

DÉDICACE

Je dédie ce mémoire a toute ma famille.

REMERCIEMENTS

Dédicace	i
Remerciements	ii
Sommaire	iv
Abréviations	v
Liste des figures	vi
Liste des tableaux	vii
Résumé	viii
Abstract	ix
Introduction	1
I État de l'art	2
1 Synthèse Bibliographique	3
1.1 Business Intelligence	3
1.2 Le Data Warehouse	5
1.3 Modélisation des données de l'entrepôt	7
1.4 Démarche de Construction d'un Data Warehouse	10
2 LES SYSTÈMES DE GESTION HÔTELIÈRE	16
2.1 LA GESTION HÔTELIÈRE ACTUELLE	16
2.2 PRÉSENTATION TECHNIQUE DE HOSTELINE	17
2.3 LE SYSTEME OPERATIONNEL ACTUEL	25
2.4 LA PROBLÉMATIQUE DU THÈME	27

2.5 SOLUTIONS PROPOSÉES	28
II Contribution	29
3 CAHIER DES CHARGES	30
3.1 EXPRESSION DES BESOINS	30
3.2 CONTRAINTES LIÉES AU PROJET	31
3.3 DÉMARCHE DE RÉALISATION	33
4 MISE EN ŒUVRE DE LA SOLUTION	35
4.1 CHOIX DES OUTILS	35
4.2 ANALYSE CONCEPTION	38
4.3 CHOIX DES OUTILS	38
4.4 IMPLÉMENTATION	38
Conclusion	39
Références	41
A titre annexe	42
A.1 Présentation de INSBI	42
A.2 DIAGRAMMES UML	43
Table des matières	46

ABRÉVIATIONS

API : Appliscation programing Interface **BI** : Business Intelligence.
BD : Database **DBA** : Data Base Administrator.
DIM : Dimension.
DW : Data Warehouse (Entrepôt de données).
EDW : Entreprise Data Warehouse.
ETL : Extract, Transform and Load (ETC).
ERD : Entity Relation Diagram (Diagramme d'entité association) **FK** : Foreign Key.
FTP : File Transfer Protocol.
HOLAP : Hybrid OnLine Analytical Process.
INSBI : Institute Business Intelligence
MOLAP : Multidimensional On Line Analytical Process.
OLAP : On Line Analytical Process.
OLTP : On Line Transactional Process.
PK : Primary Key.
PDI : Pentaho Data Integrator **ROLAP** : Relational On Line Analytical Process.
SI : Systèmes d'Information.
SID : Systèmes d'Information Décisionnels.
SID : Systèmes d'information de la distribution.
SIAD : Systèmes d'Information d'Aide à la Décision.
SGBD : Système de Gestion de Base de Données.
SMTP : Server Mail Transfer Protocol.
SQL : Structured Query Language.

LISTE DES FIGURES

1.1	Architecture generale d'un data warehouse	6
1.2	A cube with three dimensions.[Rainardi 2008].	7
1.3	Exemple d'un modèle en étoile « Dimensional star schema for campaign results data mart »[Rainardi 2008]	8
1.4	Slicing	9
1.5	Dicing	10
1.6	illustration de l'approche « Besoins d'analyse » grâce au cycle de vie dimensionnel de Kimball [Kimball, 2004].	11
1.7	Illustration de l'approche mixte.	13
2.1	Formalismes du diagramme de cas d'utilisation.	19
2.2	Diagramme de cas d'utilisation contenant deux des acteurs de Hosteline (Agrandi en Annexe XX)	20
2.3	Diagramme de cas d'utilisation contenant l'acteur "Client" (Agrandi en Annexe XX)	21
2.4	Formalismes du diagramme de classes.	22
2.5	Diagramme de classes du système (Agrandi en Annexe XX)	23
2.6	ERD du système actuel (Agrandi en Annexe XX)	24
2.7	Interface d'inscription a Hosteline	25
2.8	Interface de réservations de Hosteline	26
4.1	Problèmes et outils BI réponses.	35
4.2	Regroupement des fournisseurs de solutions BI	36
4.3	La suite Pentaho.	38

LISTE DES TABLEAUX

1.1	Résumé des différences entre le Transactionnel et décisionnel.	4
1.2	Avantages et inconvénients de l'approche « Besoins d'analyse »	11
1.3	Avantages et inconvénients de l'approche « Besoins d'analyse »	11

La startup INSBI (Institut Business Intelligence) est une entreprise nouvellement créée qui dans son (Service qu'on propose) a une liste de service qui tourne autour de solutions de digitalisation d'entreprises et Business intelligence. Dans le souci de se faire un nom sur le marché INSBI a produit et lancé son premier produit qui est Hosteline. Hosteline est une plateforme d'hôtellerie en ligne qui a pour cible les propriétaires d'établissements hôteliers mettent leurs locaux à la disposition des clients qui à leur tour peuvent réserver ces différents locaux. Hosteline est un portail web qui fédère les hôtels et donne un accès comparatif aux clients de cette plateforme. A travers cette plateforme les clients d'hôtels auront des avantages tant sur les prix que sur les facilités que pourront offrir ces établissements hôtelier grâce à son programme de fidélité. Ce système lorsqu'il sera mis en production générera une quantité importante de données par son système transactionnel. Sachant à quel point ces données sont utiles mais n'informent pas suffisamment à l'état brut, la direction de INSBI souhaite mettre en place un système décisionnel pour pouvoir exploiter ces données et en tirer les choix et décisions stratégiques pour faire grandir le produit en offrant une meilleure qualité de service à ses consommateurs. C'est là que nous intervenons afin de mettre en place un système de reporting qui servira les données issues des systèmes opérationnels dans des formats lisibles et facilement interprétables. Notre travail consistait à répondre à un besoin précis consigné dans un cahier avec des contraintes de coûts et de délais bien définies. Dans ce mémoire nous mettrons en relief, la méthode de travail, l'évolution des travaux à INSBI et le suivi qu'était le nôtre pendant toute la durée de notre stage. Ce travail nous a permis non seulement de mettre en pratique des enseignements reçus pendant notre formation, d'en apprendre davantage sur les méthodes de fonctionnement des entreprises et aussi sur des notions nouvelles.

ABSTRACT

The Startup INSBI (Institute Business Intelligence) is a newly created company which in its offer has a list of service that revolves around solutions of digitalization of companies and business intelligence. In order to make a name for itself in the market INSBI has produced and launched its first product which is Hosteline. Hosteline is an online hospitality platform that targets hotel owners and places their premises at the disposal of customers who in turn can book these different premises. Hosteline is a web portal that brings together hotels and provides comparative access to customers of this platform. Through this platform, hotel customers will have advantages both in terms of prices and the facilities that these hotels can offer thanks to its loyalty program. This system when put into production will generate a large amount of data through its transactional system. Knowing how useful these data are but not enough in the raw state, the INSBI management wants to set up a decision-making system to be able to exploit this data and to draw the strategic choices and decisions to make the product grow offering a better quality of service to its consumers. These is where we came into the game to set up a Reporting system that will serve data from operational systems in readable and easily interpretable formats. Our job was to give satisfaction to a specific need recorded in a specifications book with well-defined cost and time constraints. In this thesis we will highlight, the working method, the evolution of the work at INSBI and the follow-up that was ours throughout the duration of our internship. Thanks to this project we could put into practice lessons learned during our training and it permitted us to learn more about business intelligence and many other new concepts.

L'avènement des TIC (Technologies de l'information) vient changer notre façon de vivre en la rendant plus simple. Le secteur hôtelier n'est pas exclu de ces changements et proposer ses services sur internet n'est plus un luxe mais un besoin fondamental pour toute entreprise hôtelière qui souhaite rester compétitive. Sur la toile les sites web d'hôtels se multiplient et tout particulier ou entreprise cherchent un ou plusieurs locaux en ligne à une large gamme de sites web d'hôtels à sa portée. Le problème quand les possibilités sont larges et non exhaustives est que on n'a pas toujours le temps de visiter ces plateformes une par une et comparer leurs offres selon nos préférences. Des plateformes telles que Booking.com, Accordhotel.com et autres offrent une solution qui répond à ce problème de comparaison de prix et d'offres toujours est-il que le contexte et les particularités camerounaises et Africaines n'est pas toujours retrouvée. En effet sur les plateformes de ce type on retrouve des hôtels conventionnels qui respectent certaines normes et standards occidentaux. A titre d'exemple le ministère du tourisme reconnaît officiellement 250 hôtels accrédités au Cameroun ce qui ne représente pas 30% du nombre réel d'établissements hôtelier du pays. Parmi les hébergeurs laissés on retrouve les propriétaires d'auberges, certaines résidences hôtelières, Appartements meublés, Villa et plein d'autres encore. Hosteline s'inspire des plateformes citées plus haut en adaptant le concept au milieu et aux coutumes Camerounaises. En effet Hosteline est une plateforme ouverte à tout type d'hébergeurs et hôteliers qui souhaitent se mettre à la disposition des clients et potentiels clients par le web.

Lorsque la plateforme sera lancée et tournera à plein régime la quantité de données cumulée sur une année sera énorme et en exploitant la plateforme sur plusieurs années encore ne produira qu'une masse de données conséquente. Pour anticiper sur les besoins de s'informer grâce à ces données qui iront grandissant, Mr Kendjio le fondateur de Hosteline souhaite se munir d'un système décisionnel pour être prêt à exploiter et tirer avantage de ce volume important de données. La mise en œuvre de ce système décisionnel est le but de la rédaction du présent mémoire. **Il faut finir les parties du document**

Première partie

État de l'art

CHAPITRE 1

SYNTHÈSE BIBLIOGRAPHIQUE

1.1 Business Intelligence

1.1.1 Introduction

Dans le monde les grandes entreprises dans leurs activités journalières produisent une quantité importante de données qui a long terme devient en quantité astronomiques. Le besoin d'exploiter ces données a des fin de pilotage des entreprises a fait naitre le Business Intelligence encore appelé Intelligence Economique ou encore Informatique décisionnel qui permet d'étudier l'environnement de l'entreprise aux moyens des données qu'elle possède. Les données étant en grande quantité, il faut le nettoyer, les structures avant de les stocker dans ce qu'on a appelé Data Warehouse (Entrepôt de données en français).

Le concept de Data Warehouse, tel que connu aujourd'hui, est apparu pour la première fois en 1980 ; l'idée consistait alors à réaliser une base de données destinée exclusivement au processus décisionnel. Les nouveaux besoins de l'entreprise, les quantités importantes de données produites par les systèmes opérationnels et l'apparition des technologies aptes à sa mise en œuvre ont contribué à l'apparition du concept « Data Warehouse » comme support aux systèmes décisionnels.

1.1.2 Les systèmes décisionnels

L'entrepôt de données est au centre du système décisionnel et sa raison d'être est la mise en place de ces systèmes décisionnels. Nous allons ici rappeler quelques définitions qui serviront ont mieux expliqué la suite.

Selon Jean-Louis Le Moigne, « Le système d'information est l'ensemble des méthodes et moyens de recueil de contrôle et de distribution des informations nécessaires à l'exercice de l'activité en tout point de l'organisation. Il a pour fonction de produire et de mémoriser les

informations, de l'activité du système opérant (système opérationnel), puis de les mettre à disposition du système de décision (système de pilotage) » [Le Moigne 1977].

Selon Wikipédia, « L'informatique décisionnelle est l'informatique à l'usage des décideurs et des dirigeants d'entreprises ».

De ces définition on retient que le système opérationnel et le système décisionnel sont des parties du système d'information.

TABLE 1.1: Résumé des différences entre le Transactionnel
et décisionnel.

Critere	Transact	discisionel
par les données	<p>Orienté applications</p> <p>Situation instantanée</p> <p>Donnée détaillées et codées non redondantes</p> <p>Données changeantes constamment</p>	<p>Orienté thèmes et sujets</p> <p>Situation historique</p> <p>Informations agrégées cohérentes souvent avec redondance</p> <p>Informations stables et synchronisées dans le temps</p>
L'usage	<p>Assure l'activité au quotidien</p> <p>Pour les opérationnels</p> <p>Mises à jour et requêtes simples</p> <p>Temps de réponse immédiats</p>	<p>Permet l'analyse et la prise de décision</p> <p>Pour les décideurs</p> <p>Lecture unique et requêtes complexes transparentes</p> <p>Temps de réponse moins critiques</p>

Ces différences font ressortir la nécessité de mettre en place un système répondant aux besoins décisionnels. Ce système n'est rien d'autre que le « Data Warehouse ».

1.2 Le Data Warehouse

1.2.1 Définition

Bill Inmon définit le Data Warehouse, dans son livre considéré comme étant la référence dans le domaine “Building the Data Warehouse” [Inmon, 2002] comme suit :

« Le Data Warehouse est une collection de données orientées sujet, intégrées, non volatiles et évolutives dans le temps, organisées pour le support d'un processus d'aide à la décision. »

Les paragraphes suivants illustrent les caractéristiques citées dans la définition d'Inmon.

Orienté sujet : le Data Warehouse est organisé autour des sujets majeurs de l'entreprise, contrairement à l'approche transactionnelle utilisée dans les systèmes opérationnels, qui sont conçus autour d'applications et de fonctions telles que : cartes bancaires, solvabilité client... , les Data Warehouse sont organisés autour de sujets majeurs de l'entreprise tels que : clientèle, ventes, produits... Cette organisation affecte forcément la conception et l'implémentation des données contenues dans le Data Warehouse. Le contenu en données et en relations entre elles diffère aussi. Dans un système opérationnel, les données sont essentiellement destinées à satisfaire un processus fonctionnel et obéit à des règles de gestion, alors que celles d'un Data Warehouse sont destinées à un processus analytique.

Intégrée : le Data Warehouse va intégrer des données en provenance de différentes sources. Cela nécessite la gestion de toute incohérence.

Evolutives dans le temps : Dans un système décisionnel il est important de conserver les différentes valeurs d'une donnée, cela permet les comparaisons et le suivi de l'évolution des valeurs dans le temps, alors que dans un système opérationnel la valeur d'une donnée est simplement mise à jour. Dans un Data Warehouse chaque valeur est associée à un moment « Every key structure in the data warehouse contains - implicitly or explicitly -an element of time » [Inmon, 2000].

Non volatiles : c'est ce qui est, en quelque sorte la conséquence de l'historisation décrite précédemment. Une donnée dans un environnement opérationnel peut être mise à jour ou supprimée, de telles opérations n'existent pas dans un environnement Data Warehouse. Organisées pour le support d'un processus d'aide à la décision : Les données du Data Warehouse sont organisées de manière à permettre l'exécution des processus d'aide à la décision (Reporting, Data Mining...).

1.2.2 Historique

L'origine du concept « Data Warehouse » D.W (entrepôt de données en français) remonte aux années 80, durant lesquelles un intérêt croissant au système décisionnel a vu le jour, dû essentiellement à l'émergence des SGBD relationnel et la simplicité du modèle relationnel et la puissance offerte par le langage SQL, au début, le Data Warehouse n'était rien d'autre qu'une copie des données du système opérationnel prise de façon périodique, dédiée à un environnement de support à la prise de décision. Ainsi, les données étaient extraites du système opérationnel, stockées dans une nouvelle base de données «concept d'infocentre », le motif principal étant

de répondre aux requêtes des décideurs sans pour autant altérer les performances des systèmes opérationnels. Le Data Warehouse, tel qu'on le connaît actuellement, n'est plus vu comme une copie ou un cumul de copies prises de façon périodique- des données du système opérationnel. Il est devenu une nouvelle source d'information, alimenté avec des données recueillies et consolidées des différentes sources internes et externes.

1.2.3 ARCHITECTURE D'UN DATA WAREHOUSE

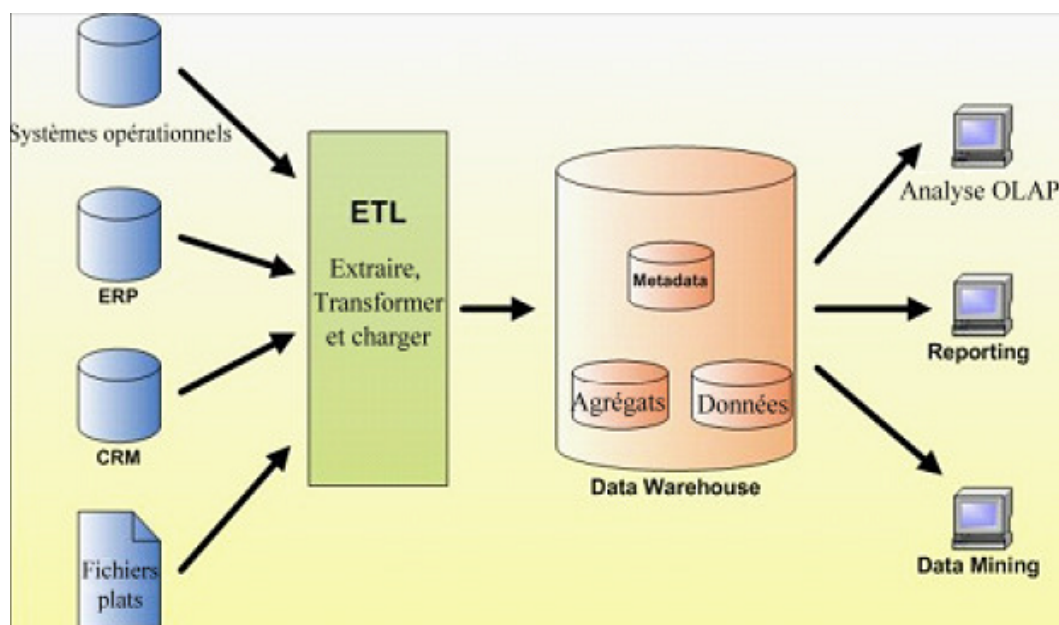


FIGURE 1.1 – Architecture generale d'un data warehouse

La figure ci-dessus illustre la forme générale d'un data Warehouse que nous allons détailler dans les paragraphes suivants.

Les Sources de données : Dans la figure, les représentations de systèmes opérationnel, ERP, CRM et Fichiers plat font office de source de données et c'est d'elles qu'on puisse les données pour alimenter la machine décisionnelle.

ETL (Extract, Transform, Load) : C'est un ensemble de méthodes et d'outils qui servent à :

- Extract : Extraire les données de sources hétérogènes
- Transform : Transformation des données pour les mettre dans un format acceptable
- Load : Charger les données dans le data warehouse

Data Warehouse : L'unité de stockage des données. Il est constitué de plusieurs éléments dont :

- **Meta données :** ce sont les informations relatives à la structure des données, les méthodes d'agrégation et le lien entre les données opérationnelles et celles du Data Warehouse. Les métadonnées doivent renseigner sur :

Le modèle de données,

La structure des données telle qu'elle est vue par les développeurs,

La structure des données telle qu'elle est vue par les utilisateurs,

Les sources des données,

Les transformations nécessaires,

Suivi des alimentations,

- **Les Agrégats** (Données agrégées) : données agrégées à partir des données détaillées.

Les derniers éléments de la figure font partie de la phase d'exploitation du data warehouse et seront détaillés plus bas dans le document.

1.3 Modélisation des données de l'entrepôt

1.3.1 La modélisation dimensionnelle et ses concepts

Les Data Warehouse sont destinés à la mise en place de systèmes décisionnels. Ces systèmes, devant répondre à des objectifs différents des systèmes transactionnels, ont fait ressortir très vite la nécessité de recourir à un modèle de données simplifié et aisément compréhensible. La modélisation dimensionnelle permet cela. Elle consiste à considérer un sujet d'analyse comme un cube à plusieurs dimensions, offrant des vues en tranches ou des analyses selon différents axes.

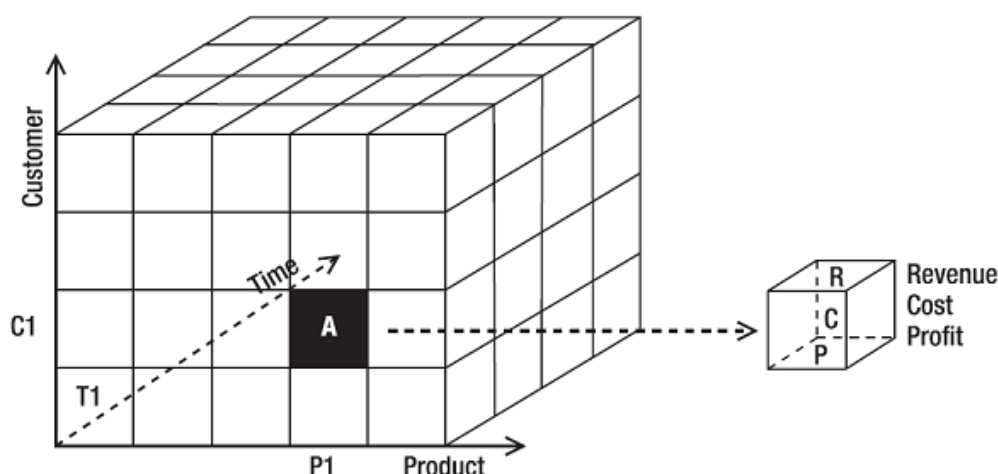


FIGURE 1.2 – A cube with three dimensions.[Rainardi 2008].

Le concept des Faits

Une table de faits est la table centrale d'un modèle dimensionnel, où les mesures de performances sont stockées. Une ligne d'une table de faits correspond à une mesure. Ces mesures

sont généralement des valeurs numériques, additives ; cependant des mesures textuelles peuvent exister mais sont rares. Le concepteur doit faire son possible pour faire des mesures textuelles des dimensions, car elles peuvent êtres corrélées efficacement avec les autres attributs textuels de dimensions.

Le concept des Dimensions

Les tables de dimension sont les tables qui raccompagnent une table de faits, elles contiennent les descriptions textuelles de l'activité. Une table de dimension est constituée de nombreuses colonnes qui décrivent une ligne. C'est grâce à cette table que l'entrepôt de données est compréhensible et utilisable ; elles permettent des analyses en tranches et en dés. Une dimension est généralement constituée : d'une clé artificielle, une clé naturelle et des attributs.

1.3.2 Différents modèles de la modélisation dimensionnelle

Modèle en étoile : comme indiqué précédemment, ce modèle se présente comme une étoile dont le centre n'est autre que la table des faits et les branches sont les tables de dimension. La force de ce type de modélisation est sa lisibilité et sa performance.

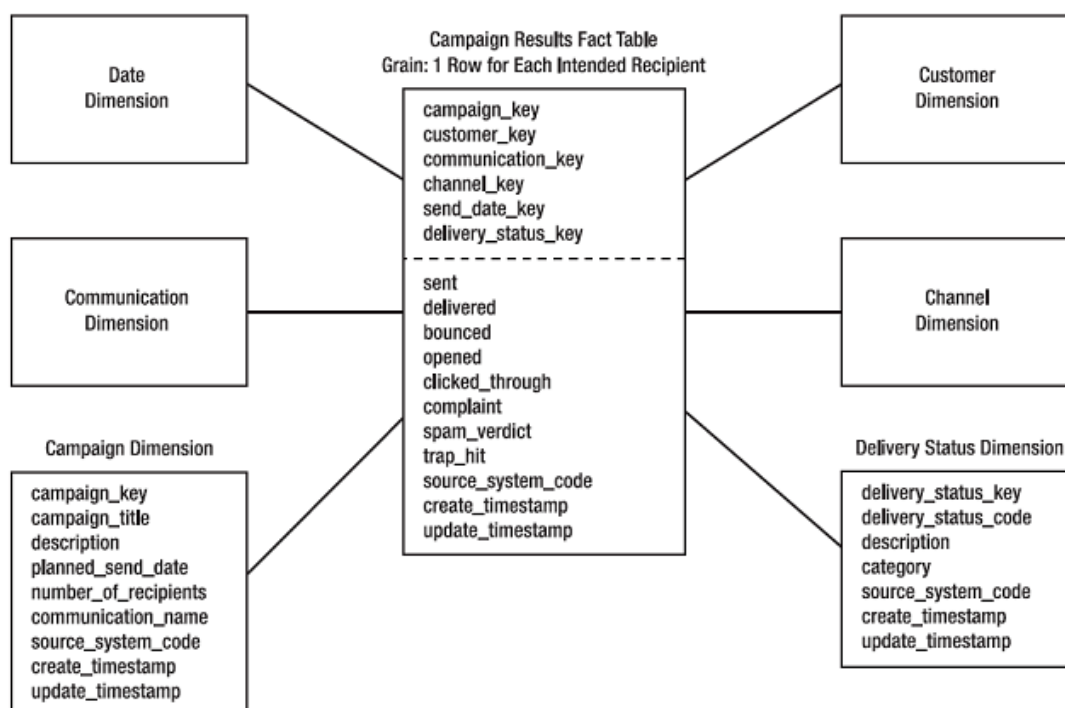


FIGURE 1.3 – Exemple d'un modèle en étoile « Dimensional star schema for campaign results data mart » [Rainardi 2008]

Modèle en flocon : identique au modèle en étoile, sauf que ses branches sont éclatées en hiérarchies. Cette modélisation est généralement justifiée par l'économie d'espace de stockage, cependant elle peut s'avérer moins compréhensible pour l'utilisateur final, et très coûteux en termes de performances.

Modèle en constellation : Ce n'est rien d'autre que plusieurs modèles en étoile liés entre eux par des dimensions communes.

1.3.3 La navigation dans les données

Une fois que le serveur OLAP a construit le cube multidimensionnel « ou simulé ce cube selon l'architecture du serveur », plusieurs opérations sont possibles sur ce dernier offrant ainsi la possibilité de naviguer dans les données qui le constituent. Ces opérations de navigation « Data Surfing » doivent être, d'une part, assez complexes pour adresser l'ensemble des données et, d'autre part, assez simples afin de permettre à l'utilisateur de circuler de manière libre et intuitive dans le modèle dimensionnel.

Afin de répondre à ces attentes, un ensemble de mécanismes est exploité, permettant une navigation par rapport à la dimension et par rapport à la granularité d'une dimension.

Slice & Dice

Le « Slicing » et le « Dicing » sont des techniques qui offrent la possibilité de faire des tranches « trancher » dans les données par rapport à des filtres de dimension bien précis, se classant de fait comme des opérations liées à la structure « se font sur les dimensions ». La différence entre eux se manifestent dans le fait que :

Slicing is the process of retrieving a block of data from a cube by filtering on one dimension [Rainardi 2008].

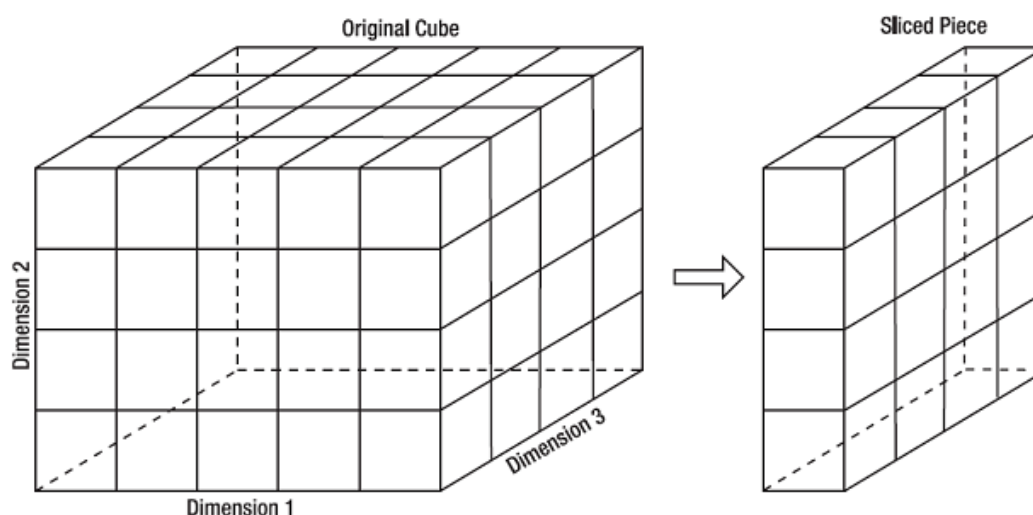


FIGURE 1.4 – Slicing

Dicing is the process of retrieving a block of data from a cube by filtering on all dimensions [Rainardi 2008].

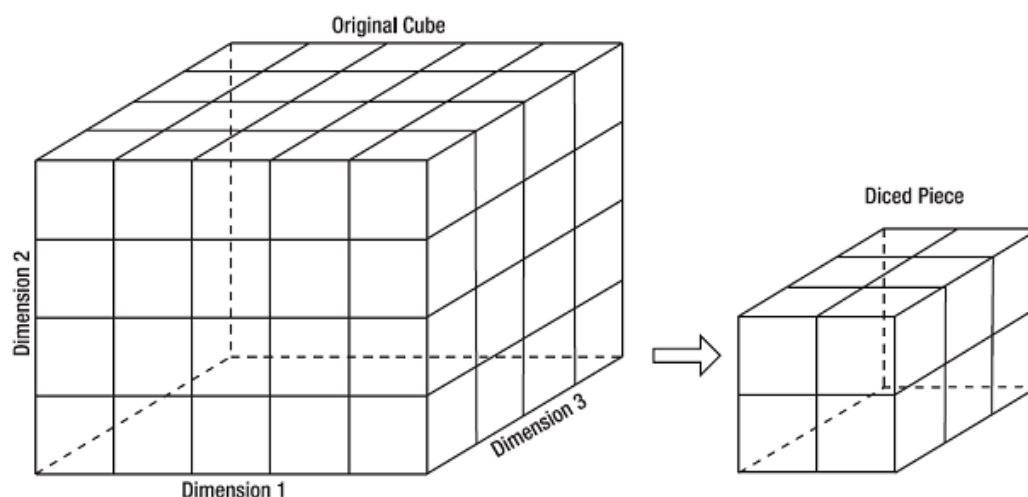


FIGURE 1.5 – Dicing

1.4 Démarche de Construction d'un Data Warehouse

Plusieurs chercheurs ou équipes de recherche ont essayé de proposer des démarches pour la réalisation d'un projet Data Warehouse, ces démarches se croisent essentiellement dans les étapes suivantes :

- Modélisation et conception du Data Warehouse,
- Alimentation du Data Warehouse,
- Mise en œuvre du Data Warehouse,
- Administration et maintenance du Data Warehouse,

1.4.1 Modélisation et conception du Data Warehouse

Les deux approches les plus connues dans la conception des Data Warehouse sont :

- L'approche basée sur les besoins d'analyse,
- L'approche basée sur les sources de données,

Aucune des deux approches citées n'est ni parfaite, ni applicable à tous les cas. Toutes deux doivent être étudiées pour choisir celle qui s'adapte le mieux à notre cas. Quel que soit l'approche adoptée pour la conception d'un Data Warehouse, la définition de celui-là reste la même. En étant un support d'aide à la décision, le Data Warehouse se base sur une architecture dimensionnelle.

Approche « Besoins d'analyse »

Le contenu du Data Warehouse sera déterminé selon les besoins de l'utilisateur final. Cette approche est aussi appelée « approche descendante » (Top-Down Approach) et est illustrée par R. Kimball grâce à son cycle de vie dimensionnel comme suit :

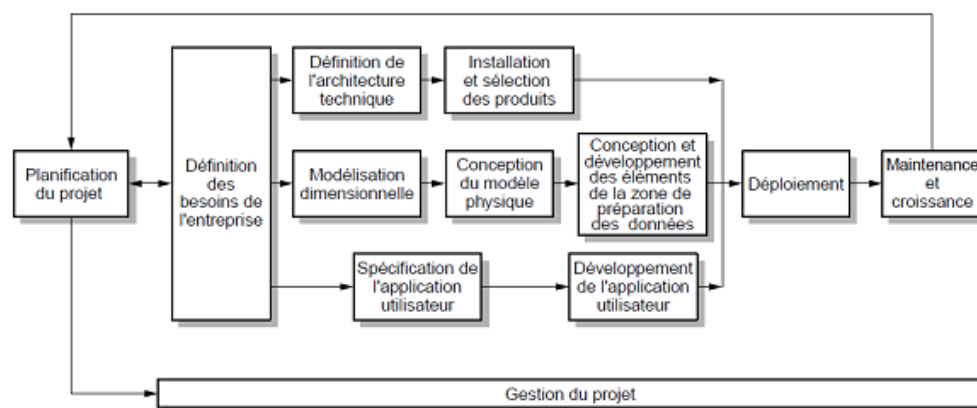


FIGURE 1.6 – illustration de l'approche « Besoins d'analyse » grâce au cycle de vie dimensionnel de Kimball [Kimball, 2004].

TABLE 1.2: Avantages et inconvénients de l'approche « Besoins d'analyse »

Avantages	Inconvénients
Aucun risque de concevoir une solution obsolète avant d'être opérationnelle	<p>Pas de prise en compte de l'évolution des besoins de l'utilisateur.</p> <p>Nécessite une modification de la structure du Data Warehouse en cas de nouveau besoin</p> <p>Négligence du système opérationnel</p> <p>Difficulté de déterminer les besoins des utilisateurs</p>

Approche « Source de données »

Le contenu du Data Warehouse est déterminé selon les sources de données. Cette approche est appelée : Approche ascendante (Bottom-up Approach).

TABLE 1.3: Avantages et inconvénients de l'approche «
Besoins d'analyse »

Avantages	Inconvénients
Meilleure prise en charge de l'évolution des besoins	Risque de concevoir une solution obsolète avant qu'elle soit opérationnelle Evolution du schéma des données source Complexité de source de données

- Inmon considère que l'utilisateur ne peut jamais déterminer ses besoins dès le départ, « Donnez-moi ce que je vous demande, et je vous dirai ce dont j'ai vraiment besoin », il considère que les besoins sont en constante évolution.

Approche mixte

Une combinaison des deux approches appelée hybride ou mixte peut s'avérer efficace. Elle prend en considération les sources de données et les besoins des utilisateurs.

Cette approche consiste à construire des schémas dimensionnels à partir des structures des données du système opérationnel, et les valider par rapport aux besoins analytiques. Cette approche cumule les avantages et quelques inconvénients des deux approches déjà citées, telles que la complexité des sources de données et la difficulté quant à la détermination des besoins analytiques.

1.4.2 Les phases de l'alimentation « E.T.L »

Les phases du processus E.T.L. représentent la mécanique d'alimentation du Data Warehouse. Ainsi elles se déroulent comme suit :

a) L'extraction des données

« L'extraction est la première étape du processus d'apport de données à l'entrepôt de données. Extraire, cela veut dire lire et interpréter les données sources et les copier dans la zone de préparation en vue de manipulations ultérieures. » [Kimball, 2005]. Elle consiste en :

- Cibler les données,
- Appliquer les filtres nécessaires,
- Définir la fréquence de chargement

Lors du chargement des données, il faut extraire les nouvelles données ainsi que les changements intervenus sur ces données. Pour cela, il existe trois stratégies de capture de changement :

- Colonnes d'audit : la colonne d'audit, est une colonne qui enregistre la date d'insertion ou du dernier changement d'un enregistrement. Cette colonne est mise à jour soit par des triggers ou par les applications opérationnelles, d'où la nécessité de vérifier leur fiabilité.
- Capture des logs : certains outils ETL utilisent les fichiers logs des systèmes sources afin de détecter les changements (généralement logs du SGBD). En plus de l'absence de cette fonctionnalité sur certains outils ETL du marché, l'effacement des fichiers logs engendre la perte de toute information relative aux transactions.
- Comparaison avec le dernier chargement : le processus d'extraction sauvegarde des copies des chargements antérieurs, de manière à procéder à une comparaison lors de chaque nouvelle extraction. Il est impossible de rater un nouvel enregistrement avec cette méthode.

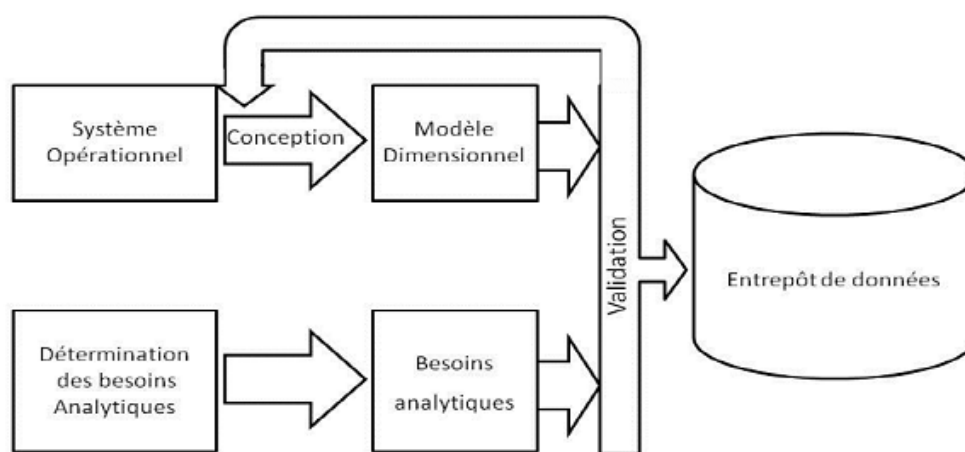


FIGURE 1.7 – Illustration de l'approche mixte.

b) La transformation des données

La transformation est la seconde phase du processus. Cette étape, qui du reste est très importante, assure en réalité plusieurs tâches qui garantissent la fiabilité des données et leurs qualités. Ces tâches sont :

- Consolidation des données.
- Correction des données et élimination de toute ambiguïté.
- Élimination des données redondantes.
- Compléter et renseigner les valeurs manquantes. Cette opération se solde par la production d'informations dignes d'intérêt pour l'entreprise et de et sont donc prêtes à être entreposées.

c) Le chargement des données

C'est la dernière phase de l'alimentation d'un entrepôt de données, le chargement est une étape indispensable. Elle reste toutefois très délicate et exige une certaine connaissance des structures du système de gestion de la base de données (tables et index) afin d'optimiser au mieux le processus.

Politiques de l'alimentation

Le processus de l'alimentation peut se faire de différentes manières. Le choix de la politique de chargement dépend des sources : disponibilité et accessibilité. Ces politiques sont :

- **Push** : dans cette méthode, la logique de chargement est dans le système de production. Il " pousse " les données vers la zone de préparation quand il en a l'occasion. L'inconvénient est que si le système est occupé, il ne poussera jamais les données.

- **Pull** : contrairement de la méthode précédente, le Pull " tire " les données de la source vers la zone de préparation. L'inconvénient de cette méthode est qu'elle peut surcharger le système s'il est en cours d'utilisation.

- **Push-pull** : c'est la combinaison des deux méthodes. La source prépare les données à envoyer et indique à la zone de préparation qu'elle est prête. La zone de préparation va alors récupérer les données. Les Processus ETL doit respecter les critères suivants

Sûr : assure l'acheminement des données et leur livraison. Rapide : la quantité de données manipulées peut causer des lenteurs. Le processus d'alimentation doit palier à ce problème et assurer le chargement du Data Warehouse dans des délais acceptables.

Correctif : le processus d'alimentation doit apporter les correctifs nécessaires pour améliorer la qualité des données.

Transparent : le processus de l'ETL doit être transparent afin d'améliorer la qualité des données.

1.4.3 Mise en œuvre du Data Warehouse

C'est la dernière étape d'un projet Data Warehouse, soit son exploitation. L'exploitation du Data Warehouse se fait par le biais d'un ensemble d'outils analytiques développés autour du Data Warehouse. Donc cette étape nécessite l'achèvement du développement, ou de la mise en place, de ces outils qui peuvent accomplir les fonctions suivantes :

a. Requêtage ad-hoc

Le requêtage ad-hoc reste très fréquent dans ce type de projet. En effet, les utilisateurs de l'entrepôt de données, et spécialement les analystes, seront amenés à interagir avec le DW via des requêtes ad-hoc dans le but de faire les analyses requises par leurs métiers et, d'élaborer aussi, des rapports et des tableaux de bords spécifiques. L'accès à ce genre de service peut se faire via différentes méthodes et outils. Cependant, les spécialistes en la matière préconisent de laisser la possibilité à l'utilisateur de choisir les outils qui lui paraissent les plus adéquats.

b. Reporting :

Destiné essentiellement à la production de rapports et de tableaux de bord, « *il est la présentation périodique de rapports sur les activités et résultats d'une organisation, d'une unité de travail ou du responsable d'une fonction, destinée à en informer ceux chargés de les superviser en interne ou en externe, ou tout simplement concernés par ces activités ou résultants* » [[http : //fr.wikipedia.org/wiki/Reporting](http://fr.wikipedia.org/wiki/Reporting)].

Ces outils de Reporting ne sont pas, à proprement parler, des instruments d'aide à la décision, mais, lorsqu'ils sont utilisés de manière appropriée, ils peuvent fournir une précieuse vue d'ensemble.

Les rapports sont alors créés par le biais d'outils de Reporting qui permettent de leur donner un format prédéterminé. Les requêtes sont constituées lors de l'élaboration des rapports qui seront ensuite diffusés périodiquement en automatique ou ponctuellement à la demande.

d. Tableaux de bord :

Les tableaux de bord sont un outil de pilotage qui donne une vision sur l'évolution d'un processus, afin de permettre aux responsables de mettre en place des actions correctives.

« Le tableau de bord est un ensemble d'indicateurs peu nombreux conçus pour permettre aux gestionnaires de prendre connaissance de l'état et de l'évolution des systèmes qu'ils pilotent et d'identifier les tendances qui les influenceront sur un horizon cohérent avec la nature de leurs fonctions » [Bouquin, 2003].

Cette forme de restitution a la particularité de se limiter à l'essentiel, c'est-à-dire la mise en évidence de l'état d'un indicateur par rapport à un objectif, tout en adoptant une représentation graphique de l'information.

e. Data Mining :

Au sens littéral du terme, le Data Mining signifie le forage de données. Le but de ce forage est d'extraire de la matière brute qui, dans notre cas, représente de nouvelles connaissances. L'idée de départ veut qu'il existe dans toute entreprise des connaissances utiles, cachées sous des gisements de données. Le Data Mining permet donc, grâce à un certain nombre de techniques, de découvrir ces connaissances en faisant apparaître des corrélations entre ces données.

Le Data Warehouse constituera alors la première source de données sur laquelle s'exécutera le processus de découverte de connaissances. Dans la majeure partie du temps, l'entrepôt de données représente un pré requis indispensable à toute fouille de données.

Le recours à ce genre de méthode est de plus en plus utilisé dans les entreprises modernes. Les applications et outils implémentant ces solutions sont rarement développés en interne. En effet, les entreprises préfèrent se reposer sur des valeurs sûres du marché afin d'exploiter au plus vite les données en leur possession.

1.4.4 Maintenance et expansion

La mise en service du Data Warehouse ne signifie pas la fin du projet, car un projet Data Warehouse nécessite un suivi constant compte tenu des besoins d'optimisation de performance et ou d'expansion. Il est donc nécessaire d'investir dans les domaines suivants :

Support : assurer un support aux utilisateurs pour leur faire apprécier l'utilisation de l'entrepôt de données. En outre, la relation directe avec les utilisateurs permet de détecter les

correctifs nécessaires à apporter.

Formation : il est indispensable d'offrir un programme de formation permanent aux utilisateurs de l'entrepôt de données.

Support technique : un entrepôt de données est considéré comme un environnement de production. Naturellement le support technique doit surveiller avec la plus grande vigilance les performances et les tendances en ce qui concerne la charge du système.

Management de l'évolution : il faut toujours s'assurer que l'implémentation répond aux besoins de l'entreprise. Les revues systématiques à certain point de contrôle sont un outil clé pour détecter et définir les possibilités d'amélioration. En plus du suivi et de la maintenance du Data Warehouse, des demandes d'expansion sont envisageables pour de nouveaux besoins, de nouvelles données ou pour des améliorations. Ces travaux d'expansion sont à prévoir de façon à faciliter l'évolution du schéma du Data Warehouse.

CHAPITRE 2

LES SYSTÈMES DE GESTION HÔTELIÈRE

2.1 LA GESTION HÔTELIÈRE ACTUELLE

Avec des besoins technologiques grandissants, les hôteliers font face à de nouveaux défis, dont celui d'investir dans des systèmes de gestion performants. C'est dans un souci de compétitivité que l'inter connectivité entre les différentes interfaces prend de l'importance, même si la tâche peut s'avérer ardue. La justification de tels investissements demeure néanmoins un perpétuel défi pour les gestionnaires

Bien que la plupart des établissements hôteliers dans le utilisent aujourd'hui des systèmes de gestion, ceux-ci sont presque tous dotés de fonctionnalités basiques comparativement à ce qui existe dans d'autres secteurs d'activité. Aujourd'hui, il existe de nouveaux outils appelés à devenir de véritables leviers de compétitivité des établissements hôteliers. On peut diviser les principaux systèmes de gestion en cinq catégories :

- base de données centrale : Property Management System (PMS) ;
- réservations : Central Reservation System (CRS) ;
- revenus : Revenue Management System (RMS) ;
- points de vente : Electronic Point of Sale (EPOS) ;
- Marketing : Customer Relationship Management (CRM)

Pour la gestion des opérations hôtelières, chaque département fonctionne avec un système différent. Le système PMS utilisé à la réception permet de gérer les activités quotidiennes (voir image ci-dessous), soit les arrivées et les départs des clients, la facturation, la gestion des chambres, etc. On s'en sert comme une centrale pour les systèmes des autres services afin de collecter toutes les informations des clients, notamment les factures du restaurant, des services aux chambres et des appels téléphoniques. Les systèmes PMS, CRS, RMS et CRM sont généralement connectés entre eux, ce qui permet de suivre les nouvelles réservations et de com-

muniquer la disponibilité et les tarifs en temps réel.

On constate également un grand nombre d'hôtels de la place qui utilise encore des systèmes de gestion manuel. De la facturation à la gestion des réservations tout est fait manuellement et sur papier.

D'autre part les gestionnaires d'hôtel les plus avertis ont eu vent des avantages du e-commerce et essayent de développer leurs activités grâce au service en ligne. Toutefois chaque hôtel possède son propre site qu'il présente individuellement à ses clients. Naviguer sur plusieurs sites web d'hôtel est extrêmement difficile et de là naît le besoin de créer des plateformes qui ressemblent à un grand nombre d'hôtel où les clients peuvent comparer les offres de chaque hôtel avant de s'engager dans une réservation. Hosteline se présente dans le registre des plateformes fédératrices d'hôtels tout comme booking.com et Jumia Travel qui vendent des produits qui appartiennent aux hôtels qui ont souscrit à leurs plateformes. On peut également lui donner la connotation de e-commerce dans le sens où des clients achètent des produits et services par Hosteline.

Le plus difficile avec les plateformes où on vend les produits des tierces personnes est de permettre au vendeur d'entretenir la relation client et de fidéliser sa clientèle. Ceci d'autant plus que la plateforme a des objectifs qui passent par la fidélisation des clients mais cette fois-ci en vers le site de vente. Pour permettre aux hôteliers d'entretenir la relation client avec leurs clients sur la plateforme, Hosteline a pensé des stratégies parmi lesquelles un système décisionnel qui servira les hôteliers et leurs clients mais Hosteline et ses clients. Cela permet de rapprocher le client et le vendeur et permettre une meilleure fourniture de service et une augmentation du chiffre d'affaire pour chacune des entreprises.

2.2 PRÉSENTATION TECHNIQUE DE HOSTELINE

Comme la majorité des projets informatiques, Hosteline a été conçu et réalisé selon un canevas avec des langages et méthodes de conception bien connus. En effet ce projet a été piloté par la méthodologie agile SCRUM (Annexe xx) avec le langage UML pour accompagner la conception. Après une description brève du langage UML, les diagrammes qui ont servi à la conception de Hosteline seront présentés et commentés.

2.2.1 LE LANGAGE UML

UML (Unified Modeling Language, que l'on peut traduire par langage de modélisation unifié) est une notation permettant de modéliser un problème de façon standard. Ce langage est né de la fusion de plusieurs méthodes existant auparavant, et est devenu désormais la référence en terme de modélisation objet, à un tel point que sa connaissance est souvent nécessaire pour obtenir un poste de développeur objet. La modélisation consiste à créer une représentation

simplifiée d'un problème : le modèle. Grâce au modèle il est possible de représenter simplement un problème, un concept et le simuler. La modélisation comporte deux composantes :

- L'analyse, c'est-à-dire l'étude du problème
- La conception, soit la mise au point d'une solution au problème

Le méta modèle UML fournit une panoplie d'outils permettant de représenter l'ensemble des éléments du monde objet (classes, objets, ...) ainsi que les liens qui les relie. Toutefois, étant donné qu'une seule représentation est trop subjective, UML fournit un moyen astucieux permettant de représenter diverses projections d'une même représentation grâce aux vues. Une vue est constituée d'un ou plusieurs diagrammes. On distingue deux types de vues :

i) **Les vues statiques**, c'est-à-dire représentant le système physiquement

- diagrammes d'objets
- diagrammes de classes
- diagrammes de cas d'utilisation
- diagrammes de composants
- diagrammes de déploiement

ii) **Les vues dynamiques**, montrant le fonctionnement du système

- diagrammes de séquence
- diagrammes de collaboration
- diagrammes d'états-transitions
- diagrammes d'activités

Mais en ce qui concerne notre cas d'étude présent, nous ne représenterons que les diagrammes utilisées lors de la conception de la plateforme Hosteline.

2.2.2 LE DIAGRAMME DE CAS D'UTILISATION

Les **diagrammes de cas d'utilisation** sont utilisés pour donner une vision globale du comportement fonctionnel d'un système logiciel. Ils sont utiles pour des présentations auprès de la direction ou des acteurs d'un projet, mais pour le développement, les cas d'utilisation sont plus appropriés. Un cas d'utilisation représente une unité discrète d'interaction entre un utilisateur (humain ou machine) et un système. Il est une unité significative de travail. Dans un diagramme de cas d'utilisation, les utilisateurs sont appelés acteurs (actors), ils interagissent avec les cas d'utilisation (use cases)[Source Wikipedia].


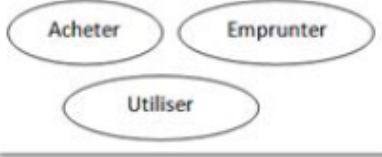
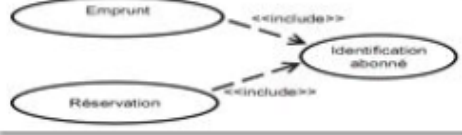
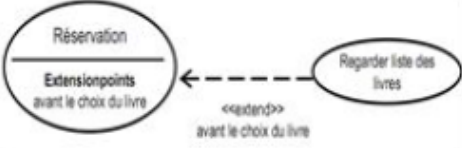
CONCEPTS	REPRESENTATION	DESCRIPTION
<i>L'acteur</i>		Un acteur représente une personne ou un périphérique qui joue un rôle (interagit) avec le système.
<i>Les cas d'utilisation</i>		Un cas d'utilisation définit une fonctionnalité utilisable par un acteur. Il exprime une suite d'interactions entre un acteur et l'application.
La relation <i>"Include"</i>		Cette relation précise qu'un cas d'utilisation contient le comportement défini dans un autre cas. Elle met en commun des comportements communs à plusieurs cas d'utilisation.
La relation <i>"Extend"</i>		Cette relation précise qu'un cas d'utilisation peut dans certains cas augmenter le comportement d'un autre cas d'utilisation. Une condition devra valider cette augmentation.

FIGURE 2.1 – Formalismes du diagramme de cas d'utilisation.

L'étude de Hosteline dans la partie diagramme de class à donner deux diagrammes représentant des groupes d'acteurs différents et ayant peu ou pas d'actions en commun comme le montrent les figures ci après.

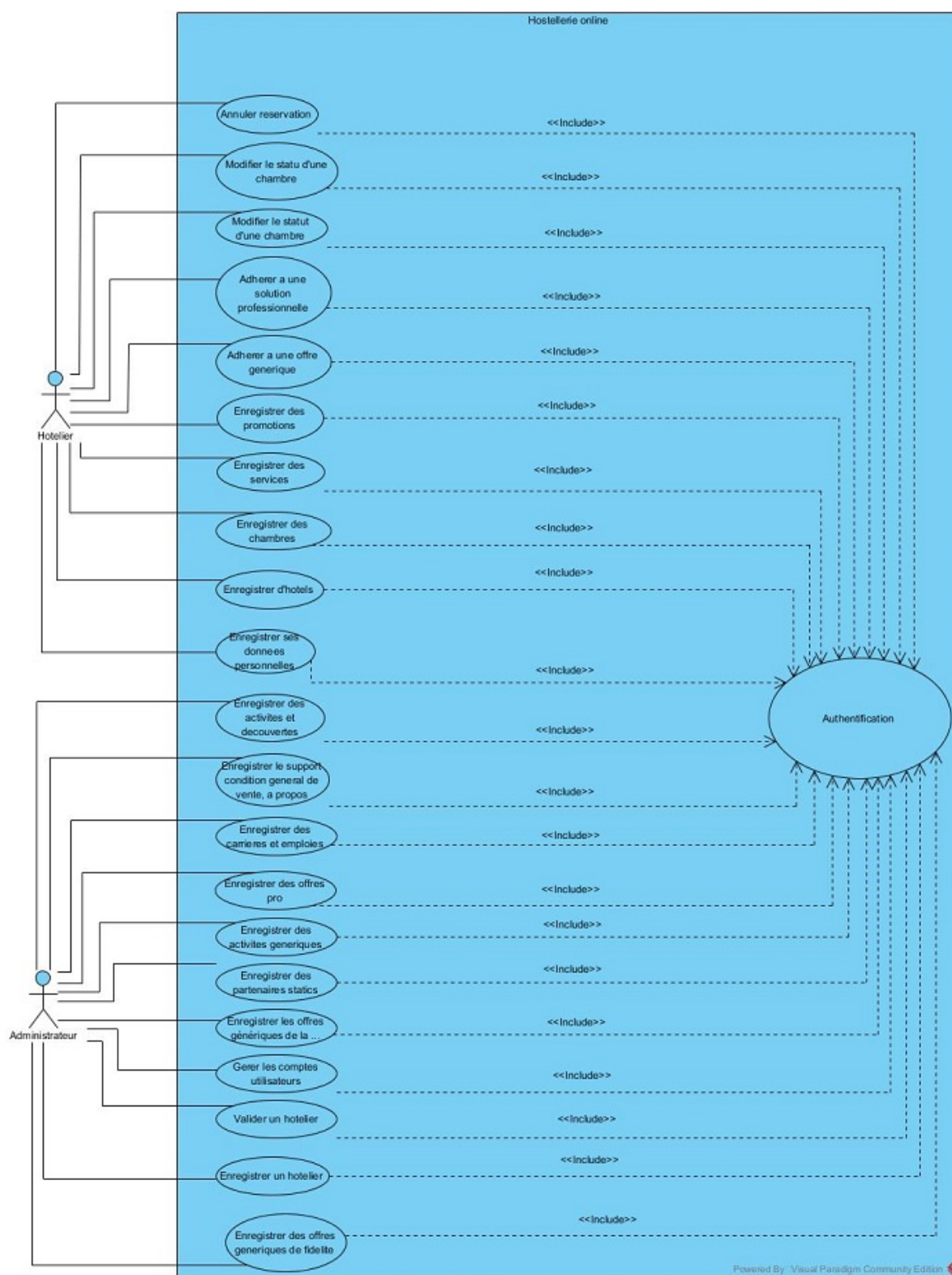


FIGURE 2.2 – Diagramme de cas d'utilisation contenant deux des acteurs de Hosteline (Agrandi en Annexe XX)

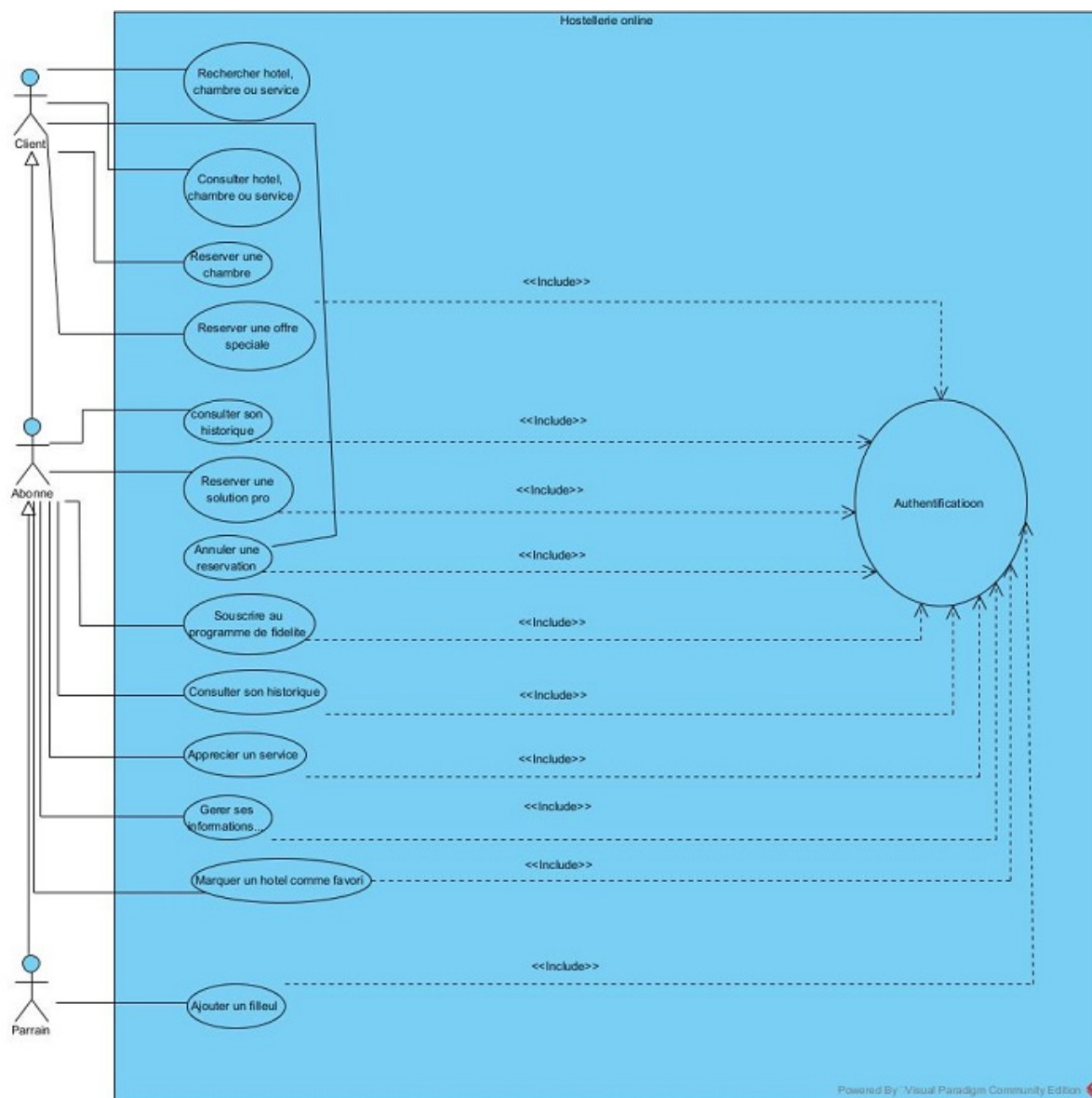


FIGURE 2.3 – Diagramme de cas d'utilisation contenant l'acteur "Client" (Agrandi en Annexe XX)

2.2.3 LE DIAGRAMME DE CLASSES

Le **diagramme de classes** est un schéma utilisé en génie logiciel pour présenter les classes et les interfaces des systèmes ainsi que les différentes relations entre celles-ci. Ce diagramme fait partie de la partie statique d'UML car il fait abstraction des aspects temporels et dynamiques. Une classe est un ensemble de fonctions et de données (attributs) qui sont liées ensemble par un champ sémantique. Les classes sont utilisées dans la programmation orientée objet. Elles permettent de modéliser un programme et ainsi de découper une tâche complexe en plusieurs petits travaux simples.

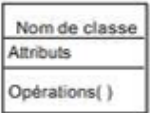
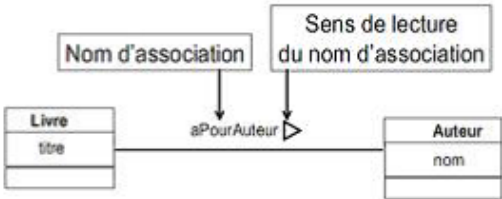
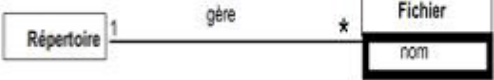
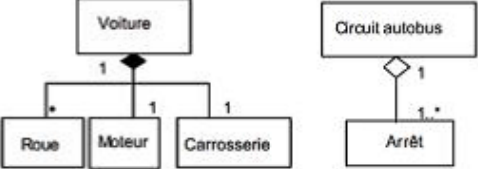
CONCEPTS	REPRESENTATION	DESCRIPTION
La classe		Elle décrit un ensemble d'objets (instances de la classe) possédant des propriétés communes.
L'association		Une association est symbolisée par une ligne continue exprime une connexion sémantique bidirectionnelle entre classes.
Les cardinalités		Les cardinalités indiquent le nombre d'instances d'une classe pour chaque instance de l'autre classe. Dans notre exemple, un répertoire gère plusieurs fichiers, et un fichier est géré par un seul répertoire
Les agrégations		C'est une association qui décrit une relation d'inclusion entre une partie et un tout (l'agrégat). Elle est symbolisée par un losange plein pour des durées de vie liées et par un losange vide du côté de l'agrégat pour des durées de vie indépendantes.

FIGURE 2.4 – Formalismes du diagramme de classes.

En ce qui concerne Hosteline, dans sa phase d'étude le diagramme de classe qui en est sorti est celui de la figure suivante (Version agrandie en annexe XX).

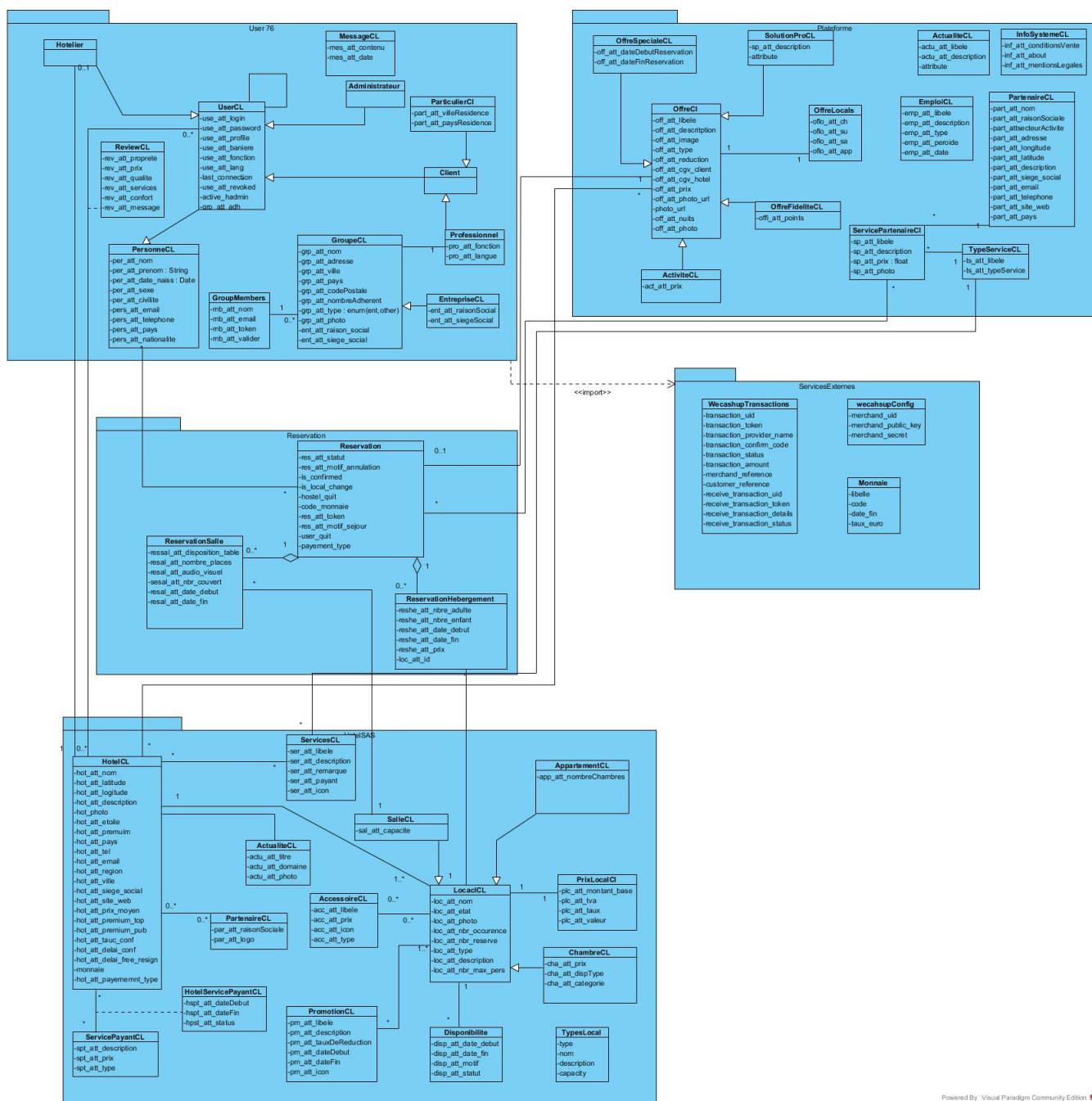


FIGURE 2.5 – Diagramme de classes du système (Agrandi en Annexe XX)

2.3 LE SYSTEME OPERATIONNEL ACTUEL

Si on parle de mettre en œuvre un système décisionnel sur Hosteline c'est parce que un système opérationnel qui va fournir les données à ce dernier existe déjà et est tout à fait fonctionnel. Dans cette partie nous allons présenter des partie du Systèmes opérationnel par les quels la collecte de données est plus forte sur Hosteline.

2.3.1 INSCRIPTION

L'inscription n'est pas une condition d'utilisation de Hosteline pour les clients qui souhaitent juste réserver. Elle est par contre le point d'entrée vers des droits et bonus supplémentaire pour le Client. Pour le Propriétaire d'hôtel ou client-hébergeur il doit s'inscrire pour avoir accès a l'administration de son établissement hôtelier. La figure en annexe XX montre comment se présente le formulaire d'inscription de Hosteline.

The image shows a web browser interface for Hosteline. A modal window for registration is centered on the screen. The modal has a title bar with 'SEGN IN' and 'REGISTER' tabs, with 'REGISTER' selected. The main heading inside the modal is 'Register for your account'. Below this, there are two options: 'Connexion' with a Google logo and 'Continue as Yannick' with a Facebook logo and a user profile picture. Below these, there is a text 'One friend uses this.' The registration form consists of several input fields: 'Login' (with a red error message 'login is required' below it), 'Email', 'Password', 'Name', and 'First name'. There are also two dropdown menus: 'Choose your sex' and 'Choose your account's type'. A red button labeled 'REGISTRATION' is positioned at the bottom right of the form. The background of the page is slightly blurred, showing a search results page with a search bar, filters, and hotel listings.

FIGURE 2.7 – Interface d'inscription a Hosteline

2.3.2 LA RESERVATION

La réservation commence dès que l'utilisateur rentre ses critères de recherche car certains de ces critères seront retenus comme informations de réservation. Le système de panier est utilisé ici et permet à l'utilisateur de sélectionner plusieurs chambres et de les réserver plus tard à sa guise. Lorsque la sélection est terminée, l'utilisateur rentre ses informations s'il n'a pas de session en cours et valide sa réservation. Toute réservation validée peut être confirmée (lorsqu'on paie) ou annulée suivant les termes et les conditions de l'hôtel et de la plateforme. La figure en Annexe XX est une capture de la fenêtre d'enregistrement des données de l'utilisateur.

The screenshot displays the Hosteline website's reservation interface. The top navigation bar is red with the Hosteline logo and links for HOME, HOTEL, VILLA, APPART, and HALL. The main content area is divided into two columns. The left column, titled 'Your Personal Informations', contains a registration form with fields for Nom, Prenom, Adress Email, Confirmez l'email, Pays, Telephone, Nationalité, Civilité, Password, Repeat Password, and Login. A checkbox for 'I want to create my account' is also present. A blue 'BACK' button is located at the top right of this section. The right column displays details for 'AKWA PALACE' in Douala, Cameroun, including a 5-star rating and a price of 46,738.00 EUR TTC for 1 suite. A 'TOTAL PRICE' of 46,738.00 EUR TTC is also shown. At the bottom of the right column, there is a contact section for Akwa Palace with a phone number and email address, and a link to activate wind.

FIGURE 2.8 – Interface de réservations de Hosteline

2.3.3 ENREGISTREMENT HOTELS

Une fois le compte hôtelier créé, la prochaine étape consiste à créer son hôtel sur la plateforme et le configurer en y ajoutant des chambres, Sales et autres selon les dispositions de son établissement. En effet, dans cette partie, l'hôtelier ou la personne en charge du profil de l'hôtel va ajouter la description de son hôtel jusqu'à la position. La figure XX en annexe présente l'interface de création du profil d'un hôtel.

2.4 LA PROBLÉMATIQUE DU THÈME

INSBI propose des services à travers Hosteline sur un marché très concurrentiel car des plateformes comme celle-ci il y en a un grand nombre sur le marché. Pour gagner sa place et être concurrentiel, il y a lieux de développer des approches qui rendrons le produits plus intéressent que les autres déjà présent sur le marché. Dans un souci d'augmentation de la qualité des services offert aux clients, Les besoins suivants se sont formulés :

- **Un tableau de bords pour chaque utilisateur :** Il s'agit de ressortir chaque profil d'utilisateur de la plateforme, revoir ses droit et ses actions sur la plateforme, et de déterminer les données génères par ce profile en suit de les traduire de manier graphique et interprétable.
- **Pouvoir personnaliser les offres au client :** Le but étant de mettre sur pied un système permettant de mémoriser et restituer les choix des clients pour pouvoir leurs faire des offres a des périodes biens précises qui sont à la portée de leurs bourses. Le programme de fidélisation de Hosteline pouvant servir comme base pour le développement de ces outils.
- **Etudie et faire de campagnes de publicité par localité :** Grace au donnée de localisation recueillit lors des réservations sur l'utilisateur et le lieu où il va passer les nuits on doit pouvoir faire une analyse croisé pour faire des promotions qui ciblent des clients d'un endroit précis ou des hôtels dans une location géographique précise.
- **Veille environnementale :** Les outils développé a la suite du projet devront pouvoir intégrer des données de sources divers pour permettre d'informer les décideurs de l'évolution de l'environnement et des tendances qui tourne au tour du secteur hôtelier.

Après avoir écouté les besoins et les problèmes cité par les responsables de Hosteline nous avons entrepris une démarche projet pour résoudre méthodiquement les problèmes qui nous étaient posés et pour pouvoir aboutir à une solution qui répond a leurs besoins. Nous avons dans un premier temps nous référer aux diagrammes et modèles utilises a la conception de la plateforme pour avoir une idée claire de la structure des données en suit nous allons concevoir un data warehouse qui sera interrogé a l'aide d'un système de requêtes pour répondre aux questions des décideurs de la plateforme.

Dans le prochain paragraphe nous proposons de solutions et évaluons des choix d'outils pour implémenter ces solutions.

2.5 SOLUTIONS PROPOSÉES

Parmi les solutions que nous avons développées pour INSBI, on pouvait dénombrer des solutions de reporting, tableaux de bord et toutes ces solutions avaient un point en commun. Il fallait au départ construire un entrepôt de données et mettre sur pied un système d'exploitation de ces données qui permet de résoudre chaque problème à sa manière. Pour chaque problème nous avons fait des propositions de solutions et celle qui ont été retenues sont citée plus bas.

- Pour le **tableau de bords par profile utilisateur** nous avons proposé de développer un système décisionnel cote à cote au système opérationnel principale. Il s'agit de concevoir un data warehouse qui est chargé en temps réel a chaque insertion dans la BD opérationnel. Grace à cette solution la mise à jour des donnée du Data warehouse est instantanée (Entrepôt de données temps réel). Ainsi les besoins explicites sont satisfaits grâce aux requêtes SQL spécialisés.
- Pour permettre de **personnaliser les offres aux clients** nous avons assimilé ces problèmes à des problématiques connues et avons adopté des solutions connues que nous avons ensuite développé. Ici il ne s'agit de rien d'autre que le Data Mining (encore appelé Forage de données). Les données de chaque profil utilisateur sont traitées séparément et on peut extrapoler pour obtenir une tendance générale et proposer des offres qui vont satisfaire des profils groupé d'utilisateurs. Bien sûr ceci passe également par la mise sur pied d'un data warehouse spécialisé.
- Face au besoin de **faire des offres localisé**, nous avons proposé de nous aider de donnée utilisateurs recueilli lors des réservations pour produire des représentations graphiques mettant en exergues les localités ou les destinations et les taux de réservations correspondants. Grace aux chiffres représente par localité, les décideurs de Hosteline pourront a tout instant donné lancer des campagnes stratégiques qui leurs permet de faire un chiffre d'affaire plus grand.
- Le problème de **veiller sur l'environnement** était le plus difficile à résoudre car celui-ci faisait intervenir des données issues de l'environnement externe a Hosteline. Apres avoir identifié ces sources de données qui étaient de nature divers nous avons décidé d'exploiter des outils de BI spécialisés dans le traitement de données. Ici il fallait consolider les données internes et externes à Hosteline pour pourvoir en informations les décideurs de Hosteline. Pour mettre ça en œuvre le choix de l'outil **Pentaho Data Integration** a fait l'unanimité. Ici contrairement aux autres problèmes plus haut, la solution est totalement détachée de la plateforme web et sera déployé sur un serveur local a l'entreprise.

Les choix des techniques et technologies pour arriver aux solutions obtenues ont fait l'objet d'une étude préalable. Dans les chapitres suivants nous allons présenter les méthodes et les raisons qui nous ont permis de faire ces choix.

Deuxième partie

Contribution

3.1 EXPRESSION DES BESOINS

Une fois le système conçu et déployé il sera accessible aux utilisateurs de la plateforme avec des niveaux d'accès selon les différents profils qui sont énoncés dans les points qui suivent. Ces besoins connus et définis sont groupés en deux types de besoins.

3.1.1 LES BESOINS FONCTIONNELS

Les besoins fonctionnels désignent les fonctions que le produit fini doit posséder pour qu'il soit considéré comme fonctionnel. Dans le cas de notre Data Warehouse, pour les profils suivant il doit pouvoir :

Pour le profile administrateur, on retrouve les exigences suivantes :

- Fournir un tableau de bord accessible en ligne qui permettra de renseigner l'administrateur sur :
 - L'activité des réservations sur des périodes modulable selon le grée de l'utilisateur sur les années, les trimestres et les mois passé et en cours
 - L'activité des hôteliers sur les mêmes périodicités que le point précédent. Ici il s'agira des détails tel que le chiffre d'affaire de chaque hôtelier le nombre de nuitées, le nombre de chambres réservés et les paiements mensuels de ce dernier.
 - L'activité des utilisateurs en général. On fournira comme informations le nombre de visites par jour, par mois, par trimestre et par années.
 - Les différentes offres de la plateforme et leurs appréciations dans le temps.
- Fournir un système accessible sur le serveur de l'entreprise qui aura un tableau de bord similaire à celui en ligne avec des fonctionnalités en plus. On aura les fonctionnalités suivant :
 - Intégrer aux informations de réservations des informations de météo du lieu de provenance de la personne qui réserve et celui de l'endroit où l'hôtel de destination grâce

à un service de météo en ligne qui sera utilisé comme source de données.

- Un comparatif de la liste d'hôtel issue du site du ministère du tourisme Camerounais. Ces informations seront reçu dans un fichier Excel et devra être utilisé comme source de données

Pour le profile Hôteleur, On dégage les fonctionnalités suivantes :

- Un tableau de bord permettant a l'utilisateur de suivre les réservations émises pour son établissement, les tendances de ses offres et promotions, les chambres les plus prisées et le chiffre d'affaire. Ces informations seront disponibles en temps réel et on pourra consulter l'historique de ces informations.
- Des indicateurs de performances qui seront disposé en temps réel sur le profil de l'utilisateur et qui reflètent l'activité hebdomadaire ou mensuelle. Pour le profile client on devra fournir :
- Sur son profile on devra afficher des indicateurs de son activité et des graphes qui permet de visualiser le chiffre d'affaire qu'il génère sur une période d'un an. Tous ces besoins ont été apprécié et pourront évoluer d'où le besoin fondamentale de développer une solution apte à subir des maintenances évolutives et correctives sans avoir à détruire ou recommencer ce qui a été fait.

3.1.2 LES BESOINS NON FONCTIONNELS

Les besoins non fonctionnels désignent les conditions à remplir pour qu'une solution logicielle fonctionne correctement. Dans le cas de notre application, ces conditions doivent être remplir :

- Un serveur de base de données pour stocker les données du data warehouse.
- Un une connexion internet pour exploiter les sources de données qui sont pour la majorité des API en ligne.

L'ensemble des besoins énuméré plus constitue une partie des questionnements qui feront du data warehouse un produit fonctionnel et exploitable par l'ensemble des décideurs de la plateforme Hosteline.

3.2 CONTRAINTES LIÉES AU PROJET

Les contraintes sont l'ensemble de la (définition). Pour ce projet nous avons fait face a des contrainte de de différentes nature que nous avons regroupé sous les points qui suivant.

3.2.1 EXIGENCES DE DOCUMENTATION

En effet la solution une fois développée et déployée devra être appuyée d'une forte documentation pour les utilisateurs et les développeurs qui prendront le relais dans la maintenance de la solution. Un manuel d'utilisateur devra être fourni pour chacun des profils utilisateur de la solution et sera accessible en un clic sur les différents profils du système selon le niveau d'accès et les fonctionnalités. Aussi un document contenant les détails de l'implémentation et de déploiement de la solution sera fourni et mis à la disposition de l'équipe de développement en charge de maintenir et faire évoluer la solution. Pour une assistance à l'utilisateur une FAQ (Foire aux questions) sera intégrée au site web ou un sous domaine du site pour faciliter la prise en main par les utilisateurs de la solution.

3.2.2 EXIGENCES DE QUALITÉ

Le but de la mise en œuvre de ce système décisionnel est l'augmentation du chiffre d'affaire de Hosteline passant par une meilleure qualité de service. La solution elle-même devra respecter des standards de qualité qui correspondent à ce type de projet.

3.2.3 CONTRAINTES TEMPORELLES

Cette contrainte s'est avérée être la plus difficile à respecter car l'équipe en charge du projet manquait d'expérience dans la mise en œuvre de solution décisionnel. Néanmoins avec un redoublement d'effort on a pu réduire le retard à juste un mois. On est passé de quatre mois que prévoyait le calendrier de départ à cinq mois. Le diagramme de gant de la figure XX présente le planning prévisionnel qui a été fixé au départ avec une équipe de trois développeurs.

3.2.4 CONTRAINTES MATÉRIEL ET LOGICIEL

Pour la mise en place du système on a dû se munir d'un matériel spécifique remplissant des conditions qui doivent satisfaire les exigences de vitesse et mémoire du système. Parmi ce matériel on a dénombré :

- Une machine qui fera office de serveur local avec des caractéristiques minimales de 4Giga de ram, un processeur 2.6 Core I3, 300 Giga de ROM avec un système d'exploitation Windows 7 ou plus.
- Le logiciel Pentaho de préférence dans sa dernière version (8.1.0 lors de la rédaction de ce mémoire)

3.3 DÉMARCHE DE RÉALISATION

3.3.1 INTERVENANTS DU PROJET

Ce qui fait en générale l'intérêt d'un projet c'est sa capacité à résoudre le problème d'un ensemble de personnes. Ainsi la pertinence du projet dépend au finale du niveau de satisfaction que peut offrir le produit fini aux parties prenantes de l'effort. Les intervenants dans un projet c'est l'ensemble des personnes qui de près ou de loin influencent le déroulement du projet. La mise en place du système décisionnel de Hosteline a comme parties prenantes :

- Les propriétaires d'établissement hôteliers et clients hôteliers de Hosteline,
- Les clients qui viennent réserver des logements sur la plateforme
- Les promoteurs de la plateforme Hosteline
- L'équipe de développement

Toutes ces groupes de personnes ci-dessus sont celles qui ont eu un apport considérable a la bonne marche du projet et sans elles ce projet n'aurait pas tout son sens car le but était de satisfaire ces parties prenantes d'une manière ou d'une autre.

3.3.2 PLANIFICATION DU PROJET

Pour garantir le bon déroulement du projet, tout en respectant les délais, élaboré une planification globale de conduite du projet. Le diagramme suivant décrit cette planification ainsi que l'ordonnancement prévu des phases du projet.

3.3.3 DIAGRAMME DE GANTT

Le diagramme de Gantt, couramment utilisé en gestion de projet, est l'un des outils les plus efficaces pour représenter visuellement l'état d'avancement des différentes activités (tâches) qui constituent un projet. La colonne de gauche du diagramme énumère toutes les tâches à effectuer, tandis que la ligne d'en-tête représente les unités de temps les plus adaptées au projet (jours, semaines, mois etc.). Chaque tâche est matérialisée par une barre horizontale, dont la position et la longueur représentent la date de début, la durée et la date de fin. Ce diagramme permet donc de visualiser d'un seul coup d'œil :

- Les différentes tâches à envisager
- La date de début et la date de fin de chaque tâche
- La durée escomptée de chaque tâche
- Le chevauchement éventuel des tâches, et la durée de ce chevauchement
- La date de début et la date de fin du projet dans son ensemble

3.3.4 ESTIMATION DU COUT DE LA SOLUTION

RESSOURCES HUMAINES

Pour chaque phase du projet nous avons fait appel à des compétences divers dans le domaine du décisionnel. Ces rôles ont été jouées en majorité par notre petite équipe de quatre développeurs mais les compétences exploité à chaque niveau étais différemment apprécié et rémunéré. Le tableau suivant récapitule.

RESSOURCES MATERIELS

CHAPITRE 4

MISE EN ŒUVRE DE LA SOLUTION

4.1 CHOIX DES OUTILS

4.1.1 CHOIX D'UN OUTIL BI

Avant de mettre en place un système décisionnel en général le besoin née d'un ensemble de question que se pose le DSI. A chacune de ces questions correspond une réponse venant d'une partie ou un processus du BI. La figure suivant résume cette question que tentent de répondre tout système décisionnel.

Question	Solution BI
Comment prendre des décisions sur les faits et les données accumulées sur le temps ?	Une solution BI end to end
Comment intégrer les données de sources diverses pour les analyser ?	ETL
Comment interpréter les données brutes de la meilleure façon ?	Data discovery/Analytics
Comment prédire la trajectoire future du business ?	Analyse prédictive/Machine learning
Quel est la meilleure façon de partager les données/information ?	Visualisation/Reporting
Comment surveiller les changements dynamiques de mon activité ?	Tableaux de bord

FIGURE 4.1 – Problèmes et outils BI réponses.

Dans le marché de produits BI il existe plusieurs critères de classement de fournisseurs. Nous les avons segmenté en deux grands groupes (Les grands Fournisseurs et les plus petits et nouveaux fournisseurs).

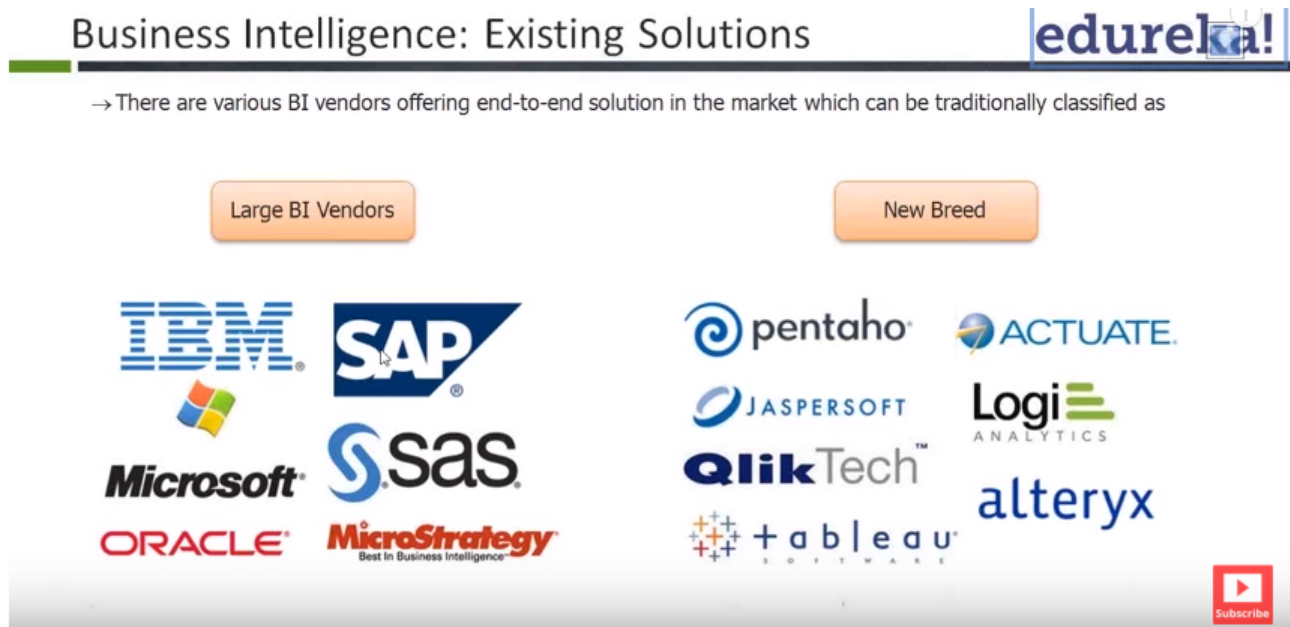


FIGURE 4.2 – Regroupement des fournisseurs de solutions BI

Toutefois il existe un grand nombre de fournisseurs qui offrent des outils spécifiques pour des processus BI précis. On a par exemple :

- INFORMATICA : Pour l'intégration de données (ETL)
- RAPIDMINER : Pour l'analyse de données (Analytics)
- TALEND : Pour l'intégration de données (ETL)

L'un des facteurs qui a motivé notre choix dans ce cas est que PENTAHO s'interface facilement avec la majorité de ces solutions qui traitent un seul processus BI.

LES ENJEUX D'UN BON CHOIX

Avec les solutions de BI qui existent déjà on fait face à un ensemble de difficultés qui cependant rendent le choix selon le contexte plus facile. Ces barrières sont les suivantes.

- Coût d'installation : Les solutions des grands fournisseurs ont un coût en infrastructure souvent très élevé car la mise en œuvre des serveurs et des architectures réseau extrêmement coûteuses à mettre en place.
- Coût des licences : Les licences d'exploitation ont un coût assez élevé comparé aux plus petits fournisseurs.

- Temps d'intégration et d'apprentissage long : Les solutions les plus élaborées en matière de BI ont une courbe d'apprentissage plus difficile et l'installation prend plus de temps que les solutions des petits fournisseurs.
- Coût de maintenance élevé : L'augmentation du coût de maintenance est due au fait que il faut une main d'œuvre qualifiée et de qualité parfois pour effectuer même les tâches les plus dérisoires.

LE CHOIX FINAL

Le plus difficile dans les plus petits et nouveaux fournisseurs est le fait de ne pas trouver facilement un logiciel qui gère l'ensemble des processus BI. Cette difficulté nous a fait converger vers un outil en particulier qui n'avait pas ce défaut. Notre choix est porté sur PENTAHO pour les raisons suivantes :

- PENTAHO est une solution tout en un pour tous problèmes d'analyse de données.
- Le coût de la licence est abordable et il existe une version community qui est presque gratuit.
- Il se superpose facilement sur des solutions partielles et même des grands fournisseurs pour l'analyse ou l'exploitation des données issues de celles-ci.
- Il bénéficie d'une grande communauté car il est open source.
- Il est facile à faire évoluer et adapter aux besoins grandissants des entreprises (en anglais Scalability)
- Les composants de Pentaho sont des projets libres individuels (Standalone) qui bénéficient chacun d'une forte communauté.

Tels sont les points qui ont confortés le choix de Pentaho en tant que solution à notre problématique. La figure suivante est une correspondance entre d'une part les outils que offre Pentaho et le processus BI qu'il gère.

4.1.2 CHOIX DU SGBD

En termes de système de gestion de base de données notre choix s'est porté sur PostgreSQL car de par son atout majeur d'être très sécurisé il offre les avantages suivants :

- Au niveau de l'architecture : PostgreSQL est un système de gestion de base de données de modèle relationnel & objet.
- Au niveau des fonctionnalités : PostgreSQL facilite la tâche des développeurs dans sa configuration et supporte plus de fonctionnalités tel que CTE (Common Table Expressions), GiST/GIN ou bien les fonctions de fenêtrage par rapport à MySQL qui hélas a du retard.

Outil Pentaho	Processus BI
Pentaho Data Intégration (PDI)	ETL
Meta data Editor (PME)	Management de Métadonnées
Pentaho BA	Analyse de données
Report Designer (PRD)	Reporting opérationnel
Saiku	Reporting ad hoc
CDE	Tableau de bord
Pentaho user console (PUC)	Gouvernance/Monitoring

FIGURE 4.3 – La suite Pentaho.

- Au niveau des migrations vers d'autres types de BDD : Il est plus simple de migrer et changer sur d'autres bases de données avec Postgresql. De même pour sauvegarder et restaurer des backups.
- Au niveau de la documentation : Postgresql est en forte expansion et sa documentation reste plus claire.

4.2 ANALYSE CONCEPTION

4.3 CHOIX DES OUTILS

4.4 IMPLÉMENTATION

CONCLUSION

RÉFÉRENCES

- [1] Laurent AUDIBERT, *UML 2*. Institut Universitaire de Technologie de Villetaneuse. 2007 - 2008, 178 pages
- [2] François Jacquenet, *Génie logiciel : Processus de développement logiciel*. Université Saint Etienne - France, Institut Universitaire de Technologie de Villetaneuse. 03 Octobre 2016
- [3] Tutorial point, *Sdlc - overview*, <[https ://www.tutorialspoint.com/sdlc](https://www.tutorialspoint.com/sdlc)>. 09 Mars 2017.
Consulté le 25/05/2017 à 17h00

ANNEXE A

TITRE ANNEXE

A.1 Présentation de INSBI

CREATION

Elle a été créée en 2009 par 2 ingénieurs ayant travaillés pendant plus de 10 ans pour les grands intégrateurs français et sur de nombreux comptes clients. La Sarl INSBI (Institut Business Intelligence) a une ligne directrice essentiellement centrée sur l'informatique décisionnelle (Business Intelligence). Elle travaille avec une dizaine de collaborateurs en réseaux. En 2013, Sarl INSBI s'associe à la SAS IFICLIDE et prend la direction et le développement du pôle business intelligence. Depuis le début d'année 2017, l'associé Rodrigue Kendjio a entrepris l'extension des activités en Afrique. Amorcé dès le second trimestre 2017 par un projet d'e-commerce, le lancement officiel des activités est prévu au Cameroun à la fin d'année 2017.

A.1.1 MISSION

Là où le contexte est en évolution permanente et les facteurs majeurs de transformation sont centrés sur les défis concurrentiels et la globalisation de l'information, nos experts interviennent pour vous accompagner dans la mise en place de projets informatiques : d'INFRASTRUCTURES, D'APPLICATIONS et de services. Nous intervenons dans le domaine bancaire, l'assurance la grande distribution et l'industrie

A.1.2 PRESENTATION

Conseil Stratégique : Accompagner les directions générales dans leur besoin de pilotage
À Métier : Guider les directions métiers dans l'expression de leurs besoins
À Technologique : Aider au choix de solution de gestion et d'aide à la décision
À Conduite du changement : Faciliter, valoriser et promouvoir le changement
Réalisation
À Audit : Analyser l'existant et réaliser l'étude d'impact
À Gestion de projet : Piloter et animer le projet
À Technique : Concevoir et

mettre en œuvre le système d'information BI À Formation : Former les utilisateurs à la nouvelle plateforme

A.2 DIAGRAMMES UML

TABLE DES MATIÈRES

Dédicace	i
Remerciements	ii
Sommaire	iv
Abréviations	v
Liste des figures	vi
Liste des tableaux	vii
Résumé	viii
Abstract	ix
Introduction	1
I État de l’art	2
1 Synthèse Bibliographique	3
1.1 Business Intelligence	3
1.1.1 Introduction	3
1.1.2 Les systèmes décisionnels	3
1.2 Le Data Warehouse	5
1.2.1 Définition	5
1.2.2 Historique	5
1.2.3 ARCHITECTURE D’UN DATA WAREHOUSE	6
1.3 Modélisation des données de l’entrepôt	7
1.3.1 La modélisation dimensionnelle et ses concepts	7
Le concept des Faits	7

Le concept des Dimensions	8
1.3.2 Différents modèles de la modélisation dimensionnelle	8
1.3.3 La navigation dans les données	9
Slice & Dice	9
1.4 Démarche de Construction d'un Data Warehouse	10
1.4.1 Modélisation et conception du Data Warehouse	10
Approche « Besoins d'analyse »	10
Approche « Source de données »	11
Approche mixte	11
1.4.2 Les phases de l'alimentation « E.T.L »	12
Politiques de l'alimentation	13
1.4.3 Mise en œuvre du Data Warehouse	14
1.4.4 Maintenance et expansion	15
2 LES SYSTÈMES DE GESTION HÔTELIÈRE	16
2.1 LA GESTION HÔTELIÈRE ACTUELLE	16
2.2 PRÉSENTATION TECHNIQUE DE HOSTELINE	17
2.2.1 LE LANGAGE UML	17
2.2.2 LE DIAGRAMME DE CAS D'UTILISATION	18
2.2.3 LE DIAGRAMME DE CLASSES	22
2.2.4 LE DIAGRAMME D'ENTITÉ ASSOCIATION	24
2.3 LE SYSTEME OPERATIONNEL ACTUEL	25
2.3.1 INSCRIPTION	25
2.3.2 LA RESERVATION	26
2.3.3 ENREGISTREMENT HOTELS	26
2.4 LA PROBLÉMATIQUE DU THÈME	27
2.5 SOLUTIONS PROPOSÉES	28
II Contribution	29
3 CAHIER DES CHARGES	30
3.1 EXPRESSION DES BESOINS	30
3.1.1 LES BESOINS FONCTIONNELS	30
3.1.2 LES BESOINS NON FONCTIONNELS	31
3.2 CONTRAINTES LIÉES AU PROJET	31
3.2.1 EXIGENCES DE DOCUMENTATION	32
3.2.2 EXIGENCES DE QUALITÉ	32
3.2.3 CONTRAINTES TEMPORELLES	32
3.2.4 CONTRAINTES MATÉRIEL ET LOGICIEL	32
3.3 DÉMARCHE DE RÉALISATION	33
3.3.1 INTERVENANTS DU PROJET	33

3.3.2	PLANIFICATION DU PROJET	33
3.3.3	DIAGRAMME DE GANTT	33
3.3.4	ESTIMATION DU COUT DE LA SOLUTION	34
4	MISE EN ŒUVRE DE LA SOLUTION	35
4.1	CHOIX DES OUTILS	35
4.1.1	CHOIX D'UN OUTIL BI	35
4.1.2	CHOIX DU SGBD	37
4.2	ANALYSE CONCEPTION	38
4.3	CHOIX DES OUTILS	38
4.4	IMPLÉMENTATION	38
	Conclusion	39
	Références	41
A	titre annexe	42
A.1	Presentation de INSBI	42
	CREATION	42
A.1.1	MISSION	42
A.1.2	PRESENTATION	42
A.2	DIAGRAMMES UML	43
	Table des matières	46