In the context of this graduation project, we were in charge of the development of a decision support system for monitoring production costs of Sonatrach in association regarding AMONT activity. To give at any moment a clear vision of the current situation of the company and its partners (in the field of production), and allow access to the information necessary for the management and decision-making system aims. Our work involves the establishment of a warehouse that can be accessed easily and customizable way (through crosstabs, graphs, reporting, dashboard ... etc.,)

Keywords: Decision-making System, data warehouse, multidimensional modeling, ETL, Dashboard, Reporting, OLAP.

ملخص :

الحصول على المعلومات الصحيحة في الوقت المناسب هو الشغل الشاغل للشركات التي تريد ضمان نموها واستدامتها . بحيث تعمل هذه الشركات في بيئة معقدة وتنافسية، مما جعل مسيرتها يتجهون أكثر فأكثر نحو تطوير نظم دعم القرار، لتزويدهم بالمعلومات اللازمة لإدارة العمليات والاستراتيجيات المناسبة .

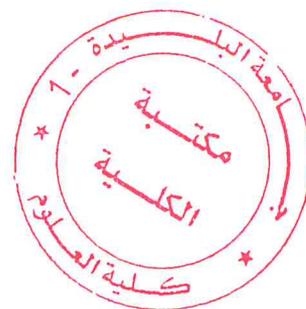
سوناطراك هو المعامل الرائد في مجال إنتاج النفط والغاز الطبيعي وتسويق النفط في الجزائر، ويهدف إلى الحفاظ على مكانته والبقاء في تطور مستمر. في هذا السياق، ومن أجل تلبية احتياجات المسيرين من حيث دعم اتخاذ القرار، الأمر الذي أدى إلى إدراج مشروع لإنجاز نظام دعم القرار .

في إطار مشروع التّخرُج كلفنا بتطوير نظام لدعم اتخاذ القرارات لمتابعة و رصد تكاليف الإنتاج لسوناطراك بالشراكة فيما يخص النشاط AMONT لإعطاء في أي لحظة رؤية واضحة للوضع الحالي للشركة وشركائها (في مجال الإنتاج) والسماح بالوصول إلى المعلومات اللازمة بهدف التحكم واتخاذ القرار.

يشمل عملنا على إنشاء مستودع بيانات يسمح لمستخدميه الاستغلال الأمثل للمعلومات بطريقة ميسرة وبسيطة من خلال الجداول المتقاطعة، الرسوم البياني، التقارير، لوحة القيادة ... الخ .

كلمات البحث: نظام دعم القرار، مستودع البيانات، نمذجة متعددة الأبعاد، استخراج-تحويل-تحميل، لوحة القيادة، التقارير، المعالجة-التحليلية-المتصلة.

Table des matières



Résumé.....	V
Abstract.....	VI
ملخص	VII
Table des matières.....	VIII
Liste des figures	XIII
Liste des tableaux.....	XV
Liste des abréviations.....	XVI

Introduction générale.....	1
-----------------------------------	----------

PARTIE 1 : ETAT DE L'ART

Chapitre I Les systèmes décisionnels	6
Introduction.....	6
I.1. Définition.....	7
I.2. Architecture d'un système décisionnel	7
I.2.1. Les sources de données.....	8
I.2.1.1. Définition.....	8
I.2.1.2. Le processus d'Extraction-Transformation-Chargement.....	9
I.2.2. Entrepôt et magasins de données.....	10
I.2.2.1. Définitions	10
I.2.2.2. Catégories de données stockées dans un entrepôt de données	12
I.2.2.3. Approches de mise en place d'un entrepôt de données	13
I.2.3. Serveurs OLAP	15
I.2.4. Outils d'analyse et de restitution	17
Conclusion.....	17
Chapitre II La modélisation multidimensionnelle.....	18
Introduction.....	18
II.1. Définition.....	19

II.2. Modèles de présentation des données.....	20
II.2.1.Modèle conceptuel.....	20
II.2.1.1.Concept de fait.....	20
II.2.1.2.Concept de dimension	20
II.2.1.3.Modèle en étoile (Star schema).....	21
II.2.1.4.Modèle en flocon (snowflake).....	22
II.2.1.5.Modèle en constellation.....	23
II.2.2.Modèle logique.....	23
II.2.2.1.R-OLAP	23
II.2.2.2.M-OLAP	24
II.2.2.3.H-OLAP	24
II.2.2.4.D-OLAP	24
II.3. Manipulation multidimensionnelle des données	25
II.3.1.Opérations classiques	25
II.3.2.Opérations agissant sur la granularité.....	26
II.3.3.Opérations agissant sur la structure.....	26
Conclusion.....	28

PARTIE 2 :

ETUDE DE L'EXISTANT ET COLLECTE DES BESOINS

Chapitre III Présentation de l'organisme d'accueil.....	30
Introduction.....	30
III.1. Présentation de Sonatrach	31
III.1.1. Sonatrach en bref	31
III.1.2. Les missions de SONATRACH	31
III.1.3. Organisation de SONATRACH	32
III.2. Contexte d'étude	33
III.2.1. L'activité Amont	33
III.2.2. La direction Contrôle de Gestion au sein de la SONATRACH	35

III.2.2.1. la dépendance hiérarchique	35
III.2.2.2. Les objectifs de la direction Contrôle de Gestion	36
III.2.2.3. Organisation de la direction Contrôle de Gestion	37
III.3. Présentation de la direction d'accueil	38
Conclusion	39
Chapitre IV Etude de l'existant et Etude des besoins.....	40
Introduction.....	40
IV.1. Etat du décisionnel.....	41
IV.2. Limites du reporting existant	44
IV.3. Sources de données.....	45
IV.2. Etude des besoins.....	46
IV.2.1. Introduction	46
IV.2.2. Définition des besoins	46
IV.2.2.1. Elaboration des questionnaires	46
IV.2.2.2. Les entretiens	48
IV.2.2.3. Recensement et validation des besoins.....	48
Conclusion.....	54

PARTIE 3 : CONCEPTION DE LA SOLUTION

Chapitre V Conception de la zone d'entreposage.....	56
Introduction.....	56
V.1. Activité de production	57
V.1.1. Présentation de l'activité	57
V.1.2. Les faits	57
V.1.2.1. Suivi des charges	57
V.1.2.2. Suivi des amortissements de développement	58
V.1.2.3. Suivi de production	59
V.1.3. Les dimensions.....	60
V.1.3.1. Dimension Temps	60

V.1.3.2. Dimension Association	61
V.1.3.3. Dimension Nature-charge	62
V.1.3.4. Dimension Périmètre	62
V.1.3.5. Dimension Produit	63
V.1.4. Dimensions participantes pour chaque fait	64
V.1.5. Schéma globale de l'activité « Production »	64
Conclusion.....	65
Chapitre VI Conception de la zone d'alimentation	66
Introduction.....	66
VI.1. Identification des sources de données	67
VI.2. Extraction des données	67
VI.3. Transformation	68
VI.4. Chargement	69
Conclusion.....	70
Chapitre VII Conception des cubes dimensionnels.....	71
Introduction.....	71
VII.1. définition des niveaux et hiérarchies des dimensions.....	72
VII.2. Listes des cubes	73
VII.3. schémas des cubes	74
VII.3.1. volet suivi des charges	74
VII.3.2. volet suivi des amortissements de développement	74
VII.3.3. volet suivi de production	75
Conclusion.....	76

PARTIE 4 : REALISATION ET DEPLOIEMENT

Chapitre VIII Réalisation et déploiement de la solution	78
Introduction.....	78
VIII.1. Environnement technique et fonctionnel	79
VIII.1.1. Matériel	79

VIII.1.2. Outils utilisés.....	79
VIII.1.2.1. Stockage de la base de données « Oracle 11g ».....	79
VIII.1.2.2. Extraction, transformation et chargement « Talend Open Studio »	80
VIII.1.2.3. Construction des cubes « Oracle OLAP 11g »	81
VIII.1.2.4. Reporting « Oracle BI Publisher Enterprise »	81
VIII.1.2.5. Plateforme de restitution « Oracle Business Intelligence (OBI) »	82
VIII.2. Réalisation du projet	83
VIII.2.1. Construction de l'entrepôt de données	83
VIII.2.2. Extraction, transformation et chargement	83
VIII.2.3. Construction et chargement des cubes	85
VIII.2.4. Mise en place du reporting	88
VIII.2.5. Tableau de bord	89
VIII.2.5.1. l'accès au tableau de bord	90
VIII.2.5.2. Visualisation des rapports prédéfinis	91
VIII.3. Utilisateurs du système	92
VIII.4. la sécurité.....	92
Conclusion.....	92
Conclusion générale.....	94
Références bibliographiques	96

Liste des figures

Figure 1 : Architecture d'un système décisionnel.....	8
Figure 2 : Structure de données d'un entrepôt de données.....	12
Figure 3 : Exemple d'un cube de données.....	19
Figure 4 : Exemple de fait.....	20
Figure 5 : Exemple de dimensions.....	21
Figure 6 : Exemple d'une modélisation en étoile.....	22
Figure 7 : Exemple d'une modélisation en flocon.....	22
Figure 8 : Exemple de modélisation en constellation.....	23
Figure 9 : Traduction R-OLAP d'un schéma étoile.....	24
Figure 10 : Principe de "slice & dice".....	25
Figure 11 : Principe du forage.....	26
Figure 12 : Principe de rotation.....	27
Figure 13 : Schéma de la macrostructure de la SONATRACH.....	33
Figure 14 : Organigramme de l'Activité AMONT.....	34
Figure 15 : Rattachement et hiérarchique du contrôle de gestion.....	36
Figure 16 : Organigramme de la Direction Contrôle de gestion.....	37
Figure 17 : Organigramme de la Direction ISI.....	38
Figure 18 : Les dimensions métier du volet charges de production.....	48
Figure 19 : Les dimensions métier du volet amortissement de développement.....	49
Figure 20 : Les dimensions métier du volet de production.....	49
Figure 21 : Le fait « Charge ».....	57
Figure 22 : Le fait « Amortissement de développement ».....	58
Figure 23 : Le fait « Production ».....	59
Figure 24 : La dimension « Temps ».....	60
Figure 25 : La dimension « Association ».....	61
Figure 26 : La dimension « Nature-charge ».....	62
Figure 27 : La dimension « Périmètre ».....	62
Figure 28 : La dimension « Produit».....	63

Figure 29 : schéma en constellation de l'activité « Production ».....	64
Figure 30 : Exemple de changement des libellés	68
Figure 31 : Diagramme d'activité de processus de chargement	69
Figure 32 : Cube dimensionnel « Charge ».....	74
Figure 33 : Cube dimensionnel « Amortissement Développement »	74
Figure 34 : Cube dimensionnel « Production ».....	75
Figure 35 : Architecture technique de la solution	83
Figure 36 : job de remplissage de la dimension Association	84
Figure 37 : job de remplissage de la dimension Temps	85
Figure 38 : job de remplissage du fait Production	85
Figure 39 : Construction des dimensions avec AWM	86
Figure 40 : Construction des cubes avec AWM	87
Figure 41 : Exploitation du cube « Production » avec AWM	87
Figure 42 : Rapport sous forme d'un tableau dynamique généré par BI Publisher	88
Figure 43 : Rapport sous forme d'un histogramme généré par BI Publisher	89
Figure 44 : Page d'accès au Tableau de Bord	90
Figure 45 : Page d'accueil du Tableau de Bord	91
Figure 46 : Exemple de rapport	91

Liste des tableaux

Tableau 1 : Avantages et inconvénients de l'approche Top-Down	13
Tableau 2 : Avantages et inconvénients de l'approche Bottom-Up	14
Tableau 3 : Comparaison des processus OLTP et OLAP	16
Tableau 4 : Questionnaire aux responsables du département Gestion «Activité Amont »	47
Tableau 5 : Description de l'indicateur « Taux de réalisation mensuel	50
Tableau 6 : Description de l'indicateur « Taux de réalisation cumulé»	50
Tableau 7 : Description de l'indicateur « Taux d'avancement »	51
Tableau 8 : Description de l'indicateur « Taux d'évolution»	51
Tableau 9 : Description de l'indicateur « Ecart »	52
Tableau 10 : Description de l'indicateur « Coût »	52
Tableau 11 : Tableau descriptif du fait « Charge ».....	58
Tableau 12 : Tableau descriptif du fait « Amortissement de développement»	59
Tableau 13 : Tableau descriptif du fait « production ».....	60
Tableau 14 : Tableau descriptif de la dimension « Temps »	61
Tableau 15 : Tableau descriptif de la dimension « Association »	61
Tableau 16 : Tableau descriptif de la dimension « Nature-charge»	62
Tableau 17 : Tableau descriptif de la dimension « Périmètre».....	63
Tableau 18 : Tableau descriptif de la dimension « Produit»	63
Tableau 19 : Tableau descriptif des dimensions participantes pour chaque fait	64
Tableau 20 : Tableau donnant les niveaux et les hiérarchies de chaque dimension.....	72
Tableau 21 : Liste des cubes	73

Liste des abréviations

BI: Business Intelligence

Dim_: Dimension

D-OLAP: Desktop Online Analytical Processing

FK: Foreign Key

H_: Hiérarchie

H-OLAP: Hybride Online Analytical Processing

Id: Identifiant

M-OLAP: Multidimensional Online Analytical Processing

OLAP: OnLine Analytical Processing

OLTP: Online Transaction Processing

PK: Primary Key

R-OLAP: Relational OnLine Analytical Processing

SGBD : Système de Gestion de Base de Données

SQL : Structured Query Language

SONATRACH : Société Nationale pour la Recherche, la Production, le Transport, la Transformation, et la Commercialisation des Hydrocarbures

DCG : Direction de coordination Groupe

ISI : Informatique et Système Information

TRC : Transport par Canalisation

PDF : Portable Document Format

TEP : Tonne Equivalent Pétrole

KDA : Kilo Dinar Algérien

ETL : Extraction Transformation Load

AWM : Analytic Workspace Manager

XML : eXtensible Markup Language

OBI : Oracle Business Intelligence

3FN : Troisième Forme Normale

PED : Petroleum Engineering et Développement.

ENC : Engineering et Construction

Introduction générale

Contexte :

Face à la mondialisation, la concurrence et la crise subie par le système financier mondial, les entreprises ont été incitées à examiner de près l'ensemble de leurs projets et de leurs dépenses. En outre, la prise de décision est devenue encore plus cruciale pour les dirigeants d'entreprises.

A cet effet, les directions financières doivent pouvoir offrir aux entités dirigeantes une vision synthétique et suffisamment détaillée de la situation économique de l'entreprise.

Pour cela, il ne suffit pas d'enrichir son système d'information de données multiples, correspondant aux flux générés par l'ensemble des événements commerciaux, matériels, juridiques et économiques, il faut aussi pouvoir les analyser, les corréler et fournir à l'ensemble des décideurs les moyens de prendre les bonnes décisions à court, moyen et long terme.

L'efficacité de la prise de décision repose sur la mise à disposition du décideur des informations pertinentes au moment opportun, et des outils adaptés à l'analyse. Des exigences auxquels ne répondent pas les systèmes opérationnels, et qui ont donné lieu à l'élaboration de nouveaux systèmes, dits décisionnels.

Les systèmes décisionnels reposent sur un espace de stockage centralisé, appelé entrepôt de données, et proposent des fonctionnalités d'extraction et d'analyse. Ils permettent notamment de collecter des informations provenant de sources différentes afin de les exploiter aux travers d'interfaces de représentation.

Dans un tel contexte économique, les systèmes décisionnels représentent une opportunité pour les entreprises d'optimiser le pilotage de leurs activités. C'est ce qui a poussé Sonatrach, opérateur historique dans le domaine des hydrocarbures et hydrocarbures gazeux en Algérie, à vouloir se doter d'un tel système pour optimiser sa gestion financière, et ainsi maintenir sa position de leader.

A cet effet, Sonatrach nous a confié dans le cadre de notre projet de fin d'étude, de concevoir et de réaliser un système décisionnel pour le suivi des coûts de production de la Sonatrach en association (Activité AMONT)

Problématique

Selon un sondage réalisé par la société d'entreposage de données « Teradata¹ », près des trois quarts des professionnels de la finance ont du mal à recueillir des données provenant de sources multiples, d'effectuer les analyses nécessaires pour répondre aux questions sur la performance financière de l'entreprise, et de faciliter la prise de décision.

SONATRACH est en association avec 30 partenaires dans l'exploration et la production des hydrocarbures et hydrocarbures gazeux, disposant chacun d'eux son propre champ pétrolier.

Ce nombre important d'associés engendre une complexité dans l'analyse et le suivi des coûts de production au niveau de la direction de contrôle de gestion (activité AMONT), et rend ainsi la prise de décision pour un bon pilotage difficile et lent. En effet les contrôleurs de gestion ont exprimé un certain nombre de problèmes, que nous citons ci-dessous :

- la direction du groupe dépend des rapports périodiques fournis par les associés, les données sont décentralisées, alors l'information n'est pas disponible en temps voulu.
- Non historisation des données, seule les opérations des trois derniers exercices comptables sont maintenues, et les exercices antérieurs sont juste archivés.
- Manque d'outils de reporting : les différentes analyses sont basées sur des tableaux croisés proposés par Excel qui fournissent quelques indicateurs, mais qui ne permettent pas de synthétiser la situation financière de l'entreprise ni de faciliter la prise de décision.

¹ Teradata (<http://www.teradata.com/>) est un leader mondial dans la construction et l'édition de solutions informatiques spécialisées en matière d'entrepôt de données et d'applications analytiques.

Objectifs :

Dans le cadre de notre projet et afin de répondre aux problèmes mentionnées dans la problématique, l'objectif de ce travail est de concevoir et implémenter un système décisionnel qui se propose de :

- Réduire le temps consacré à l'analyse.
- Permettre un accès facile et rapide à l'ensemble des données en les centralisant.
- Permettre d'établir un diagnostic spécifique aux besoins des décideurs en donnant l'accès à des indicateurs pertinents, fiables, facilement interprétables et mesurés dans le temps.
- Mettre à la disposition des utilisateurs, dans des délais très courts, des états de sorties multiples et personnalisés (rapport, graphes, tableau de bord ...).
- Offrir à l'utilisateur la possibilité de naviguer dans les données de manière interactive à l'aide des cubes OLAP.

Contenu du mémoire :

Pour présenter au mieux notre travail, nous avons structuré ce document en quatre parties. Avant d'entamer ces quatre parties nous avons introduit le contexte de notre étude et fixé la problématique ainsi que les objectifs du projet.

Le contenu des quatre parties peut être résumé comme suit :

- **Partie 1 Etat de l'art :** Il s'agit d'une partie théorique où nous présentons les notions relatives à notre projet dans les domaines des systèmes décisionnels et de la modélisation multidimensionnelles.
- **Partie 2 Etude de l'existant et collecte des besoins :** Nous abordons la première étape de la partie pratique de notre projet en présentant l'organisme d'accueil, l'étude de l'existant en terme de décisionnel puis nous exposons notre démarche de collecte des besoins.
- **Partie 3 Conception de la solution :** Dans cette partie, nous présentons les différentes étapes de la conception de la solution qui consiste en un système décisionnel qui s'articule autour d'un entrepôt de données. Nous abordons donc la conception de l'entrepôt, de sa zone d'alimentation et la conception des cubes dimensionnels.
- **Partie 4 Réalisation et déploiement :** C'est la dernière partie de notre projet où nous procédons à la réalisation et au déploiement de la solution conçue. Nous débutons avec la présentation de l'environnement technique et fonctionnel et des outils utilisés puis nous présentons les différentes étapes de la réalisation qui aboutiront à un système fonctionnel.

Nous clôturons notre document avec une conclusion générale où nous synthétisons notre travail.

PARTIE 1 : ETAT DE L'ART

Nous présenterons dans cette partie une synthèse bibliographique sur les systèmes décisionnels, ainsi que la modélisation multidimensionnelle.

Chapitre I

Les systèmes décisionnels

Introduction

L'information représente un capital immatériel dont la bonne gestion est un facteur primordial pour la réussite de toute entreprise.

Une entreprise moderne brasse dans son système d'information d'immenses volumes de données, or trop d'information tue l'information, donc il est très difficile de donner un sens à ces données, de comprendre ce qu'elles expriment : des tendances sous-jacentes, des faiblesses ou des forces cachées.

Le besoin d'une exploitation efficace des données dans une perspective décisionnelle a donné lieu à l'élaboration de nouveaux systèmes, dits systèmes décisionnels, facilitant le stockage et le traitement synthétique de grands volumes de données.

I.1. Définition

Il existe deux grandes familles de systèmes d'Information (SI):

- Les systèmes d'information opérationnels et,
- Les systèmes d'information décisionnels.

Les premiers, les SI opérationnels, sont conçus spécialement pour les opérations quotidiennes reliées au fonctionnement d'une entreprise, et souvent associés à des progiciels ou des applications orientées métier, fonctionnant dans un environnement multi-accès. Leur objectif principal est la saisie puis les traitements rapides de données, ainsi que la production de résultats en sortie (Naoum, 2006).

Les seconds, les systèmes d'information décisionnels appelés aussi systèmes décisionnels (en anglais Business Intelligence : BI) représentent l'ensemble des moyens, outils et méthodes qui supportent le processus de collecte, consolidation, modélisation, analyse et restitution des données issues des systèmes d'information opérationnels dans le but de faciliter la prise de décision (Lau et al., 2009).

Les systèmes décisionnels constituent une synthèse d'informations opérationnelles choisies pour leur pertinence et leur transversalité fonctionnelles.

I.2. Architecture d'un système décisionnel

Le processus d'un système décisionnel vise à récupérer des données brutes issues des différentes sources, internes ou externes, à les transformer en information afin de les diffuser sous forme de rapport ou de tableaux de bord (Lau et al., 2009).

Afin d'assurer ce processus, l'architecture d'un système décisionnel peut être représentée en quatre niveaux comme l'illustre la figure.

- Le premier niveau est celui des sources des données, internes ou externes, de l'entreprise,
- Le deuxième niveau concerne la récupération, la transformation des données, puis l'alimentation d'un entrepôt de données,

- Le troisième niveau est celui de l'exploitation de l'entrepôt de données en question, ou les données sont organisées d'une manière multidimensionnelle afin de supporter et de faciliter les requêtes analytiques de type OLAP,
- Le quatrième niveau fournit à l'utilisateur des outils de restitutions et d'analyse qu'ils lui permettent d'interroger et d'analyser les données sous une forme adaptée.

Dans la suite nous détaillons chaque niveau de cette architecture.

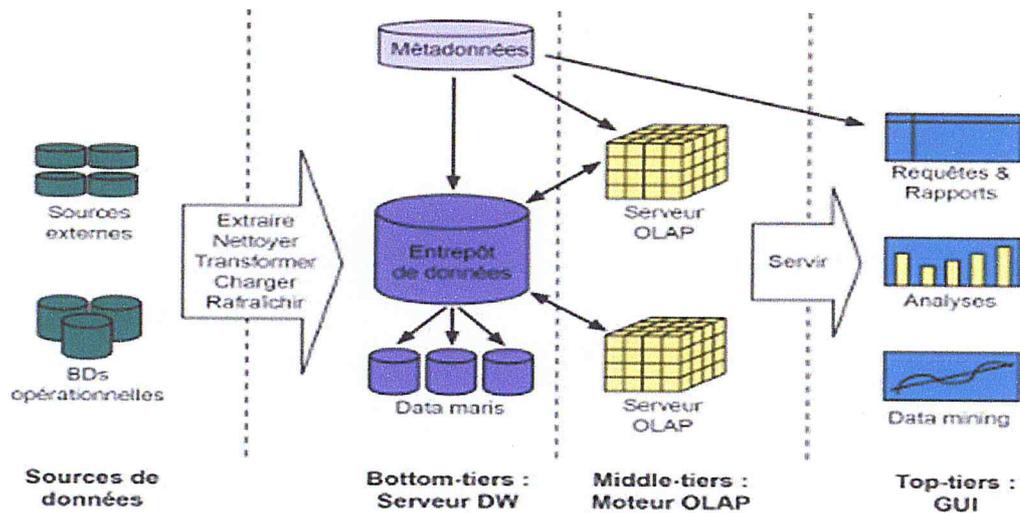


Figure 1 : Architecture d'un système décisionnel (Naoum, 2006)

I.2.1. Les sources de données

I.2.1.1. Définition

Afin d'alimenter les entrepôts, les informations doivent être identifiées et extraites de leurs emplacements originels. Ces données peuvent être regroupées en quatre catégories :

(Ponniah, 2001)

- **Les données de production** : ce sont les données provenant des différents systèmes opérationnels, Elles sont contenues dans des bases de données transactionnelles et elles sont caractérisées par leur disparité (plusieurs technologies, environnements différents, ...).
- **Les données internes** : il s'agit des données concernant l'entreprise mais détenues par des employés ou des départements dans leur fichiers et base de données personnels, leur acquisition est de ce fait complexe.

- **Les données archivées** : les données anciennes que le système opérationnel stocke périodiquement dans des fichiers ou base de données.
- **Les données externes** : données externes à l'entreprise, il peut s'agir de statistiques d'agences externes, d'indicateurs de performance... nécessaires pour se situer par rapport à la concurrence et repérer les tendances du marché.

I.2.1.2. Le processus d'Extraction-Transformation-Chargement

Les données sources vont alimenter l'entrepôt de données, mais il faut d'abord les préparer pour les stocker dans un format approprié pour l'analyse et le requêtage, et ainsi les rendre exploitables. Pour cela on passe par trois étapes connues sous l'acronyme ETL (Extraction, Transformation and Loading) (Gruau, 2004) :

- **Extraction** : on identifie les données pertinentes, destinées à être exploitées et en accord avec le résultat à obtenir. On les extrait à partir de multiples sources selon une fréquence déterminée. Les données extraites, provenant de sources hétérogènes, nécessitent d'être transformées.
- **Transformation**, qui regroupe les opérations de :
 - Nettoyage (filtrage des données manquantes, correction des erreurs, suppression des doublons),
 - Formatage et standardisation, (types de données, longueur des champs, résolution des synonymes et homonymes)
 - Fusion ou éclatement des informations composites
 - Agrégation des données (la somme, la moyenne, le comptage, la somme cumulée, le minimum, le maximum). Ces opérations sont à considérer compte tenu du niveau de granularité de l'entrepôt.
 - Affectation des clés de substitution on (surrogatekeys).
 - Dénormalisation.

Les données sont alors prêtes à être chargées.

- **Chargement** des données propres dans l'entrepôt de données. Il y a deux cas à considérer :
 - Le chargement initial : est effectué au tout premier chargement de l'entrepôt et dans des cas spéciaux comme après la perte des données de l'entrepôt. Dans ce cas qui est généralement facile, on charge toutes les données de l'entreprise dans l'entrepôt.
 - Le chargement incrémentiel : est le fait d'ajouter des données à un entrepôt existant, c'est l'opération qui va se répéter périodiquement (chaque jour par exemple) et automatiquement. Il faudra faire attention dans ce cas à ne charger que les informations nouvelles, et ne pas charger deux fois la même information et ne pas écraser les informations existantes. La validation des chargements et la possibilité de restauration des données en cas d'échec sont aussi des points importants à prendre en compte.

La fréquence à laquelle les phases ETL sont opérées doit permettre d'historiser les données avant qu'elles ne soient purgées des bases de production. Il est à noter que la réalisation de l'ETL représente la plus grande partie d'un projet décisionnel. Ce système est complexe et ne doit rien laisser s'échapper, sous peine d'avoir une mauvaise information dans l'entrepôt, donc des données fausses, donc inutilisables.

I.2.2. Entrepôt et magasins de données

A l'issue du processus de l'ETL, les données sont stockées et organisées dans un entrepôt de données. Le concept d'entrepôt de données a été introduit car les bases de données transactionnelles ne répondent pas aux besoins d'analyse.

I.2.2.1. Définitions

Ralph Kimball définit l'entrepôt de données comme un espace de stockage centralisé sur lequel repose un système décisionnel, son rôle est d'intégrer et de stocker l'information utile aux décideurs et de conserver l'historique des données pour supporter les analyses effectuées lors de la prise de décision. (Kimball and Ross, 2002).

Bill Inmon définit un entrepôt de données par ses caractéristiques : « *collection de données orientées sujet, intégrées, non volatiles et historisées, organisées pour le support d'un processus d'aide à la décision.* »(Inmon, 2002)

- **Données orientées sujet** : Contrairement au système opérationnel où les données sont organisées par processus fonctionnel, le Data Warehouse est organisé par thème.
- **Données intégrées** : pour assurer la cohérence, les données provenant des différents systèmes opérationnels doivent être mises en forme et unifiées avant d'être intégrées dans le Data Warehouse. Cela nécessite une forte normalisation, un référentiel unique et de bonnes règles de gestion.
- **Données non volatiles** : pour assurer la traçabilité, les données introduites dans l'entrepôt ne peuvent être ni modifiées, ni supprimées.
- **Données historisées** : les données sont stockées de manière à être réutilisées pour des fins d'analyse. Donc chaque donnée est étiquetée par la période à laquelle elle se rapporte et elle n'est jamais mise à jour : on n'écrase pas l'ancienne valeur lors d'une nouvelle insertion, mais on crée une nouvelle instance.

Les entrepôts étant, en général, très volumineux et très complexes à concevoir, ils ont été divisés en petits entrepôts, faciles à créer et à entretenir, appelés **data marts** ou magasins de données. La division peut se faire par fonction (ventes, les commandes, les ressources humaines) ou par sous-ensemble organisationnel (un data mart par succursale). Un magasin de données est un extrait de l'entrepôt des données. Les données extraites sont adaptées à une classe de décideurs ou à un usage particulier (recherche de corrélation, logiciel de statistiques) rendant très rapide les temps de réponses aux requêtes (Tranchant, 2012).

Pour gérer physiquement et sémantiquement l'ensemble des données, l'entrepôt de données doit nécessairement disposer de "données sur les données", à savoir de métadonnées.

Les métadonnées d'un entrepôt de données se présentent sous trois catégories (Ponniah, 2001) :

- **Métadonnées opérationnelles** : contiennent toutes les informations sur les sources de données opérationnelles.
- **Métadonnées d'extraction et de transformation** : les fréquences d'extraction, les méthodes d'extraction, règles d'extraction, informations sur toutes les transformations opérées sur les données.

- **Métadonnées de l'utilisateur final** : permet à l'utilisateur final de retrouver l'information dans l'entrepôt de données.

I.2.2.2. Catégories des données stockées dans un entrepôt de données

Un entrepôt de données peut se structurer en quatre classes de données, organisées selon un axe historique et un axe de synthèse. **Axe synthétique** qui établit une hiérarchie d'agrégation comprenant :

- **Les données détaillées** : elles sont issues des systèmes transactionnels et représentent les événements les plus récents. Seules les données susceptibles d'avoir un sens pour le processus d'aide à la décision doivent être stockées dans l'entrepôt.
- **Les données agrégées** : elles synthétisent les données détaillées et correspondent à des éléments d'analyse représentatifs des besoins des utilisateurs. Elles ont pour but de faciliter la navigation suivant les besoins décisionnels et la restitution d'un résultat d'analyse ou de synthèse.
- **Les données fortement agrégées** : elles synthétisent à un niveau supérieur les données agrégées.

Axe historique qui comprend les données détaillées historisées représentant les événements passés.

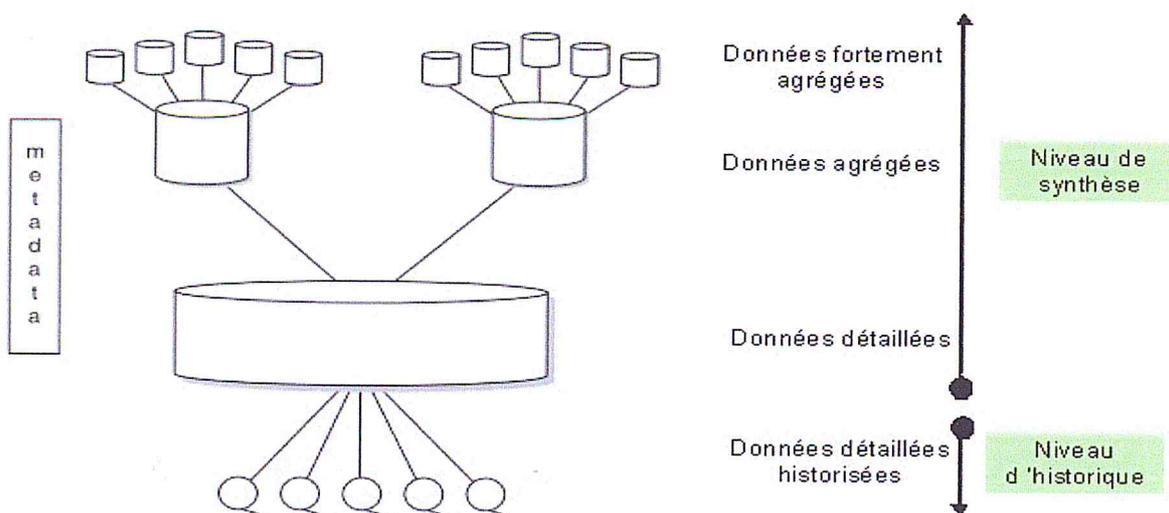


Figure 2 : Structure de données d'un entrepôt de données (Inmon, 2002)

I.2.2.3. Approches de mise en place d'un entrepôt de données

Plusieurs approches existent pour la mise en place d'un entrepôt de données. Par contre seulement trois approches sont répandues. Il s'agit de l'approche "Top-Down" (de haut en bas) prônée par **Inmon**, l'approche "Bottom-up" (de bas en haut) de **Kimball** et de l'approche "Hybride" qui dérivent des deux premières approches. (**Ponniah, 2001**).

- L'approche Top-Down

C'est une approche globale dans laquelle un grand entrepôt de données à l'échelle de toute l'entreprise est construit. L'entrepôt va alimenter des magasins de données locaux ou départementaux.

Avantages	Inconvénients
- Vision claire et globale des données de l'entreprise	- Longue et coûteuse à développer
- intégration et consolidation complète des données de l'entreprise	- Risque d'échec important
- Stockage unique et centralisé des données	- Résultats non immédiats
- Contrôle centralisé	
- Maintenance facile	

Tableau 1 : Avantages et inconvénients de l'approche Top-Down (Ponniah, 2001).

- L'approche Bottom-Up

L'approche Bottom-Up consiste à construire des magasins de données indépendants un par un, puis à les regrouper par des niveaux intermédiaires jusqu'à obtention d'un entrepôt.

Avantages	Inconvénients
<ul style="list-style-type: none"> - Implémentation rapide, facile et peu coûteuse. - Résultats rapides - Risque d'échec minime - Approche incrémentale (magasins importants d'abord) 	<ul style="list-style-type: none"> - Chaque magasin de donnée indépendant a sa propre vision limitée des données, - Incohérences et redondance des données entre les magasins de données, - Fragmentation des données, - Performance non optimale des analyses impliquant plusieurs magasins de données, - Volume de travail d'intégration important pour obtenir un entrepôt de données.

Tableau 2 : Avantages et inconvénients de l'approche Bottom-Up (Ponniiah, 2001)

- L'approche Hybride

Les deux approches précédentes ont toutes deux leurs avantages et leurs inconvénients, l'approche hybride tire le meilleur des deux précédentes sans avoir leurs contraintes.

Les étapes de cette approche sont :

- Planifier et définir les exigences à l'échelle de l'entreprise
- Etablir une architecture pour un entrepôt de données complet
- Concevoir un modèle de donnée et standardiser les données selon celui-ci
- Déterminer les données de chaque magasin pour découper l'entrepôt
- Implémenter l'entrepôt comme une série de magasins de données un à la fois.

I.2.3. Serveurs OLAP

Comme le montre la figure 1, le troisième pilier d'un système décisionnel est celui des serveurs OLAP.

On-Line Analytical Processing est une « *Activité globale de requêtage et de présentation de données textuelles et numériques contenues dans l'entrepôt de données ; Style d'interrogation spécifiquement dimensionnel* » (Kimball, 2005). Le terme OLAP fut introduit par E.F.Codd en 1993, le père des bases de données relationnelles, dans un document de présentation technique pour la compagnie Arbor Software, devenue aujourd'hui Hyperion, où il a défini douze règles que doivent satisfaire les modèles OLAP (Codd et al., 1993). Mais le terme OLAP ne donne ni une définition ni une description claire de sa signification. Cependant, en 1995, Nigel Pendse et Richard Creeth proposent un autre terme et récapitule la définition de l'OLAP en cinq mots : **Fast Analysis of Shared Multidimensional Information (FASMI)**, ce terme est traduit par *Analyse Rapide d'Information Multidimensionnelle Partagée* (Laurent, 2002).

- Analyse : le système doit fournir des outils d'analyse numérique et statistique,
- Rapide : le système doit offrir un temps de réponse rapide aux requêtes des utilisateurs,
- Information : le système doit assurer la disponibilité des données quelques soit leur volume et leur mode de stockage,
- Multidimensionnelle : le système doit fournir une vue multidimensionnelle des données. La notion de données multidimensionnelles sera expliquée en détails dans le chapitre suivant,
- Partagée : les données doivent être sécurisées en vue d'une architecture multi-utilisateurs.

Cette technologie OLAP a été proposée pour répondre aux limites que présentaient les systèmes transactionnels en matière d'analyse et de support pour la prise de décision.

❖ **OLTP Vs OLAP**

Les processus transactionnels en ligne OLTP, sont dédiés aux métiers de l'entreprise pour les assister dans leurs tâches de gestion quotidiennes et pour gérer les informations contenues dans leurs systèmes opérationnels. Ils permettent l'insertion, la modification et l'interrogation rapide des données, tout en maintenant leur intégrité dans un environnement multi-accès. Les données sont détaillées, à jour et habituellement normalisées (3FN). Seulement ces systèmes OLTP ne peuvent pas répondre aux besoins spécifiques des entreprises pour analyser l'information et supporter efficacement leurs processus d'aide à la décision.

Contrairement aux systèmes OLTP, les systèmes OLAP sont dédiés au management de l'entreprise pour l'aider au pilotage de ses activités. Ils offrent au décideur une vision transversale de l'entreprise. Les systèmes OLAP sont caractérisés par un volume relativement important de données mais petit en transaction, et les objectifs principaux sont : regrouper, organiser les informations provenant des systèmes transactionnels, les intégrer et les stocker pour offrir une vue orientée métier à l'utilisateur, retrouver et analyser l'information facilement et rapidement (Metais, 2007).

Le tableau ci-dessous récapitule la différence entre OLTP et OLAP.

Caractéristiques	Processus OLTP	Processus OLAP
Objectifs	Interrogation, Mise à jour	Analyse, Décision
Nature des données	Individuelles	Multidimensionnelles, agrégées, orientées utilisateur
Fraîcheur des données	Récents, dynamique	Historiques, statiques
Traitement	Simple	Complexe, semi-automatique
Utilisateurs	Tout type	Décideurs
Temps de réponse	Rapide	Moins rapide

Tableau 3 : Comparaison des processus OLTP et OLAP (Naoum, 2006)

I.2.4. Outils d'analyse et de restitution

Le dernier niveau correspond à la partie visuelle du système décisionnel. C'est l'élément le plus important pour l'utilisateur, car il lui permet d'exploiter, d'analyser et de restituer les données stockées. On distingue à ce niveau plusieurs types d'outils différents :

- **Les outils de reporting** : permettent d'assurer la réalisation, la publication et la diffusion de rapports d'activités périodiques selon un format prédéterminé et paramétrable par les utilisateurs. Les outils de reporting permettent aussi de produire *des tableaux de bord* à l'aide des indicateurs les mieux adaptés à évaluer la performance de l'entreprise (Fernandez, 2012).
- **Les outils d'analyse OLAP** : permettent de traiter les données et les afficher sous forme de cubes multidimensionnels et de naviguer sur ces dimensions en utilisant des opérations de navigation tel qu'expliqué dans le Chapitre 2.
- **Les outils de datamining** : permettent l'exploration et l'analyse de données, de façon automatique, en vue de mettre en évidence des tendances ou des corrélations cachées parmi des masses de données, ou encore pour détecter des informations stratégiques ou découvrir de nouvelles connaissances (Naoum, 2006). De manière générale, c'est extraire, à partir d'une quantité volumineuse de données, de la connaissance, auparavant inconnues, potentiellement utiles, en s'appuyant sur des méthodes de traitement statistiques comme les arbres de décision, ou sur l'intelligence artificielle avec entre autres les réseaux de neurones (Kuonen, 2001).

Conclusion

Dans ce chapitre nous avons introduit le système décisionnel et nous avons décrit les quatre modules autour desquels il s'articule, en expliquant leurs rôles respectifs et l'interaction existant entre eux. Il s'agit d'alimenter, grâce aux outils d'ETL, un entrepôt de données à partir de données sources, ces données deviennent alors exploitables à l'aide d'outils adaptés. Et pour supporter d'une manière efficace le processus décisionnel, les données de l'entrepôt doivent être organisées d'une manière multidimensionnelle. Le chapitre suivant est consacré à l'étude de la modélisation multidimensionnelle.

Chapitre II

La modélisation multidimensionnelle

Introduction

Contrairement aux bases de données relationnelles, l'intérêt des bases de données décisionnelles ne se situe pas au niveau de l'individu mais plutôt au niveau de l'identification des tendances dans un ensemble ou un groupe.

Les bases de données relationnelles, modélisées selon les principes classiques de normalisation, s'adaptent très mal à un contexte analytique (OLAP). En analyse, l'utilisateur doit disposer d'un modèle relativement intuitif et capable de stocker le résultat de nombreux calculs d'agrégation. De plus, les données sont orientées sujet pour refléter la vision des décideurs selon plusieurs axes d'analyse.

Par conséquent les données doivent correspondre à une structure adaptée à l'aspect multidimensionnel. En effet, la modélisation dimensionnelle est la plus adaptée pour répondre aux exigences de rapidité et de facilité d'analyse. Elle permet, en outre, de rendre les données d'un entrepôt cohérentes, lisibles, intelligibles et faciles d'accès.

II.1. Définition

La modélisation multidimensionnelle consiste à considérer les données comme des points projetés dans un espace à plusieurs dimensions, qui représentent les différents axes d'analyse. Ces données sont organisées de manière à mettre en évidence le sujet analysé et les différentes perspectives de l'analyse. (Gray et al., 1996).

❖ La métaphore du cube

La vision multidimensionnelle des données peut être vue comme un *cube de données*. Chaque donnée représente une cellule du cube et les arêtes représentent les axes d'analyse des données. Sur chaque arête, une graduation est choisie afin d'observer les données à un niveau adéquat de détail.

Exemple : la figure 3 représente un exemple de cube de données qui permet l'analyse des publications scientifiques.

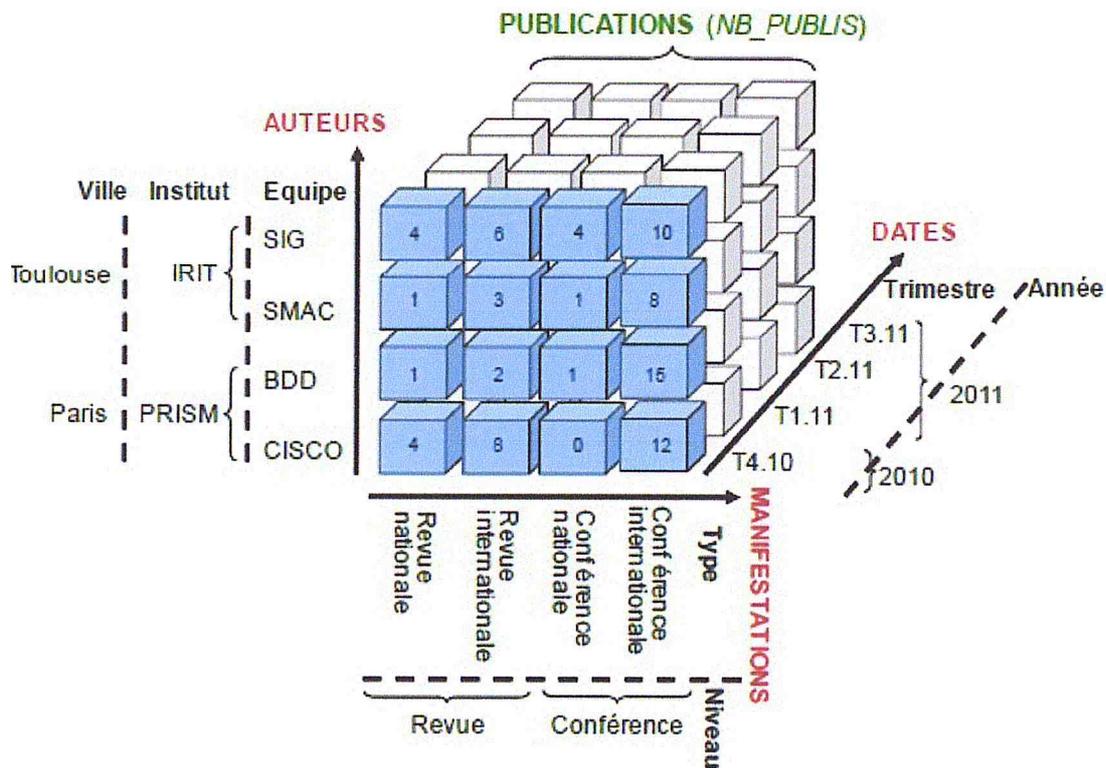


Figure 3 : Exemple d'un cube de données (Teste, 2009)

II.2. Modèles de présentation des données

La modélisation des données au sein d'un système OLAP peut être représentée conformément aux niveaux d'abstraction classiques définis dans les méthodes de conception des bases de données (Teste, 2009) :

- Modèle conceptuel : c'est la description de la base de données indépendamment des choix technologiques.
- Modèle logique : c'est la définition de la base de données en utilisant une technologie informatique (objet, relationnel,...).

II.2.1. Modèle conceptuel

La modélisation multidimensionnelle a donné naissance à différents concepts, qui reposent sur la métaphore des cubes (Gray et al., 1996).

II.2.1.1. Concept de fait

Un fait modélise le sujet d'analyse et regroupe un ensemble d'indicateurs appelés *mesures*.

Exemple : soit un fait de vente constitué des mesures suivantes : quantité des produits vendus et montant des ventes. Le fait est représenté par un rectangle englobant les différentes mesures qu'il contient. En outre le symbole d'un cube estampille le fait.

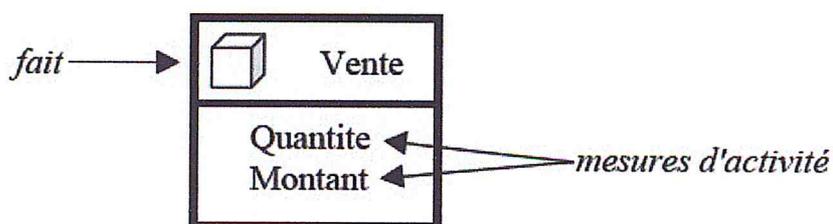


Figure 4 : Exemple de fait (Teste, 2000)

II.2.1.2. Concept de dimension

Une dimension modélise une perspective d'analyse, composée de *paramètres* qui peuvent faire varier les mesures.

Exemple : le fait de l'exemple précédent peut être analysé suivant trois dimensions qui représentent les perspectives d'analyse : le temps, la catégorie et la situation géographique.

Une dimension est représentée par un rectangle englobant les différents paramètres qu'elle contient. En outre un symbole représentant trois axes estampille les dimensions.

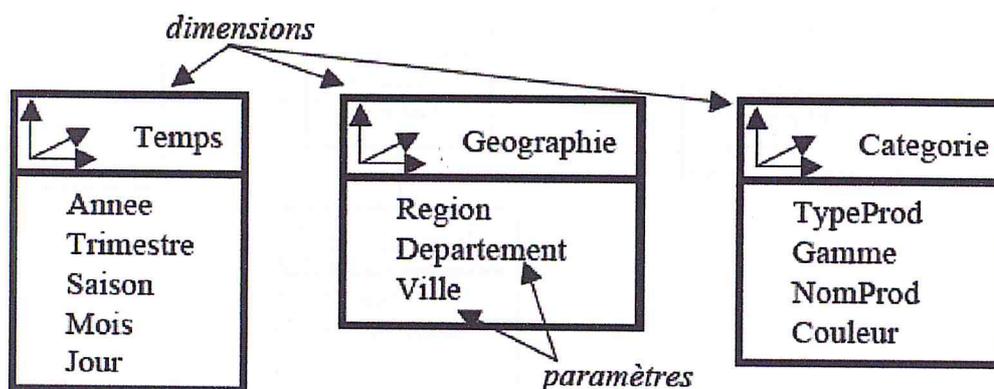


Figure 5 : Exemple de dimensions (Teste, 2000)

Chaque dimension est munie d'une ou plusieurs hiérarchies qui permettent d'organiser les paramètres conformément à leur niveau de détail.

Exemple : la dimension Géographie est organisée suivant la hiérarchie suivante :

Ville \longrightarrow Département \longrightarrow Région.

Ainsi chaque ville appartient à un département qui est situé dans une région.

L'association de ces concepts compose un modèle. Il existe trois types de modèles :

II.2.1.3 Modèle en étoile (Star schema)

Constitué d'un fait entouré de plusieurs dimensions formant ainsi une sorte d'étoile. Les dimensions sont représentées de manière dénormalisée ce qui permet d'avoir un temps de réponse rapide aux requêtes d'analyse, mais une redondance des données.

Exemple : la figure 6 correspond à un schéma en étoile modélisant une analyse de vente (quantité et montant) des médicaments dans une pharmacie selon trois dimensions : le temps, la catégorie et la situation géographique.

II.2.1.5. Modèle en constellation

Consiste à interconnecter plusieurs schémas en étoile ayant des dimensions communes.

Exemple : la figure 8 illustre une modélisation en constellation constituée de deux schémas en étoile : l'un correspond aux ventes des pharmacies et l'autre aux prescriptions des médecins.

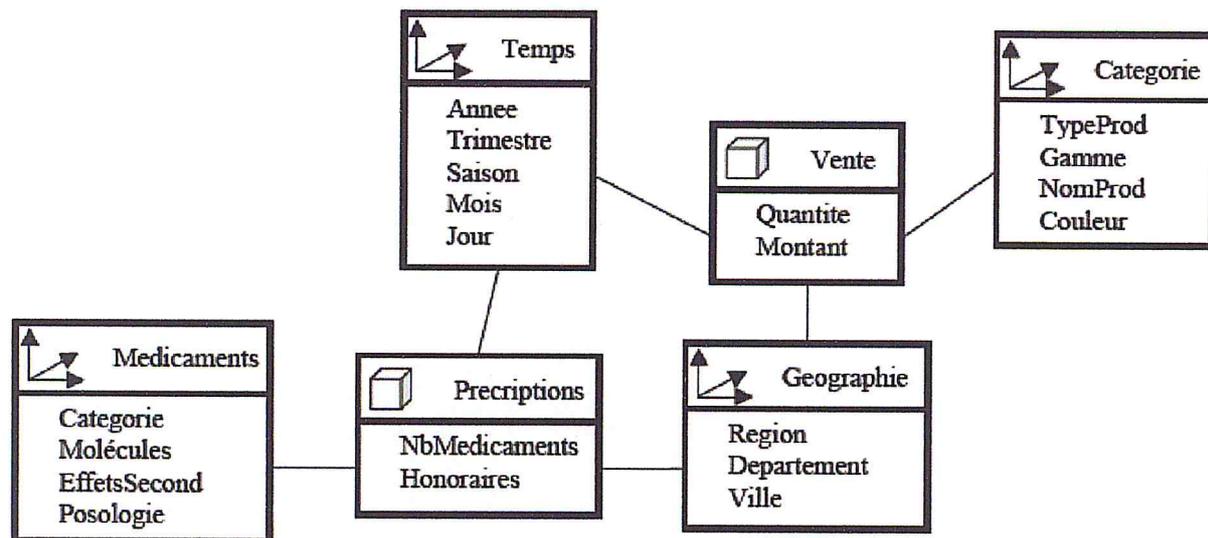


Figure 8 : Exemple de modélisation en constellation (Teste, 2000)

II.2.2. Modèle logique

La modélisation multidimensionnelle logique est possible selon un modèle R-OLAP ou un modèle M-OLAP, qui diffèrent essentiellement dans la manière du stockage physique des données. D'autres modèles existent comme H-OLAP et D-OLAP mais leurs fonctionnements découlent des modèles précédents.

II.2.2.1. R-OLAP

Relational OnLine Analytical Processing consiste à stocker physiquement les données multidimensionnelles dans un environnement relationnel (Système de Gestion de Base de Données Relationnelle SGBDR). Dans l'approche R-OLAP, le modèle dimensionnel est traduit par des tables relationnelles, ainsi chaque fait représente une table de fait et chaque dimension une table de dimension (Naoum, 2006), et les opérations de navigation OLAP (roll-up, drill-down, pivot,...) sont transformées en requête SQL classique (GRIM, 2008). Cette approche procure la capacité de gérer des volumes très importants de données et utilise

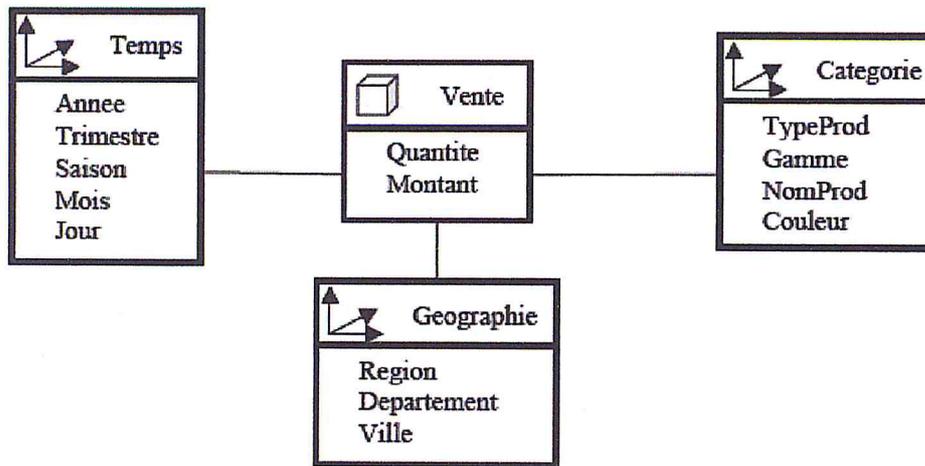


Figure 6 : Exemple d'une modélisation en étoile (Teste, 2000)

II.2.1.4. Modèle en flocon (snowflake)

Consiste à conserver le fait du schéma en étoile et d'éclater les dimensions conformément à la hiérarchie des paramètres. L'intérêt principal du modèle en flocon réside dans le gain en espace de stockage, car il permet d'éviter le problème de redondance.

Exemple : la figure 7 correspond à un modèle en flocon ; le fait de la figure 6 est conservé, et les dimensions sont éclatées en formant ainsi une sorte de flocon.

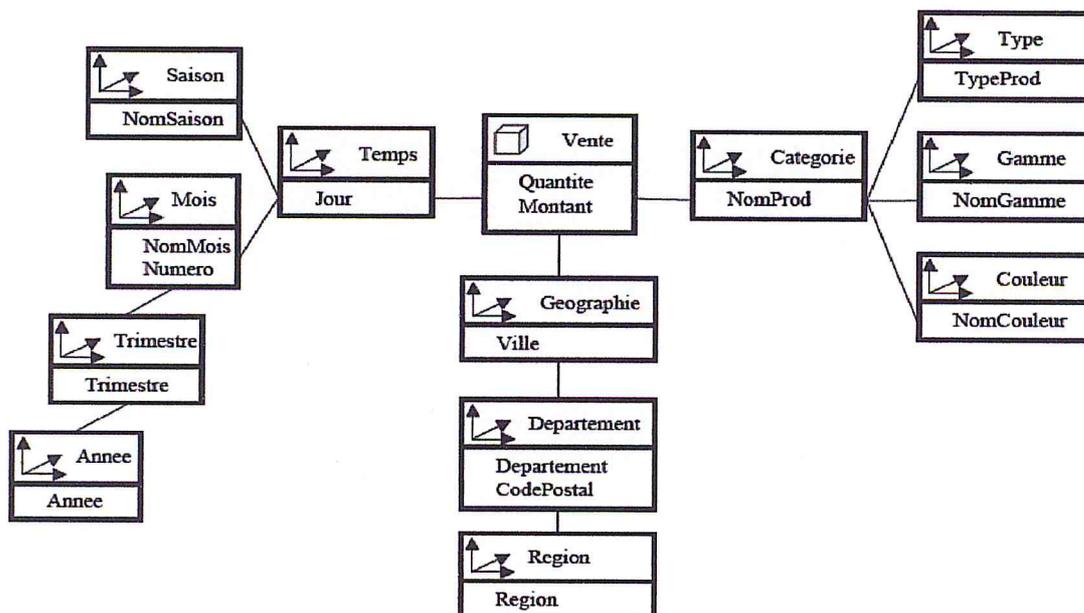


Figure 7 : Exemple d'une modélisation en flocon (Teste, 2000)

les mécanismes de gestion des données approuvés depuis des décennies, mais le traitement est souvent lent et peu performant (Jerbi, 2012).

II.2.2.2. M-OLAP

Multidimensional On Line Analytical Processing consiste à stocker les données sous une structure multidimensionnelle : il s'agit d'une application physique du concept du cube. La structure multidimensionnelle utilisée est un tableau à n dimensions. Nous trouvons également l'appellation d'*hyper-cube* ou de *cube* pour désigner cette structure (NAOUM, 2006). L'intérêt de cette approche réside dans l'optimisation du temps d'accès. Mais, cette approche reste limitée dans la capacité de gérer d'importants volumes de données, et nécessite de redéfinir tous les mécanismes des SGBD pour manipuler les structures multidimensionnelles (Jerbi, 2012).

II.2.2.3. H-OLAP

Hybride On Line Analytical Processing consiste à faire cohabiter l'approche R-OLAP pour les données détaillées et l'approche M-OLAP pour les données agrégées au sein du même système. L'objectif est donc de bénéficier des avantages des deux approches tout en minimisant leurs faiblesses.

II.2.2.4. D-OLAP

Desk On Line Analytical Processing est un mode de fonctionnement consistant à récupérer en local une partie d'une base de données multidimensionnelle. Cette utilisation est intéressante pour continuer à faire des analyses sur des données de manière nomade et déconnectée.

Exemple : la figure 9 présente une traduction d'un modèle conceptuel (schéma étoile) en modèle logique (R-OLAP).

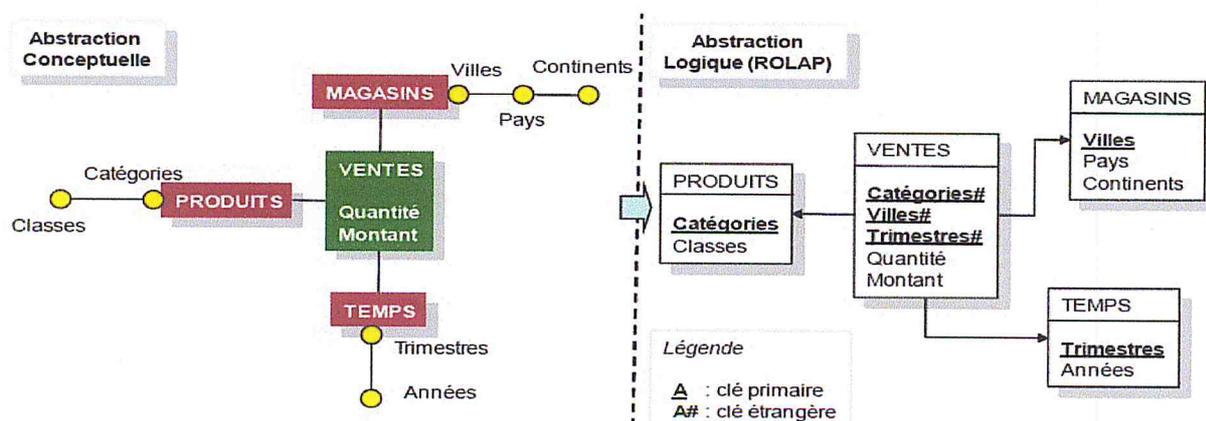


Figure 9 : Traduction R-OLAP d'un schéma étoile

II.3. Manipulation multidimensionnelle des données

Il existe de nombreuses propositions concernant la définition d'opérations de manipulation OLAP dans laquelle les structures manipulées sont les cubes de données et les tranches du cube de données. Cependant, il n'existe pas de consensus sur la définition d'un ensemble minimum d'opérateurs assurant l'intégralité des opérations de manipulation OLAP, mais la plupart des propositions offrent un support partiel des différentes catégories d'opérations suivantes :

II.3.1. Opérations classiques

- **Opérations de sélection.** Ces opérations permettent à un utilisateur de restreindre l'ensemble des données analysées. (Tournier, 2007)
 - **La spécification d'une « tranche de cube » (slice)** consiste à exprimer une restriction sur une des données de l'un des axes d'analyse. Cela correspond à une projection selon une dimension du cube.
 - **La spécification d'un « sous-cube » (dice)** consiste à exprimer une restriction sur les données d'un indicateur d'analyse.

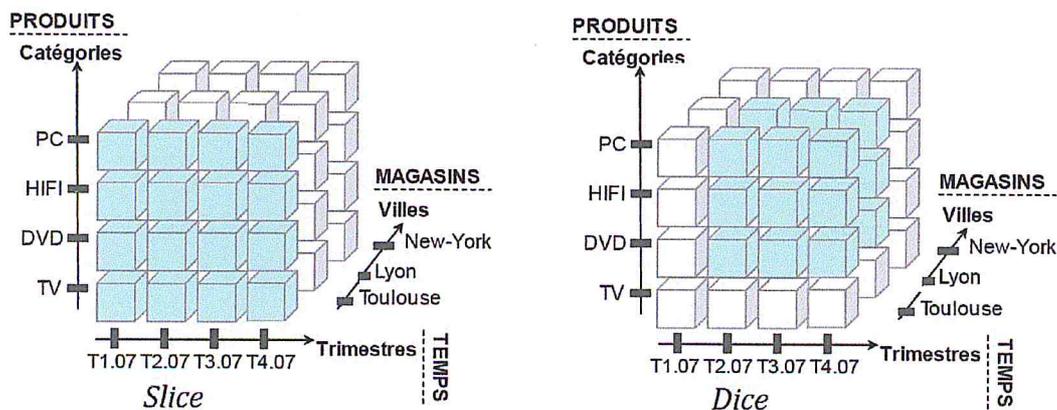


Figure 10 : Principe de "slice & dice"

Des auteurs ont aussi proposé des opérations additionnelles :

- **Opérateurs binaires.** Certains auteurs proposent aussi l'emploi des opérations binaires ensemblistes : union, différence et intersection. Certains travaux ont aussi proposé la notion de jointure (join) inspirée de la jointure relationnelle, mais d'un intérêt limité dans un environnement multidimensionnel. Il est intéressant de noter que les opérateurs binaires sont des opérateurs nécessitant de très fortes contraintes. Par

- exemple, l'union entre deux structures multidimensionnelles nécessite une compatibilité presque complète des deux structures.

II.3.2. Opérations agissant sur la granularité

- **Opérations de forage.** Ces opérations permettent la navigation au moyen de la structure hiérarchique des axes d'analyses, afin de permettre l'analyse d'un indicateur avec plus ou moins de précision. (Jerbi, 2012)
 - **Le forage vers le haut (roll-up)** consiste à analyser les données en fonction d'un niveau de granularité supérieur conformément à la hiérarchie définie sur la dimension.
 - **Le forage vers le bas (drill-down)** à l'inverse du roll-up permet d'analyser les données avec un niveau de granularité inférieur donc sous une forme plus détaillée.

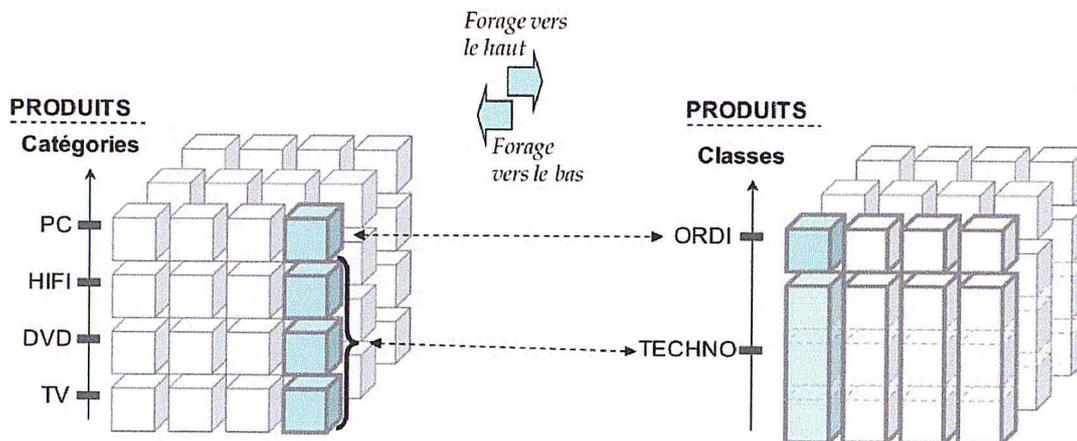


Figure 11 : Principe du forage (Teste, 2009)

II.3.3. Opérations agissant sur la structure

Opérations de rotation. Ces opérations permettent la réorientation d'une analyse. Elles permettent de changer l'un des axes d'analyse en cours d'utilisation (rotation de dimension), de changer le sujet de l'analyse (rotation de fait ou drill-across) et de changer de perspective d'analyse (rotation de hiérarchie) (Laurent, 2002).

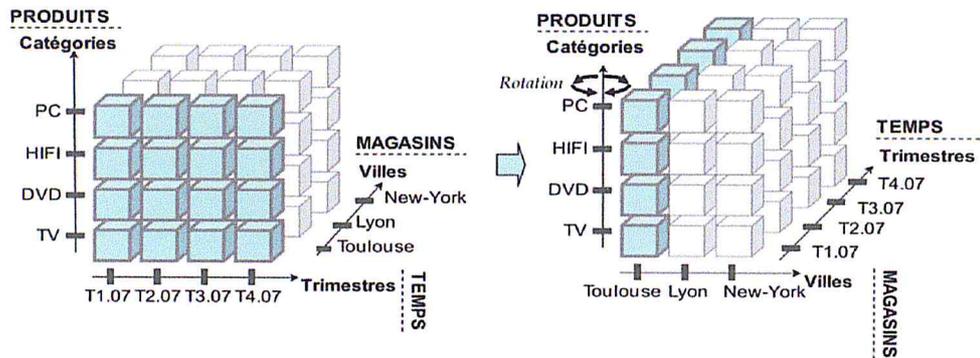


Figure 12 : Principe de rotation (Teste, 2009)

- La **permutation (switch)** consiste à inverser deux dimensions, de manière à permuter deux tranches du cube
- La **division (split)** consiste à présenter chaque tranche du cube et de passer de sa présentation tridimensionnelle à sa présentation sous la forme d'un ensemble de tables. Plus généralement, cette opération permet de réduire le nombre de dimensions.
- L'**emboîtement (nest)** permet d'imbriquer les valeurs d'un paramètre d'une dimension avec un autre paramètre. L'intérêt de cette opération est qu'elle permet de représenter de manière bidimensionnelle toutes les informations du cube quel que soit le nombre de dimensions.
- L'**enfocement (push)** consiste à combiner les valeurs d'un paramètre d'une dimension aux mesures du fait et donc de transformer un paramètre en mesure.
- L'opération inverse de **retrait (pull)** permet de transformer une mesure en paramètre en changeant le statut de certaines mesures du cube pour constituer une nouvelle dimension.
- La **factualisation (fold)** consiste à transformer une dimension en mesure(s); cette opération permet de transformer en mesure l'ensemble des paramètres d'une dimension.
- La **paramétrisation (unfold)** permet de transformer une mesure en paramètre dans une nouvelle dimension.
- L'opération **Cube** permet de calculer des sous-totaux et un total final dans le cube.

Conclusion

Dans ce chapitre nous avons introduit la modélisation multidimensionnelle des données qui repose sur les concepts de fait et de dimension. L'association de fait et de dimension compose un schéma qui peut être soit en étoile, en flocon ou en constellation. L'implémentation de ces schémas se fait principalement selon des modèles R-OLAP et M-OLAP, qui diffèrent essentiellement dans la manière de stockage. Plusieurs opérations permettent la navigation et la manipulation multidimensionnelle des données stockées.

Le modèle multidimensionnel est la structure de données la plus utilisée et la plus appropriée aux requêtes et analyses des utilisateurs d'entrepôts de données.

PARTIE 2 : ETUDE DE L'EXISTANT ET COLLECTE DES BESIONS

Cette deuxième partie portera sur la présentation de l'organisme d'accueil et l'existant en matière de décisionnel afin d'aborder les besoins des utilisateurs dans l'intention de proposer une solution.

Chapitre III

Présentation de l'organisme d'accueil

Introduction

Le marché algérien a récemment assisté à la naissance d'un grand nombre de petites et grandes entreprises, qui diffèrent entre elles en termes d'activité et de volume de travail, et qui veulent imposer leurs présences à la lumière des entreprises nationales dominante sur l'économie algérienne depuis l'indépendance. Parmi ces dernières l'entreprise au sein de laquelle s'est déroulé notre stage « la Sonatrach ». Dans ce chapitre nous allons s'intéresser à son historique, son organisation et ses activités dans sa globalité puis nous détaillerons l'activité « Amont », la direction « Contrôle de Gestion » qui sera le client de ce projet, ainsi que la direction d'accueil « ISI » qui encadrera le projet.

III.1. Présentation de Sonatrach

III.1.1. Sonatrach en bref...

SONATRACH est la plus importante compagnie d'hydrocarbures en Algérie et en Afrique. C'est une entreprise publique économique à caractère industriel et commercial dénommée Société Nationale de Transport, Transformation et Commercialisation des Hydrocarbures. Elle a été créée par le décret présidentiel N°63-49, du 31 décembre 1963.

Adoptant une stratégie de diversification, SONATRACH se développe dans les activités de génération électrique, d'énergies nouvelles et renouvelables, de dessalement de l'eau de mer, de recherche et d'exploitation minière. Poursuivant sa stratégie d'internationalisation, SONATRACH opère en Algérie et contractante avec plusieurs régions du monde : en Afrique (Mali, Niger, Libye, Egypte), en Europe (Espagne, Italie, Portugal, Grande Bretagne), en Amérique Latine (Pérou) et aux USA.

Avec un chiffre d'affaires à l'exportation de près de 56,1 milliards de dollars réalisé en 2010, SONATRACH est classée 1ère compagnie en Afrique et 12ème compagnie dans le monde. Elle est également 4ème exportateur mondial de GNL, 3ème exportateur mondial de GPL, et 5ème exportateur de Gaz Naturel [Web 9].

III.1.2. Les missions de SONATRACH

SONATRACH se présente actuellement sous l'aspect d'une entreprise intégrée, intervenant directement dans l'ensemble des activités du secteur des hydrocarbures.

Elle conserve la charge des opérations liées à la recherche, la production, le transport par canalisation des hydrocarbures, le conditionnement des hydrocarbures gazeux ainsi que leur commercialisation. Elle emploie plus de 120000 personnes.

Les principales missions de SONATRACH sont :

- ✓ La recherche et l'exploitation des gisements.
- ✓ Le traitement et la transformation des hydrocarbures liquides et gazeux.
- ✓ Le transport des hydrocarbures.
- ✓ L'approvisionnement énergétique à moyen et long terme.

- ✓ La commercialisation des hydrocarbures ainsi que leurs substances connues tant en Algérie qu'à l'étranger.
- ✓ L'étude et le développement de toute nouvelle forme d'énergie.
- ✓ La participation dans le créatif
- ✓ on d'entreprises nouvelles en la matière au niveau national ou l'étranger.
- ✓ L'acquisition et la maîtrise des nouvelles technologies dans le domaine ... etc.

III.1.3. Organisation de SONATRACH

Avec le parachèvement de la macrostructure du groupe, SONATRACH dispose désormais d'une organisation simplifiée, cohérente et rationalisée.

Le schéma de la macrostructure s'articule, en fait, autour :

- ❖ **De la direction générale** : elle est assurée par le Président Directeur Générale.
- ❖ **Des activités opérationnelles** : exercent les métiers du groupe et développent son potentiel d'affaire tant en Algérie qu'en international, ces activités liées chacune à l'autre sous la forme d'une chaîne, à la tête de la chaîne on trouve l'Activité en **Amont** qui sera bien détaillée par la suite, suivie par :
 - **Activité TRC** (Transport par Canalisation) : cette activité assure l'acheminement des hydrocarbures (pétrole brut, gaz naturel, GPL et condensat) et dispose d'un réseau de canalisations de près de 16 200 Km.
 - **Activité Aval** : qui prend en charge le développement et l'exploitation de la liquéfaction de gaz naturel, la séparation de GPL et deraffinage.
 - **Activité Commercialisation** : consiste à recenser toutes les exportations en volumes (Tep), en valeurs (US\$), et les importations des hydrocarbures ainsi que les quantités vendues sur le marché national.
- ❖ **Des directions fonctionnelles** (directions centrales, directions de coordination groupe) : élaborent et veillent à l'application des politiques et stratégies du groupe.

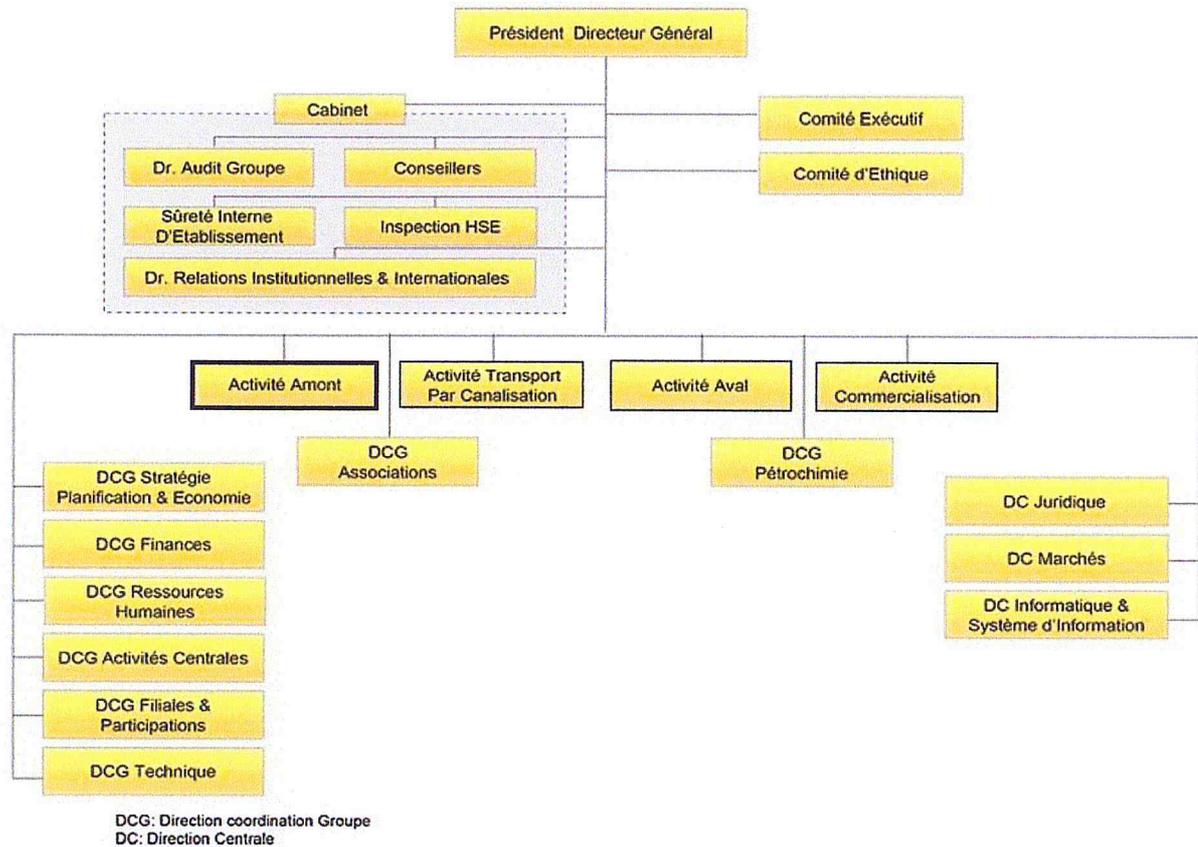


Figure 13 : Schéma de la macrostructure de la SONATRACH [Web 9].

III.2. Contexte d'étude

L'étude réalisée au cours de ce mémoire s'est déroulée dans le contexte général d'un suivi des coûts de départ champ (coût de production, amortissement de développement), et de production vendue.

D'une manière plus ciblée, ce travail est consacré à l'activité Amont de la Sonatrach, dont on a mis au point seulement la production des hydrocarbures de Sonatrach en association en excluant l'exploration.

III.2.1. L'activité Amont

L'activité Amont est l'activité la plus importante, et comme son nom l'indique, c'est le point de départ dans les domaines de la recherche, de l'exploration, de développement et de production d'hydrocarbures. Celles-ci sont assurées par Sonatrach seule ou en association avec d'autres compagnies pétrolières.

L'Activité Amont est organisée selon les structures suivantes :

- ❖ **Structures opérationnelles** : concernent toutes les divisions de l'Activité Amont, à savoir : Exploration, Forage, Production, PED, Laboratoires et ENC.
- ❖ **Structures fonctionnelles** : Ces structures constituent les différentes directions dont relève la « **direction des finances** », objet de notre prospection.

L'organigramme qui suit illustre globalement cette distribution :

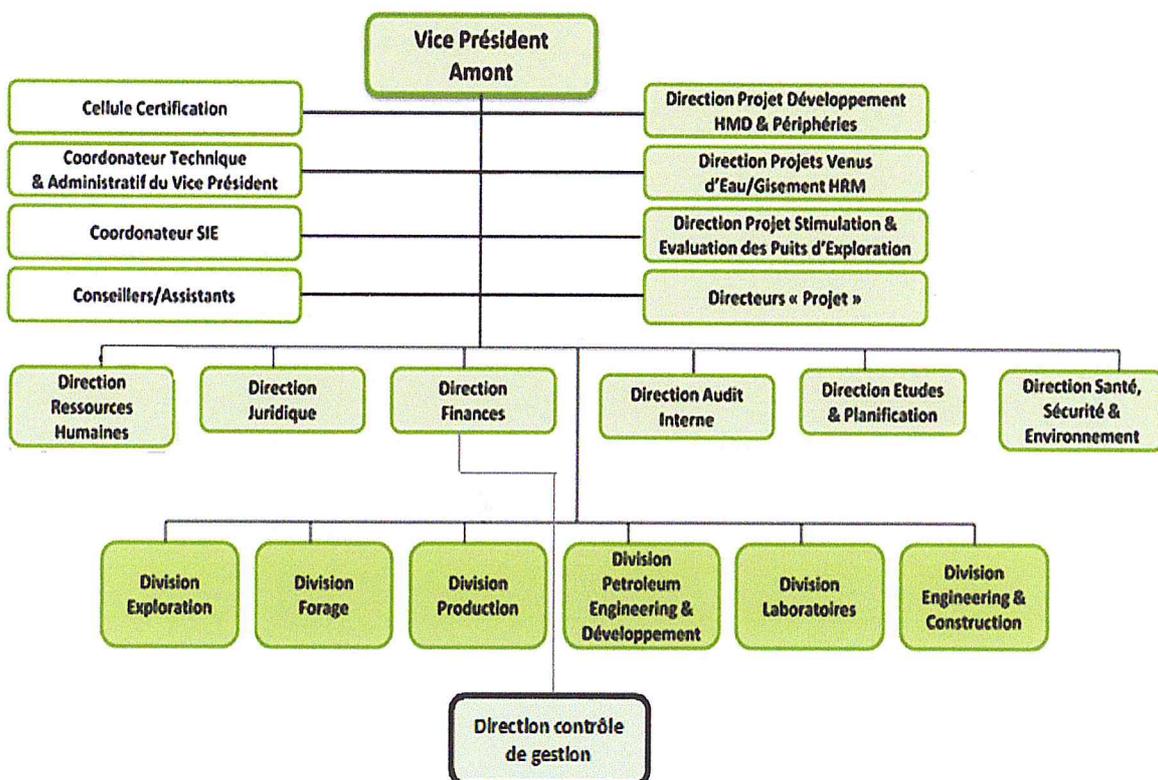


Figure 14 : Organigramme de l'Activité AMONT [web9].

L'Activité « AMONT » a pour missions essentielles

- ✓ L'élaboration et l'application des politiques et stratégies de développement et d'exploitation de l'amont pétrolier et gazier en effort propre.
- ✓ La conduite et le développement des activités de prospection et de recherche des hydrocarbures.
- ✓ Le développement et l'exploitation des gisements pour la valorisation optimale des ressources d'hydrocarbures.

- ✓ Les études et la réalisation de projets de développement, en coordination avec la « Direction Coordination Groupe Engineering et Développement » pour les nouveaux projets, ainsi que les études et la réalisation de projets de réhabilitation des installations de traitement des hydrocarbures.
- ✓ La recherche, la négociation et le développement de nouveaux projets en interne et la contribution au développement à l'international des métiers relevant de son domaine.

III.2.2. La direction Contrôle de Gestion au sein de la SONATRACH

III.2.2.1. la dépendance hiérarchique

Le contrôle de gestion au sein du groupe est du ressort de la direction contrôle de gestion de la direction coordination groupe Finance (DCG Finance).

La direction Finance qui a pour mission essentielle :

L'élaboration et la mise en œuvre, des politiques et stratégie du contrôle de Groupe Sonatrach en matière de :

- ✓ Financement d'ingénierie financière ;
- ✓ Gestion de la trésorerie ;
- ✓ Gestion de la dette ;
- ✓ Assurances et patrimoine ;
- ✓ Centralisation et consolidation comptable ;
- ✓ Gestion de la fiscalité ;
- ✓ Contrôle de gestion opérationnelle, gestion prévisionnelle et budget de fonctionnement ;
- ✓ L'exercice de la fonction de contrôle de gestion en matière d'exploitation et de financement ;
- ✓ La contribution au reporting général du Groupe dans les domaines concernés.

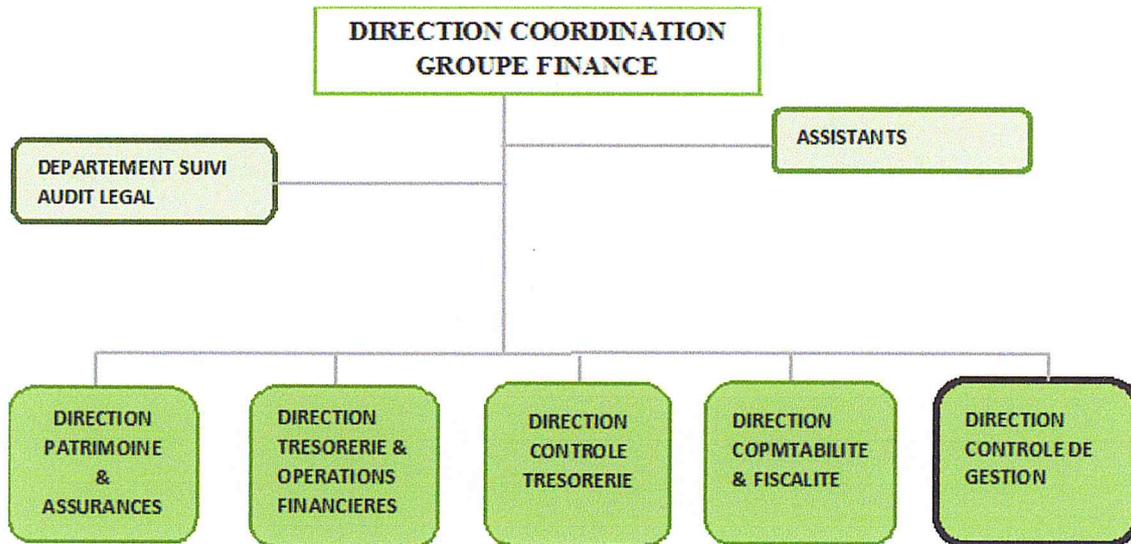


Figure 15 : Rattachement et hiérarchique du contrôle de gestion [DCG/Finances, 2012].

III.2.2.2. Les objectifs de la direction Contrôle de Gestion

Les principaux objectifs de la direction Contrôle de Gestion sont les suivants :

- ✓ L'organisation des activités permettant à l'entreprise d'évoluer et d'apprécier le chemin parcouru : mesure des réalisations par rapport aux objectifs assignés ;
- ✓ La définition de la politique de procédure de la gestion budgétaire ;
- ✓ L'initiation et le suivi de la mise en œuvre de la politique et des procédures de gestion budgétaire ;
- ✓ L'orientation et la coordination des travaux d'élaboration du budget d'exploitation ;
- ✓ Vérification, l'analyser et l'évaluation des performances par rapport aux objectifs assignés.
- ✓ L'élaboration du reporting des données financières et économique consolidées.

III.2.2.3. Organisation de la direction Contrôle de Gestion

La direction de contrôle de gestion est structurée en 6 départements comme suit :

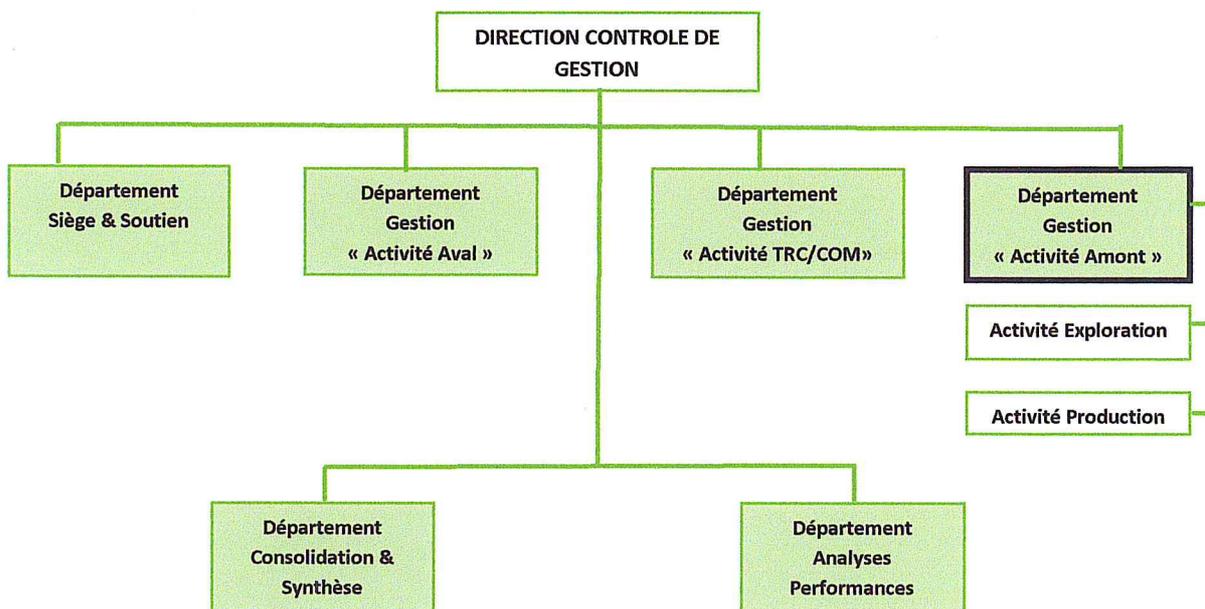


Figure 16 : Organigramme de la Direction Contrôle de gestion [DCG/Finances, 2012].

Les missions attribuées à la direction Contrôle de Gestion qui sont en relation avec notre objet d'étude sont organisées comme suit :

❖ Département Gestion « Activité Amont » :

Le département Gestion «Activité Amont» est composé de deux principales activités :

• L'activité d'exploration

Cette activité a pour missions essentielles :

- ✓ La mise en œuvre de la stratégie de la société en matière d'exploration ;
- ✓ La préparation, l'établissement et la recommandation des programmes techniques d'exploration et leur suivi ;
- ✓ La constitution d'un fichier sur les données relatives aux périmètres opérés sur le domaine minier national ;
- ✓ La gestion et le suivi des contrats en effort propre ;
- ✓ Le développement d'expertise dans le domaine de l'exploration.

- **L'activité de production**

Cette activité a pour mission essentielles :

- ✓ La détermination des besoins financiers de la division de production ;
- ✓ La consolidation, le contrôle et l'analyse des budgets et bilans globaux ;
- ✓ La consolidation et l'actualisation des plans annuels et pluriannuels d'investissement et d'exploitation de la division.
- ✓ Le suivi de la production des hydrocarbures et leurs charges.

III.3. Présentation de la direction d'accueil :

Nous effectuons notre stage au niveau de la direction Centrale Informatique et Système D'Information (ISI), au sien de la direction système d'information.

La direction système d'information s'organise en quatre départements comme suit :

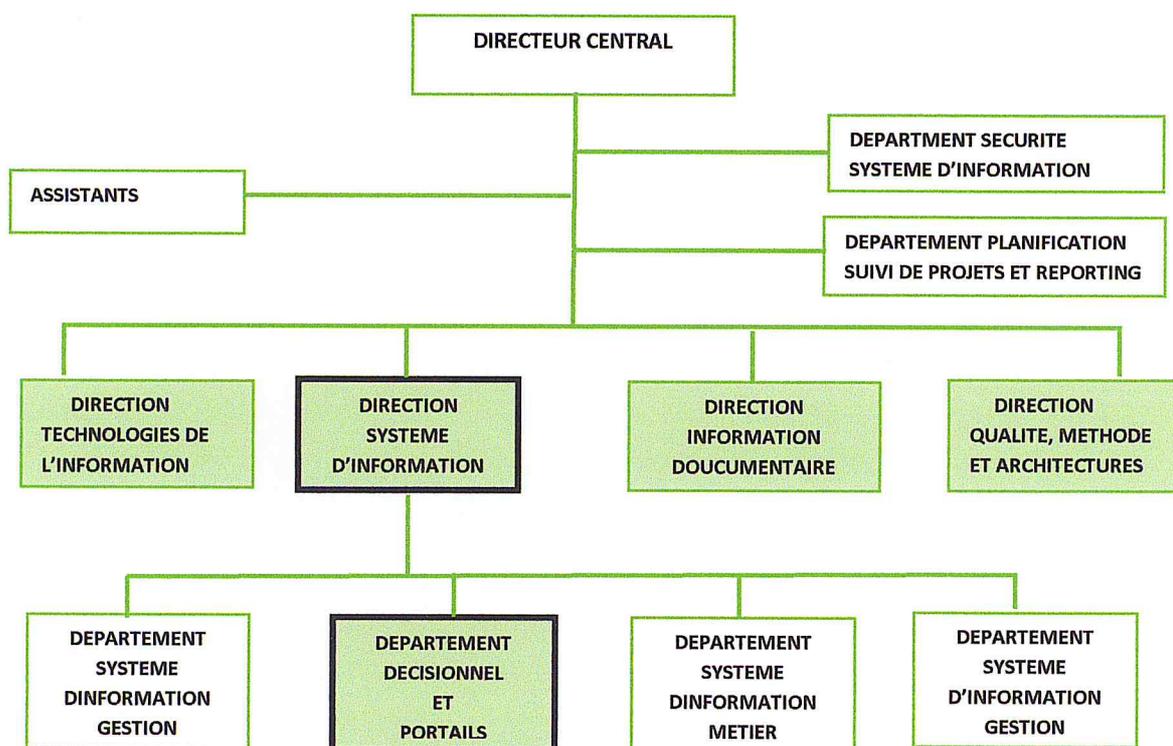


Figure 17 : Organigramme de la Direction ISI [DCG/ISI, 2012].

Conclusion

A travers ce chapitre nous avons vu une présentation du groupe Sonatrach, ses activités dont celle de l'Amont, les objectifs qui ont menés aux différentes structures et aboutis a l'organisation actuelle. Cette présentation nous a également fait comprendre la structuration et l'organisation de la fonction contrôle de gestion qui est la fonction cible du présent projet. Cette compréhension est nécessaire pour bien appréhender l'étude de l'existant objet de prochain chapitre.

Chapitre IV

Etude de l'existant

Et

Etude des besoins

Introduction

Afin de répondre mieux aux exigences des décideurs, une bonne compréhension de l'existant est primordiale. Il est en effet nécessaire de faire un état des lieux en matière de décisionnel et comprendre le système existant (fonctionnalités offertes, reporting, limites et données manipulées), car cela constituera le point de départ du système à mettre en place. De plus cela nous permettra de cerner les ressources (les sources de données par exemple) qui nous seront utiles parmi celles disponibles.

IV.1. Etat du décisionnel

Le groupe Sonatrach ne dispose pas d'un système décisionnel pour la direction contrôle de gestion. Et pour remédier à ce manque, la direction a confié à des gestionnaires la responsabilité de consolider, de façon manuelle (tableaux croisés de Excel), les rapports périodiques, et procéder par la suite à la diffusion des résultats aux acteurs internes (PDG-SH, VP-AMT, DCG-SPE) de Sonatrach. En effet au niveau de la direction contrôle de gestion, il existe plusieurs gestionnaires liés chacun à un domaine d'activité :

➤ Les Gestionnaires Activités Amont, Aval et TRC/COM

Ont pour missions, chacun dans son portefeuille d'activités :

- ✓ L'animation et le suivi de l'élaboration des projets de budget des structures appartenant au portefeuille d'activités ;
- ✓ La réalisation des analyses de cohérence des projets de budget ;
- ✓ La participation aux travaux de consolidation du budget de la société initiés et coordonnés par le département consolidation et synthèse ;
- ✓ L'organisation et la réalisation des opérations liées au contrôle budgétaire ;
- ✓ L'établissement des décisions d'exécution du budget d'exploitation ;
- ✓ La réalisation des analyses des coûts d'exploitation ;
- ✓ La constitution d'une banque de données ;
- ✓ L'analyse et le suivi de la cohérence des flux physique et financiers ;
- ✓ La contribution à la réalisation du reporting de la Société (tableaux de bord, situations économiques synthétiques,...) ;
- ✓ La réalisation des analyses de performances économiques et financières des entités appartenant au portefeuille d'activité.

➤ **Les rapports générés à partir de département gestion « Activité Amont » :**

Durant la phase d'analyse portée sur l'étude de l'existant nous avons constaté l'absence d'un système d'information, la gestion de l'Activité Amont (processus d'analyse, prise de décision) est actuellement assurée à la base de deux rapports distincts réalisés à base d'EXCEL :

a) Rapport Mensuel :

Ce rapport renseigne d'un part sur les sous activités de l'activité Amont dont la recherche, l'exploration, la production et d'un autre part sur les principaux agrégats.

- ❖ **La recherche** : consiste à donner le nombre et la situation géographique des découvertes d'hydrocarbures, qui sont réalisées par « Sonatrach » en effort propre et en association. Ces découvertes sont réparties par gisement comme le montre l'exemple ci-dessous.

Situation des 29 découvertes à fin décembre 2013 :

- **Berkine : 15**
- **Amguid Messaoud : 01**
- **Illizi : 09**
- **Oeud Maya : 01**
- **Ahnet : 03**

- ❖ **L'exploration** : il s'agit d'un ensemble d'histogrammes et diagrammes circulaires qui donnent les réalisations physiques dont forage, sismique terrain 2D et sismique terrain 3D.

Les informations affichées par l'histogramme de l'activité forage sont : Unités de forage (10^3 Mètres), les prévisions (nombre de forages prévus en termes de métrage), les réalisations (nombre de forages réalisés en termes de métrage), le taux de réalisation durant un mois donné et cumul à fin du mois.

Les informations affichées par l'histogramme de sismique terrain 2D et sismique terrain 3D sont : Unités (10^3 Km), les prévisions (nombre de

kilomètres prévus), les réalisations (nombre de kilomètres réalisés), le taux de réalisation au cours du mois et cumul à fin du mois.

Les diagrammes circulaires (camemberts) permettent d'afficher la répartition de la sismique 2D et 3D par gisements.

- ❖ **La Production** : se matérialise sous la forme des histogrammes et des diagrammes circulaires qui déterminent la production globale réalisée par Sonatrach en effort propre et en association.

Les informations affichées par l'histogramme de la production globale sont : Unité de production (Million de Tep), les prévisions (la production globale prévue), les réalisations (la production globale réalisée), le taux de réalisation au cours du mois et cumul à fin du mois

Le diagramme circulaire affiche la répartition de la production globale par produits (Pétrole brut, Gaz naturel, Condensat, GPL).

- ❖ **Les principaux agrégats** : ces principaux indicateurs se présentent sous forme de tableaux, qui donnent les informations spécifiques à la production globale, et au développement des puits :

- **La production globale** : présenté sous forme d'un tableau qui donne la production globale d'hydrocarbures réalisée en effort propre durant un mois donné.

Il affiche les informations suivantes : Unités de production de chaque produit, les Prévisions, les Réalisations, le Taux de Réalisation par le Total de production dont association, et produit (Pétrole brut, Gaz naturel, Condensat, GPL).

- **Développement des puits** : c'est un tableau qui donne le nombre de puits achevés durant un mois donné.

Les champs de ce tableau sont : Unités (nombre de puits), les Prévisions, les Réalisations, le Taux de Réalisation **selon** les Puits terminés.

b) Tableau de Bord :

Ce rapport reprend les principaux indicateurs sous la forme des tableaux croisés, avec une analyse qui détaille chaque tableau.

Les informations fournies par chaque tableau sont :

Les indicateurs (Taux de Réalisation), les prévisions, les réalisations, les objectifs de l'année courante par rapport aux différents paramètres de l'Activité Amont :

- La recherche (découvertes d'hydrocarbures).
- L'exploration (sismique 2D, sismique 3D, forage d'exploration, forage de développement).
- La production d'hydrocarbures (production globale, production par produit).
- Le développement (principaux projets en développement).

IV.2. Limites du reporting existant

Certes que le reporting, offert par les gestionnaires de l'Activité Amont, fournisse quelques informations aux décideurs ; or il présente aussi des limites :

- Il n'est pas du tout évident de comparer les données entre elles comme par exemple : comparer entre les associations ou bien comparer les données d'exercices différents.
- Le temps de réponses est ainsi très lent, si le gestionnaire souhaite effectuer plusieurs analyses, il doit naviguer entre plusieurs fichiers excel et comparer les tableaux entre eux pour ensuite confronter les différentes informations.
- Les données sont présentés selon deux axes d'analyses métier au plus.
- Les données sont présentées sous un format tabulaire qui offre une lecture limitée et une interprétation difficile (difficulté de voir les variations, les évolutions...).

IV.3. Sources de données

Les données qui vont alimenter le futur système décisionnel que nous nous proposons de le réaliser sont issues de fichiers EXCEL, ces derniers comportent les informations liées au :

➤ **Récapitulatif des charges par nature :**

Ce fichier regroupe les 30 associations avec les prévisions et les réalisations des charges de production (Marchandises Consommées, Matières & Fournitures Consommées, Services, Impôt & Taxe, Frais Financiers, Frais Divers).

➤ **Récapitulatif de production vendue :**

Ce fichier donne la production vendue des hydrocarbures par produit pour chaque association.

IV.2. Etude des besoins

IV.2.1. Introduction

L'étude des besoins des utilisateurs est l'étape la plus importante dans un projet décisionnel. La réussite d'un tel projet dépend d'une bonne compréhension de ces besoins, un système décisionnel qui ne répond pas à cette exigence est obsolète.

Un projet décisionnel requiert du temps et des ressources importantes, avec un risque potentiel d'échec si cette étape est négligée ou mal effectuée.

Dans la définition des besoins des utilisateurs, nous avons opté pour une approche hybride, prenant en compte les besoins des utilisateurs et les données sources, afin de s'assurer que les besoins définis correspondent à ceux des utilisateurs et que les données existantes permettent de répondre à tels besoins.

IV.2.2. Définition des besoins

Afin de recenser les besoins pertinents, et de répondre aux attentes des futurs utilisateurs, et aux limites existantes en termes de consolidation et reporting décisionnel, nous avons tout au long de notre projet, travaillé d'une façon régulière avec les deux chefs de départements « Gestion Activité AMONT » et « Décisionnel & Portails ». Nous avons ainsi opté pour la démarche suivante :

- Elaboration des questionnaires.
- Organisation des entretiens.
- Recensements et validation des besoins.

IV.2.2.1. Elaboration des questionnaires

Dans le cadre de notre enquête, on a élaboré un questionnaire qui dépend de la fonction de chaque personne enquêtée, afin de recueillir les informations nécessaires et d'extraire l'ensemble d'indicateurs pertinents à la réalisation de notre projet.

Le tableau suivant présente les principales questions que nous avons élaborées pour le responsable de département Gestion « Activité Amont » :

Question	Réponse
Quels sont les processus métiers que vous pilotez dans votre secteur ?	<ul style="list-style-type: none"> • Un seul processus composé de 3 maillons : • Exploration • Développement • Production
Quel est votre rôle dans votre secteur d'activité ?	<ul style="list-style-type: none"> • Calculer le montant des charges, la production vendue (prévu et réalisé). • Consolider des prévisions et des suivis en matière de coûts et prix de revient.
Quels sont les principaux problèmes liés à l'analyse et à la prise de décision ?	<ul style="list-style-type: none"> • Problème de production de l'information par rapport aux délais. • Lenteur dans la collecte de l'information • Absence d'un système d'information fiable
Quels sont les types d'analyse effectuée et à partir de quelle source ?	<ul style="list-style-type: none"> • Suivre la production des hydrocarbures et leurs charges. • Fichiers, rapport (Excel, PDF).
Quels sont les bons indicateurs pour la prise de décision ?	<ul style="list-style-type: none"> • Les indicateurs de l'efficacité et de l'efficience et ses évolutions (taux de réalisation, les écarts).
Quels sont les types d'analyse souhaitez-vous au futur de notre système finale ?	<ul style="list-style-type: none"> • Suivre le coût de production (coût opératoire, coût départ champ) et l'amortissement de développement pour chaque association. • Suivre la production globale

Tableau 4 : Questionnaire aux responsables du département Gestion «Activité Amont »

IV.2.2.2. Les entretiens

Contrairement au questionnaire, l'entretien se révèle comme le moyen le plus efficace pour recueillir les informations nécessaires à notre travail. Il laisse plus au moins de liberté au choix de réponses des enquêtés avec leurs mots et des détails faisant sens selon eux.

L'entretien avec le chef de département Gestion «Activité Amont » nous a permis, tout en étant centré sur le sujet interrogé, de déduire un ensemble d'informations servant de base pour l'identification des besoins et de fixer l'ensemble des axes d'analyse, qui peuvent nous aider à la conception de la solution proposée.

IV.2.2.3. Recensement et validation des besoins

Cette étude nous a permis de recenser les besoins issues des questionnaires et des entretiens concernant l'activité de production. Qui peut être classée par volets spécifiques de la manière suivante :

IV.2.2.3.1. Volet suivi des charges de production

- ✓ Utilisateurs : les gestionnaires du département gestion « activité AMONT ».
- ✓ Besoins : les utilisateurs auront la possibilité de :
 - Suivre les montants accordés aux charges de production (prévues, réalisées).

Les différents suivis se font selon plusieurs axes et mesures d'analyses comme le montrent les dimensions métier, dont les flèches représentent les axes d'analyses et ce qui a été encadré représente les mesures d'analyses.

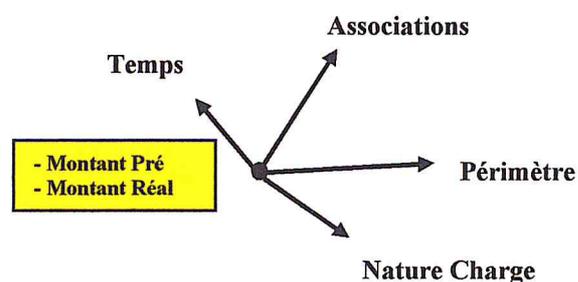


Figure 18 : Les dimensions métier du volet charges de production

IV.2.2.3.2. Volet suivi des amortissements de développement

- ✓ Utilisateurs : les gestionnaires du département gestion « activité AMONT ».
- ✓ Besoins : les utilisateurs auront la possibilité de :
 - Suivre la valeur de l'amortissement de développement pour chaque association.

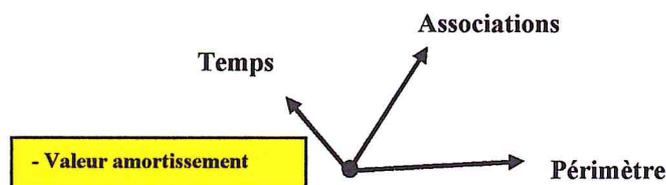


Figure 19 : Les dimensions métier du volet amortissement de développement

IV.2.2.3.3. Volet suivi de production vendue

- ✓ Utilisateurs : les gestionnaires du département gestion « Activité Amont ».
- ✓ Besoins : les utilisateurs auront la possibilité de suivre la production vendue par produit pour chaque association.

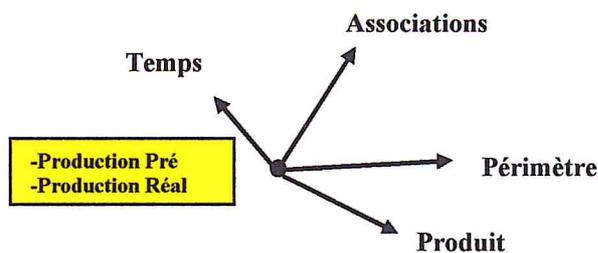


Figure 20 : Les dimensions métier du volet de production

Après avoir récupéré les besoins d'analyse on peut les classer sous forme d'indicateurs de performance comme suit :

❖ **Indicateur « Taux de réalisation mensuel »**

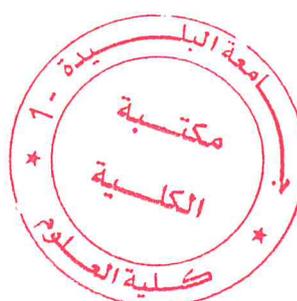
Indicateur	Taux de réalisation mensuel
Rôle	Ce taux permet de mesurer le montant des charges réalisé par rapport à ce qui est prévu.
Mode de calcul	$\sum (\text{charges mois } x \text{ réalisé}) / \sum (\text{charges mois } x \text{ prévu})$
Unité de mesure	%
Source de données	Fichier Excel « charge de production par nature »
Fréquence	Chaque fin du mois

Tableau 5 : Description de l'indicateur « Taux de réalisation mensuel »

❖ **Indicateur « Taux de réalisation cumulé »**

Indicateur	Taux de réalisation cumulé
Rôle	Ce taux permet de mesurer le montant des charges réalisé par rapport à ce qui est prévu.
Mode de calcul	$\sum (\text{charges mois } x_n, x_{n-1} \dots x_1 \text{ réalisé}) / \sum (\text{charges mois } x_n, x_{n-1} \dots x_1 \text{ prévu})$
Unité de mesure	%
Source de données	Fichier Excel « charge de production par nature »
Fréquence	Chaque fin du mois

Tableau 6 : Description de l'indicateur « Taux de réalisation cumulé »



❖ **Indicateur « Taux d'avancement »**

Indicateur	Taux d'avancement
Rôle	Ce taux permet de calculer les charges des mois réalisées par rapport au total des charges prévues pour une meilleure visibilité des mois restants.
Mode de calcul	$\sum (\text{charges mois } x_1, x_2 \dots x_n \text{ réalisé}) / \text{Total des charges prévu}$
Unité de mesure	%
Source de données	Fichier Excel « charge de production par nature »
Fréquence	Chaque fin du mois

Tableau 7 : Description de l'indicateur « Taux d'avancement ».

❖ **Indicateur « Taux d'évolution »**

Indicateur	Taux d'évolution
Rôle	Ce taux permet de mesurer l'évolution des charges entre 2 mois successifs durant le même exercice N ou le même mois de deux exercices (N, N+1).
Mode de calcul	$\frac{\sum (\text{charges mois } x_n \text{ réalisé}) - \sum (\text{charges mois } x_{n-1})}{\sum (\text{charges mois } x_{n-1})}$
Unité de mesure	%
Source de données	Fichier Excel « charge de production par nature »
Fréquence	Chaque fin du mois

Tableau 8 : Description de l'indicateur « Taux d'évolution ».

❖ Indicateur « Ecart »

Indicateur	Ecart
Rôle	Désigne l'écart entre les réalisations et les prévisions des charges.
Mode de calcul	$\sum (\text{charges mois } x \text{ réalisé}) - \sum (\text{charges mois } x)$
Unité de mesure	Kilo Dinar (KDA)
Source de données	Fichier Excel « charge de production par nature »
Fréquence	Chaque fin du mois

Tableau 9 : Description de l'indicateur « Ecart ».

❖ Indicateur « Coût »

Indicateur	Coût
Rôle	Permet de calculer le rapport entre le total des charges réalisées et la production globale réalisée. Plus le coût est élevé plus les charges sont importantes.
Mode de calcul	$\sum (\text{charges mois } x \text{ réalisé}) + \text{Amortissement} / \sum (\text{production mois } x \text{ réalisé})$
Unité de mesure	KDA (Kilo Dinar) / Tep (Tonne équivalent Pétrole)
Source de données	Fichier Excel « charge de production par nature » Fichier Excel « Production vendue »
Fréquence	Chaque fin du mois

Tableau 10 : Description de l'indicateur « Coût ».

Remarque :

La fréquence détermine le nombre de calculs de l'indicateur dans une période.

Exemple : chaque fin de mois indique cet indicateur est calculé une fois chaque fin de mois.

Dans cette partie on a rencontré plusieurs problèmes tels que :

- Les préoccupations quotidiennes et l'indisponibilité des responsables en raison des missions, des formations et des réunions successives, ce qui a rendu la collecte de besoins plus difficile à cerner.
- L'absence d'une base de données, et la grande confidentialité des informations. Cela est accompagné d'un ensemble de fichiers Excel incomplets.

En revanche, malgré ce nombre de limites qui s'est présentés, on a pu identifier les besoins des utilisateurs en restructurant les fichiers Excel, puis les compléter, ce qui permet de faciliter l'alimentation de notre entrepôt de données.

Cela n'était pas la seule amélioration, on a aussi pu apporter des nouveaux indicateurs tels que le taux d'avancement, le taux d'évolution et le coût, à partir d'indicateur préexistant comme le taux de réalisation. Nous signalons que cette étape nous a pris beaucoup de temps.

. Conclusion

A cette étape cruciale du projet, nous avons veillé à bien comprendre les besoins des utilisateurs à travers des questionnaires, des interviews des utilisateurs clés, et la détermination des axes d'analyse ainsi que des processus à modéliser. Puis nous avons confronté notre définition des besoins aux utilisateurs pour s'assurer que cela correspond bien à ce qu'ils recherchent, nous nous sommes assurés de la possibilité de répondre à ces besoins avec les données disponibles.

Cette définition des besoins permettra la conception du système, abordée dans le chapitre suivant.

PARTIE 3 : CONCEPTION DE LA SOLUTION

La solution proposée va être conçue en plusieurs étapes. La solution étant un système décisionnel s'articulant autour d'un entrepôt de données, on va donc concevoir cet entrepôt de données, sa zone d'alimentation, sa zone de restitution en passant par la conception des cubes.

Chapitre V

Conception de la zone d'entreposage

Introduction

Nous allons dans ce chapitre présenter la conception de la zone d'entreposage qui constitue la première étape de notre conception. À l'issue de cette étape, nous aurons les mesures avec les tables de faits qui les contiennent, ainsi que les dimensions selon lesquelles les mesures vont être analysées. La démarche de **Ralph Kimball**, que nous avons suivie dans cette conception, comprend les étapes suivantes :

- sélectionner le processus d'activité à modéliser (étape effectuée lors de l'étude des besoins),
- déclarer le grain de ce processus (niveau de détail des données),
- choisir les dimensions participantes à ce processus (les axes selon lesquels le processus sera analysé),
- identifier les faits numériques (ce que nous voulons mesurer dans l'analyse).

Nous utiliserons dans cette démarche la modélisation multidimensionnelle.

V.1. Activité de production

V.1.1. Présentation de l'activité

La production fait partie de la vie de l'entreprise depuis sa création. Pour maintenir ou augmenter sa production, Sonatrach doit nécessairement assurer une bonne gestion de son capital et mettre en place de nouveaux moyens. Pour faire face à la demande future en hydrocarbures, il est donc important de connaître les besoins des décideurs à évaluer, qui s'articulent autour de trois volets : volet « suivi des charges », volet « suivi des amortissements de développement », et volet « suivi de production ».

V.1.2. Les faits

- **Le grain :** La déclaration du grain permet de spécifier ce qu'une ligne de la table de faits représente. Le grain doit être le plus fin possible, ce qui veut dire qu'il serait avantageux de retenir les données les plus atomiques possibles.

Dans notre cas, le grain correspond à : connaître le montant des charges prévues, le montant des charges réalisées, et le coût opératoire par *Mois*, par *Association*, par *Périmètre*, et par *Nature de charge*.

- **Les mesures :**
 - Le montant prévu accordé aux charges de production : <montant_prev>
 - Le montant réalisé accordé aux charges de production : <montant_real>
- **Détails du fait :**

fait_charge		
unite_compt	integer	<fk1>
code_temps	integer	<fk2>
code_compt	integer	<fk3>
id_perimetre	integer	<fk4>
montant_prev	numeric(22,2)	
montant_real	numeric(22,2)	

Figure 21 : Le fait « Charge »²

² Toutes les figures de ce chapitre correspondent à des captures d'écran réalisées à partir du logiciel de conception Power Designer 16.5

Charge	
Attributs	Description
unite_compt	Unité de comptable, concaténation d'un code et d'un numéro séquentiel
code_temps	Code de la date, concaténation du numéro du mois et de l'année
code_compt	Code comptable, concaténation du numéro de la classe comptable et d'un numéro séquentiel
id_perimetre	Numéro séquentiel
montant_prev	Montant prévisionnel accordé aux charges
montant_real	Montant réalisé accordé aux charges

Tableau 11 : Tableau descriptif du fait « Charge »

V.1.2.2. Suivi des amortissements de développement

- **Le grain :** Le grain correspond à : connaître l'amortissement de développement d'une association par Mois, Périmètre et Association.
- **Les mesures :**
 - Valeur de l'amortissement d'une association : <amortissement>
- **Détail du fait :**

fait_amortissement		
unite_compt	integer	<fk1>
code_temps	integer	<fk2>
id_perimetre	integer	<fk3>
amortissement	numeric(22,2)	

Figure 22 : Le fait « Amortissement de développement »

Amortissement de développement	
Attributs	Description
unite_compt	Unité de comptable, concaténation d'un code et d'un numéro séquentiel
code_compt	Code comptable, concaténation du numéro de la classe comptable et d'un numéro séquentiel
id_perimetre	Numéro séquentiel
amortissement	Valeur de l'amortissement

Tableau 12 : Tableau descriptif du fait « Amortissement de développement »

V.1.2.3. Suivi de production :

- **Le grain :** Le grain correspond à : connaître la quantité de production prévue et réalisée de chaque produit par *mois*, par *Périmètre*, *Produit*, et *Association*.
- **Les mesures :**
 - Production prévue : <production_prev>
 - Production réalisée : <production_real>
- **Détail du fait :**

fait_production		
id_perimetre	integer	<fk1>
unite_compt	integer	<fk2>
code_temps	integer	<fk3>
id_produit	integer	<fk4>
production_real	numeric(22,2)	
production_prev	numeric(22,2)	

Figure 23 : Le fait « Production »

Production	
Attributs	Description
id_perimetre	Numéro séquentiel
unite_compt	Unité de comptable, concaténation d'un code et d'un numéro séquentiel
code_temps	Code de la date, concaténation du numéro du mois et de l'année
id_produit	Numéro séquentiel
production_prev	Production prévue
production_real	Production réalisée

Tableau 13 : Tableau descriptif du fait « Production »

V.1.3. Les dimensions

V.1.3.1. Dimension Temps :

C'est une dimension triviale dans un système décisionnel, elle permet de faire des regroupements temporels selon le mois et l'année, les utilisateurs ont besoin de suivre leur activité d'un mois à un autre et d'en garder l'historique

dim_temps		
<u>code_temps</u>	integer	<pk>
mois	varchar(2)	
design_mois	varchar(15)	
annee	varchar(15)	

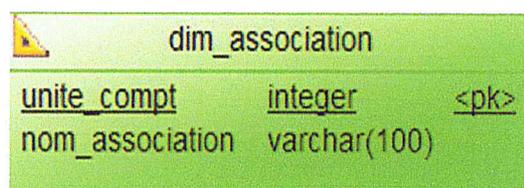
Figure 24 : La dimension « Temps »

Temps	
Attributs	Description
code_temps	Code de la date, concaténation du numéro du mois et de l'année.
mois	Numéro du mois
design_mois	Désignation du mois
annee	Année

Tableau 14 : Tableau descriptif de la dimension « Temps »

V.1.3.2. Dimension Association :

Cette dimension décrit toutes les associations de la Sonatrach, dans le domaine de production des hydrocarbures, elle contient l'unité comptable de l'association ainsi que son nom.



```

dim_association
unite_compt    integer    <pk>
nom_association varchar(100)
  
```

Figure 25 : La dimension « Association »

Association	
Attributs	Description
unite_compt	Unité de comptable, concaténation d'un code et d'un numéro séquentiel
nom_association	Nom de l'association

Tableau 15 : Tableau descriptif de la dimension « Association »

V.1.3.3. Dimension Nature-charge :

Cette dimension permet de regrouper les différentes charges de production qui sont enregistrées au compte comptable de la classe 6 (60 marchandise consommées, 61 matière et fournitures consommées, 62 services...).

Le compte comptable est un compte dans lequel une entreprise enregistre ses actions financières. Il est composé de plusieurs chiffres, le premier chiffre du compte représente le numéro de la classe dans lequel doit s'inscrire l'opération concernée.

dim_nature charge		
<u>code_compt</u>	integer	<pk>
design_charge	varchar(100)	

Figure 26 : La dimension « Nature-charge »

Nature-charge	
Attributs	Description
code_compt	Code comptable, concaténation du numéro de la classe comptable et d'un numéro séquentiel
design_charge	Désignation de la nature de charge

Tableau 16 : Tableau descriptif de la dimension « Nature-charge »

V.1.3.4. Dimension Périmètre :

Elle permet d'assurer un suivi par périmètre ou gisement. En effet, chaque gisement (In Saleh, In Amenas, Hassi Berkine...), est décomposé en bloc (403a, 403d...).

dim_perimetre		
<u>id_perimetre</u>	integer	<pk>
n_bloc	varchar(15)	
nom_gisement	varchar(100)	

Figure 27 : La dimension « Périmètre »

Périmètre	
Attributs	Description
id_perimetre	Numéro séquentiel
n_bloc	Numéro du bloc
nom_gisement	Nom du gisement

Tableau 17 : Tableau descriptif de la dimension « Périmètre »

V.1.3.5. Dimension Produit

Cette dimension contient les produits des hydrocarbures, on distingue quatre types de produit : Pétrole brute, Gaz naturel, Condensat, GPL (gaz de pétrole liquifié).

dim_produit		
<u>id_produit</u>	integer	<pk>
design_produit	varchar(6)	

Figure 28 : La dimension « Produit »

Produit	
Attributs	Description
id_produit	Numéro séquentiel
design_produit	Désignation de produit

Tableau 18 : Tableau descriptif de la dimension « Produit »

V.1.4. Dimensions participantes pour chaque fait

Dimension Fait	Temps	Association	Nature charge	Périmètre	Produit
Charge	✓	✓	✓	✓	
Amortissement	✓	✓		✓	
Production	✓	✓		✓	✓

Tableau 19 : Tableau descriptif des dimensions participantes pour chaque fait

V.1.5. Schéma globale de l'activité « Production »

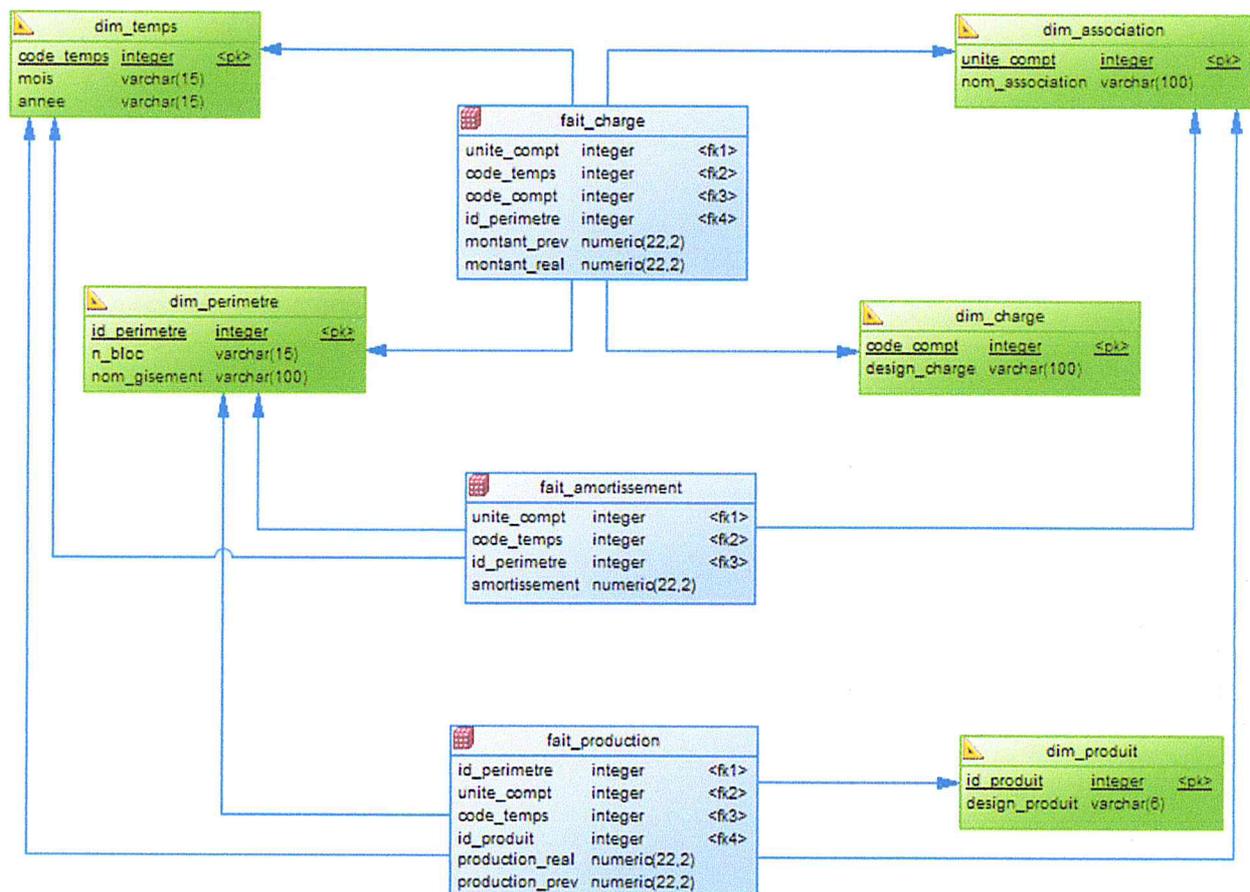


Figure 29 : schéma en constellation de l'activité « Production »

Conclusion

Au terme de cette première étape, nous avons identifié les mesures, les faits qui leur sont associés et les dimensions selon lesquelles seront analysés ces faits. Et cela en tenant compte des besoins des utilisateurs définis précédemment. Conformément à la définition des besoins, nous avons trois volets : suivi des charges, suivi des amortissements, et suivi de production. Ces deux volets correspondent un schéma en constellation qui va constituer l'entrepôt.

L'entrepôt de données que nous avons conçu va être exploité par les utilisateurs, mais avant cela il est nécessaire de l'alimenter.

Nous passons donc à l'étape suivante qui est la construction de la zone d'alimentation de l'entrepôt, abordé dans le prochain chapitre.

Chapitre VI

Conception de la zone d'alimentation

Introduction :

L'alimentation de l'entrepôt de données est une étape importante dans le projet décisionnel, car elle garantira la pertinence et la qualité des données que contiendra l'entrepôt et assurera au décideur l'accès à la bonne information. L'alimentation se fait grâce à l'opération d'ETL, il s'agira d'extraire les données pertinentes ce qui nécessite une bonne connaissance des sources de données, puis de transformer les données au format voulu, pour enfin les charger selon une périodicité précise et de la manière adaptée. Nous présentons dans ce chapitre les différentes étapes que nous avons suivi pour l'alimentation de notre entrepôt.

VI.1. Identification des sources de données

Lors de l'étape de l'identification des besoins, nous avons adopté une démarche qui assure que les besoins définis correspondent à ceux des utilisateurs et que les données existantes permettent de répondre à de tels besoins. De plus l'étape de conception de l'entrepôt a abouti un schéma en constellation qui indique les données nécessaires.

Comme vu précédemment, nos sources de données sont les fichiers Excel issues de département gestion activité « Amont ». Chaque fichier correspond aux réalisations et prévisions des charges et à la production vendue des hydrocarbures.

L'outil utilisé dans le processus d'ETL est Talend Open Studio qui est à la fois gratuit et puissant pouvant supporter une charge importante de données.

VI.2. Extraction des données

Une fois les données identifiées, nous procédons à leur extraction, nous allons récupérer les données jugées pertinentes, destinées à être exploitées et en accord avec le résultat à obtenir. Il est à noter que les fichiers sources ont subis plusieurs modifications pour les rendre plus structurés et organisés sous forme de colonnes avec des champs d'attributs :

- à partir du fichier Excel récapitulatif des charges on a pu extraire les fichiers suivant :
 - fichier associations (UC, association),
 - fichier gisements (id_gisement, gisement, UC),
 - fichier blocs (bloc, id_gisement, UC),
 - fichier charges (UC, association, id_perimetre, gisement, bloc, date, num_compte, design_charge, objectif, réalisation),
 - fichier amortissement (UC, association, id_perimetre, gisement, bloc, date, amortissement).
- à partir du fichier Excel récapitulatif de production vendue on a pu extraire les fichiers suivant :
 - fichier produit (id_produit, produit),
 - fichier production (UC, association, id_perimetre, gisement, bloc, date, id_produit, produit, objectif, réalisation).

VI.3. Transformation

Les données extraites subissent des transformations avant d'être chargées dans l'entrepôt, nous distinguerons les opérations de :

- **Filtrage** : le filtrage nous a permet de faire des restrictions sur les données nulles, les champs vides... etc., dans notre cas l'utilisation de filtrage est d'éliminer le chargement :
 - des associations ou le montant des charges non présent
 - des produits non fournis par certaines associations.
- **Conversion** : la conversion consiste à changer le type de données de certains attributs, les unités de mesures... etc., par exemple :
 - **Date**: date (dd-mm-yyyy) \longrightarrow String (ddmmyyyy).
 - **Production vendue** :
KDA (kilo dinars) \longrightarrow TEP (tonne équivalent pétrole).
 - **Montant charges** : Integer \longrightarrow BigDecimal.
- **Changement des libellés des champs** : le changement des libellés permet d'éviter des conflits qui se produisent lorsqu'on utilise des noms différents pour le même attribut, ainsi pour rendre les champs plus significatifs. Par exemple l'unité comptable des associations se trouve avec des différents libellés dans plusieurs fichiers sources :

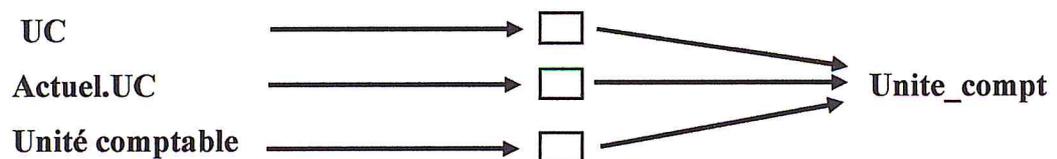


Figure 30 : Exemple de changement des libellés

De plus de ces opérations on a utilisé les opérations de sélection (sélectionner le mois ou l'année d'une date donnée), de concaténation (concaténer le mois et l'année pour avoir le code date), et de regroupement des données.

Ces transformations se font généralement au moment du mapping (correspondance entre les données).

VI.4. Chargement

Nous distinguerons le chargement des dimensions et le chargement des tables de faits, mais aussi un chargement initial et des chargements périodiques. Le chargement initial ne comporte pas de contraintes particulières, cependant les chargements périodiques doivent prendre en compte différents critères :

- La fréquence de mise à jour des fichiers sources,
- Le volume de données à charger, le temps nécessaire et le temps disponible pour le chargement.
- Niveau d'actualisation des données demandé par les utilisateurs.

Le chargement des tables de faits doit venir après le chargement des dimensions car les premières référencent les secondes à l'aide des clés étrangères.

L'opération de chargement de la source vers l'entrepôt sera exécutée à chaque fin du mois, car le plus détaillé en termes de période d'analyse est le mois.

La figure suivante illustre le processus de chargement.

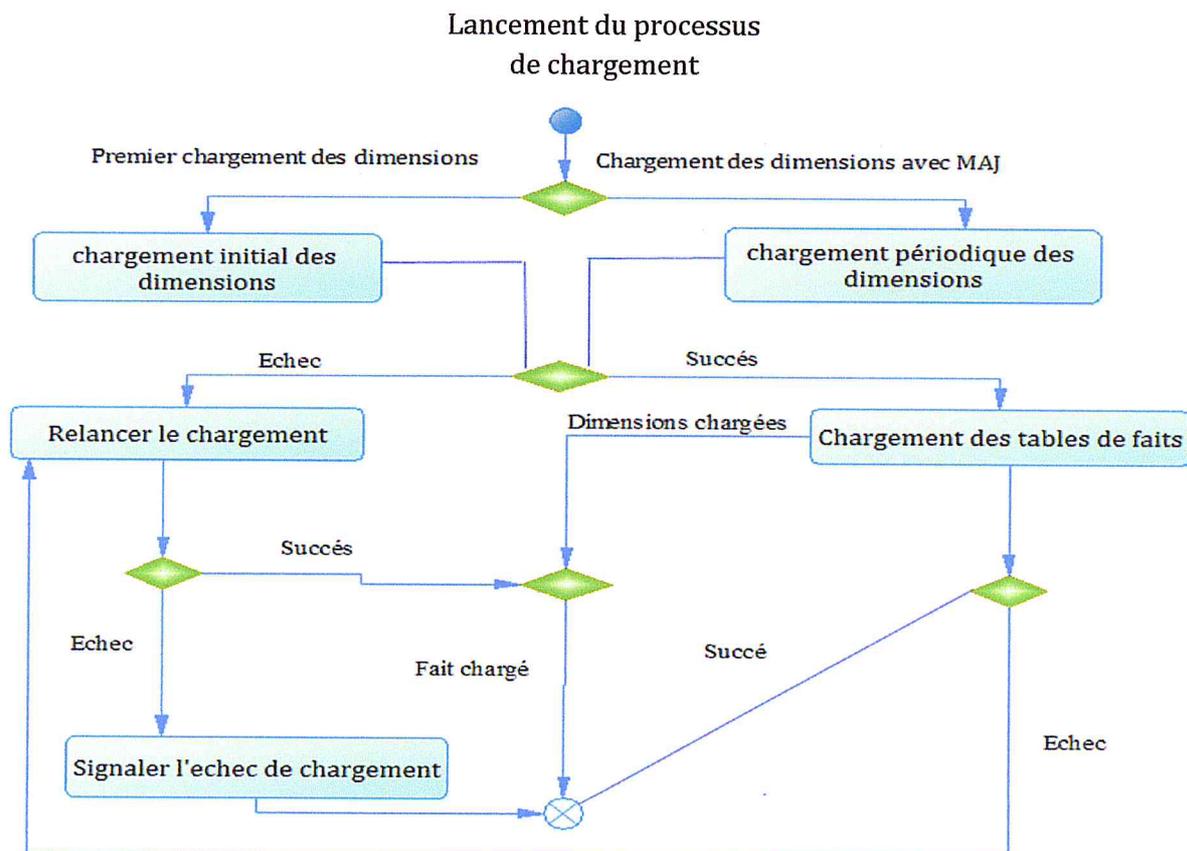


Figure 31 : Diagramme d'activité de processus de chargement.

Conclusion

Dans ce chapitre nous avons abordé la conception de la zone d'alimentation qui va nous assurer des données fiables et de qualité dans l'entrepôt de données. Nous avons vu les différentes étapes : identification des sources de données, extraction, transformation et enfin chargement. Pour la périodicité des chargements, elle a été déterminée en fonction de la fréquence de mise à jour des tables sources et des besoins des utilisateurs (chaque fin du mois).

Cette étape achevée nous passons dans le chapitre suivant, à la conception des cubes de données.

Chapitre VII

Conception des cubes dimensionnels

Introduction

Dans ce chapitre nous abordons la conception des cubes dimensionnels, une structure très souple à l'interrogation des données offrant une navigation adaptée et une exploitation. Un cube correspond à un sujet d'analyse : un fait décrit par les besoins des utilisateurs, il contient un ensemble d'indicateurs en hiérarchies, les hiérarchies étant décomposées en niveaux. La navigation peut se faire de manière croisée (interroger la valeur d'un indicateur selon trois axes par exemple), de manière ascendante ou descendante en jouant sur le niveau de granularité des axes d'analyse et de manière spécifique à l'aide de filtres.

L'expression des besoins des utilisateurs a conduit à la conception entre autres un schéma en constellation concernant les trois volets : suivi des charges et suivi des amortissements, et suivi de production. Ce schéma correspond (dimensions, mesures) aux cubes qui seront conçu. Nous détaillons, dans ce qui suit la conception de chaque cube.

VII.1. définition des niveaux et hiérarchies des dimensions

Les hiérarchies de chaque dimension ainsi que les niveaux correspondant à chaque hiérarchie vont définir une structure hiérarchique. Cette structure permettra la navigation à l'intérieur de la dimension à travers les opérations de « forage vers le bas » (accès à un niveau de détail supérieur) ou de « forage vers le haut » (accès à un niveau de détail inférieur).

Par exemple, l'utilisateur peut analyser la production par année (sommet de la hiérarchie) puis par mois (opération de forage vers le bas).

Le tableau suivant regroupe les dimensions de l'ensemble des cubes, avec la définition des hiérarchies, niveaux et colonnes correspondantes.

Dimension	Hiérarchies	Niveaux	Colonnes
Dim_temps	h_temps	Niveau 1 = N1	annee
	ALL → N1 → N2	Niveau 2 = N2	code_date
			mois
			design_mois
Dim_association	h_association	Niveau 1 = N1	unit_compt
	ALL → N1		nom_association
Dim_nature charge	h_nature_charge	Niveau 1 = N1	code_compt
	ALL → N1		design_chareg
Dim_perimetre	h_perimetre	Niveau 1 = N1	nom_gisement
	ALL → N1 → N2	Niveau 2 = N2	id_perimetre
			n_bloc
Dim_produit	h_produit	Niveau 1 = N1	id_produit
	ALL → N1		design_produit

Tableau 20 : Tableau donnant les niveaux et les hiérarchies de chaque dimension

VII.2. Listes des cubes

Le tableau suivant liste tous les cubes à réaliser avec, pour chacun, les mesures correspondantes et les dimensions participantes :

Nom de cube	Mesures	Dimensions
Volet suivi des charges :		
Charge	<ul style="list-style-type: none"> ▪ Montant_prev ▪ Montant_real 	Dim_temps
		Dim_association
		Dim_nature_charge
		Dim_peimetre
Volet suivi des amortissements :		
Amortissement de développement	<ul style="list-style-type: none"> ▪ Amortissement 	Dim_temps
		Dim_association
		Dim_peimetre
Volet suivi de production :		
Production	<ul style="list-style-type: none"> ▪ Production_prev ▪ Production_real 	Dim_temps
		Dim_produit
		Dim_association
		Dim_perimetre

Tableau 21 : Liste des cubes

VII.3. schémas des cubes

VII.3.1. volet suivi des charges

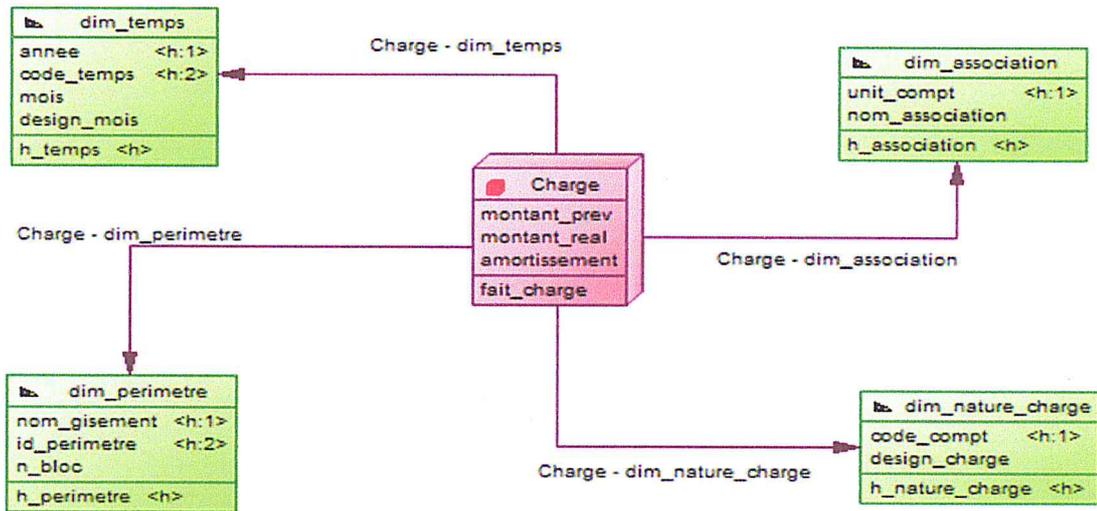


Figure 32 : Cube dimensionnel « Charge »

VII.3.2. volet suivi des amortissements de développement

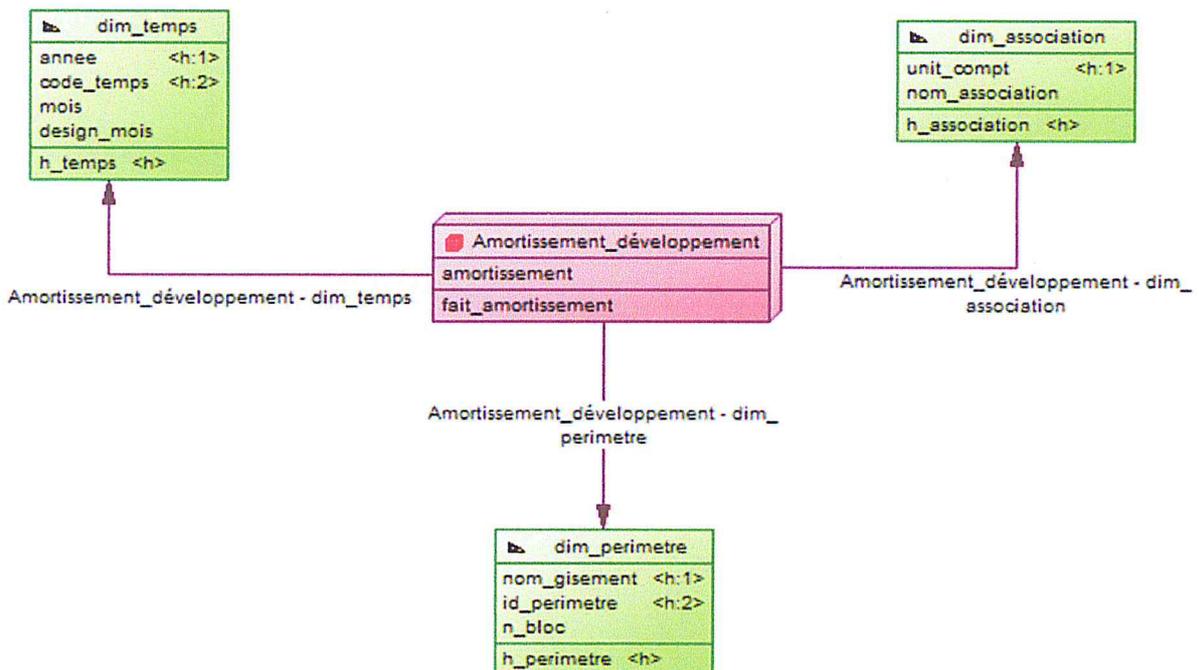


Figure 33 : Cube dimensionnel « Amortissement Développement »

VII.3.3. volet suivi de production

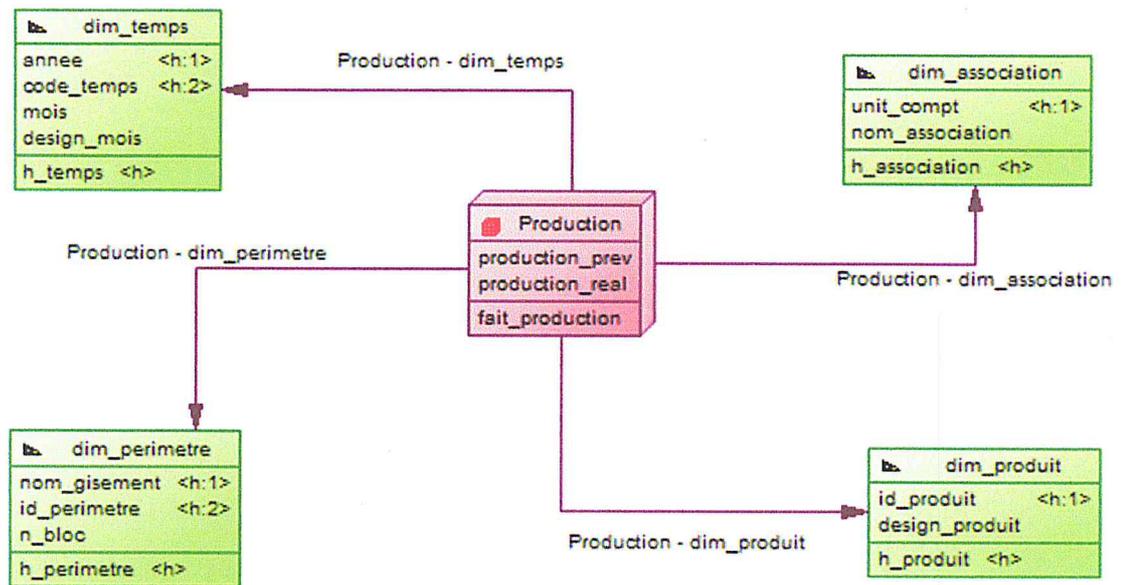


Figure 34 : Cube dimensionnel « Production »³

³ Toutes les figures de ce chapitre correspondent à des captures d'écran réalisées à partir du logiciel de conception Power Designer 16.5

Conclusion

Dans ce chapitre nous avons conçu les cubes dimensionnels qui vont permettre d'effectuer des analyses multidimensionnelles sur les données. Nous avons défini les mesures, les dimensions participantes ainsi que la structure hiérarchique pour chaque dimension. Ces cubes sont relatifs aux trois volets : suivi des charges, suivi des amortissements de développements et suivi de production. Ils offrent la possibilité aux décideurs d'exploiter les informations relatives à ces deux volets, grâce à une navigation rapide, intuitive et personnalisable (choix des axes d'analyse, du niveau du détail...).

La conception des cubes achevée, nous devons les implémenter. Cette implémentation est une des étapes de la réalisation et du déploiement, que nous aborderons dans la partie suivante

PARTIE 4 : REALISATION ET DEPLOIEMENT

La dernière phase du projet, à savoir la réalisation de la solution conçue. En utilisant les outils appropriés conformément à notre conception, nous aboutissons à un système décisionnel fonctionnel conforme aux attentes des utilisateurs.

Chapitre **VIII**

Réalisation et déploiement de la solution

Introduction

La conception de notre système étant achevée, nous passons à la phase de réalisation.

Nous commençons par décrire l'environnement existant ainsi que les outils que nous avons utilisés. Puis nous abordons la réalisation proprement dite avec la construction de l'entrepôt, la construction de la partie ETL, la construction et chargement des cubes et la mise en place de l'outil de restitution. Enfin nous évoquons les aspects : gestion des utilisateurs et sécurité, tous deux d'une importance capitale.

Nous obtenons à l'issue de cette étape un système fonctionnel et nous donnons un aperçu visuel de ce dernier à travers quelques captures d'écran.

VIII.1. Environnement technique et fonctionnel

Comme première étape de la réalisation, nous avons considéré l'environnement technique et fonctionnel en place afin d'éviter d'éventuels problèmes de comptabilité entre les différentes plateformes exécutantes, ainsi que de compatibilité des outils utilisés avec ces dernières.

VIII.1.1. Matériel

On retrouve des machines tournant sous les systèmes d'exploitation suivants :

- Système d'exploitation Windows XP SP3 et Seven.
- Windows Server 2003.
- UNIX-AIX.
- LINUX.
- La disponibilité de licences Oracle.

- Des machines HP :

- ✓ HP 9000-RP 5470 équipées d'un processeur *PA-RISC*
- ✓ HP-9000/800 équipées d'un processeur *PA-RISC 8700*
- ✓ HP RX 4640 équipées d'un processeur *Itanium2*

- Des machines DELL :

- ✓ *PC DELL GX 620 équipées d'un processeur Intel pentium4.*

VIII.1.2. Outils utilisés

VIII.1.2.1. Stockage de la base de données « Oracle 11g⁴ »

Oracle Database est un système de gestion de bases de données relationnelles (**SGBDR**) qui depuis l'introduction du support du modèle objet dans sa version 8 peut être aussi qualifié de système de gestion de bases de données relationnelles-objet (**SGDRO**). Proposé par le

⁴ <http://www.ouestdecision.fr>

géant américain Oracle Corporation, il a été développé par Larry Ellison accompagné de Bob Miner.

Caractéristiques principales d'Oracle Database 11g :

- Possibilité de stocker de gros volumes de données.
- Partitionnement des grosses tables.
- Sa disponibilité sur un grand nombre de systèmes et de plateformes (Unix, Linux, Windows, ...).
- PL/SQL, langage de programmation propre à Oracle, utilisé pour créer des triggers lors de l'insertion, la modification ou l'effacement des éléments.
- Accès aux données système via des vues, bien plus aisément manipulable que des procédures stockées.
- Un système de droits et mots de passe très souples et sécurisés, qui vérifie aussi les hôtes se connectant. Les mots de passe sont bien protégés, car tous les échanges de mot de passe sont chiffrés, même lors des connexions.

VIII.1.2.2. Extraction, transformation et chargement « Talend Open Studio⁵ »

La société Talend est le premier fournisseur d'outil d'intégration de données Open Source. Son logiciel, distribué sous licence GNU GPL⁶, permet de procéder à des opérations d'ETL, pour la construction d'entrepôt de données.

C'est un ETL générateur de code vu qu'il permet de créer des processus de manipulation et de transformation des données. À l'aide d'une interface graphique *job Designer* puis générer l'exécutable correspondant en JAVA qui sera déployé par la suite sur le serveur d'exécution.

Cet outil offre plusieurs avantages notamment :

- Un grand nombre de fonctionnalités à travers une gamme très large de composants,
- Une documentation complète, riche et gratuite,
- Compatibilité avec la majorité des SGBD et système d'exploitation,
- Une communauté très réactive et participe dans son développement.

⁵ <http://fr.talend.com/>

⁶ La licence GNU GPL est une des licences définissant le statut de logiciel libre

VIII.1.2.3. Construction des cubes « Oracle OLAP 11g⁷ »

Oracle OLAP est un moteur d'analyse multidimensionnelle (ROLAP) de classe mondiale intégrée dans Oracle Database 11g. Oracle OLAP cubes offrent des calculs sophistiqués utilisant des simples requêtes SQL, produit des résultats avec un temps de réponse très rapide comme Oracle OLAP est intégré dans Oracle Database, il permet une gestion centralisée des données et les règles métier dans une plate-forme sécurisée, évolutive et prête pour l'entreprise.

Oracle OLAP utilise **Analytic Workspace Manager (AWM 11g)⁸** pour construire et gérer des espaces de travail analytiques, il permet de concevoir un modèle dimensionnel, créer un espace de travail, puis le remplir à partir d'une série de tables relationnelles. AWM est devenu l'un des outils nécessaires pour gérer nos espaces de travail, Ces métadonnées servent un objectif semblable à Oracle Discoverer "couche de l'utilisateur final" qui organise vos tables dans les dimensions, les niveaux et les hiérarchies.

VIII.1.2.4. Reporting « Oracle BI Publisher Enterprise⁹ »

Oracle BI Publisher (précédemment nommé XML Publisher) est une solution de publication de document (rapport, états, tableaux de bord, ...) proposé par Oracle. BI Publisher permet d'afficher les rapports en ligne ou les programmer et livrer une dizaines de milliers de documents à travers un serveur FTP, serveur SMTP, ...

Oracle BI Publisher est capable de formater les données XML ou les générer. Il peut s'intégrer avec n'importe quelle base de données, capable de générer des données XML ou service web, il permet également de récupérer des données provenant de multiple source et de les présenter sur un même document avec des différents formats PDF, Word, Excel, et HTML.

⁷ <http://www.oracle.com/technetwork/database/enterprise-edition/awm-readme-112010-098372.html>

⁸ <http://www.vlamis.com/storage/papers/collab2009>

⁹ <http://www.absodia.com/T20-bi-publisher.html>

VIII.1.2.5. Plateforme de restitution « Oracle Business Intelligence (OBI) ¹⁰»

Oracle Business Intelligence est une plateforme intégrée BI complète et innovante qui répond à l'ensemble des besoins décisionnels. Possède également un portail web permettant aux utilisateurs d'accéder aux différents outils BI. Offrant une gamme étendue de fonctionnalités : Tableaux de bord interactifs, analyse à la demande, reporting interne et reporting financier, tableaux d'indicateur, exports et publication vers les outils bureautique (Office et Adobe) à partir d'un référentiel unique (couche de présentation). etc...

Composants

Oracle BI utilise plusieurs composants décisionnels :

- Reporting : Oracle BI Publisher (*utilisé dans ce projet*),
- Tableaux de bord : Oracle Dashbord (*utilisé dans ce projet*),
- Analyse OLAP : Oracle BI Answers (*utilisé : renforcé par Oracle OLAP 11g car offrant plus de fonctionnalités et une meilleur ergonomie*),
- Oracle BI Office : *Non utilisé.*

Oracle BI compte principalement les modules :

- Oracle BI Server : le serveur de Business Intelligence qui accueille les fonctionnalités centrales et analytiques, il gère des modèles décisionnels et permet aussi la gestion des multi-sources.

La plateforme Oracle BI est composée des moteurs d'exécution de chacun de ses composants, qui fonctionnent de façon individuelle (Console WebLogic, Oracle Enterprise Manager,...). Il est donc possible de paramétrer chacun des moteurs de traitement via l'interface dédiée. Cette souplesse offre des possibilités intéressantes, car l'utilisateur avancé peut se permettre d'ajouter ou d'enlever les moteurs, et il peut les configurer à sa façon.

La gestion des comptes utilisateurs et des groupes d'utilisateurs est une composante importante de la plateforme. L'administrateur a en effet la possibilité de créer et modifier les groupes d'utilisateurs qui vont interagir avec la plateforme en leur attribuant des rôles.

¹⁰ http://www.ouestdecision.fr/contenu/les_editeurs/fiches/fiche-produit-obiee.htm



Figure 35 : Architecture technique de la solution

VIII.2. Réalisation du projet

VIII.2.1. Construction de l'entrepôt de données

Nous allons créer, en utilisant SQL Developer (composant d'Oracle Database 11g), une base de données relationnelle à partir des schémas en étoiles (composés de faits et de dimension) obtenus dans la phase de conception de l'entrepôt. Dans la base de données que nous allons créer :

- Les dimensions sont des tables avec des clés primaires créées en utilisant un code spécifique ou générées à l'aide de concaténation,
- Les tables de faits sont des tables avec des clés étrangères de toutes les tables de dimension associées au fait.

VIII.2.2. Extraction, transformation et chargement

Une fois la base de données créée, nous procédons à l'ETL à l'aide de l'outil Talend Open Studio.

La première étape consiste à créer les jobs de remplissage (processus de manipulation et de transformation des données), chaque job correspondant à un ou plusieurs fichiers Excel source.

Dans les métadonnées, nous définissons les connexions aux fichiers sources ainsi que la base de données de l'entrepôt (paramètres de connexion). Nous importons le schéma des tables (correspondant aux connexions définies) dont nous avons besoin, ces tables seront utilisées dans les jobs. Les fichiers sources seront utilisés dans les composants d'entrée et les tables de l'entrepôt dans les composants de sortie, un composant de type « tmap » sera utilisé pour faire les correspondances entre les deux composants.

Les transformations effectuées sur les données sont généralement faites :

- à l'aide de filtres (composant « tFilter ») effectuant des restrictions sur les données.
- à l'aide de conversions (composant « tConvert »),
- au niveau du composant « tMap », à l'aide de correspondances, d'expressions simples (concaténation, opération arithmétique), ou bien à l'aide de routine (classes java effectuant une opération particulière) définies précédemment.

Nous créons les jobs de remplissage des tables de dimensions et des tables de faits. Par la suite, nous exécuterons les jobs de remplissage des tables de dimensions avant ceux des tables de faits, car les premiers référencent les seconds à l'aide de clés étrangères.

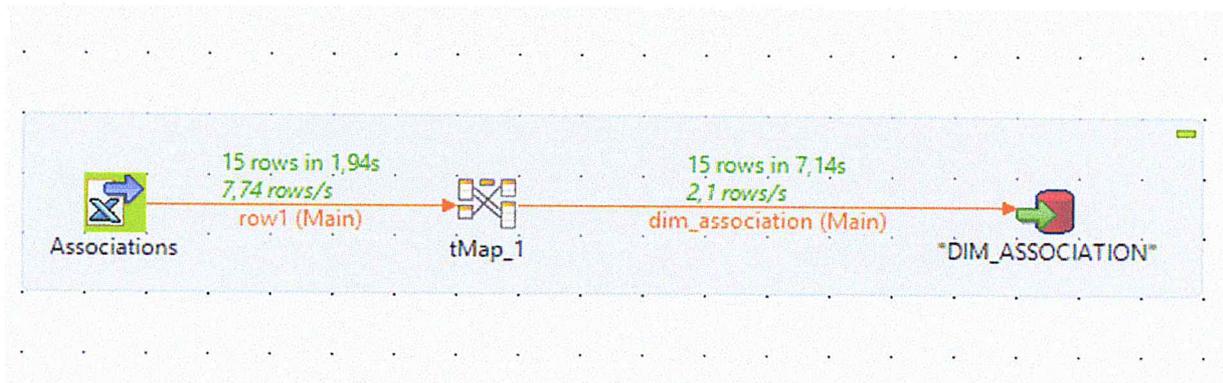


Figure 36 : job de remplissage de la dimension Association

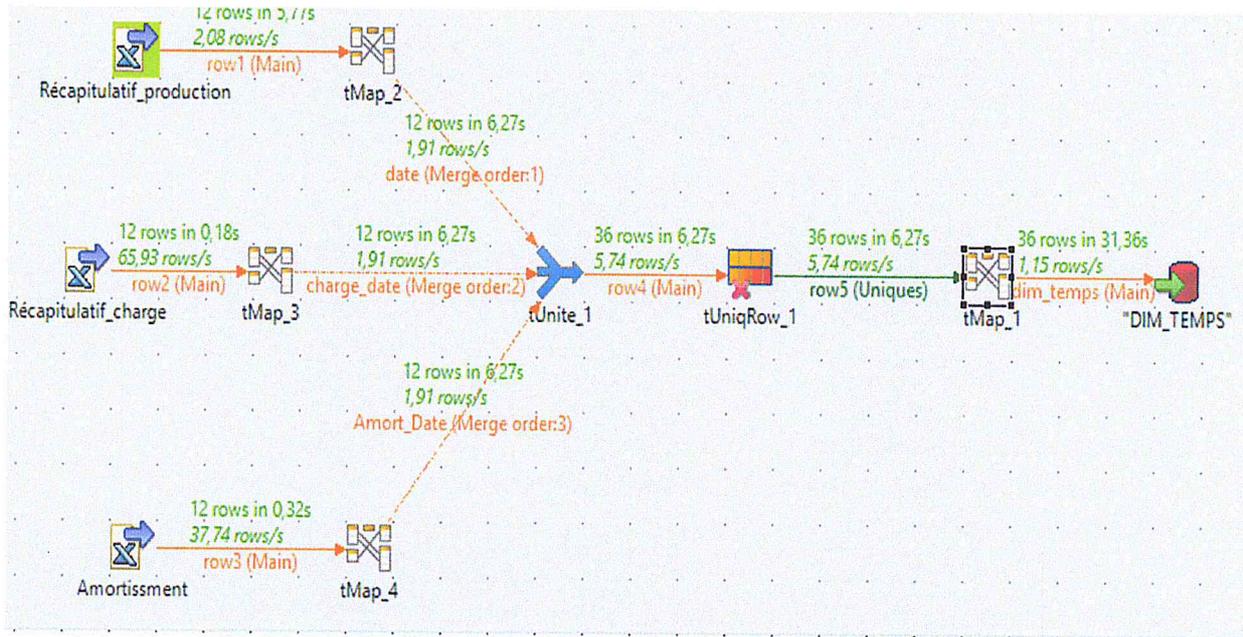


Figure 37 : job de remplissage de la dimension Temps

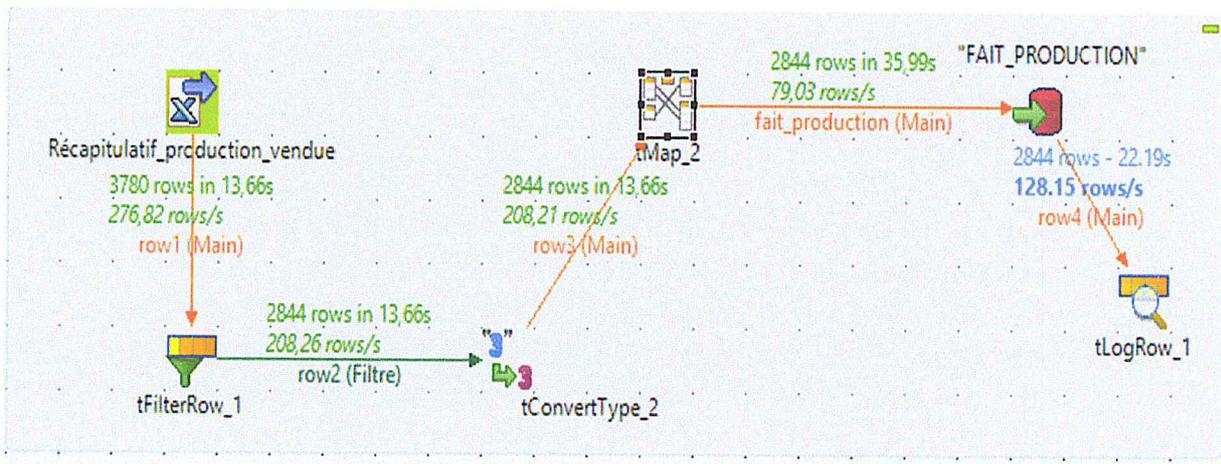


Figure 38 : job de remplissage du fait Production

VIII.2.3. Construction et chargement des cubes

Pour la construction des cubes, nous avons suivi les étapes suivantes :

- Création d'un nouveau schéma (nom utilisateur, mot de passe) au niveau de SQL Developer pour contenir l'ensemble des cubes relatifs au volet d'analyse correspondant.

- Création de la connexion à l'entrepôt de données sous Analytic Workspace Manager avec un utilisateur précédemment défini, pour obtenir l'ensemble des tables qui serviront à la création des cubes et leurs dimensions.
- Création d'un nouvel espace de travail analytique.
- Création de l'ensemble des dimensions, avec les hiérarchies et niveaux correspondant à chacune d'elles.
- Création des cubes à partir des dimensions déjà définies, avec les mesures correspondant à chaque cube.
- Mettre à jour les dimensions et les cubes à partir de l'entrepôt de données afin de visualiser les cubes.

Après avoir créé les cubes nous les exploitons à l'aide :

- Des tableaux croisés personnalisables à travers la sélection du cube, la sélection des mesures et des dimensions (à l'aide de « glisser-déposer »).
- Graphes associés aux tableaux croisés précédent avec la possibilité de choisir le type de graphe, d'inverser les axes d'analyse...

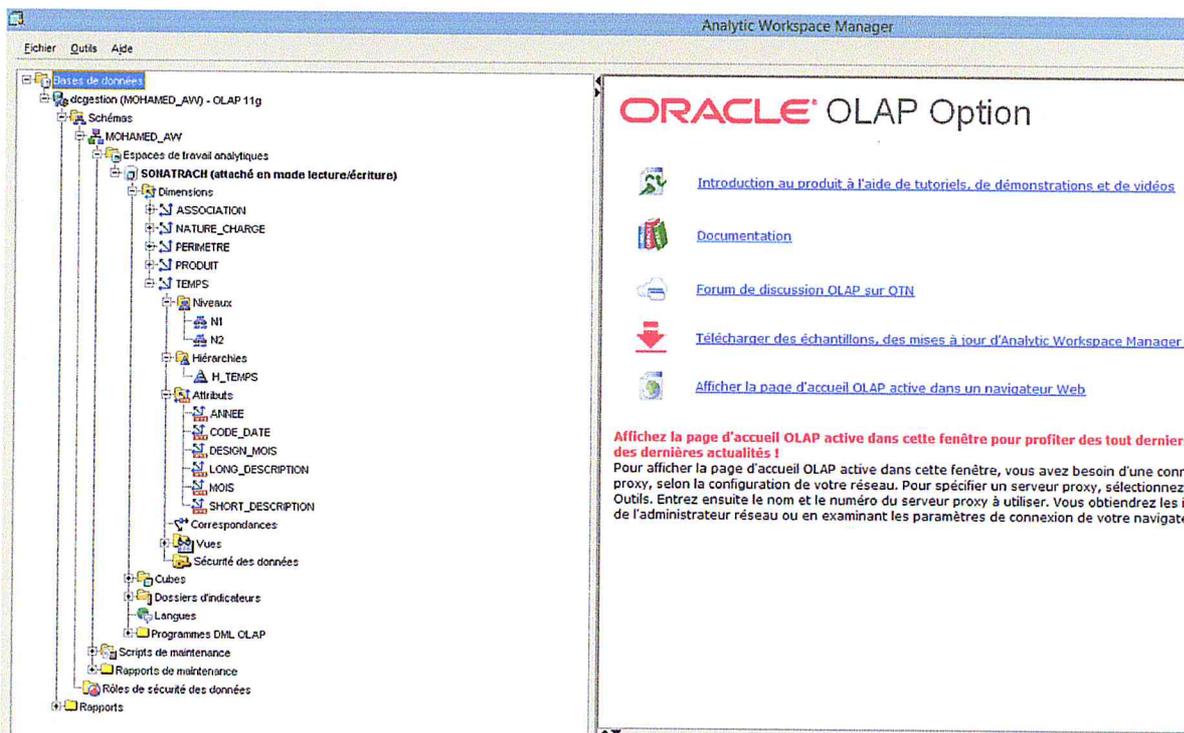


Figure 39 : Construction des dimensions avec AWM

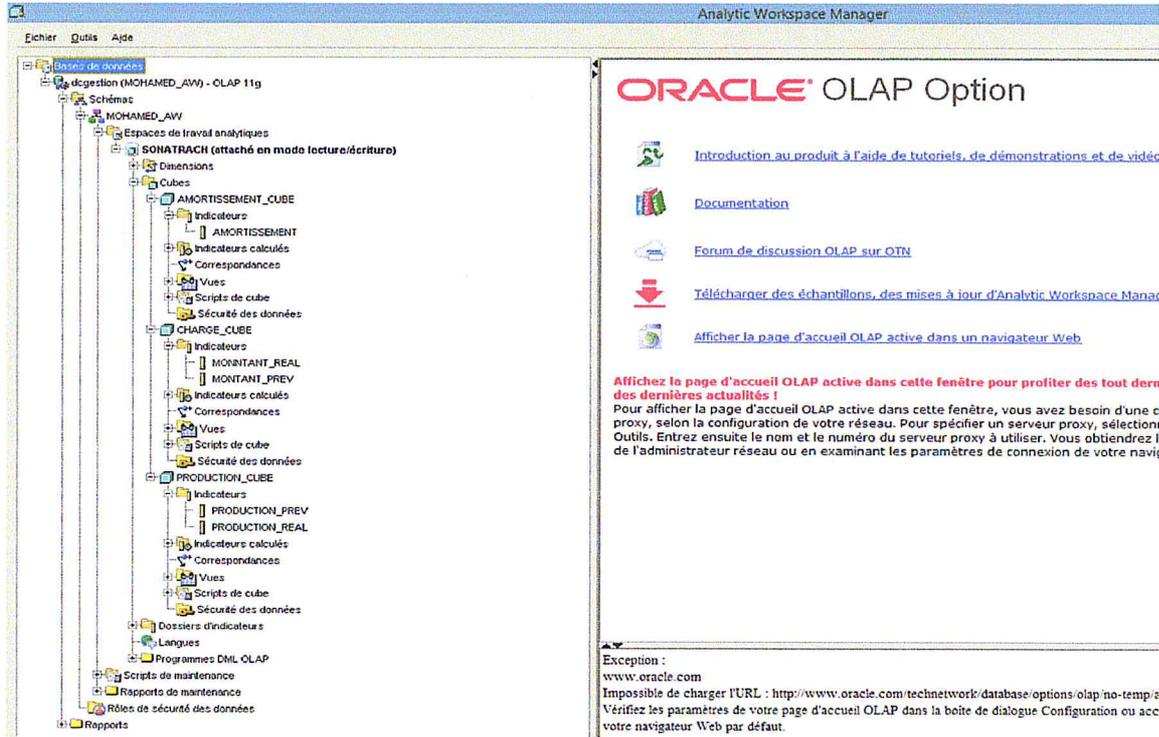


Figure 40 : Construction des cubes avec AWM

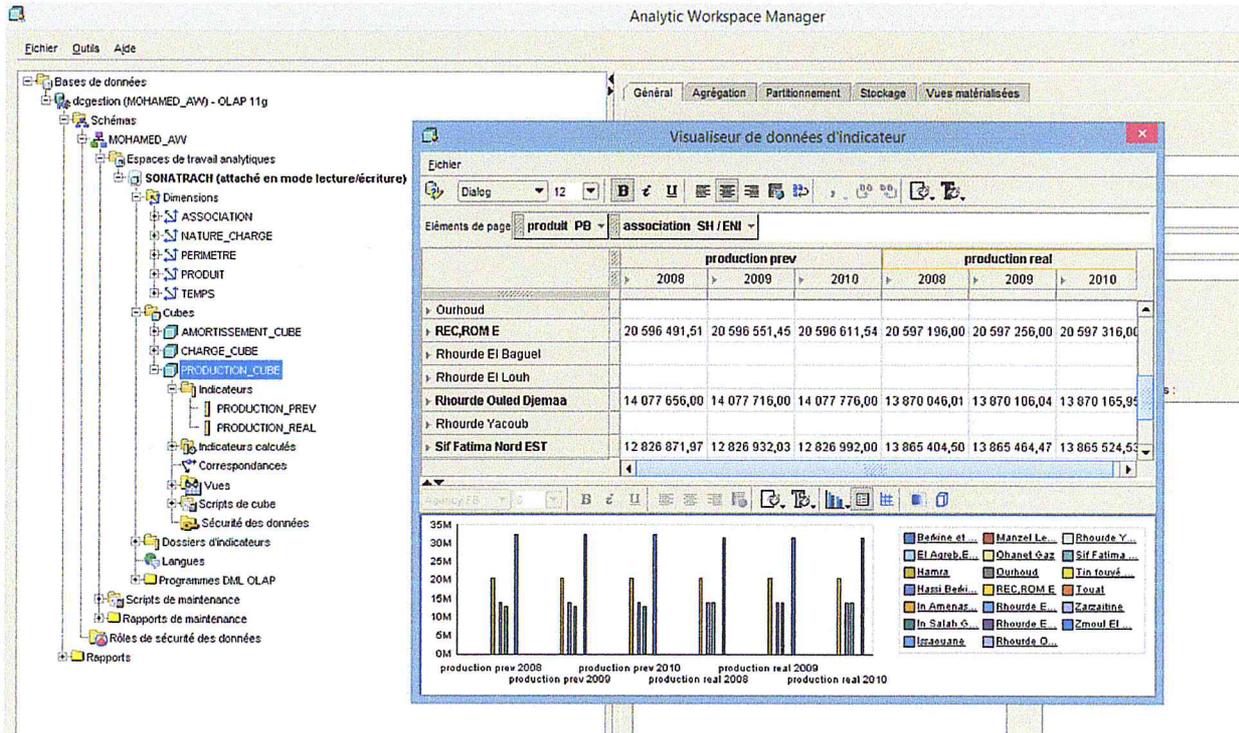


Figure 41 : Exploitation du cube « Production » avec AWM

VIII.2.4. Mise en place du reporting

La conception et la création du rapport se fera au moyen de l’outil BI Publisher, la restitution, se fera dans Oracle BI.

Nous procédons à la création de l’état du rapport, nous commençons par définir l’accès au référentiel du l’entrepôt de données, qui est déjà créé grâce à l’outil BI Administration. Puis nous disposons par « glisser-déposer » les éléments graphiques et les zones de données de manière à obtenir le résultat voulu.

Les résultats seront sous forme de graphes ou de tableaux ou les deux. Nous obtenons alors notre rapport exportable sous différents formats.

SONATRACH
 Direction Coordination Groupe Finance
 Direction Contrôle de Gestion
 Département Gestion Activité "AMONT"

Récapitulatif des coûts opératoires par Nature des charges
 En Milliers de dinars Algériens

		2008		2009		2010	
		Objectif	Réalisation	Objectif	Réalisation	Objectif	Réalisation
SH / ANADARKO (AAC)	SERVICES	3.6405792E7	5.995366776E7	3.6405864E7	5.995373976E7	3.6405936E7	5.995381176E7
	FRAIS FINANCIERS	5922.0	6486.0	5994.0	6558.0	6066.0	6630.0
	MAT.& FOURNITURES CONSOMMEES	7263084.0	5106193.02	7263156.0	5106265.02	7263228.0	5106337.02
	IMPÔT ET TAXES	1451.94	3403.26	1523.94	3475.26	1595.94	3547.26
	FRAIS DIVERS	331002.0	384762.0	331074.0	384834.0	331146.0	384906.0
	Coûts Opérateur	4.400725194E7	6.545451204E7	4.400761194E7	6.545487204E7	4.400797194E7	6.545523204E7
SH / AAC - ENI	SERVICES	1.7355516E7	2.88598272E7	1.7355588E7	2.69236371E7	1.7355654E7	2.69237091E7
	FRAIS FINANCIERS	1236.0	1686.0	1308.0	1758.0	1380.0	1830.0
	MAT.& FOURNITURES CONSOMMEES	3462474.0	2085534.0	3462546.0	2085606.0	3462618.0	2085678.0
	IMPÔT ET TAXES	6657136.26	8164518.0	6657214.26	8164590.0	6657274.26	8164662.0
	FRAIS DIVERS	126414.0	91770.0	126486.0	91842.0	126558.0	91914.0
	Coûts Opérateur	2.760277625999998E7	3.92033352E7	2.760314225999998E7	3.72674331E7	2.760348425999998E7	3.72677931E7

Figure 42 : Rapport sous forme d’un tableau dynamique généré par BI Publisher

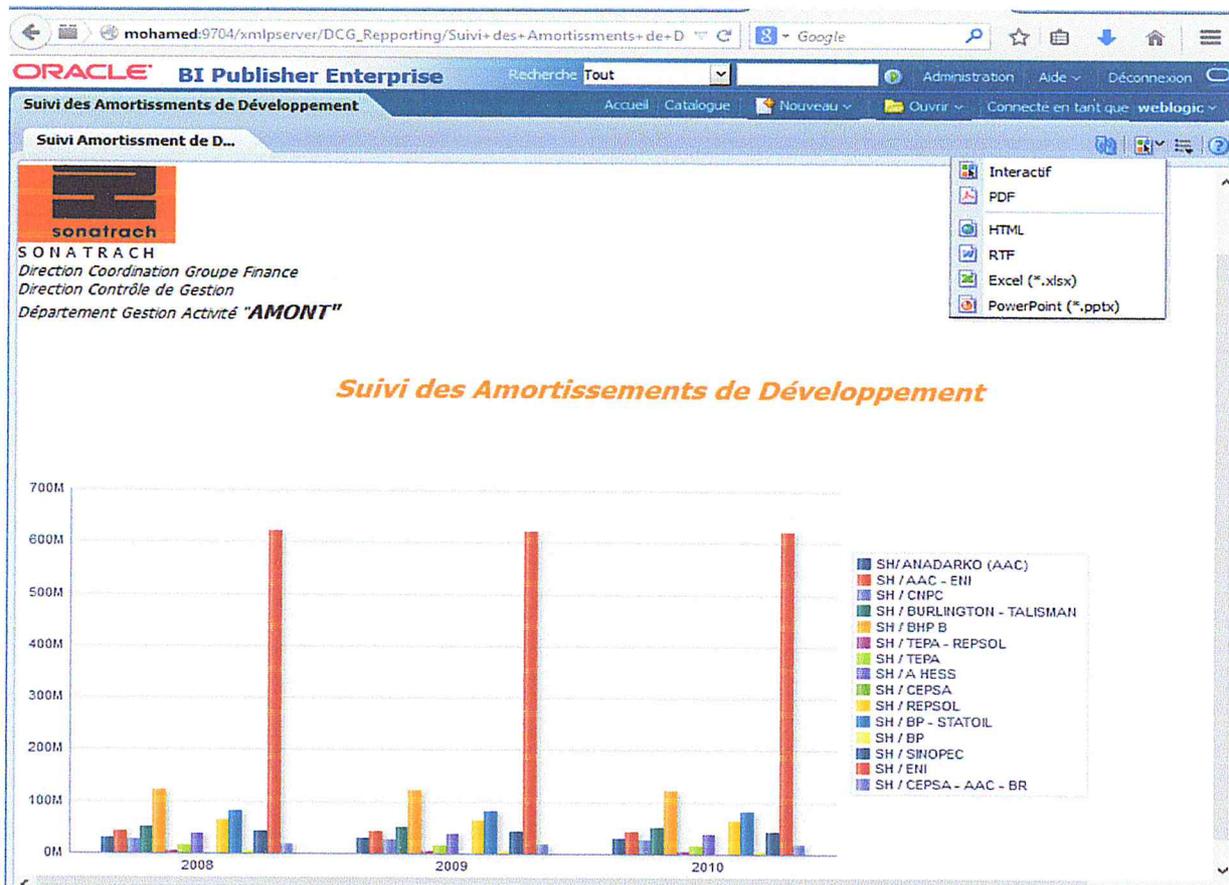


Figure 43 : Rapport sous forme d'un histogramme généré par BI Publisher

VIII.2.5. Tableau de bord

Le tableau de bord de la plateforme OBI offre aux utilisateurs toute une palette de méthodes de visualisation très riche. S'appuyant sur une architecture web, permet de présenter des données de façon intuitive, pertinente et facile à comprendre.

Le tableau de bord que nous avons mis en place offre des listes interactives de contenus analytiques, on trouve :

- Des rapports prédéfinis, sous forme des tableaux croisés dynamiques et des graphes,
- Des listes de contrôle des KPI (Indicateur clé de performance)
- Des analyses OLAP.

VIII.2.5.1. l'accès au tableau de bord

Pour accéder au tableau de bord il suffit juste de faire entrer le nom d'utilisateur et le mot de passe, par la suite une page d'accueil est ouverte.

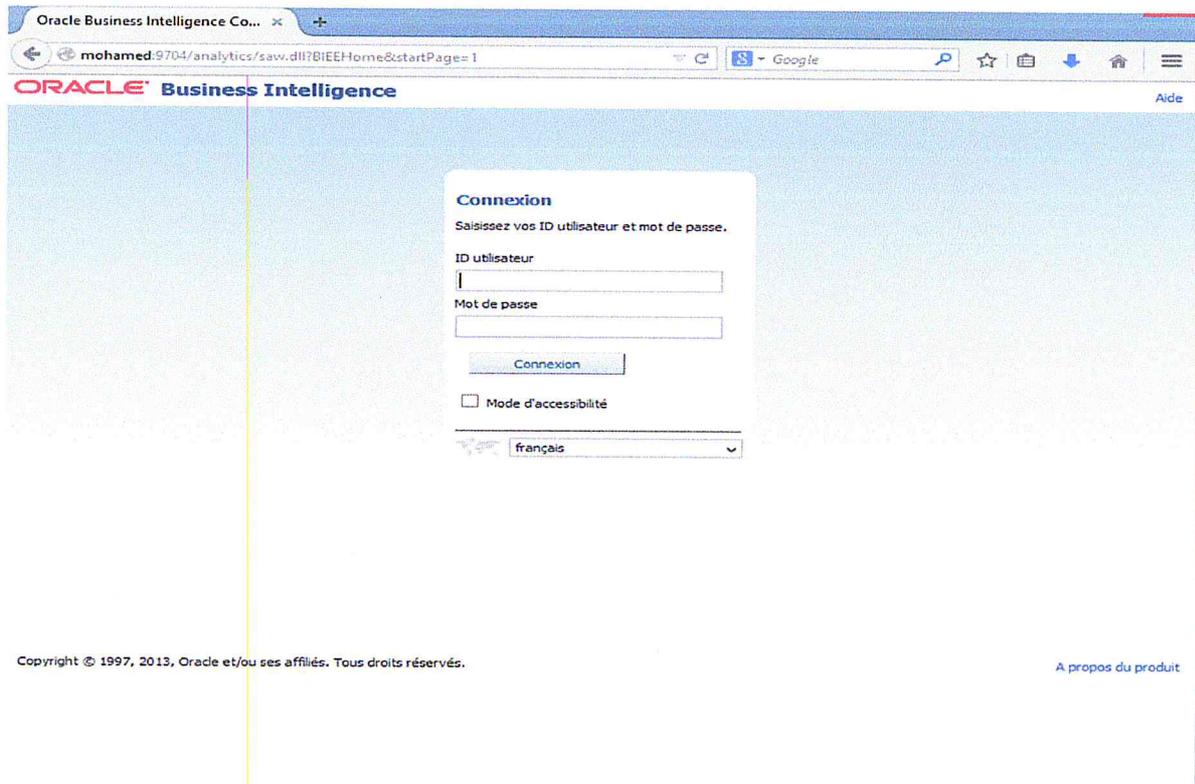


Figure 44 : Page d'accès au Tableau de Bord

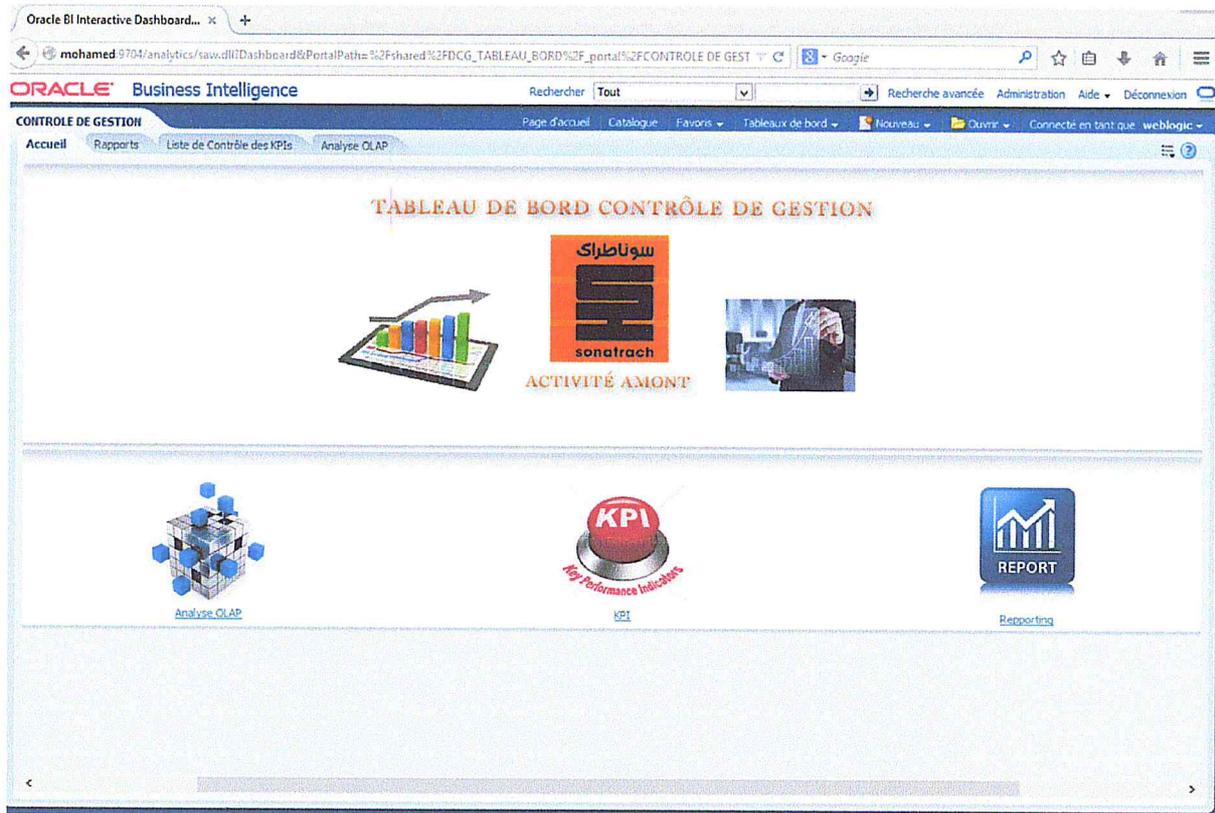


Figure 45 : Page d'accueil du Tableau de Bord

VIII.2.5.2. Visualisation des rapports prédéfinis

La figure suivante présente le rapport « suivi des amortissements » et « suivi des charges »

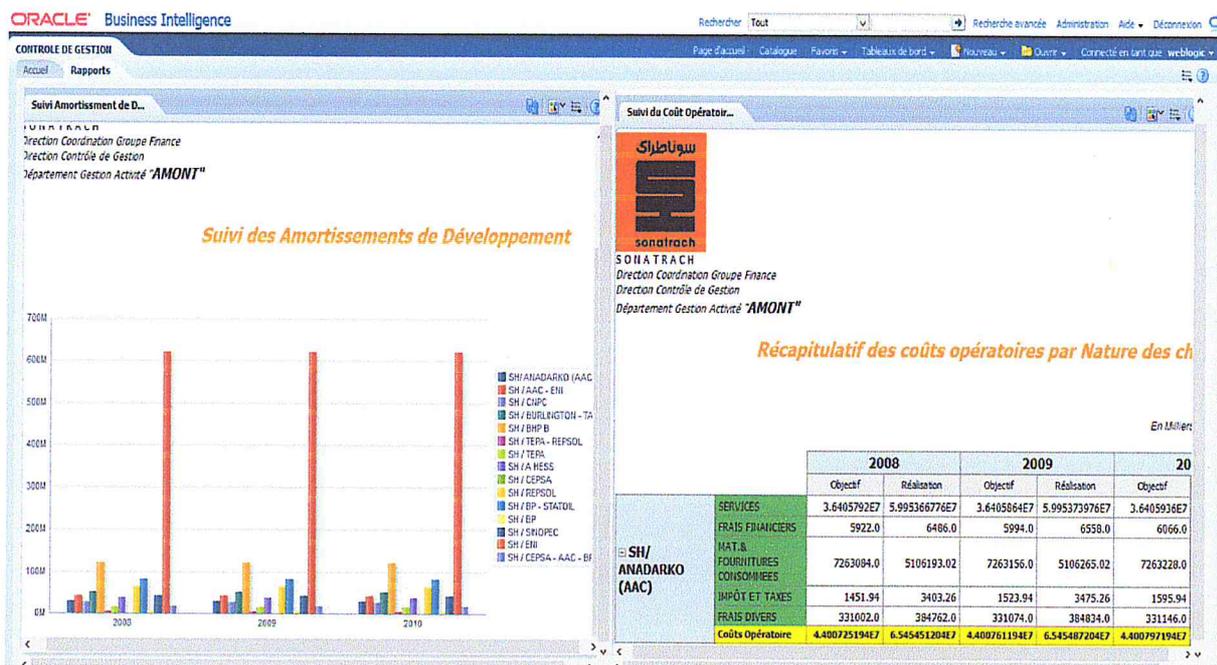


Figure 46 : Exemple de rapport

VIII.3. Utilisateurs du système

Les utilisateurs de notre système décisionnel sont les gens de métier tels que : les analystes, les gestionnaires ... etc., de la direction de contrôle de gestion de la SONATARACH. Ils utilisent de manière poussée le système en exploitant toutes ses possibilités (rapports, tableau de bord...).

VIII.4. la sécurité

- L'accès au système nécessite une authentification et utilise le gestionnaire des utilisateurs d'Oracle BI. Il permet la gestion des groupes d'utilisateur, des utilisateurs, la création des nouveaux rôles et de l'affecter aux groupes. Un utilisateur d'un groupe n'accède donc qu'aux informations correspondant à son rôle assurant.
- Pour assurer la sécurité de l'ensemble des bases de données, nous utilisons la sécurité liée au serveur e la base de données Oracle. Mécanisme d'autorisation et d'authentification permettant de sécuriser l'accès des utilisateurs à la base de données.

Conclusion

L'implémentation et le déploiement du système requiert une connaissance de l'environnement technique et des outils à utiliser, nous avons donc commencé par la présentation de ces derniers. Nous avons ensuite recensé toutes les étapes de la réalisation depuis la construction de l'entrepôt jusqu'à son exploitation en passant par l'ETL et la construction des cubes multidimensionnels. Nous avons également détaillé les différentes possibilités offertes pour l'exploitation de l'entrepôt (cube, rapport, graphe, KPI) car c'est l'aspect qui intéresse en premier lieu l'utilisateur final. Enfin nous terminons avec la gestion des utilisateurs du système et de leur accès aux données, et les mesures garantissant la sécurité de ces données.

Nous avons abouti au terme de ce travail à un résultat conforme aux attentes des utilisateurs, et à un système qui reste extensible et évolutif, ce qui permettra aux développeurs d'ajouter facilement de nouveaux axes d'analyses ou de l'adapter à d'autres volets de l'entreprise.

CONCLUSION GENERALE

Conclusion générale

Les entreprises évoluent dans des environnements complexes et compétitifs dans lesquels l'accès à la bonne information est d'un enjeu capital. Le groupe Sonatrach n'échappe pas à cette règle, c'est pour cela qu'il décide de mettre en place des systèmes décisionnels qui lui permettront de conforter et d'améliorer sa position de leader dans le domaine des hydrocarbures et hydrocarbures gazeux et d'être mieux préparé face à la concurrence. L'un des domaines concernés par un système décisionnel est celui de production. En effet, les dirigeants ont exprimé un besoin important d'informations spécifiques à l'analyse et au pilotage de l'entreprise dans ce domaine.

Nous avons réalisé, dans le cadre de notre projet de fin d'études, un système décisionnel pour le suivi des coûts de production de Sonatrach en association.

Ce système décisionnel s'articule autour d'un entrepôt de données alimenté à partir des fichiers Excel bien structurés. Ce système décisionnel doit répondre aux attentes des utilisateurs, ce qui fait que de la définition des besoins des utilisateurs est une étape cruciale du projet, une mauvaise définition des besoins rendra le système obsolète.

Nous avons donc procédé, en premier lieu, à une étude du système existant et du domaine métier (le contrôle de gestion au sein de Sonatrach) puis à une collecte d'information auprès des futurs utilisateurs, à travers des questionnaires et notamment des interviews. Nous avons opté, tout au long de cette étape pour la méthode mixte qui considère à la fois les besoins des utilisateurs et les données disponibles. A l'issue de cette étape nous avons défini les besoins des utilisateurs et nous les avons répertoriés en trois volets : suivi des charges, suivi des amortissements de développement et suivi de production.

Ensuite nous entamons la conception de la solution, avec la conception de l'entrepôt, à l'aide de la modélisation multidimensionnelle, où nous obtenons des schémas en constellation chaque schéma correspondant à un volet d'analyse. Puis nous abordons la conception de la zone d'alimentation, où nous définissons les détails du processus d'ETL qui assurera que les données stockées dans l'entrepôt soient fiables et mises à jour. Nous enchaînons avec la conception des cubes dimensionnels, ces cubes offrent une structure adaptée à l'interrogation multidimensionnelle et seront exploités pour l'analyse OLAP.

Enfin la dernière étape consiste en la réalisation et le déploiement de la solution. L'utilisation de différents outils, nous a permis d'implémenter les étapes sus citées, pour aboutir à un système décisionnel s'articulant autour d'un entrepôt de données pertinentes et fiables, exploitable de manière facile, rapide et intuitive à l'aide des cubes, des graphes, des rapports et des tableaux de bord. Le système intègre la gestion des utilisateurs qui offre une interface et un accès adapté à chacun, et les différentes mesures de sécurité assurent, la protection des données.

Au final, les utilisateurs possèdent une solution adaptée à leurs besoins et cela constitue l'objectif que nous nous sommes fixé.

Notre système bien que fonctionnel, il offre ainsi les perspectives suivantes :

- Amélioration des fonctionnalités disponibles en recueillant les remarques des utilisateurs,
- Ajout d'autres schémas couvrant d'autres activités (Aval, TRC, COM),
- Utilisation des méthodes et algorithmes de Data Mining (fouille de données) pour des analyses plus poussées.

La réalisation de ce projet nous a été bénéfique à tous les niveaux et nous a permis de découvrir le domaine de l'informatique décisionnelle, domaine vaste et en pleine expansion.

Références bibliographique

(Codd et al., 1993) CODD, EF, CODD, SB et SALLEY, CT, 1993. Providing OLAP (On-Line Analytical Processing) to User-Analysis: An IT Mandate. In: 1993.

(Fernandez, 2012) FERNANDEZ, Alain, 2002, La création et la publication de rapport d'activité. In : *Piloter la performance* [en ligne]. 5 mai 2012.

Disponible à l'adresse :

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

(Gray et al., 1996) GRAY, Jim, BROSWORTH, Adam, LYMAN, Andrew, REICHART, Don et PIRAHESH, Hamid, 1996. Data cube: A relational aggregation operator generalizing group-by, cross-tab, and sub-totals. In : New Orleans : s.n. mars 1996. pp. 152-159.

(GRIM, 2008) GRIM, Yazid, 2008. OLAP, les fondamentaux. In : *Developpez* [en ligne].

Mai 2008. Disponible à l'adresse :

<http://grim.developpez.com/articles/concepts/olap/>.

(Gruau, 2004) GRUAU, Cyril, 2004. *SQL Server 2000, Analysis Services et DTS*.

4 mars 2004.

(Inmon, 2002) INMON, William H., 2002. *Building the Data Warehouse*. Third Edition.

United States of America : WILEY. ISBN 0-471-08130-2.

(Jerbi, 2012) JERBI, Housseem, 2012. *Personnalisation d'analyses décisionnelles sur des données multidimensionnelles* [en ligne]. THESE : Université Toulouse 1 Capitole.

- (Kimball, 2005)** KIMBALL, Ralph, 2005. *Le Data Warehouse : guide de conduite de projet* : Eyrolles. ISBN 9782212116007.
- (Kimball and Ross, 2002)** KIMBALL et ROSS, Margy, 2002. *The Data Warehouse Toolkit*
United States of America: WILEY. ISBN 0-471-20024-7.
- (Kuonen, 2001)** KUONEN, Diego, 2001. *Introduction au Data Mining avec R : Vers la reconquête du « Knowledge Discovery in Database » par les statisticiens*. 2001.
- (Lau et al., 2009)** LAU, Stéphane, SABATIER, Julien et BOUTEILLER, Sophie, 2009. *Business Intelligence. Place de la BI et pilotage des projets décisionnels dans les grandes organisations françaises*. octobre 2009 : CIGREF.
- (Laurent, 2002)** LAURENT, Anne, 2002. *Bases de données multidimensionnelles floues et leur utilisation pour la fouille de données*. Paris : Paris 6.
- (Métais, 2007)** METAIS, Elisabeth, 2007. SYSTÈMES INFORMATIQUES - Systèmes d'aide à la décision. In : *Encyclopædia Universalis* [en ligne]. 2007.
Disponible à l'adresse :
<http://www.universalis.fr/encyclopedie/systemes-informatiques-systemes-d-aide-a-la-decision/>.
- (Naoum, 2006)** NAOUM, Lamiaa, 2006. *Un modèle multidimensionnel pour un processus d'analyse en ligne de résumés flous*. Nante : Université de Nante : Ecole polytechnique
- (Ponniah, 2001)** PONNIAH, Paulraj, 2001a. *Data Warehousing Fundamentals*. United States : WILEY. ISBN 0-471-22162-7.
- (Teste, 2000)** TESTE, Olivier, 2000. *Modélisation et manipulation d'entrepôts de données complexes et historisées* : Université Paul Sabatier - Toulouse III.

(Teste, 2009) TESTE, Olivier, 2009. *Modélisation et manipulation des systèmes OLAP : de l'intégration des documents à l'utilisateur*. HDR : Université Paul.

(Tournier, 2007) TOURNIER, Ronan, 2007. *Analyse en ligne (OLAP) de document* : Université Paul Sabatier - Toulouse III.

(Tranchant, 2012) TRANCHANT, Michael, 2012. Qu'est-ce que l'informatique

décisionnelle ? In : *www.developpez.com* [en ligne]. 27 avril 2012.

Disponible à l'adresse :

<http://mtranchant.developpez.com/tutoriels/Business-Intelligence/qu-est-que-informatique-decisionnelle/>.

