

REPUBLIQUE DU BENIN

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

(MESRS)



Université d'Abomey-Calavi (UAC)



Ecole Nationale d'Economie Appliquée et de Management (ENEAM)

MASTER PROFESSIONNEL

Option : Gestion

Filière : Informatique de Gestion

MEMOIRE DE FIN DE FORMATION

THEME

**MISE EN PLACE D'UN DATAWAREHOUSE POUR UN
MEILLEUR SUIVI DES PROGRAMMES D'INVESTISSEMENTS
PUBLICS AU BENIN**

Réalisé et présenté par

Marlène Roseline Sèhouédé GBAGUIDI

Sous la direction de

Maître de stage

M. Kossi AMESSINOU

Chef du Service Informatique au
MDAEP (Ministère du Développement,
de l'Analyse Economique et de la
Prospective)

Directeur de mémoire

Dr Jamâl-Dine DISSOU

Professeur assistant à la FASEG
(Faculté des Sciences Economiques et de
Gestion)

Année : 2013

IDENTIFICATION DU JURY

PRESIDENT: M. DAGBA Théophile

VICE-PRESIDENT: M. GAGNON Bonaventure

DECLARATION D'ENGAGEMENT

L'Ecole Nationale d'Economie Appliquée et de Management (ENEAM) n'entend donner aucune approbation ni improbation aux opinions émises dans ce mémoire. Ces opinions doivent être considérées comme propres à son auteur.

DEDICACE

A mes enfants, Sènamì AGBANGBATIN, Sètondji AGBANGBATIN.

Pour tous les sacrifices que vous avez dû consentir par mes absences dans le cadre de la réalisation de ce travail.

Marlène R. S. GBAGUIDI

REMERCIEMENTS

Un profond merci à tous ceux qui de loin ou de près ont contribué à ce travail, en particulier :

- à tous les **professeurs** de l'Ecole Nationale d'Economie Appliquée et de Management (ENEAM) et au **personnel administratif**, pour avoir contribué à divers niveaux à notre formation ;
- à M. **Jamâl-Dine DISSOU**, mon directeur de mémoire, pour sa disponibilité et ses inestimables conseils;
- à M. **Kossi AMESSINO**, mon tuteur de stage, et à tout le personnel du Ministère du Développement, de l'Analyse Economique et de la Prospective (MDAEP) pour leur accueil et leur assistance ;
- aux **honorables membres du jury** pour avoir accepté apporter leur contribution à ce travail ;
- à mon père **Euloge A. GBAGUIDI** et à ma mère **Isabelle VILLAÇA** pour tous vos sacrifices ;
- à mon époux **Davy Thierry AGBANGBATIN**, pour son soutien indéfectible et ses conseils qui ont été d'un grand secours.

Marlène R. S. GBAGUIDI

LISTE DES SIGLES ET ABREVIATIONS

ADPG	: Autorité de Développement du Périmètre de Glo-Djigbé
AS	: Analysis Services
BEAP	: Bureau d'Evaluation de l'Action Publique
CASPPP	: Cellule d'Analyse et de Synthèse des Performances des Projets et Programmes
CCAD	: Cellule de Coordination de l'Aide au Développement
CePED	: Centre de Partenariat et d'Expertise pour le Développement Durable
CPI	: Centre de Promotion des Investissements
CPMP	: Cellule de Passation des Marchés Publics
DDD	: Directions Départementales du Développement Direction Générale des Investissements et du Financement du
DGIFD	: Développement
DGPD	: Direction Générale des Politiques de Développement
DGSPP	: Direction Générale du Suivi des Projets et Programmes
DN/PAM	: Direction Nationale du Programme Alimentaire Mondial
DPFCAD	: Direction de la Prospection des Financements et de la Coordination de l'Aide au Développement
DPIP	: Direction de la Programmation des Investissements Publics
DPP	: Direction de la Programmation et de la Prospective
DRFM	: Direction des Ressources Financières et du Matériel
DRH	: Direction des Ressources Humaines
DSPEI	: Direction du Suivi des Projets Economiques et d'infrastructures
DSPS	: Direction du Suivi des Projets Sociaux
IDH	: Indice de Développement Humain
IGM	: Inspection Générale du Ministère
INSAE	: Institut National de la Statistique et de l'Analyse Economique
IS	: Integration Services
MDAEP	: Ministère du Développement, de l'Analyse Economique et de la Prospective
OCS	: Observatoire du Changement Social
OMD	: Objectifs du Millénaire pour le Développement
OSD	: Orientations Stratégiques de Développement
PCC	: Plans de Consommation des Crédits
PIP	: Programmes d'Investissements Publics
PPBS	: Chaîne de Planification, de Programmation, de Budgétisation et de Suivi des investissements publics
PPM	: Plans de Passation des Marchés
PTA	: Plans de Travail Annuel
PTF	: Partenaires Techniques Financiers
RS	: Reporting Services
SA	: Secrétariat Administratif

SCRP : Stratégies de Croissance Pour la Réduction de la Pauvreté
SI : Service Informatique
SP/CTD : Secrétariat Permanent de la Commission Technique de Dénationalisation
SPD : Service de Pré-archivage et de la Documentation
SRU : Service des Relations avec les Usagers

LISTE DES TABLEAUX

TABLEAU 1 : REPARTITION ET CLASSIFICATION PAR ORDRE DECROISSANT DU PIP GESTION 2010 PAR OSD	12
TABLEAU 2 : TABLEAU DE COMPARAISON ENTRE LES MODELES OLAP ET OLTP	27
TABLEAU 3 : LISTE DES FICHIERS UTILISES PAR SOURCE	58
TABLEAU 4 : BESOINS FONCTIONNELS PAR GROUPE D'UTILISATEURS.....	60
TABLEAU 5 : BESOINS REGROUPES PAR THEMES D'ANALYSE ET PROCESSUS CONCERNES	61
TABLEAU 6 : DESCRIPTION DES TABLES DE DIMENSION	64
TABLEAU 7 : MESURES DE LA TABLE DES FAITS	65
TABLEAU 8 : DESCRIPTION DES TABLES DE DIMENSION DU DATAMART SUR L'EVOLUTION DES PREVISIONS D'INVESTISSEMENTS PUBLICS.....	68
TABLEAU 9 : DESCRIPTION DES MESURES DU DATAMART SUR L'EVOLUTION DES PREVISIONS D'INVESTISSEMENTS PUBLICS	69

LISTE DES FIGURES

FIGURE 1 : DIFFERENTES ETAPES DU PROCESSUS DE PLANIFICATION DU DEVELOPPEMENT AU BENIN	7
FIGURE 2 : EVOLUTION DES PREVISIONS D'INVESTISSEMENTS PUBLICS DE 2005 A 2010 EN MILLIONS DE FCFA	10
FIGURE 3 : REPARTITION DU PIP GESTION 2010 PAR SOURCE DE FINANCEMENT	11
FIGURE 4 : REPARTITION DU PIP GESTION 2010 PAR GRAND SECTEUR.....	11
FIGURE 5 : ARCHITECTURE DECISIONNELLE	25
FIGURE 6 : ARCHITECTURE MATERIELLE SMP	32
FIGURE 7 : ARCHITECTURE MATERIELLE MPP	33
FIGURE 8 : ARCHITECTURE MATERIELLE NUMA.....	34
FIGURE 9: SCHEMA DU CUBE DIMENSIONNEL.....	42
FIGURE 10 : SCHEMA GENERIQUE DE L'ETOILE	46
FIGURE 11 : CLASSIFICATION DES ENTITES	49
FIGURE 12 : CYCLE DE VIE DIMENSIONNELLE DE KIMBALL	53
FIGURE 13 : CARTOGRAPHIE DES SOURCES	57
FIGURE 14 : BUX MATRIX	62
FIGURE 15 : TABLEAU RECAPITULATIF DES DIMENSIONS DU DATAMART SUR L'EXECUTION DES PROJETS.....	66
FIGURE 16 : SCHEMA EN ETOILE SUR L'EXECUTION DES PROJETS.....	67
FIGURE 17 : TABLEAU DES DIMENSIONS ET FAIT DU DATAMART EVOLUTION DES PREVISIONS D'INVESTISSEMENTS PUBLICS	70
FIGURE 18 : SCHEMA EN ETOILE SUR L'ÉVOLUTION DES PREVISIONS D'INVESTISSEMENTS PUBLICS.	71
FIGURE 19 : ARCHITECTURE RETENUE DANS LE CADRE DE NOTRE ETUDE	74
FIGURE 20 : PAGE D'ACCUEIL DU GESTIONNAIRE DE RAPPORT	80
FIGURE 21 : AFFICHAGE DES RAPPORTS DISPONIBLES DANS LE RAPPORT PRS_RAPPORT	80

FIGURE 22 : AFFICHAGE DU RAPPORT DANS LE GESTIONNAIRE DE RAPPORTS.....	81
FIGURE 23 : SAISIE DU PARAMETRE DE RAPPORT DANS LE GESTIONNAIRE DE RAPPORT	81
FIGURE 24 : AFFICHAGE DU RAPPORT REPARTITION PAR SOURCE DE FINANCEMENT	82
FIGURE 25 : SELECTION DU FORMAT D'EXPORT	82
FIGURE 26 : EXECUTION DE LA COMMANDE D'EXPORT SOUS EXCEL	83
FIGURE 27 : AFFICHAGE DU RAPPORT EXPORTE SOUS EXCEL	83
- FIGURE 28: CONNEXION A LA BASE DE DONNEES DIMENSIONNELLE	84
FIGURE 29 : CHOIX DU CUBE A UTILISER COMME SOURCE DE DONNEES	85
FIGURE 30 : ENREGISTREMENT DU FICHER DE CONNEXION	85
FIGURE 31 : CREATION D'UN TABLEAU DE DONNEES MONTRANT QUELQUES INDICATEURS PAR STRUCTURE.....	86

RÉSUMÉ

Les pays en voie de développement peinent à trouver des solutions adéquates à leurs divers problèmes de développement. Ils constituent des lieux où différentes stratégies et théories du développement ont été appliquées ou essayées. La difficulté apparente de progrès amène l'esprit critique à se demander s'il n'est pas temps de repenser ces stratégies. Pour repenser ces stratégies, il est primordial de disposer des informations fiables traitées à temps. Ceci demande alors à nos pays tant au niveau organisationnel qu'au niveau informatique, de mettre en place des outils de collectes d'informations, de traitement et de présentation de rapports spécifiques pour un meilleur suivi des stratégies. Actuellement, la plupart des pays passent par la mise en place d'un planning de projets d'investissements publics et d'un outil pour son évaluation. Les Investissements Publics constituent l'ensemble des dépenses engagées par l'Etat et les collectivités locales en équipement collectif, également désigné comme infrastructures publics.

Au Bénin, l'instrument d'opérationnalisation des politiques nationales et sectorielles de développement économique et social est le Programme d'Investissements Publics (PIP). Le PIP est constitué de l'ensemble des projets que l'Etat envisage de réaliser au cours d'une année. Il est évalué à l'aide d'un rapport donnant l'état d'avancement des projets qui y sont inscrits. La réalisation du rapport de suivi est confrontée à des difficultés se rapportant tant à la collecte des données qu'à leurs traitements. Entre autres, nous pouvons relever l'accès difficile aux informations présentes dans les bases opérationnelles et l'inexistence d'une structure permettant la réalisation des situations standards et avancées pour des prises de décision. Afin de permettre aux dirigeants du Bénin, d'accéder facilement aux informations existantes et de les traiter suivant les besoins d'analyses pour un meilleur suivi du PIP, la construction d'un "datawarehouse" est proposée. Cet instrument permet de regrouper un ensemble d'informations dans une base unique accessible aux non informaticiens, et d'assurer ainsi l'accès facile aux données pour la réalisation de divers états d'analyse. Il propose aussi l'exportation des différents états via des formats voulus comme Microsoft Excel, Microsoft Word etc.

ABSTRACT

Developing countries are struggling to find suitable solutions to their development problems. They are places where various development strategies and theories have been tested or implemented. The obvious difficulty to progress leads the critical mind to wonder if it isn't about time we rethought these strategies. To do so, it is crucial to have reliable information processed on time. This requires from our countries to both at their organizational level and at the IT level to set up tools so as to collect information, to process information and to present specific reports for a better monitoring strategies. Currently, majority of African countries through the establishment of public investment projects planning and a tool for their evaluation. Public Investment includes all the expenses incurred by the State and local authorities in public facilities, also called a public infrastructure.

In Benin, the operationalization instrument of national and sectoral policies for economic and social development is the Public Investment Program (PIP). PIP include of all the projects that the state intends to implement in a year. It is evaluated with a report that gives the state of projects that are enrolled. The realization of this tracker is facing difficulties related to both the data collection and their treatments. Among those problems, we distinguish the difficulty to access information in the operational databases on the one hand; and the other hand the lack of a structure that helps to realize the standard and advanced situations for decision making. In other to enable the Benin authorities, to easily access to existing information and process then according to analyzes needs for a better monitoring of PIP, the building of a " data warehouse " is suggested. This instrument is used to gather a set of information in a single accessible to non-computer based users and to therefore activate easy access to information for the realization of various analytical statements. It also provides export of various reportings via formats like Microsoft Excel, Microsoft Word etc.

SOMMAIRE

INTRODUCTION GENERALE	1
PARTIE I : CONTEXTE DE L'ETUDE	4
CHAPITRE I : INVESTISSEMENT PUBLIC	5
<i>Section 1 : Généralités sur l'investissement public.....</i>	<i>5</i>
<i>Section 2 : Etude du système</i>	<i>12</i>
CHAPITRE II : ETAT DE L'ART	23
<i>Section 1 : Présentation générale du datawarehouse</i>	<i>23</i>
<i>Section 2 : Etudes antérieures et modélisation dimensionnelle.....</i>	<i>37</i>
PARTIE II : ETUDE CONCEPTUELLE.....	52
CHAPITRE III : TRAVAUX REALISES.....	53
<i>Section 1 : Conception dimensionnelle</i>	<i>53</i>
<i>Section 2 : Conception physique</i>	<i>72</i>
CHAPITRE IV : CREATION DU MODELE PHYSIQUE	75
<i>Section 1 : Mise en place des bases de données</i>	<i>75</i>
<i>Section 2 : Construction de l'entrepôt de données</i>	<i>78</i>
CONCLUSION GENERALE	94
BIBLIOGRAPHIE	97

INTRODUCTION GENERALE

Le développement économique nécessite la création de richesses et entraîne généralement, une progression du niveau de vie des habitants. On mesure le développement d'un pays par des indices statistiques tels que l'Indice de Développement Humain (IDH), le revenu par habitant, le taux d'analphabétisme, l'accès à l'eau, etc.

Pour relancer l'Economie et parvenir au développement rapide et durable au Bénin, le gouvernement élabore des politiques nationales et sectorielles d'investissements publics sur le long, moyen et court terme. L'investissement public représente l'ensemble des dépenses engagées par l'Etat et les collectivités locales en équipement collectif, que l'on désigne également par infrastructures publiques.

L'opérationnalisation de ces politiques se fait grâce au Programme d'Investissement Public (PIP). Il permet d'avoir à court terme, l'état d'exécution des politiques sectorielles définies et par ricochet de connaître le niveau de développement de l'Etat béninois. Il est suivi et évalué par le Ministère du Développement, de l'Analyse Economique et de la Prospective (MDAEP) qui élabore un rapport d'avancement des projets et programmes d'investissements. Cette tâche rencontre des difficultés tant au niveau de la collecte des données qu'au niveau de leurs traitements. Les problèmes rencontrés au niveau des traitements des données s'expriment entre autres par :

- l'accès difficile aux données existantes dans des bases opérationnelles hétérogènes qui ne communiquent pas entre elles;
- la lourdeur remarquée dans l'élaboration des situations spécifiques demandées et l'impossibilité à réaliser des analyses avancées ;
- la vétusté et le manque de matériel informatique.

L'étude des différentes difficultés rencontrées a permis de proposer la mise en place d'un datawarehouse ou Entrepôt de Données (ED) qui permettra de :

- regrouper toutes les données nécessaires issues des sources opérationnelles diverses dans un entrepôt de données;
- élaborer les états spécifiques et d'analyses avancées dans des délais réglementaires ;

- archiver les données issues des bases opérationnelles sur une période d'au moins cinq (5) ans.

Etant donné qu'« un entrepôt de données ne s'achète pas, elle se construit »¹ (Inmon, 1994), notre travail se consacrera à l'étude de la mise en place d'un datawarehouse pour l'Etat béninois, qui lui permettra d'effectuer un meilleur suivi des PIP. Le datawarehouse est une base géante permettant de regrouper des données issues des sources diverses à des fins d'analyse. Il est destiné à la prise de décision. L'alimentation de cette base sera coordonnée par le service informatique du Ministère du Développement, de l'Analyse Economique et de la Prospective (MDAEP).

Ainsi, tout en tenant compte de l'importance de la construction d'un datawarehouse pour le MDAEP, nous avons choisi de focaliser notre recherche sur le thème : "MISE EN PLACE D'UN DATAWAREHOUSE POUR UN MEILLEUR SUIVI DES PROGRAMMES D'INVESTISSEMENTS PUBLICS AU BENIN".

Pour mieux circonscrire la thématique objet du présent mémoire, notre étude portera sur deux (2) parties qui seront consacrées :

- en première partie, à la présentation du contexte d'étude et à la présentation du datawarehouse ;
- en deuxième partie, aux travaux effectués dans la mise en place du datawarehouse et à la présentation du prototype réalisé.

¹ Cité par Antonio Boko (2008) : « Etude et Mise en œuvre d'un datawarehouse pour l'analyse du trafic de communication GSM de MOOV BENIN » Mémoire professionnel en Informatique de Gestion, ENEAM.

PARTIE I : CONTEXTE DE L'ETUDE

Cette partie est consacrée en premier lieu à la présentation du sujet choisi. Les différentes informations nécessaires pour la compréhension du concept de datawarehouse sont ensuite développées dans un second temps.

CHAPITRE I : INVESTISSEMENT PUBLIC

Ce chapitre est consacré à la présentation des différents éléments qui permettront la compréhension du sujet.

Section 1 : Généralités sur l'investissement public

Qu'est-ce-que l'investissement public du point de vue général et au Bénin en particulier ? Comment est-t- il suivi ?

Cette partie est élaborée pour répondre aux différentes questions posées ci-dessus.

Paragraphe 1 : Définition de l'investissement public

1 Définition général de l'Investissement public

Les **investissements publics** sont constitués d'une part des dépenses publiques dites de « consommation collective », investissements considérés comme « non productifs » et en tout cas comme relevant d'une production « non marchande » (réseau routier, justice, police, éclairage public, enseignement public et recherche, défense nationale, etc.), et d'autre part par la subvention d'une partie de l'accumulation du capital technique des entreprises, qu'elles soient publiques ou privées.²

2 Investissement Public au Bénin

2.1 Planification du développement au Bénin

Dans le souci d'améliorer la planification de son développement, le Bénin a adopté une approche à quatre (4) niveaux articulés les uns aux autres à savoir :

- les études prospectives pour le long terme ;

² Trader Finance, Guide Economie, <http://economie.trader-finance.fr/investissement+public/>,
Copyright © 2008-2012

- la planification à moyen terme ;
- les programmes d'investissement triennaux glissants ;
- et la budgétisation.

La présentation des différentes phases de la planification du développement au Bénin sera réalisée dans les lignes qui suivent.

2.1.1 Etudes prospectives

Les études prospectives offrent à moyen terme un cadre stratégique consensuel, tandis que les plans à moyen terme s'exécutent à travers des politiques et des programmes d'investissements, annuellement réajustés. Le rôle de la prospective est d'aider les différents acteurs de développement à comprendre les possibilités de choix et à prendre des décisions averties. Ainsi, grâce aux Etudes Nationales de Perspectives à Long Termes (NLTPS), le document de politique à long terme "Bénin-ALAFIA 2025" a été élaboré et présente le projet de société du Bénin de 2000 à 2025. " Bénin-ALAFIA 2025" est une vision, un projet de société qui inspire tous les acteurs de développement (gouvernement, secteur privé, société civile et partenaires au développement) à élaborer leurs programmes d'actions.

2.1.2 Planification

La planification à moyen terme permet au gouvernement de disposer d'un cadre stratégique de référence en vue d'offrir à tous les acteurs les repères nécessaires pour l'action. Ce niveau permet également de faire des choix fondamentaux, de natures structurelles qui sont inscrits dans une durée de 10 ou 15 ans. Convaincu de cette nécessité, le gouvernement du Président Boni YAYI a décidé d'élaborer les Orientations Stratégiques de Développement (OSD) pour la période 2006-2011. Un nouveau document des OSD est en cours d'élaboration. De même, des stratégies sectorielles sont conçues et représentent aussi d'autres outils à moyen terme. Les Objectifs du Millénaire pour le Développement (OMD) constituent un document international qui présente des directives pour impulser un meilleur développement aux pays en voie de développement. Le Bénin l'ayant adopté, dans le

but de relever les défis de la compétitivité de l'économie, de l'attractivité de l'espace national, a considéré certains points dans l'élaboration des OSD.

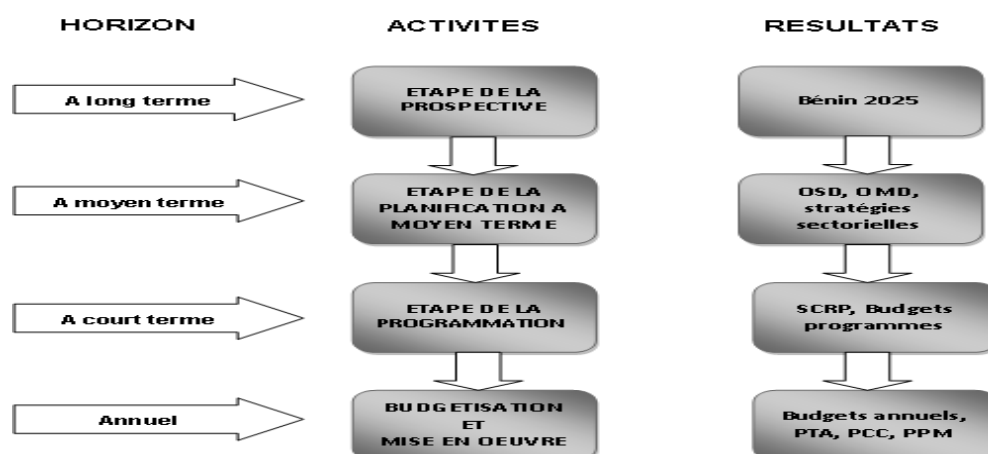
2.1.3 Programmation

La programmation permet de décliner les outils de planification sur une période de 3 à 5 ans. Les outils programmatiques élaborés par le Bénin sont les Stratégies de Croissance pour la Réduction de la Pauvreté (SCRP) et les Budgets Programmes Triennaux (BPT).

Chaque SCRП constitue un cadre de référence du gouvernement en matière de dialogue avec les Partenaires Techniques Financiers (PTF) et se veut être un cadre d'opérationnalisation des OSD. De plus, la SCRП contribue également à la réalisation des OMD et servira de mécanisme de mobilisation et de coordination des ressources conformément à la « Déclaration de Paris » (Cf. Annexe 1). Le BPT permet de définir les actions sur une période glissantes de trois (3) ans. C'est à ce niveau que les programmes d'investissements sont élaborés pour le court terme à travers la définition de projets à réaliser chaque année sur la période retenue.

Le **Programme d'Investissements Publics** (PIP) est un instrument d'opérationnalisation des politiques nationales et sectorielles de développement économique et social du Bénin. Il est constitué de l'ensemble des programmes et projets que l'Etat envisage de réaliser au cours d'une période donnée. La figure 1 présente un récapitulatif de la planification de développement du Bénin.

Figure 1 : Différentes étapes du processus de planification du développement au Bénin



Source : DGPD/DEPS, 2010

1.1.1 Budgétisation

Pour mieux exécuter la SCRP, une étape de budgétisation et de mise en œuvre annuelle est faite et permet d'obtenir les budgets annuels, les Plans de Travail Annuel (PTA), les Plans de Consommation des Crédits (PCC) et les Plans de Passation des Marchés (PPM).

2.2 Acteurs de l'élaboration de la planification du développement

Les différentes activités élaborées pour la planification du développement au Bénin sont réalisées au Ministère du Développement, de l'Analyse Economique et de la Prospective (MDAEP). Il centralise toutes les informations venant des différents ministères pour effectuer des analyses. Il a pour mission d'impulser et de piloter le développement économique et social.

En nous référant au décret n°2008-107 du 10 mars 2008, le MDAEP est chargé entre autres de:

- d'animer la réflexion stratégique, d'élaborer les politiques et stratégies de développement national, de suivre leur mise en œuvre et d'assurer leur évaluation ;
- de centraliser et de promouvoir les projets de développement intégrateurs ;
- de contribuer à l'élaboration des stratégies de promotion du développement régional et local ;
- etc.

Ces différentes attributions font de lui un ministère de référence. Le MDAEP, à travers sa mission et ses attributions, se retrouve au centre de toutes les questions de développement. De ce fait, il travaille en étroite collaboration avec les autres ministères, les institutions et autres organismes à travers les projets et politiques de développement. L'organigramme du ministère se trouve en annexe 2. Dans le fonctionnement de ce ministère, les directions techniques jouent un rôle primordial. Elles assurent l'essentiel des tâches nécessaires pour permettre l'analyse économique et par ricochet l'étude prospective. Il traite ainsi de tous les projets, fait leur suivi et

élabore des rapports sur l'état d'avancement des projets inscrits au programme du gouvernement au cours d'une année. De plus, il analyse l'évolution des indicateurs de performance retenus dans les programmes dans le but de mieux définir les actions à mettre en œuvre dans les prochains programmes à élaborer.

Paragraphe 2 : Base de l'étude sur l'investissement public

1 Présentation générale du rapport d'avancement

Dans le souci d'atteindre les objectifs visés par la mise en œuvre des projets, il est élaboré par le MDAEP un rapport périodique de leur avancement. Le rapport sur l'état d'avancement des projets et programmes inscrits au Programme d'Investissements Publics (PIP) est l'un des instruments techniques d'évaluation des niveaux de réalisation des projets de développement au Bénin. Il permet de prendre des décisions idoines à temps pour corriger les insuffisances notées dans la mise en œuvre des projets ou Programmes d'Investissements Publics en vue de l'atteinte des objectifs de développement fixés. Il présente, à la fin du trimestre, le point de l'exécution des actions inscrites au Programme d'Investissements Publics (PIP). Les indicateurs de suivi s'énumèrent entre autres par :

- le nombre de projets et programmes anciens et nouveaux par année ;
- le coût global des projets inscrits au PIP de l'année en cours ;
- le coût prévisionnel des projets inscrits au PIP ;
- le montant engagé par rapport au coût prévisionnel ;
- le taux d'engagement ;
- le montant ordonnancé ;
- le taux d'exécution du montant ordonnancé ;
- le taux de réalisation physique ou taux d'exécution physique ;
- le taux d'exécution financière, ou taux de consommation financière ;
- la réalisation physique ;
- le montant exécuté par trimestre ;
- le montant réalisé ;
- les sources de financement.

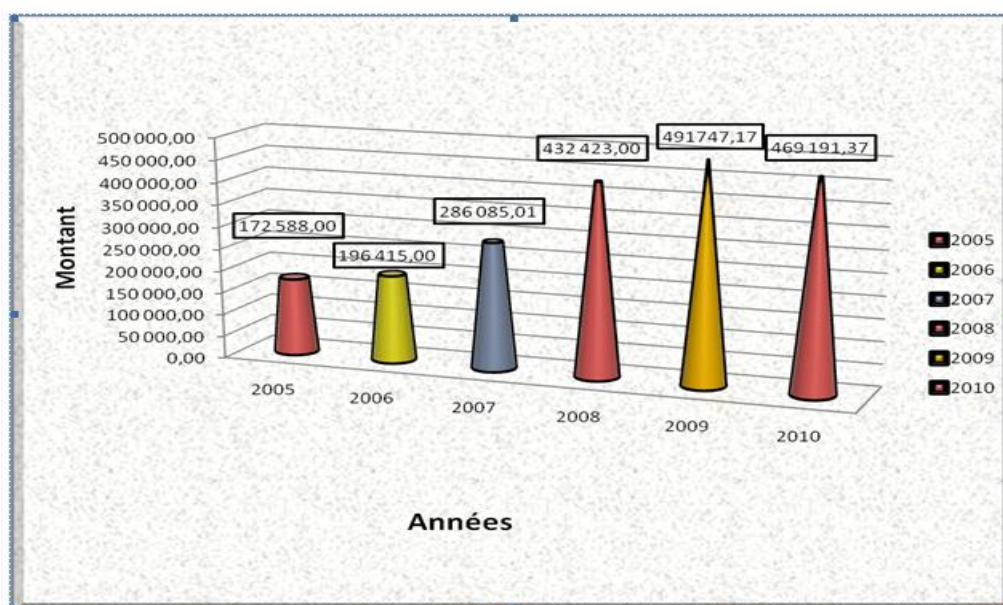
2 Présentation technique du rapport d'avancement

Le dépouillement et le traitement des informations consistent à recenser les fiches par ministère, à saisir les informations sur les réalisations physiques et financières suivant les tableaux conçus à cet effet, à vérifier les montants et les taux, à les réajuster le cas échéant. Les informations physiques et financières sont analysées et illustrées par des tableaux et graphiques. Nous vous présenterons quelques états représentés dans le rapport d'avancement.

2.1 Présentation du portefeuille actif

Cette situation renseigne sur le nombre de projets et programmes, le coût global des projets inscrits pour la période et le coût prévisionnel de la période. Elle se présente comme suit :

Figure 2 : Evolution des prévisions d'investissements publics de 2005 à 2010 en millions de FCFA



Source : DGSPP/MPDEPP-CAG, Juillet 2010

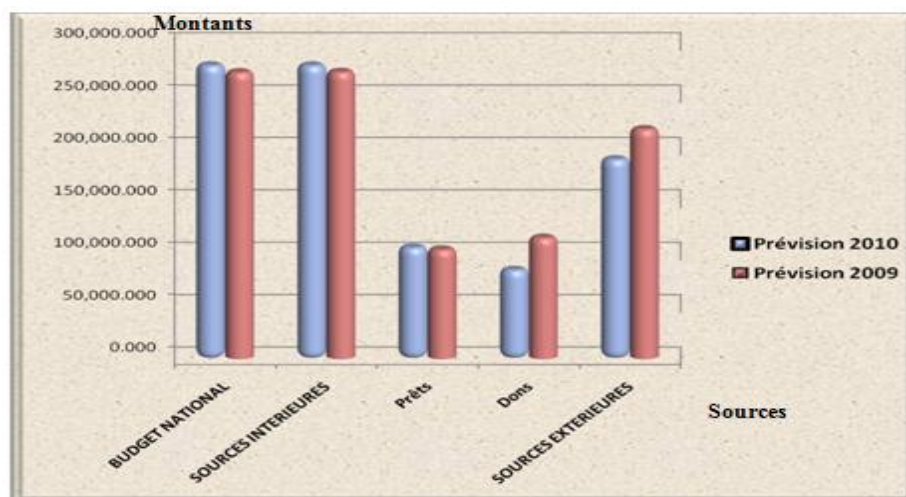
2.2 Répartition du PIP par gestion

Cet état présente les prévisions financières d'investissements publics au titre de la gestion en cours par source de financement, par secteur, par ministère et institution de l'Etat et par OSD.

2.2.1 Par source de financement

Cet état est représenté sous forme de tableau et de graphe. Le tableau est présenté en annexe 3. Le graphe est présenté à la figure 3.

Figure 3 : Répartition du PIP gestion 2010 par source de financement

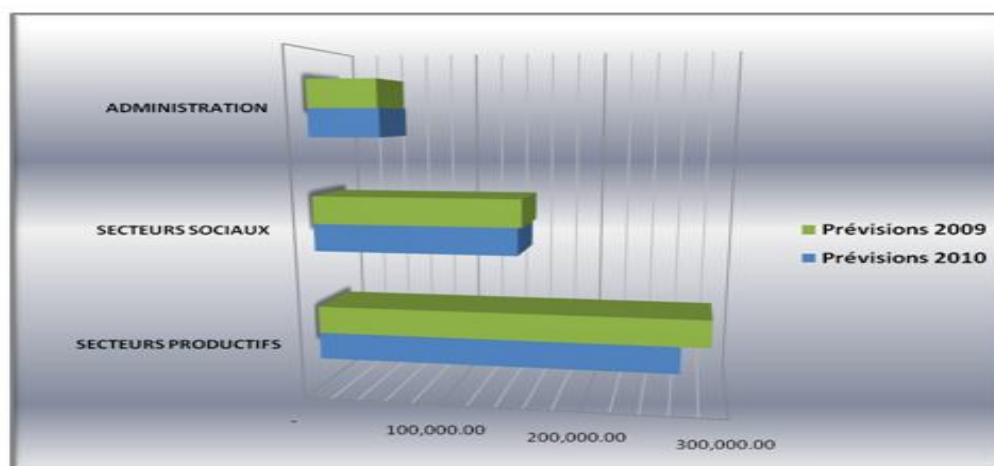


Source : DGSP/MPDEPP-CAG, Juillet 2010

2.2.2 Par secteur

Il se présente aussi sous forme de tableau et de graphe. Le tableau est présenté en annexe 4. Le graphe est représenté à la figure 4.

Figure 4 : Répartition du PIP gestion 2010 par grand secteur



Source : DGSP/MPDEPP-CAG, Juillet 2010

2.2.3 Par OSD

Le tableau suivant présente la répartition du PIP gestion 2010 par Orientation Stratégique de Développement (OSD). Il se présente aussi sous forme graphique.

Tableau 1 : Répartition et classification par ordre décroissant du PIP gestion 2010 par OSD

N° D'ORDRE	ORIENTATIONS STRATEGIQUES DE DEVELOPPEMENT (OSD)	N° de l'OSD	PREVISIONS 2010 EN MILLIARD DE FCFA	POIDS (en %)
1	Développer les infrastructures économiques et sociales	OSD 4	201,40	43,47
2	Promouvoir le renouveau économique	OSD 3	119,70	25,84
3	Renforcer le capital humain	OSD 5	70,50	15,22
4	Reconstruire une Administration au service de l'intérêt général et du développement du secteur privé	OSD 1	34,10	7,36
5	Assurer un développement équilibré et durable de l'espace national à travers le développement à la base	OSD 6	33,20	7,17
6	Assainir le cadre macro économique et maintenir sa stabilité	OSD 2	4,40	0,95
	TOTAL		463,30	100,00

Source : DPIP/DGIFD/MPDEPP-CAG, Juillet 2010

2.3 Point d'exécution financière et réalisations physiques du portefeuille actif

Cette situation présente l'exécution financière par trimestre des projets du portefeuille actif par source de financement, le bilan du trimestre du PIP par secteur, ou par ministère et institution de l'Etat. Il présente aussi les dépenses d'investissements réalisées hors PIP. A l'annexe 6, le point d'exécution financière par source de financement et le bilan du trimestre du PIP par secteur sont représentés.

Le point des réalisations physiques renseigne sur les réalisations physiques effectuées, le montant réalisé et le taux d'exécution physique.

Section 2 : Etude du système

La présentation des difficultés rencontrées dans l'élaboration du suivi du Programme d'Investissements Publics (PIP) fait l'objet de cette section. La solution proposée en vue de résoudre ces obstacles, est ensuite abordée.

Paragraphe 1 : Difficultés rencontrées et approches de solution

1 Difficultés rencontrées dans le suivi du PIP

L'élaboration du rapport d'avancement du PIP fait partie des attributions de la Direction Générale du Suivi des Projets et Programmes (DGSP). En effet, selon ses attributions, elle est chargée de coordonner la mise en œuvre de tous les projets et programmes des ministères, de suivre l'exécution des projets de développement et d'assurer l'évaluation de leur efficacité, de leur efficacité et de leur performance.

Trimestriellement, une tournée de collecte auprès des Directions de la Programmation et de la Prospective (DPP) de tous les ministères et des institutions de l'Etat est effectuée et permet de disposer des données sur l'exécution physique et financière des projets et programmes du PIP de la gestion en cours. De plus, les informations ayant trait à la réalisation financière sont également collectées auprès d'autres acteurs de la chaîne de Planification, de Programmation, de Budgétisation et de Suivi des investissements publics (PPBS), notamment la DGIFD du MDAEP, la Direction Générale du Budget (DGB) et la Caisse Autonome d'Amortissement (CAA) du Ministère de l'Economie et des Finances (MEF).

Les difficultés rencontrées dans l'élaboration de ce rapport auprès des acteurs de la chaîne PPBS sont d'ordres multiples.

1.1 Problèmes au niveau du traitement des données du PIP

L'élaboration du PIP s'effectue au niveau de la DGIFD précisément par la DPIIP. Le PIP réalisé et voté par les députés de l'assemblée, est mis à disposition de la DGSP pour le suivi.

1.1.1 Planification et programmation du PIP

La DPIIP s'est dotée d'applications informatiques, lui permettant d'accélérer ses travaux. Elle utilise actuellement le logiciel SIAPIP (Système Intégré d'Analyse et de Programmation des Investissements Publics) pour faire l'analyse et la programmation du PIP.

Il faut rappeler que la DPIP avait initialement utilisé le logiciel ANAPROG. C'est un logiciel développé sous ACCESS 2000 par un cabinet Ivoirien. Il a servi pendant neuf (9) ans à faire la programmation et le suivi de Projets d'Investissements Publics. Face aux difficultés d'accès aux codes sources et à la nécessité de créer une nouvelle base chaque année, les dirigeants ont décidé de mettre en place le SIAPIP.

Le logiciel SIAPIP est orienté web et développé sous php5.2.8. Il fournit toutes les informations sur les projets retenus pour le PIP de l'année en cours. Sa base de données est sous MySQL, version 3.1.1. Il permet aussi de faire l'analyse et la programmation de tous les Projets d'Investissements Publics des institutions et ministères de l'Etat par année et donne plus de possibilité pour des modifications éventuelles. L'analyse des données du PIP dans SIAPIP peut s'étendre sur une période de trois (3) ans.

Cependant, les informations contenues dans la base Access n'ont pas pu être intégrées dans la base MySQL. Aujourd'hui, il est difficile de fournir des informations pour produire des comptes rendus d'analyse avancés sur une période de plus de cinq (5) ans. De plus, les informations récupérées au niveau de SIAPIP sont envoyées sur fichier Excel à la DGSP qui reprendra les saisies dans différents fichiers Excel. Ce processus entraîne une répétition dans la mise en œuvre des tâches à exécuter et rallonge les délais de traitement des données. Le traitement des informations se fait donc à partir des données disponibles dans des bases de données hétérogènes, ne communiquant pas entre elles et rendant répétitives les tâches à exécuter.

1.1.2 Suivi du PIP

Pour le suivi du PIP, la DGSP utilise un fichier Excel dans lequel les synthèses de toutes les informations nécessaires sur le PIP de la gestion en cours sont effectuées. Les tâches de saisie des informations reçues à la DGSP rallongent les temps d'exécution du processus. De plus, elle effectue la saisie des informations de suivi récupérées auprès de la DGB et des ministères sectoriels. Il faut noter qu'il existe un logiciel nommé SIGFIP (Système Intégré de Gestion des Finances Publiques) gérant l'exécution du budget national et est disponible au MEF. D'autres

outils applicatifs sont utilisés dans certains ministères. La DGSPP ne possédant aucun outil applicatif en liaison avec les diverses sources d'informations, se trouve obligé de reprendre la saisie des informations indispensables pour le suivi.

De plus, l'archivage des fichiers Excel réalisés trimestriellement ne s'effectue pas pour faciliter les analyses dans le temps. A ce niveau aussi, établir une situation du PIP sur une période de cinq (5) ans constitue un problème. Cependant, le besoin d'analyse est toujours présent.

1.2 Problèmes au niveau du matériel

Les difficultés recensées au niveau matériel dans toutes les structures concernées, sont dues à l'insuffisance d'outils informatiques (ordinateurs, serveurs, etc.) et à la vétusté du matériel existant. La performance dans les traitements de données, repose aussi sur l'efficacité du matériel.

1.3 Problèmes au niveau du personnel

Le personnel informatique de tout le MDAEP, souffre d'un manque de formation continue dans les nouvelles approches technologiques. Il est nécessaire de former le personnel informatique pour le rendre performant et prêt à assister les autorités dans leurs différents besoins de solutions informatiques.

2 Analyse du système et approche de solution

2.1 Analyse du système

L'étude des processus de planification, de programmation et de suivi du PIP montre que leur réalisation est confrontée à des obstacles. Nous pouvons énumérer :

- manque de communication entre les applications ou sources de données clés ;
- difficultés au niveau de l'analyse de données collectées et existantes ;
- accès difficile à l'information, allongeant ainsi le temps des traitements ;

- mauvaise stratégie d'historisation des données et conservation des informations de manière disparate ;
- difficulté au niveau de la collecte des informations ;
- insuffisance d'outils informatiques ;
- manque de formation continue des informaticiens.

Il est indispensable de résoudre tous ces différents problèmes pour réaliser des comptes-rendus performants. Les divers problèmes énumérés peuvent se regrouper en 3 questions de réflexions essentielles que sont :

- 1- une refonte totale du matériel informatique et la formation continue du personnel informatique ne contribuerait-t-il pas à l'amélioration du suivi du PIP?
- 2- une mise en place d'une intégration d'applications d'entreprise regroupant toutes les applications existantes et la création de modules supplémentaires n'améliorerait-t-il pas le suivi du PIP?
- 3- une mise en place d'outil décisionnel pouvant intégrer toutes les informations de sources diverses à des fins d'analyses ne contribuerait-t-il pas à une amélioration du suivi du PIP ?

La réponse à chacune de ces questions permettra de connaître la meilleure solution qui s'appliquera dans le temps et contribuera à résoudre le besoin pressant d'un meilleur suivi du PIP en vue d'améliorer les conditions de vie de la population.

2.2 Approche de solution

Cette partie de notre étude représente un point essentiel. Elle permet de proposer une solution pour la contribution de l'amélioration des tâches exécutées pour la production des rapports de suivi du PIP. Nous trouverons des réponses aux différentes questions de réflexions ci-dessus citées.

2.2.1 Améliorer le suivi du PIP par renouvellement du parc et formation continue du personnel informatique.

Le matériel informatique permet de nos jours, d'accélérer les traitements quotidiens effectués. Disposer d'applications informatiques nécessite la possession d'un parc informatique performant. Ainsi, la configuration technique des différents postes ou matériels informatiques contribue à faire fonctionner les applications avec rapidité au cours de leur utilisation ou à ralentir leur fonctionnement.

La formation continue du personnel informatique assure sa montée en compétence. De ce fait, le personnel informatique acquiert une base supplémentaire pour assurer une assistance de qualité aux utilisateurs en leur proposant de meilleurs produits.

De tout ceci, la question du matériel et du personnel informatique ne résout pas directement la question d'amélioration du suivi du PIP. Cependant, résoudre le problème de suivi du PIP sans disposer de matériel ni de personnel informatique performant rendrait, cette solution désuète.

2.2.2 Améliorer le suivi du PIP par intégration d'applications d'entreprise

L'Intégration d'applications d'entreprise ou IAE (en anglais *Enterprise Application Integration, EAI*) est une architecture intergicielle permettant à des applications hétérogènes de gérer leurs échanges. On la place dans la catégorie des technologies informatiques d'intégration métier (*Business Integration*) et d'urbanisation. Sa particularité est d'échanger les données en pseudo temps réel. (WIKI, 2012).

La mise en place d'un IAE nécessite une volonté d'unification de l'intégration des systèmes d'information de l'entreprise. Ses données seront représentées dans l'IAE sous forme d'objet métier (OM). Cela permet ainsi de construire des flux d'information métiers unifiés par lesquels chaque application spécifique peut partager ses informations avec les autres au sein d'une étape de l'organisation de l'entreprise.

L'IAE est utilisée au sein d'une entreprise pour connecter des applications dans le but de créer une communication entre elles. Si l'IAE permet de réaliser une communication entre les applications, elle ne permet pas d'archiver, ni d'analyser les données pour la prise de décision. Elle contribue surtout à réunir des applications opérationnelles pour une meilleure communication permettant l'accès aux données opérationnelles. Elle ne permet pas d'obtenir des données agrégatives pour des analyses décisionnelles. Ainsi, une IAE ne résoudra pas immédiatement le besoin pressant d'amélioration du suivi du PIP. Néanmoins, réaliser une IAE contribue à résoudre la répétition des tâches de récupération des données entre les différentes applications opérationnelles. Elle permettra la mise à jour automatique des nouvelles données dans les applications.

Il est à noter que le MDAEP en collaboration avec le Millenium Challenge Account (MCA), procède à la mise en place d'un système d'information qui favorisera la gestion informatisée du suivi et l'évaluation des Budgets-Programmes à travers l'exploitation des outils d'opérationnalisation (PTA, PCC, PPM) en vue d'atteindre les objectifs fixés. Il s'agit de SHISEPIP (Système Harmonisé et Intégré de Suivi et Evaluation des Projets et Programmes d'Investissements Publics). Ce système regroupera ainsi toutes les informations nécessaires pour réaliser le suivi et l'évaluation de l'ensemble des projets et programmes de développement. Sa mise en place permettra d'assurer une meilleure collecte et traitement des informations sur le PIP. Il constituera une application opérationnelle possédant des informations réelles et régulièrement mises à jour.

2.2.3 Améliorer le suivi du PIP par la mise en place d'un datawarehouse

Une entreprise moderne brasse dans son système d'information d'immenses volumes de données, mais bien souvent, du fait même de ces volumes trop importants, 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, toutes choses que l'on doit connaître pour *prendre de bonnes décisions*. On appelle « aide à la décision », ou bien « le décisionnel », ou encore « *business*

intelligence », un ensemble de solutions informatiques permettant l'analyse des données de l'entreprise, afin d'en dégager les informations qualitatives nouvelles qui vont fonder des décisions, qu'elles soient tactiques ou stratégiques. Ainsi, les outils d'aide à la décision, avant d'aider à la prise de décision, aident d'abord à analyser les données afin *d'identifier les informations macroscopiques pertinentes cachées dans de grands volumes de données*.

Les informations, les données, qui seront à la base du système décisionnel, proviennent souvent de multiples systèmes d'information de l'entreprise. Le datawarehouse est une base de données de consolidation, souvent une base relationnelle relativement standard, mais qui doit accueillir et manipuler de gros volumes d'information. Le datawarehouse concentre de l'information décisionnelle issue de différents systèmes d'information de l'entreprise.

L'analyse des données se fait à travers les rapports proposés. Un rapport (*report* en anglais), est la présentation de données de manière synthétique et lisible, généralement à des fins d'impression. Le rapport fait apparaître des tableaux de chiffres, en gérant entêtes, pieds de pages, sous-totaux, etc. Un *générateur de rapports* est un programme qui permet de définir un rapport selon ses besoins, avec un minimum de programmation, uniquement au moyen d'une interface interactive. Une fois défini, le rapport peut être généré à échéance régulière³.

On peut distinguer deux modes de travail en décisionnel : le mode 'rapport statique' et le 'mode interactif', avec une analyse qui répond interactivement aux demandes des utilisateurs. Ils correspondent à des besoins différents, et parfois complémentaires :

- obtenir une information récurrente, correspondant à une analyse prédéfinie;
- chercher une information en effectuant différentes analyses successives, les résultats de l'une amenant des questions nouvelles, qui demandent une nouvelle restitution de données.

³ Extrait de Smile OpenSource, Livre blanc, Décisionnel Solutions Open Source, Pages9 et11, Edition 2011. <http://www.smile.fr>

Comme évoqué précédemment, cette nouvelle base de données au sein de l'entreprise n'est pas créée pour les bienfaits d'une application opérationnelle directement liée à l'activité de l'entreprise (ex : système comptable, base de données RH, base de données des commerciaux, etc.). Le datawarehouse ne va stocker que les informations clefs de l'entreprise et ne sera dédié qu'aux requêtes d'analyse et de reporting. L'entreprise pourra ainsi analyser ses données sans diminuer les performances de ces outils de production courants. La base décisionnelle va être modélisée dans le but de faciliter les « requêtes ».

Au vue de tout ce qui précède, le datawarehouse permet l'accès facile aux données indispensables pour réaliser les différents rapports de suivi du PIP standard et spécifiques.

Paragraphe 2 : Problématique et objectifs de l'étude

1 Problématique de l'étude

Le MDAEP à travers ses fonctions, procède à la planification, à la programmation, et au suivi des actions ou projets du Bénin. Les différents rapports de suivi et d'analyse de ces actions sont obtenus à partir de diverses bases opérationnelles sous MySQL, sous Access, sous Excel et sous format .Txt. Il n'est donc pas possible à l'utilisateur, de formuler directement des interrogations sauf en s'adressant à la Direction Générale des Investissements et du Financement du Développement (DGIFD) et à la Direction Générale du Suivi de Projets et Programmes (DGSP). Ces directions n'exécutent ces demandes qu'en se rapprochant de leur service informatique qui procède aux manipulations nécessaires pour fournir les résultats. Les contraintes s'expriment en termes de temps de réponse ou de traitements des requêtes pouvant prendre deux à plusieurs jours, voire des semaines pour la livraison des résultats attendus. A cela, il faut ajouter les risques d'incohérence de données ou une impossibilité à satisfaire la requête.

En dehors de requêtes difficiles à satisfaire, il faut noter l'absence d'outils pouvant permettre d'évaluer efficacement l'impact des différentes actions exécutées sur une période donnée au niveau des PIP du fait de la mauvaise stratégie

d'archivage et de sauvegarde des données dans le temps. En effet, le MDAEP, du fait de sa position de centralisation des données liées à l'exécution des projets, brasse d'importants volumes d'informations archivées annuellement. Les informations archivées ne sont pas regroupées ensemble dans un système unique. Ceci ne permet toujours pas de cerner avec efficacité les éléments cachés dans ce grand volume d'information sur une période donnée pour des fins d'analyses avancées.

Il faut rajouter au manque de mécanisme de suivi des éléments précédemment cités, le manque d'automatisation de certains états statistiques trimestriels et annuels adressés aux cadres et décideurs du MDAEP.

D'autre part, face à l'indisponibilité de certaines informations liées aux sources de données, le problème de qualité des données est notable et leur enrichissement s'avère nécessaire. En prenant connaissance de ces différentes difficultés, nous réalisons qu'il y a inexistence d'un outil permettant les de résoudre.

Du point de vue de l'importance accordée au processus de suivi du PIP par les dirigeants de notre pays le Bénin, nous avons proposé la mise en place d'un entrepôt de données ou datawarehouse.

Dans le cadre de notre étude, l'appréhension de la thématique formulée nécessitera la mise en pratique des connaissances en architecture logicielle. Particulièrement, il s'agira de mettre en contribution nos expériences dans la conception et la mise en place de base de données décisionnelle sous Microsoft SQL Server.

Cette étude trouve son intérêt dans l'amélioration qu'elle apporte au système de suivi des PIP grâce à la mise en place de cet outil qui facilitera l'accès aux données existantes dans les différentes bases éparses. Aussi, permettra-t-elle l'autonomie des utilisateurs, leur procurant les moyens de manipuler les données par eux-mêmes (tableaux croisés).

2 Objectifs de l'étude

L'objectif général de cette étude est de mettre en place un datawarehouse en vue d'améliorer le système de suivi des programmes d'investissements publics au Bénin.

Autour de cet objectif général gravitent deux (2) objectifs spécifiques à savoir :

- améliorer l'obtention de reporting varié et exhaustif ;
- raccourcir les délais de production par l'automatisation de la publication régulière par mail et/ou portail d'états statistiques prédéfinis pour une meilleure prise de décision.

CHAPITRE II : ETAT DE L'ART

Ce chapitre sera consacré à présenter, les différentes informations nécessaires pour la compréhension du concept de datawarehouse. Notre objectif est de ressortir les points essentiels à connaître pour une bonne construction d'un entrepôt de données.

Section 1 : Présentation générale du datawarehouse

Pourquoi un datawarehouse ? Qu'est-ce qu'un datawarehouse ? représentent quelques questions auxquelles cette partie donnera des réponses.

Paragraphe 1 : Généralités sur le datawarehouse

Ce paragraphe présente les causes de l'avènement du concept de datawarehouse et sa présentation de manière générale.

1 Datawarehouse (Entrepôt de Données ED)

1.1 Définition

"A data warehouse is a subject-oriented, integrated, non volatile, and time-variant collection of data in support of management's decisions". (B. Inmon, 1996)

En terme clair, un ED est une collection de données **thématiques, intégrées, non volatiles** et **historisées**, organisées pour la prise de décision."

- Le datawarehouse est dit **thématique** lorsque les données collectées doivent être orientées « métier » donc par thèmes ou sujets ;
- le datawarehouse est composé de données **intégrées**, c'est-à-dire qu'un « nettoyage » préalable des données est nécessaire dans un souci de rationalisation et de normalisation;
- les données du datawarehouse sont **non volatiles** ce qui signifie qu'une donnée entrée dans l'entrepôt l'est pour de bon et ne sera pas supprimée ;
- les données du datawarehouse doivent être **historisées**, donc datées. L'organisation des données est conçue pour que les personnes intéressées aient

accès rapidement et sous forme synthétique à l'information stratégique dont elles ont besoin pour la prise de décision.

Une autre définition : *"A data warehouse is a system that extracts, cleans, conforms, and delivers source data into a dimensional data store and then supports and implements querying and analysis for the purpose of decision making."* (Kimball et Caserta, 2004)

- Un datawarehouse est un système qui extrait c'est-à-dire collecte des informations de sources diverses ;
- un datawarehouse nettoie et rend conformes les données ;
- un datawarehouse historise les données dans des bases de données dimensionnelles pour implémenter les requêtes et analyses pour la prise de décision.

Les deux auteurs cités B. Inmon et R. Kimball s'accordent sur le fait que le datawarehouse intègre les données à partir de différents systèmes sources opérationnelles. D'un point de vue architectural, il existe deux manières de l'appréhender :

L'architecture "de haut en bas" : selon Bill Inmon, l'entrepôt de données est une base de données au niveau détail, consistant en un référentiel global et centralisé de l'entreprise. En cela, il se distingue du Datamart, qui regroupe, agrège et cible fonctionnellement les données.

L'architecture "de bas en haut" : selon Ralph Kimball, l'entrepôt de données est constitué peu à peu par les Datamarts de l'entreprise, regroupant ainsi différents niveaux d'agrégation et d'historisation de données au sein d'une même base.

La définition la plus communément admise est un mélange de ces deux points de vue. Le terme "datawarehouse" englobe le contenant et le contenu : il désigne d'une part la base détaillée qui est la source de données à l'origine des Datamarts, et d'autre part, l'ensemble constitué par cette base détaillée et ses Datamarts. De la même manière, les méthodes de conception actuelles prennent en compte ces deux approches, privilégiant certains aspects selon les risques et les opportunités inhérents à chaque entreprise.

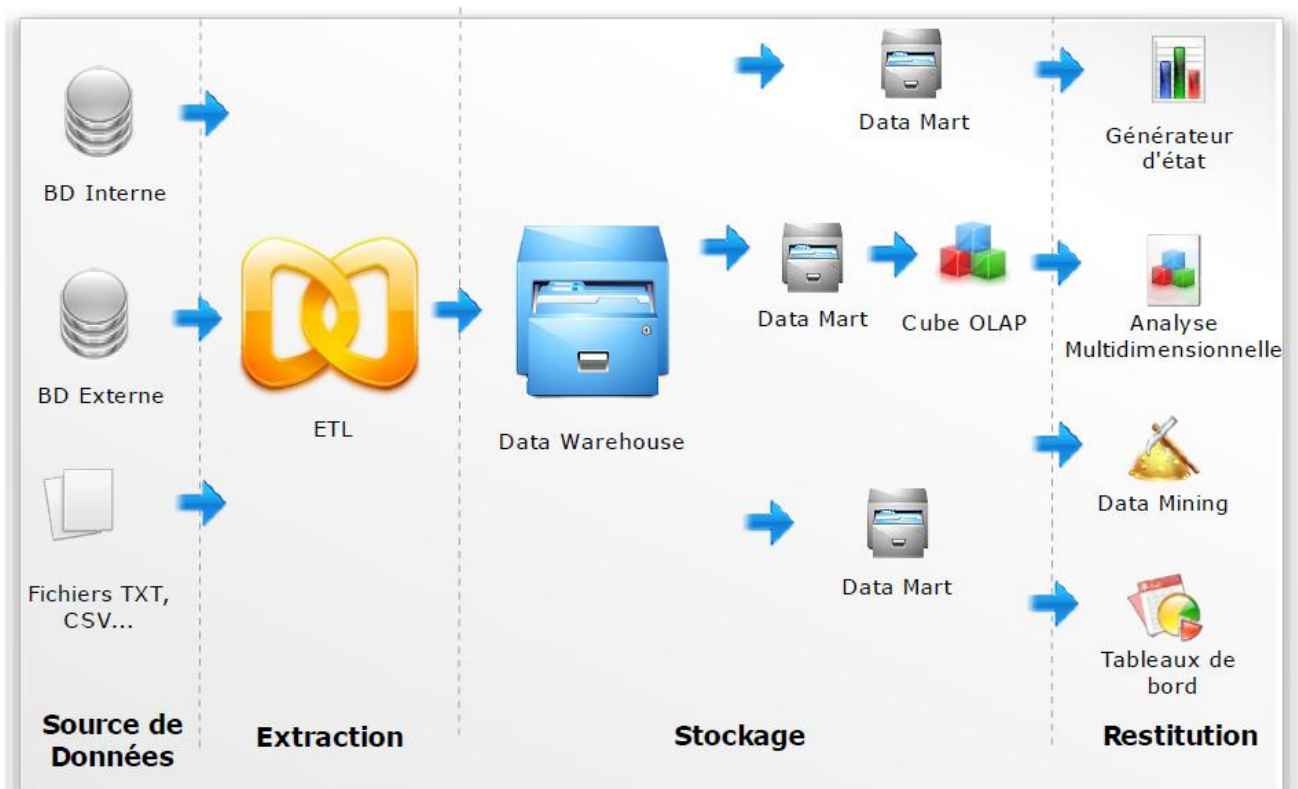
1.2 Architecture d'un datawarehouse

Les systèmes opérationnels, bases de données indispensables, à la vie d'une entreprise, permettent d'avoir une activité journalière (gestion de stocks, base de fournisseurs/clients, etc.). Ceci n'est pas le rôle d'un datawarehouse, il a pour unique but de faciliter la prise de décision en apportant une vue synthétisée de l'ensemble des données de l'entreprise éparpillées dans toutes ces bases opérationnelles.

L'environnement de l'entrepôt de données comporte quatre parties différentes⁴ :

- les applications opérationnelles sources ;
- la préparation des données ;
- la présentation des données ;
- les outils d'accès aux données.

Figure 5 : Architecture décisionnelle



Source : Adullact, 2008

⁴ Ralph Kimball et Margy Ross (2002) : « Entrepôts de données, Guide pratique de modélisation dimensionnelle » (2^e édition), pp 7

1.2.1 Les applications opérationnelles sources

Ce sont les applications opérationnelles qui « capturent » les transactions de l'entreprise. On doit se les représenter comme extérieures à l'entrepôt de données, car nous n'avons vraisemblablement guère ou pas du tout de moyens d'influencer le contenu ou le format des données qu'ils traitent. Les principales priorités de ces applications sources sont la performance des traitements et la disponibilité.

1.2.2 La préparation des données

La partie préparation de l'entrepôt de données est à la fois une zone de stockage et un ensemble de processus couramment appelé ETC, extraction/transformation/chargement (en anglais, ETL :Extract Transform Load). La préparation inclut tout ce qu'il y a entre les applications opérationnelles sources et la présentation 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.

1.2.3 La présentation des données

La zone de présentation des données est le lieu où les données sont organisées, stockées et offertes aux requêtes directes des utilisateurs, aux programmes de reporting et autres applications d'analyse.

1.2.4 Les outils d'accès aux données

Le dernier composant majeur d'un environnement d'entrepôt de données est l'ensemble des outils d'accès aux données. Nous appliquons ce terme général à un ensemble de moyens fournis aux utilisateurs pour exploiter la zone de présentation en vue de prendre des décisions basées sur des analyses. Par définition, tous les outils d'accès aux données font des requêtes sur les données de la zone de présentation. Les requêtes sont évidemment la raison d'être de l'entrepôt de données.

2 Autres concepts

2.1 Analyse multidimensionnelle

OLAP (On-Line Analytical Processing) désigne les bases de données multidimensionnelles ou cubes destinées à l'analyse. Ce terme s'oppose à OLTP (On Line Transaction Processing) qui désigne les systèmes transactionnels. OLAP a été défini par Ted Codd en 1993 au travers de 12 règles (Cf. annexe 7) que doit respecter une base de données si elle veut adhérer au concept OLAP. OLAP est un mode de stockage prévu pour l'analyse statistique des données. Une base de données OLAP peut se représenter comme un cube à N dimensions où toutes les intersections sont pré-calculées.

Afin de se rendre compte des différences entre ces deux modèles, OLAP et OLTP, le tableau suivant est présenté :

Tableau 2 : Tableau de comparaison entre les modèles OLAP et OLTP

Caractéristiques	OLAP	OLTP
Données	Historisées	Courantes
Utilisation	Analyse de l'entreprise	Opérations quotidiennes
Nombre de n-uplets accédés	Millions	Dizaines
Unité de travail	Requête	Transaction
Mode d'accès aux données	Principalement en lecture	Lecture/Ecriture
Type d'utilisateur	Gestionnaire	Employé
Nombre d'utilisateurs	Cent	Mille

Source : Extrait de Hervé COLASUONNO (2001)⁵

2.2 Les différentes approches d'OLAP

2.2.1 Implémentation

Le concept de l'hypercube OLAP peut être utilisé par le biais de différentes interfaces selon les besoins et les capacités. Il en existe actuellement quatre.

⁵ Hervé COLASUONNO (2001) : « Conception d'entrepôt de données » Examen probatoire en informatique, CNAM (Conservatoire Nationale des Arts et Métiers)

- *M-OLAP* : La forme la plus classique car la plus rapide. Elle utilise des tables multidimensionnelles pour sauvegarder les informations et réaliser les opérations.
- *R-OLAP* : Celle qui demande le moins d'investissement. Elle travaille sur des tables relationnelles. Une nouvelle table est créée pour contenir chaque agrégat.
- *H-OLAP (Hybrid OLAP)* : Elle utilise à la fois les tables relationnelles pour stocker les informations brutes, et des tables multidimensionnelles pour les agrégats d'informations prédictives.
- *D-OLAP (Dynamic ou Desktop OLAP)*.

2.2.2 Utilisation

OLAP est un mode de stockage prévu pour l'analyse statistique des données. Une base de données OLAP peut se représenter comme un cube à N dimensions où toutes les intersections sont pré-calculées. La restitution d'information se fait à l'aide de simples requêteurs utilisant le langage MDX (Multidimensional Expressions), d'Executive Information System (EIS), d'applications spécialisées (logiciels d'entreprise), ou dans un tableur (muni d'un plugin de navigation).

La force d'OLAP tient dans le fait que comme toutes les intersections du cube sont calculées ; l'accès à l'information voulue est une opération très rapide, quasiment instantanée. En effet, le résultat voulu se trouve aux croisements des différentes dimensions.

2.3 Autres outils du décisionnel

Au nombre des outils utilisés, nous pouvons citer :

- Les tableaux de bord : les tableaux de bord constituent un outil de pilotage qui porte une vision sur l'évolution d'un processus, afin de permettre aux responsables de mettre en place des actions correctives.
- Le Balanced Score Card : le Balanced Score Card (BSC) ou tableau de bord prospectif diffère du tableau de bord évoqué dans la partie précédente. Le BSC

s'adresse uniquement aux dirigeants de l'entreprise et procure une vision très synthétique de l'activité de l'entreprise.

- Datamining : ou forage de données, consiste à forer dans un grand volume de données afin d'en extraire des informations pertinentes pour le décideur.
- Générateur d'état : permet de réaliser des états, appelés également reporting, qui sont des rapports présentant de manière synthétique et lisible des données, sous forme de tableaux de chiffres, tout en gérant la mise en page (en-tête, pied de pages, etc.).

Paragraphe 2 : Considérations pratiques pour la mise en place d'un datawarehouse

Ce paragraphe présente les différentes approches de mise en place d'un ED dans un premier temps, puis son infrastructure dans un second temps.

1 Les grandes approches de mise en place d'un système datawarehouse

Il existe plusieurs approches pour mettre en place un Datawarehouse.

Il serait utopique de croire qu'il existe une approche idéale. Il faut rechercher l'approche la plus adaptée à son contexte.

1.1.1 Datamarts Indépendants

Selon Kimball la première étape est la conception du modèle dimensionnel pour les datamarts, c'est-à-dire ayant une vue métier. Celui-ci placera les datamarts au centre de l'architecture. Le reste sera composé d'un staging area temporaire. Dans cette approche, on dit que les datamarts sont centraux car ils peuvent contenir à la fois des données atomiques et agrégées, et qu'ils offrent la possibilité de fournir une vue entreprise et une vue métier. Il est à noter que l'implantation des datamarts se fait d'une façon incrémentale et intégrée. Pour finir, les utilisateurs ne peuvent effectuer des requêtes sur le staging area.

1.1.2 Datamarts dépendants ODS

Selon B.Inmon, c'est le datawarehouse qui est au centre de l'architecture. La première étape est donc de concevoir un modèle de datawarehouse mais au niveau entreprise cette fois-ci et non plus métier. Cette architecture sera de type multi-tiers composée du staging area, du datawarehouse et des datamarts. Le staging area sera permanent et les datamarts seront interdépendants. C'est à dire qu'ils communiqueront entre eux.

La principale caractéristique de cette approche réside dans le fait que c'est le datawarehouse qui contient les données atomiques. Il utilise un modèle normalisé de toute l'entreprise. Il est donc orienté entreprise, à l'opposé des datamarts qui eux sont orientés métier. Ces derniers utilisent des modèles dimensionnels orientés métier. Les données y sont stockées sous formes agrégées. Pour finir, les utilisateurs peuvent effectuer des requêtes sur le datawarehouse et sur les datamarts.

1.1.3 Hybride

Cette approche, comme son nom l'indique, est un message des deux premières approches. On commence par concevoir un modèle de données de l'entreprise en même temps que les modèles spécifiques. Ensuite, on crée un modèle normalisé d'entreprise de haut niveau qui génère les modèles des premiers datamarts qui seront chargés avec les données atomiques en utilisant un staging area temporaire. Un outil ETL sera utilisé pour charger les datamarts et pour échanger le métadonnées avec ces derniers.

2 Infrastructure du datawarehouse

L'infrastructure d'un entrepôt de données inclut le matériel, les réseaux et les fonctions de bas niveaux, telles que la sécurité, que les composants de haut niveau considèrent acquises. Plusieurs facteurs doivent être combinés pour déterminer l'infrastructure adaptée à une implémentation donnée, ils ne sont pas forcément tous techniques. Même dans les couches techniques, les plus profondes de l'entrepôt, les besoins métiers restent les éléments déterminants de la mise en œuvre.

L'infrastructure matérielle d'un entrepôt englobe les plates-formes matérielles de chaque magasin de données, de chaque serveur d'application et des postes de travail. La question essentielle à se poser à ce niveau est de déterminer les magasins de données à mettre en œuvre et parmi eux, combien devront disposer de leur propre plate-forme. Le schéma de l'infrastructure d'un datawarehouse est illustré à l'annexe 8.

2.1 Infrastructure des outils d'arrière-plan (back room)

Dans tout processus de sélection de plates-formes, la première étape consiste à bien assimiler les besoins. Le simple fait de comprendre ce qu'une plate-forme doit faire et la façon dont elle doit le faire du point de vue technique ne suffit pas. Il est essentiel de prendre aussi en compte les besoins métiers. Dans presque tous les projets de datawarehouse, le serveur de la base de données est peut-être la décision la plus délicate en matière de matériel.

2.1.1 Plates-formes matérielles et systèmes d'exploitation

Les principales catégories de combinaisons matériel/système d'exploitation sont présentées dans les lignes qui suivent.

- **Gros système.** En général, le gros système n'est pas la plate-forme idéale pour un entrepôt de données et les quelques réussites en la matière sont des exceptions. Les ED sur gros système est en général peu rentable. Les coûts relatifs à l'administration, au matériel et à la programmation sont plus élevés que ceux des systèmes dits ouverts en partie parce que le gros système dispose d'une infrastructure de traitement des transactions poussées qui ne présente aucun intérêt dans le cadre du datawarehouse.
- **Serveurs de systèmes ouverts.** Les serveurs de systèmes ouverts Unix ou Linux, sont aujourd'hui les plates-formes les plus courantes pour les entrepôts de données de moyenne ou de grande dimension. Unix est généralement assez robuste pour gérer correctement les applications de production et pratique le traitement parallèle depuis plus de 10 ans.

- **Serveur NT.** Bien qu'étant de loin, le SE connaissant la plus forte croissance sur le marché, NT vient seulement d'atteindre les capacités nécessaires à l'implémentation de l'ED de taille moyenne. A l'heure actuelle, ce système n'est toutefois pas le mieux adapté aux ED de moyenne et de grande dimension. Il est en revanche rentable dans le cadre de datawarehouses modestes ou de datamarts peuplés de données atomiques.

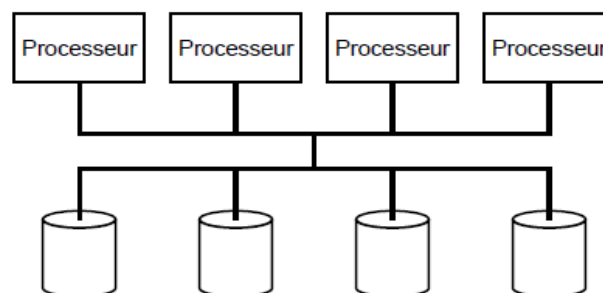
2.1.2 Architectures de traitement en parallèle

Les constructeurs se sont toujours montrés créatifs en matière de sigles et continuent d'en créer de nouveaux. Le marché de serveurs offre trois architectures matérielles de traitement parallèle. Il s'agit de SMP (Symmetric Multi Processing), MPP (Massive Parallel Processing) et NUMA (Non-Uniform Memory Architecture).

2.1.2.1 Architecture SMP (Fonctionnement en multiprocesseurs symétrique)

L'architecture SMP présente une machine unique équipée de plusieurs processeurs, chacun étant géré par un SE et accédant à son propre disque et à sa zone de mémoire. Une machine SMP équipée de 8 à 32 processeurs, une base de données parallèle, beaucoup de mémoire (au moins 2 GO), un bon disque et une conception adaptée conviennent parfaitement à un ED de taille moyenne. Quelques exemples de systèmes SMP sont SQL Server, Oracle, DB/2, Informix and Sybase.

Figure 6 : Architecture matérielle SMP



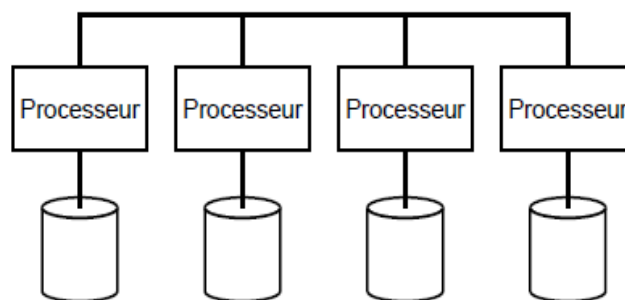
Source : Ralph Kimball (2000)⁶

⁶ « Concevoir et déployer un data warehouse », Ralph KIMBAL, Editions Eyrolles, 2000

2.1.2.2 Architecture MPP (Traitement massivement parallèle)

Les considérations MPP sont fondées sur des chaînes d'ordinateurs relativement indépendants les uns des autres équipés chacun de son propre SE et de son disque dur, de sa mémoire, le tout étant coordonné par des échanges de message. La force des MPP réside dans sa capacité à connecter des centaines de nœuds (machines) en vue de leur soumettre un problème selon l'approche par la force. Ces systèmes sont bien adaptés aux entrepôts de données de grande taille (au-delà du téraoctet) et aux applications qui accèdent aux données de manière intensive (datamining). Quelques exemples de systèmes MPP sont Teradata, Neoview, Netezza and DATAlegro.

Figure 7 : Architecture matérielle MPP

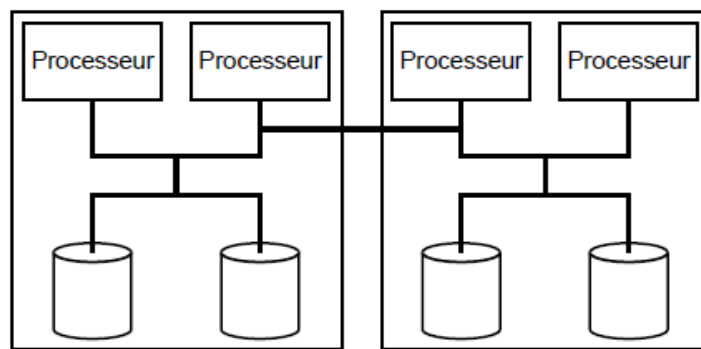


Source : Ralph Kimball (2000)⁷

2.1.2.3 Architecture NUMA (Architecture de mémoire non conforme)

L'architecture NUMA est une combinaison de SMP et de MPP qui vise à aligner la souplesse du partage des disques du premier aux performances du traitement en parallèle du second. Du point de vue conceptuel, l'architecture NUMA reprend l'idée de clusters des machines du SMP mais des connexions plus « serrées » de la bande passante supplémentaire et une meilleure coordination des nœuds. Quelle que soit la plate-forme, il est nécessaire de s'interroger sur la disponibilité du logiciel et sur la complexité de l'administration des systèmes.

⁷ Ralph Kimball (2000), Op cit

Figure 8 : Architecture matérielle NUMA

Source : Ralph Kimball (2000)⁸

2.1.3 Facteurs stimulant les performances matérielles

En matière de datawarehouse, le débit des disques et de la mémoire sont importants car les requêtes peuvent solliciter fortement les données.

➤ Disques

Les lecteurs de disques influent fortement sur les performances, la flexibilité et l'évolutivité d'une plate-forme matérielle.

➤ Mémoire

Plus un datawarehouse dispose de mémoire, mieux c'est ; voici une différence supplémentaire entre l'aide à la décision et le traitement transactionnel. Les requêtes sur les transactions sont généralement peu gourmandes en mémoire. Les requêtes d'aide à la décision sont exigeantes et impliquent souvent plusieurs passes dans des tables volumineuses.

2.1.4 Niveau de service attendu

Le type et la puissance du matériel requis dépendent du degré de disponibilité qu'on veut offrir. Si les données doivent être accessibles au monde entier, les machines en parallèle et une forte redondance des composants seront nécessaires.

⁸ Ralph Kimball (2000), Op cit

2.1.5 Stockage secondaire

On doit s'assurer que la configuration permet la gestion de sauvegarde et d'archivage. Si possible, opter pour un système de sauvegarde assez rapide pour effectuer son travail pendant la durée impartie au chargement.

2.1.6 Autres facteurs

Les environnements serveur Unix et NT sont à ce jour, les plateformes les mieux adaptés aux ED. Unix est la meilleure option pour les systèmes de moyenne et de grande dimension.

2.2 Infrastructure des outils frontaux (front room)

L'infrastructure des outils frontaux (front room) dépend plus fortement de l'activité et des outils que celle des outils d'arrière plan (back room) et les choix à faire y sont plus nombreux.

2.2.1 Serveur d'application

Les serveurs d'application tendent à proliférer à toute vitesse. Les uns gèrent l'accès aux données via le web, les autres gèrent les requêtes, les états standard, l'authentification, les bases de données de métadonnées etc. les critères de choix portent sur la mémoire, le disque, le partage de plate-forme et les goulots d'étranglement du système.

2.2.2 Poste de travail

La puissance du poste de travail dépend de son utilisateur et de ses besoins en matière d'outils. Les critères de choix portent sur :

- le support inter-plate-forme ;
- le système d'exploitation du poste de travail ;
- la distribution des logiciels ;
- les outils web ;

- la mémoire.

2.3 Connectivité et réseau

La connectivité et le réseau relient les outils d'arrière-plan et les outils frontaux. En règle générale, la connectivité est un composant de l'infrastructure.

2.3.1 Bande passante

Il est souvent judicieux d'isoler la base de données et les serveurs d'application sur un réseau local à haut débit. Cette configuration procure la bande passante nécessaire aux transferts de gros volumes de données avec des performances optimales.

2.3.2 Accès à distance

Il faut prévoir une connexion à large bande passante fiable, entre le réseau local des utilisateurs distants et celui qui héberge la base de données et les serveurs d'application.

2.3.3 Passerelles

La plupart des constructeurs des bases de données proposent des passerelles, qui permettent d'établir des liens avec des bases de données concurrentes et avec les sources de données de production.

2.3.4 Transfert de fichiers

Il existe un large éventail de protocoles de transfert de fichiers et de programmes chargés de les implémenter. Le principal est le protocole FTP (File Transfer Protocol), qui est un utilitaire de transfert de données universel. L'un des protocoles les plus récents, SSL (Secure Sockets Layer), émane de Netscape. Il présente l'avantage d'inclure une fonction de cryptage des données transmises, qui permet de sécuriser les informations sensibles.

2.3.5 Connectivité des bases de données

La connectivité des bases de données fait généralement partie de l'offre des outils frontaux. Il existe quelques standards en matière de connectivité de base de données notamment ODBC (Open Database Connectivity), originellement développé par Microsoft, et JDBC (Java Database Connectivity), initialement conçu par JavaSoft.

2.3.6 Service d'annuaire

L'infrastructure de réseau doit prévoir des fonctionnalités destinées à attribuer des noms aux hôtes et à assurer l'indépendance des adresses.

Section 2 : Etudes antérieures et modélisation dimensionnelle

Le concept d'entrepôt formalisé au début des années 1990⁹, a connu un succès fulgurant de part le monde. Divers travaux ont été effectués pour rendre l'information décisionnelle à la disposition des décideurs. Dans les pays en voie de développement, le progrès technologique ne suit pas à la même fréquence que dans les pays développés. Il a fallu du temps pour que les étudiants commencent à être formés à ces nouveaux concepts technologiques. Les travaux de recherche effectués par ces derniers sur le terme d'entrepôt de données permettent d'obtenir des maquettes qui présentent des propositions de résolutions de difficultés rencontrées par les décideurs africains.

Cette partie présentera quelques travaux effectués en Afrique pour la résolution de différents problèmes liés à la prise de décision dans un premier temps et à la méthodologie à adopter dans la modélisation dimensionnelle dans un second temps.

⁹ W.H. Inmon, *Building the Data Warehouse* (1992)

Paragraphe 1 : Etudes de recherches antérieures effectuées sur le thème de datawarehouse

Notre objectif dans cette partie est de présenter les travaux effectués par des étudiants dans la construction d'un datawarehouse pour l'amélioration des prises de décision par les dirigeants des pays africains.

1 Travaux effectués en général

1.1 Au Maroc

Msougar Younès (Dec, 2003)¹⁰, a pensé mettre en place un datamart pour la société Lafarge Maroc pour un meilleur suivi des finances. En effet, au vu de l'importance des projets, ainsi que de la recherche de la performance, l'organisation du contrôle de gestion s'articule autour de deux pôles :

- Pôle "Performance et Reporting" chargé de la production des reportings et de l'analyse de la performance.
- Pôle "Projets" chargé de mettre en place les outils de développement du contrôle de gestion.

Dans le cadre du pôle "Projets", il a reçu pour mission la mise en place du Data Mart Finance au sein de Lafarge Maroc. Lafarge Maroc a adopté, pour la conception du Datawarehouse, une démarche progressive en utilisant le concept du Data Mart.

Le Data Mart finance aura pour finalité d'être un outil de pilotage et d'aide à la décision destiné aux responsables des finances (contrôle de gestion et comptabilité). Il sera de ce fait, le réceptacle de l'information financière. Ainsi, Lafarge Maroc qui est confrontée à un environnement de plus en plus concurrentiel (risques d'importations...) se dotera d'outils performants destinés à assister les managers dans leurs prises de décisions, aussi bien sur le plan financier, que sur les plans commercial, achats et industriel.

¹⁰ Msougar Younès (2003) « Mise en place d'un Datawarehouse : Data Mart Finance Lafarge Maroc », INSTITUT SUPERIEUR DE COMMERCE ET D'ADMINISTRATION DES ENTREPRISES (I S C A E), CASABLANCA

1.2 Au Burkina-faso,

La mise en place d'un data mart pour la paie du personnel de l'Etat a été réalisé par w. Stanislas KABORE¹¹(2005). Son objectif principal est de profiter des avancées technologiques pour mettre à la disposition des responsables de la solde toute la manne d'informations qu'ils avaient à travers l'ancien système de paie qui était au CENATRIN et le système actuel. Ainsi les données existants au niveau du DPS 7000/250, couplées à celles du même genre produites par le système opérationnel actuel SIGASPE (Système Intégré de Gestion Administrative et Salariale du Personnel de l'Etat), pourraient être orientées pour un sujet « les éléments de salaire payés », intégrées, historisées et organisées pour aider les responsables dans la prise de décision. Le data mart devra permettre aux utilisateurs d'avoir des états de synthèse, des statistiques et aussi de pouvoir à partir d'échantillons assez représentatifs faire des projections, des prévisions.

Les travaux réalisés ont permis de faire une étude théorique sur les entrepôts de données ou datawarehouse, les data marts, la modélisation dimensionnelle et le concept OLAP. Cette étude était indispensable pour la mise place d'un data mart concernant la paie des agents de l'Etat.

1.3 En Algérie

En République Algérienne Démocratique et Populaire, Ahmed HACIANE (2005 – 2006)¹², à travers la conception d'un datawarehouse Orienté CRM (*Customer Relationship Management*) ou GRC en français (*Gestion de la Relation Client*), a voulu révolutionner les décisions à prendre par les décideurs pour fidéliser les clients. Ainsi affirme t-il que le CRM a trouvé son origine dans plusieurs études américaines qui ont démontré que fidéliser un client coûtait jusqu'à cinq fois moins cher que d'en conquérir de nouveaux.

¹¹ w. Stanislas KABORE (2005) « Mise en place d'un data mart concernant la paie du personnel de l'Etat », ECOLE SUPERIEURE D'INFORMATIQUE, UNIVERSITE POLYTECHNIQUE DE BOBO-DIOULASSO

¹² Ahmed HACIANE (2005 – 2006), « Conception d'un datawarehouse Orienté CRM », Institut National de formation en Informatique

La mise en place d'une démarche CRM est une stratégie qui va mettre le client au centre de l'entreprise et qui a pour objectif d'en améliorer la rentabilité et de le fidéliser. Le but de sa recherche est de traiter les éléments liés à la conception d'un datawarehouse orienté CRM pouvant constituer l'élément de base de toute infrastructure CRM efficace. Dans ce mémoire, une méthodologie complète de conception et d'implémentation de datawarehouses orienté CRM est présentée. Ce mémoire traite de tous les aspects liés à la conception d'un datawarehouse orienté CRM. Dans ce mémoire, nous avons traité les éléments nécessaires à la conception et l'implémentation d'un datawarehouse orienté CRM. Nous avons focalisé notre travail sur la dimension « Clients » ; car pour le reste, le projet diffèrera peu des datawarehouses classiques. La différence entre les datawarehouses de première génération et ceux de la deuxième génération est la manière d'aborder la conception de la dimension « *Clients* ».

Dans les datawarehouses orientés CRM, les données « *Clients* » ne se limitent pas à une unique table dimension. La dimension « *Client* » devient un modèle complet de données qui est le centre du modèle global du datawarehouse. Ce modèle intègre les données sur le comportement, les circonstances et la segmentation des clients. Dans son mémoire, il a principalement traité la modélisation conceptuelle et logique des données « *Clients* » et proposé une architecture d'implémentation physique du datawarehouse.

2 Au Bénin

Antonio ¹³(2007-2008) quant à lui a voulu étudier la possibilité de mettre en place un datawarehouse pour améliorer l'obtention des reporting sur le trafic de communication GSM de MOOV BENIN. L'entrepôt de données a permis :

- une mise en ligne d'un reporting varié et exhaustif. Ceci facilitera l'accès aux données, garantissant de fait, l'autonomie des utilisateurs, leur procurant les moyens de manipuler les données par eux même (tableaux croisés).

¹³ Antonio BOKO (2007-2008), « L'étude et mise en œuvre d'un datawarehouse pour l'analyse du trafic de communication GSM de MOOV BENIN », ENEAM (Ecole Nationale d'Economie Appliquée et de Management)

- Le raccourcissement des délais de production par l'automatisation de la publication régulière par mail et/ou portail d'état statistiques prédéfinis pour une meilleure prise de décision.

Aujourd'hui au Bénin, la réalisation d'entrepôt de données est presque inexistante surtout dans le domaine des investissements publics. Notre objectif est de doter notre gouvernement d'un outil décisionnel (datawarehouse) permettant d'obtenir des données décisionnelles avec facilité et d'améliorer aussi les prises de décision. Ainsi, allons-nous réaliser une maquette permettant aux décideurs de mieux comprendre les facilités offertes par ces technologies et ainsi amorcer la phase de réalisation de l'entrepôt.

Paragraphe 2 : Modélisation dimensionnelle

L'entrepôt de données est aujourd'hui l'application la plus importante dans la technologie des bases de données. La modélisation d'un entrepôt de données est une étape cruciale car il faut concevoir un modèle qui permettra d'historiser des données et de répondre à des questions que les décideurs se posent. Comment doit-on procéder ? Doit-on faire évoluer un modèle existant ? Doit-on tout repenser à zéro?

1 Généralités sur les modèles dimensionnels

1.1 Problématiques des Entrepôts de Données

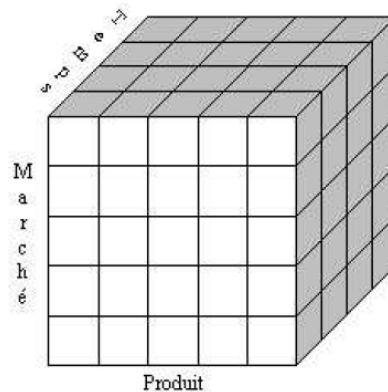
Les quatre grandes problématiques auxquelles sont confrontés les entrepôts de données sont :

- la conception ;
- l'alimentation en données ;
- la consultation, l'optimisation des requêtes ;
- la présentation des résultats.

1.1.1 Conception

La **conception** de l'entrepôt est une tâche critique et difficile à mettre en œuvre. Elle doit répondre de façon spécifique aux besoins de l'entreprise et apporter les bonnes réponses. La modélisation dimensionnelle permet de voir un entrepôt sous la forme d'un **cube** de données dont les arêtes portent une étiquette.

Figure 9: Schéma du cube dimensionnel



Source : Extrait de Hervé COLASUONNO (2001)¹⁴

Chaque donnée est donc modélisée en plusieurs (n) dimensions et l'on va pouvoir extraire les informations par ligne, par colonne ou par tranche. Le modèle dimensionnel comporte une table de **faits** et plusieurs tables de **dimensions**. Les tables de dimension sont de petites tables contenant un minimum de champs, décrivant chacune un axe de l'activité. Ainsi, la table de faits est la table centrale qui va contenir tous les enregistrements qui seront analysés par les utilisateurs.

1.1.2 Alimentation

L'alimentation du datawarehouse est la phase la plus critique. En effet, importer des données inutiles apportera de nombreux problèmes. Personne ne remplit l'entrepôt au sens de la saisie de données car ce dernier est alimenté par les données extraites des bases opérationnelles. Cela consommera des ressources système et du temps. De plus, cela rendra les services d'analyses plus lents. C'est un ETL (Extract

¹⁴ Hervé COLASUONNO (2001), op cité.

Transform Load), qui est utilisé pour alimenter le datawarehouse à partir des bases de données de production. Un ETL se définit comme suit :

- Extract : extrait les données à partir de différentes sources ;
- Transform : transforme ces dernières afin de les unifier sous un même format ;
- Load : charge les données dans le Data Warehouse.

Les intérêts d'un ETL sont multiples :

- il peut prendre en charge différentes natures de sources (SGBD relationnels, flux XML, fichiers CSV...), que ce soit en entrée comme en sortie ;
- l'intégration d'un nouveau flux ne nécessite pas de développement spécifique, une configuration interactive, par le biais d'interface graphique, des 3 étapes vues précédemment suffit ;
- l'intégration d'outil de planification, au sein même des ETL, permet d'éviter le développement de programmes batch spécifiques, ainsi que leur maintenance.

Un autre point à prendre en compte est la périodicité d'extraction des données. Effectivement, le plus souvent, les opérations de collecte de données sont coûteuses en ressource pour la base accédée. Il faut donc trouver un équilibre entre le délai acceptable entre deux mises à jour des tables du datawarehouse et les ressources consommées.

1.1.3 Interrogation

L'interrogation de l'entrepôt pose aussi un problème. En effet, il s'agit de bien cibler le type de requêtes à implémenter, car sur des millions, voire des milliards d'enregistrements, une requête peut prendre plusieurs heures avant d'afficher un résultat. Ainsi, il faudra faire la différence entre l'interrogation prédéfinie et l'interrogation ponctuelle. Une interrogation prédéfinie doit se dérouler en un minimum de temps et est faite périodiquement. A l'inverse, la requête ponctuelle peut être réalisée afin d'approfondir la tendance de certaines mesures lors de l'étude d'une activité. En général, lors de la conception physique d'un entrepôt, on ne prend en compte que les requêtes prédéfinies. En conséquence, le temps de réponse des requêtes ponctuelles est en général bien supérieur à celui des requêtes prédéfinies.

1.1.4 Présentation

La **présentation** est vue comme l'élément essentiel de l'entrepôt par les personnes qui vont l'interroger. En effet, les utilisateurs ne sont pas des informaticiens et ne connaissent pas le langage SQL ; ils doivent néanmoins poser des questions à l'entrepôt et il n'est plus question d'avoir des pages de résultats comme lors de l'interrogation des bases de données. Un état schématique, avec des mots précis, le tout sous forme graphique, est attendu par les usagers. Il existe plusieurs façons de faire : soit l'entrepôt possède son propre outil de présentation (reporting), soit on pourra faire appel à des logiciels annexes qui permettront ce travail.

1.2 Concepts des modèles dimensionnels

1.2.1 Objectifs

Le modèle dimensionnel est plus simple qu'un modèle relationnel normalisé, en vue de faciliter la compréhension et l'interrogation de l'utilisateur final. L'objectif des bases normalisées est d'éviter les redondances afin de maximiser les performances et de garantir la cohérence des données lors des mises à jour. Poser une question non prédéfinie à une base de données normalisée relève du service informatique, à cause de la complexité du schéma de la base. L'entrepôt de données est mis à jour par des programmes d'alimentation à partir des bases sur lesquelles l'entrepôt s'appuie, et non par des utilisateurs. Il n'y a donc pas de problème lors des mises à jour. Par contre, il doit être conçu pour répondre aux requêtes des utilisateurs, qu'elles soient nombreuses ou non. Ainsi, les objectifs d'un entrepôt de données sont de produire une structure compréhensible par les utilisateurs finaux et d'augmenter la qualité et les performances des requêtes.

1.2.2 Concepts fondamentaux

1.2.2.1 Dimension

Lorsqu'on fait un schéma de base de données pour un système d'informations classique, on parle en termes de table et de relation. Une table étant une

représentation d'une entité, en business on parle de dimension et de faits. C'est une autre approche des données.

On entend par dimension, les axes avec lesquels on veut faire l'analyse. On peut par exemple avoir une dimension Client, une dimension Produit et une dimension Géographie pour faire les analyses par secteurs géographiques. Une dimension est tout ce qu'on utilisera pour faire des analyses.

1.2.2.2 Fait

Les faits en complément à la dimension sont ce sur quoi va porter l'analyse. Ce sont des tables qui contiennent des informations opérationnelles et qui relatent la vie de l'entreprise. On peut avoir par exemple les tables de fait pour les ventes (Chiffre d'affaire, la quantité, le montant commandé, quantité facturé etc.) ou sur les stocks ou sur les ressources humaines. La table des faits est la clé de voute du modèle dimensionnel ou sont stockées les tables des indicateurs de performances. Le terme fait est utilisé pour représenter une mesure économique.

1.2.2.3 Indicateurs

Les indicateurs sont les informations d'un processus d'entreprise dans un Système d'Information. Les indicateurs les plus utiles d'une table de faits sont numériques et additifs. L'additivité des attributs d'une table de faits est cruciale pour les outils décisionnels. Tous les attributs utiles ne sont pas additifs. Certains sont semi-additifs et ne peuvent être additionnés que pour certaines dimensions. D'autres sont non additifs et ne peuvent être additionnés par dimension. Pour cette dernière catégorie, on utilise les fonctions comme la moyenne de prix de vente des produits, le ratio ou le comptage des lignes.

1.2.2.4 Structure de la base de données

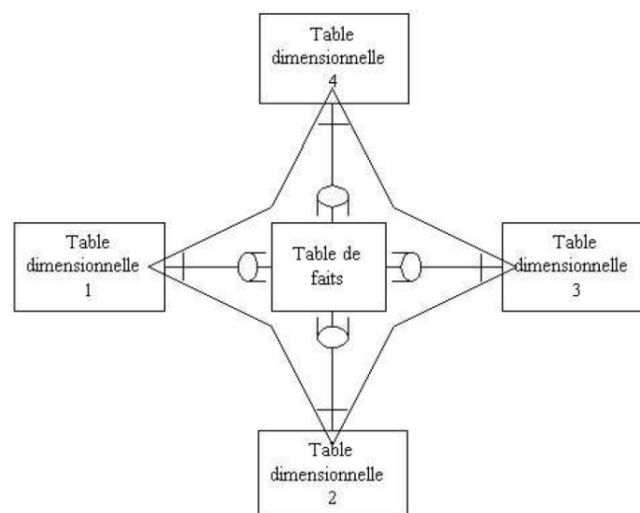
Au sein d'un ED, les données sont redondantes et dénormalisées. Nous sommes loin de la modélisation en troisième forme normale et pour cause : cela permet de faciliter l'utilisation et d'améliorer les performances lors de l'analyse des données. Trois types de schémas sont fréquemment rencontrés : le schéma en étoile, le schéma en flocon et le schéma en constellation de faits.

1.2.3 Modèles dimensionnels

1.2.3.1 Schéma en étoile

La première structure utilisée en modèle dimensionnel est l'étoile. Elle est composée d'une table centrale, appelée la table des faits, qui est reliée à des tables dimensionnelles de taille plus petite, les points de l'étoile. La table des faits contient les mesures (prix, nombre d'unités vendues...). Les tables dimensionnelles contiennent les conditions de regroupement des mesures. La table des faits est liée aux tables dimensionnelles par des associations, à tout élément d'une table de dimensions sont associés plusieurs éléments de la table centrale, éventuellement aucun ; à tout élément de la table centrale est associé un élément d'une table de dimension et un seul.

Figure 10 : Schéma générique de l'étoile



Source : Extrait de Hervé COLASUONNO (2001)¹⁵

Avantages:

- Facilité de navigation ;
- Performances : nombre de jointures limité; gestion des données creuses ;
- Gestion des agrégats ;
- Fiabilité des résultats.

¹⁵ Hervé COLASUONNO (2001), op cité.

Inconvénients:

- Toutes les dimensions ne concernent pas les mesures ;
- Redondances dans les dimensions ;
- Alimentation complexe.

On considère classiquement deux formes dérivées du schéma en étoile : le schéma en constellation et le schéma en galaxie.

1.2.3.2 Schéma en constellation

Le schéma en constellation consiste à placer plusieurs schémas en étoile avec des tables de faits reliées hiérarchiquement. Les liens entre les différentes tables de faits permettent de visionner les différents niveaux de détail. Les hiérarchies définies peuvent être géographiques ou organisationnelles.

Plusieurs schémas en étoile peuvent être combinés en galaxie. La galaxie est un ensemble d'étoiles avec des tables de dimension partagées. Contrairement à la constellation, les tables de faits d'une galaxie n'ont pas forcément de liens entre elles.

1.2.3.3 Le schéma en flocon de neige

Dans une étoile, les dimensions sont dénormalisées : les hiérarchies d'agrégation sont implicites. Un flocon de neige est une étoile où les hiérarchies sont respectées et explicites. Un flocon peut être formé depuis une étoile en normalisant les dimensions. Le flocon diffère de l'étoile de par sa constitution : un flocon reste relativement normalisé, alors qu'une étoile est dénormalisée. Un flocon comme une étoile peut être combiné pour former une tempête par agrégation de flocons.

1.2.3.4 Le schéma en grappe :

Ce schéma est apparu car il n'existe pas de schéma en étoile ou de schéma en flocon parfait. Le schéma en grappe est une dérivation de ces deux schémas pour en former un troisième.

Kimball (Kimball, 1996) déclare qu'un schéma en flocon n'est pas optimal, car il est trop complexe. Toujours pour les mêmes raisons de simplification des tables, afin

de pouvoir trouver facilement les informations dans l'entrepôt, le schéma en grappe apparaît alors comme un compromis entre le schéma en étoile et le schéma en flocon.

2 Méthodologie de conception

La méthode de construction considérée permet de réaliser un modèle dimensionnel en partant de son modèle relationnel.

D'après (Kortink et al, 1999), on peut arriver à un modèle dimensionnel en partant du modèle d'entreprise s'il existe. Mark A.R. Kortink et Daniel L. Moody (Kortink et al, 1999) présentent aux concepteurs d'entrepôts et aux décideurs, une méthode de conception basée sur un modèle existant. Ils proposent de définir le modèle de l'entrepôt de données à partir du modèle d'entreprise, dont la mise en place est en général très onéreuse.

2.1 Définition du modèle de datawarehouse

2.1.1 Classification des entités

La première étape pour créer un modèle dimensionnel à partir du modèle d'entreprise est de classer ses entités. Elles se décomposent en trois catégories.

2.1.1.1 Entité de transaction

Elles enregistrent les détails d'événements particuliers comme les salaires, les réservations d'hôtels. Ce sont ces événements, entre autres, que les décideurs veulent analyser. Les caractéristiques des entités de transaction sont :

- elles décrivent un événement qui se produit à un instant donné ;
- elles contiennent les mesures, comme le montant en devise, le poids, les volumes.

Ces mesures forment la base des résultats que l'entrepôt permet d'étudier. Les entités de transaction forment le noyau à partir duquel la table des faits d'un schéma en étoile peut être créée. Toutes les entités de transaction ne sont pas bonnes pour l'aide à la décision, et l'on doit choisir et identifier les plus intéressantes.

2.1.1.2 Entité Composante

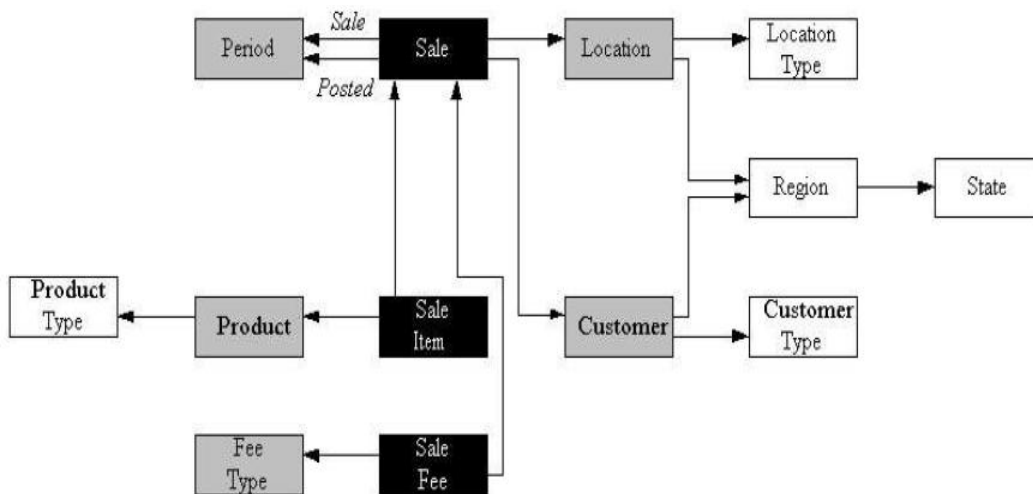
Elle est directement liée à l'entité de transaction par une relation multivaluée. Ces entités définissent la finesse des détails ou composants de chaque transaction. Les composants répondent aux questions qui, quoi, quand, où, combien et pourquoi, d'un événement commercial. Les entités composantes forment la base des tables de dimension dans les schémas en étoile. Chaque table de dimension correspond à une ou plusieurs entités composantes.

2.1.1.3 Entité de classification

Ce sont des entités qui sont apparentées à des entités composantes par une chaîne de relations multivaluées. Ces entités représentent la hiérarchie entre les entités dans le schéma en étoile.

La figure suivante montre les hiérarchies.

Figure 11 : Classification des entités



Source : Extrait de Hervé COLASUONNO (2001)¹⁶

Les entités en noir représentent les entités de transaction.

¹⁶ Hervé COLASUONNO (2001), op cité.

Les entités en gris représentent les entités composantes.

Les entités en blanc représentent les entités de classification.

2.1.1.4 Ambiguïtés

Parfois, des entités sont présentes dans plusieurs catégories. Les auteurs définissent la priorité des hiérarchies comme suit :

- transaction : la plus haute ;
- classification ;
- composante : la plus basse.

2.1.2 Hiérarchies

2.1.2.1 Identification des hiérarchies

La hiérarchie est un concept extrêmement important dans le modèle dimensionnel, car c'est par la hiérarchie que l'on va pouvoir passer du modèle relationnel au modèle dimensionnel. Une hiérarchie est un modèle d'entité-association qui est identifié par des séquences d'entités reliées par des relations multivaluées, toutes alignées dans le même sens. Nous parlons alors d'un schéma normalisé. La figure ci-dessous illustre un exemple de hiérarchie.

2.1.2.2 Réduction de la hiérarchie

Des entités de rang supérieur peuvent être concaténées à des entités inférieures en respectant la hiérarchie. Cela introduit une redondance et forme une dépendance transitive, ce qui viole la troisième forme normale de Codd écrite en 1970. On introduit alors le concept de base dénormalisée, terme désigné par Tonkin en 1991.

2.1.2.3 L'agrégation : une fonction des entités de transaction

Une agrégation peut être appliquée à une entité de transaction pour former une nouvelle entité de transaction, constituée d'une synthèse de données. La clé de cette nouvelle entité est une combinaison des attributs utilisés par l'agrégation.

Dans le but de contribuer à la résolution des différents problèmes rencontrés dans l'élaboration du suivi du PIP, nous avons proposé la mise en place d'un datawarehouse. Elle est une base de données destinées à recueillir, à organiser, à mettre en forme et à diffuser l'ensemble des informations de l'entreprise de façon cohérente pour faciliter l'analyse et la prise de décision. Cette première partie a permis de présenter de manière globale les entrepôts de données ainsi que les approches et modèles qui ont pu être proposés depuis quelques années. Les différentes étapes de la construction de l'entrepôt de données proposé seront abordées dans la suite.

PARTIE II : ETUDE CONCEPTUELLE

L'étude conceptuelle présente les travaux effectués pour la construction du datawarehouse. Cette partie est consacrée à la présentation de la méthodologie adoptée et sa mise en pratique dans la réalisation de la solution retenue. Ensuite, le prototype réalisé à partir des travaux effectués est présenté.

CHAPITRE III : TRAVAUX REALISES

Cette partie présente la méthodologie utilisée pour la construction du datawarehouse. La modélisation dimensionnelle de l'entrepôt est abordée dans un premier temps et sa conception physique dans un second.

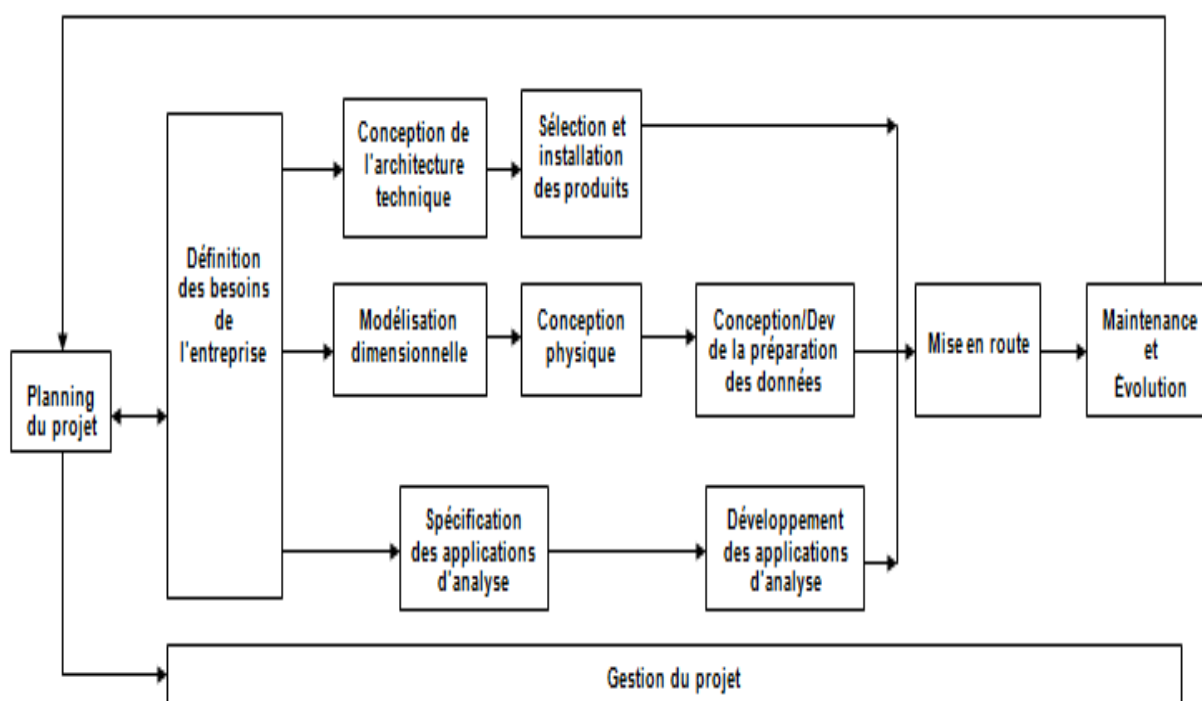
Section 1 : Conception dimensionnelle

Paragraphe 1 : Méthodologie adoptée

1 Présentation

Les démarches pour la conduite d'un projet datawarehouse sont multiples et il s'avère indispensable d'opérer un choix. Au nombre de celles présentées plus haut, nous nous sommes inspirés de l'approche de Ralph Kimball dont la méthodologie (approche Bottom-up ou datamarts indépendants) est décrite au travers du célèbre schéma dénommé " Business Dimensional Life cycle " qui est une référence dans la stratégie décisionnelle et qui se présente comme suit :

Figure 12 : Cycle de vie dimensionnelle de KIMBALL



Source : KIMBALL, 2000

En effet, ce modèle permet la conception du modèle par vue métier. Cette approche a été retenue pour les raisons suivantes :

- l'obligation de tenir compte d'un existant de rapports statistiques au sein de ministère, expression du besoin des utilisateurs ;
- cette approche est orientée analyse des besoins utilisateurs. Elle permet de construire un schéma d'entrepôt calqué sur les besoins d'analyse exprimés ;
- cette approche réduit considérablement le risque de concevoir une solution obsolète avant d'être opérationnelle ;
- elle permet d'utiliser l'approche de conception MOLAP pour la construction de l'entrepôt. Cette dernière garantit une meilleure performance d'exécution des requêtes.

Nous avons élaboré un planning du projet. Ce planning a permis entre autres, une étude du domaine métier d'activités, la collecte des besoins fonctionnels et non fonctionnels. Avant la mise en œuvre de ce planning, nous avons choisi un échantillon et construit un questionnaire (Cf. Annexe 09) nous permettant de mieux cibler les interviews.

2 BESOINS DE L'ENTREPRISE

Les travaux de recherche effectués ont permis de recenser les besoins des utilisateurs pour le datawarehouse futur. Ces besoins représentent les attentes des utilisateurs pour l'entrepôt de données.

2.1 Besoins fonctionnels

2.1.1 DGSP : Direction Générale du Suivi des Projets et Programmes

1. Besoin d'avoir le point d'exécution financière pour une période donnée des PIP par secteur dans le but de voir le montant proportionnel consommé du PIP pour chaque secteur ;

2. Besoin d'avoir le point d'exécution financière pour une période donnée par ministère et institution de l'état pour obtenir le montant proportionnel consommé du PIP pour chaque ministère et institution;
3. Besoin d'avoir l'exécution financière sur une période donnée, base ordonnancement, du portefeuille actif de l'état par secteur dans le but d'obtenir le montant réel ordonné à être consommé par secteur ;
4. Besoin d'analyser l'évolution des prévisions d'investissements publics au cours d'une période. Ceci permet de voir comment l'apport des ressources évolue par rapport aux investissements et pour savoir si l'économie du pays se porte bien ;
5. Besoin de répartir les PIP par source de financement de l'année N à l'année N+i au cours d'une période ; dans le but de connaître l'apport des différents types de bailleurs et les types de financements comme prêts et dons et voir si l'Etat est autonome financièrement ou non ;

2.1.2 DGIFD : DIRECTION GENERALE DES INVESTISSEMENTS ET DU FINANCEMENT DU DEVELOPPEMENT

1. Besoin d'analyser les actions ou projets mis en œuvre par secteur, OSD, domaine, ministère ou institution et par axe sur une période (par exemple sur les 6 dernières années). Ceci donne le nombre de projets par critères énumérés sur une période ;
2. Besoin d'analyser la répartition des crédits du PIP par OSD sur une période dans le but d'observer la consommation des crédits du PIP par OSD périodiquement ;
3. Besoin d'analyser l'accroissement du PIP an N avec le PIP an N+i par secteur dans le but d'avoir la croissance du PIP par période ;
4. Besoin d'analyser les types de financements (extérieur ou national) reçus qui ont contribué au développement sur une période;
5. Besoin d'analyser les financements des bailleurs par secteur, par OSD, et suivant le mode de financement sur une période.

2.1.3 Ensemble des directions :

- 1- Besoin de manipulation des informations par les utilisateurs (requêtes, cross-checking) ;
- 2- Pertinence de l'information ;
- 3- Besoin d'interroger les données au moins sur les 4 derniers trimestres.

2.2 Besoins non fonctionnels

Les exigences fonctionnelles permettent de déterminer les fonctionnalités du système tandis que les besoins non fonctionnels fournissent les guides et contraintes de l'architecture du système. C'est un ensemble de contraintes à respecter liées au temps, budget, accessibilité, sécurité et ressources relatives au projet. Ils présentent le plan qualité du projet.

1. Tous les utilisateurs de l'entrepôt de données doivent être identifier pour avoir accès aux données par les applications frontales (reports et OLAP);
2. Les applications frontales des entrepôts de données ne doivent être accessibles en dehors de l'entreprise sauf cas autorisé ;
3. Toutes les applications frontales sont idéalement basées sur le web et accessibles à partir de n'importe où dans le réseau de l'entreprise ;
4. Tous les utilisateurs accèderont aux applications frontales à partir d'un portail central. La présentation de l'interface utilisateur et la conception doivent se conformer aux normes d'entreprise et des lignes directives ;
5. Les accès et manipulations de la base doivent être liés à un profil utilisateur ;

3 Etude de faisabilité et risques associés

Cette étude vise à explorer les données sources afin de les analyser plus en détails, de mesurer les risques associés à une éventuelle indisponibilité (par exemple une exigence pourrait préciser deux années de données historiques mais seulement six mois de données sont disponibles) et d'identifier les solutions alternatives afin de satisfaire les besoins énoncés par les futurs utilisateurs.

L'étude des différentes sources des données nous a amené à interroger par requêtes les bases des données et leurs structures, à établir la documentation associée pour mieux comprendre la signification et la qualité des données obtenues ainsi que leur localisation.

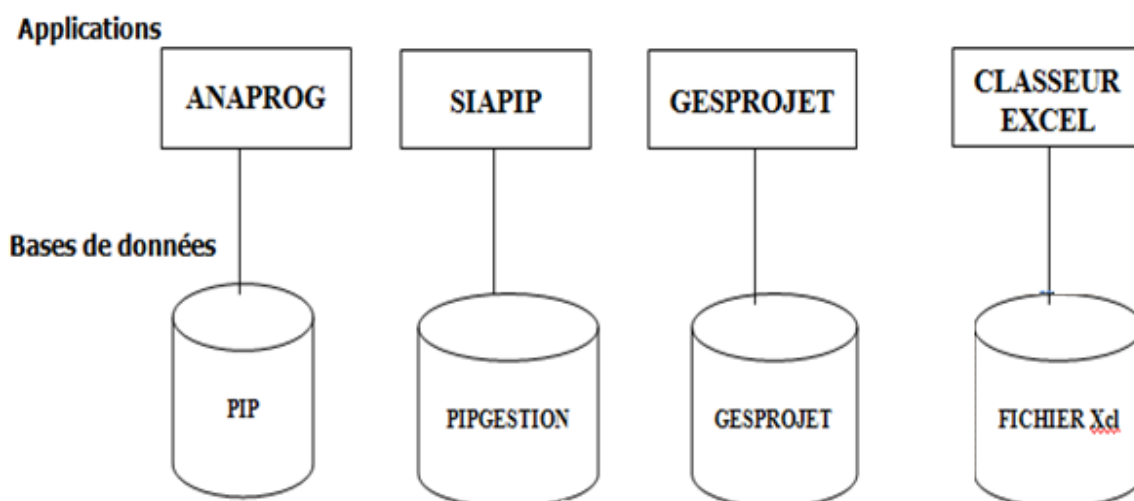
Les points suivants ont été recensés et traités :

- Etablissement de la cartographie des sources ;
- Collecte des documents et/ou informations sources (modèles des données, tables, dictionnaire de données) ;
- Interrogation des sources par requêtes SQL ;
- Identification des problèmes potentiels (bases de données, réseau, faisabilité, qualité, exhaustivité) ;
- Etude des fiches des identificateurs existants et règles de gestion techniques ou fonctionnelles ;
- Recensement des tableaux de bords existants par direction.

3.1 Cartographie des sources

Les documents d'élaboration de chaque PIP retraçant les indicateurs et valeurs à atteindre, les Stratégies de Croissance Pour la Réduction de la Pauvreté (SCRP) ont constitué aussi des sources de données.

Figure 13 : Cartographie des sources



Source : Conçus par nous-mêmes

3.2 Documentation des sources

➤ Les Fichiers

Les fichiers utilisés par les différents systèmes opérationnels se transmettent par des fichiers World ou Excel. Chaque structure se charge de saisir les informations indispensables pour elle. Voici une liste de fichiers utilisés.

Tableau 3 : Liste des fichiers utilisés par source

Fichier	Source	Description	Destination	Document de référence
Etat.PointParSecteur.pdf	GESPROJET	Fichier donnant un récapitulatif du nombre de projets, par nature et le montant correspondant	Classeur Excel	Rapport d'avancement à fin XXXX.doc
Etat.ProjetParDepart.pdf	GESPROJET	Récapitulatif du coût de chaque projet par secteur avec leur date de démarrage et de fin probable	Classeur Excel	Rapport d'avancement à fin XXXX.doc
Etat.ProjetParSecteur.pdf	GESPROJET	Présentation par secteur des projets par commune et département	Classeur Excel	Rapport d'avancement à fin XXXX.doc
PIPGestion.Annee XXXX.pdf	PIPGESTION	Liste des projets retenus au cours de l'année en renseignant sur le montant alloué, les ressources financées et le coût actualisé	Classeur Excel	Rapport d'avancement à fin XXXX.doc
BilinduPIPXXXa uJJ/MMAAAA.xlsx	Fichier Excel	Evolution financière et physique du PIP de l'année	SIAPIP	Présentation du PIP gestion de l'année XXXX

Source : Conçus par nous-mêmes

➤ Les tables

A l'instar de toutes les autres sources de données, un filtrage des champs utiles de données se fera sur la base des résultats attendus. L'étude des tables du SIAPIP a permis de recenser une liste des tables. (Annexe 10)

3.3 Identification des problèmes potentiels et filtrage des données

La structure de la base de données ne permet pas de récupérer toutes les informations en fonction des résultats attendus. A cet effet, nous recensons :

- absence des bénéficiaires du projet ;
- absence des noms des coordonnateurs des projets ;
- redondance au niveau des données (par exemple pour les montants alloués par commune et département, car montant département est un cumul de communes) ;
- redondance des montants alloués, ordonnés etc. dans la plupart des tables du système.

3.4 Risques liés à l'étude

Dans le cadre de cette étude, les risques représentent tout élément ou fait, de nature à perturber, freiner ou bloquer l'atteinte des objectifs poursuivis.

- A cet effet, suivant les objectifs visés, de nombreuses données sont non identifiées et/ou absentes des sources existantes ;
- Certaines informations existent sur support papier mais ne sont pas renseignées dans les systèmes existants ;
- Les difficultés rencontrées au niveau de la collecte d'information dans les divers ministères ou institutions fragilisent la qualité des données. Ceci ne permet donc pas d'avoir toujours une situation réelle sur l'exécution des projets.

Une approche modulaire itérative du projet a été retenue du fait du nombre important d'indicateurs à produire. Elle se traduit par la mise en place de magasin de données ou datamart (magasin de données orienté métier). Ainsi à court terme, chaque secteur métier pourra valider et commencer l'utilisation des outils mis à sa disposition. Cette démarche s'inscrit dans le plan de management de ce projet de construction de datawarehouse notamment la gestion des changements. La méthode itérative permet une validation à chaque étape de la réalisation favorisant de fait, une participation active et un regain d'intérêt pour les utilisateurs.

4 Thèmes analytiques et processus métiers

A travers les différentes séances de travail, nous avons dégagé les différents Business Process (processus métiers) à implémenter. Il s'agit en fait de classer les besoins en thème ou sujet d'analyse.

Des différents besoins fonctionnels évoqués plus haut, nous obtenons le tableau 4 regroupant les besoins par structure et suivant les datamarts proposés.

Tableau 4 : Besoins fonctionnels par groupe d'utilisateurs

DIRECTIONS		DSPG, DSPEI, DSPS, CASPPP, DPIP	DGIFD, DPIP, CASPPP, DPF	DSPG, DSPEI, DSPS, CASPPP
DATAMARTS		Datamart sur le suivi financier des projets	Datamart sur l'évolution des prévisions d'investissements publics	Datamart sur l'évaluation des indicateurs de programmes
BESOINS FONCTIONNELS	1	Montant des projets exécutés par PIP, par secteur, par OSD, par ministère et par axe pour une période.	Répartition par secteur du PIP sur une période.	Etat sur le taux d'exécution financière des projets sur une période.
	2	Etat financier des actions exécutées par PIP dans le temps.	Répartition des crédits PIP gestion par ministère et institution.	Situation des programmes mis en œuvre avec le taux d'exécution des indicateurs de performance.
	3	Top de meilleurs projets exécutés par secteurs et par bailleurs.	Etat des prévisions du PIP alloués par secteur et par OSD.	
	4	Etat des sources de financement par PIP au cours d'une période.	Etat des projets exécutés par type de financement reçu par OSD et par pôle sur une période.	
	5	Répartition du financement du PIP par source de financement au cours d'une période.		

Source : Conçus par nous-mêmes

Le tableau suivant présente les besoins regroupés par thèmes d'analyse et processus concernés.

Tableau 5 : Besoins regroupés par thèmes d'analyse et processus concernés

THEMES D'ANALYSE	PROCESSUS CONCERNES	COMMENTAIRE
Planification et programmation des projets	- Programmation et planification des actions, - Suivi des actions.	Ce sujet concerne tous les projets initiés par les différents ministères ou institutions. Il permet d'avoir une liste réelle des projets autorisés par le gouvernement.
Suivi financier des projets	- Suivi des projets, - La Programmation et la Planification, - La coordination de la coopération pour la recherche du financement.	Ce sujet est conçu pour accéder aux informations sur les financements des projets, leurs bailleurs, le type de financement, etc.
Suivi physique des projets	- Suivi des projets, - La Programmation et la Planification, - La coordination de la coopération pour la recherche du financement.	Ce sujet concerne l'exécution physique des projets. Il permet de savoir l'état d'exécution de certains projets et leur évaluation par rapport à l'exécution financière obtenue.
Suivi des indicateurs de performance des programmes d'investissement	- Suivi des projets, - Etude des stratégies	Ce sujet concerne l'état d'avancement des programmes d'investissement. A partir des suivi financier et physique, il rend disponible les proportions atteintes par chaque indicateurs de performances.

Source : Conçus par nous-mêmes

Paragraphe 2 : Modélisation dimensionnelle

« Un SID comporte donc en réalité au moins *deux* Modèles Conceptuels de Données. L'un des deux représente l'*intégration* des sources opérationnelles à partir desquelles s'alimente le système. Il se conçoit et se normalise selon une démarche traditionnelle de génie logiciel, qui n'a pas lieu d'être développée ici. L'autre, correspond à la structure informationnelle destinée à supporter les *requêtes* des utilisateurs. C'est le MCD de *diffusion*. C'est ce dernier qui représente la structure selon laquelle l'information doit être mise à disposition ; il constitue la *spécification fonctionnelle* du SID. » (Jean-Marie Gouarné, 1997).

1 Bus Matrix

La modélisation dimensionnelle repose d'abord sur un modèle conceptuel de données qui représente en fait le datawarehouse. Notre modèle conceptuel est présenté en annexe 11.

Le modèle dimensionnel quant à lui est réalisé en fonction des informations utilisateurs à produire. La construction de ce modèle nous permettra d'atteindre nos objectifs en terme de :

- présentation des besoins d'informations aux utilisateurs le plus simplement possible ;
- présentation des résultats des requêtes utilisateurs renvoyées aussi rapidement que possible ;
- production des informations pertinentes relatives à une meilleure lecture des processus métiers.

Les processus métiers (« business process ») constituent des unités de travail dans l'élaboration de la couche d'information de la BI. Le Bus Matrix contient les processus métiers et les dimensions nécessaires à leur gestion.

Figure 14 : Bux Matrix

		date	projet	baillleur	structure	sous-programme	pole développement	execution financière	se deroule	OSD	mode financement	domaine	secteur	filière économique	beneficiaire	execution physique	Indicateurs
1	Planification et programmation des projets	X	X	X	X	X	X		X	X	X	X	X	X	X		
2	Suivi financier des projets	X	X	X	X	X	X	X	X	X	X	X	X	X	X		
3	Suivi physique des projets	X	X	X	X	X			X	X	X	X	X	X	X	X	
4	Suivi des indicateurs de performances des programmes d'investissement	X	X	X	X	X	X			X	X	X	X		X	X	X

Source : Conçu par nous-mêmes

Les datamarts obtenus sont :

- Datamart sur le suivi financier des projets ;
- Datamart sur le suivi physique des projets ;
- Datamart sur l'évolution des prévisions d'investissements publics ;
- Datamart sur l'évaluation des indicateurs de programmes.

Un modèle dimensionnel est réalisé pour chaque datamart obtenu. Nous retenons de réaliser le modèle dimensionnel de deux datamarts. Il s'agit du datamart sur le suivi financier des projets et le datamart sur l'évolution des prévisions d'investissements publics.

2 Conception du datamart sur le suivi financier des projets

2.1 Présentation du datamart

Le datamart sur le suivi financier des projets a pour but de mettre à disposition des utilisateurs, toutes les informations relatives à un projet, du point de vue financier et détail sur une période.

2.2 Grain du datamart

La granularité est la plus petite information analysable. Le grain de ce datamart est le suivi individuel des projets par trimestre.

2.3 Modèle logique

2.3.1 Tables de dimensions de l'Exécution des projets

D'après la figure 21, nous avons la liste des dimensions que nous pouvons obtenir au niveau de ce datamart. La description de chaque table de dimension est présentée dans le tableau suivant.

Tableau 6 : Description des tables de dimension

N°	Table de Dimension	Signification
1	Projet_Dimension	cette table regroupe tous les projets mis en œuvre par le gouvernement béninois.
2	Bailleur_Dimension	Cette table renseigne sur les bailleurs des projets.
3	Structure_Dimension	Elle permet d'avoir les ministères, institutions et autres organismes qui exécutent la mise en œuvre des différents projets.
4	SousProgramme_Dimension	Cette entité donne le sous-programme concerné de chaque projet.
5	PoleDeveloppement_Dimension	Elle donne le pôle de développement couvert par le projet.
6	ExécutionFinanciere_Dimension	Cette table donne les informations sur l'exécution financière de chaque projet
7	SeDeroule_Dimension	Elle renseigne sur les communes et les départements concernés par l'exécution des projets.
8	OSD_Dimension	Cette entité permet de renseigner les Orientations Stratégiques de Développement que chaque projet met en œuvre.
9	ModeFinancement_Dimension	Elle précise le type de financement du projet.
10	Domaine_Dimension	Elle donne le domaine dans lequel s'exécute le projet.
11	Secteur_Dimension	Cette table donne le secteur dans lequel le projet est mis en œuvre.
12	FiliereEconomie_Dimension	Elle spécifie les filières économiques du pays dans lequel le projet est élaboré.
13	Date_Dimension	Cette table représente la table de date.

Source : Conçus par nous-mêmes

2.3.2 Table des faits

Les éléments mesurables dans la table des faits se présentent dans le tableau 7.

Tableau 7 : Mesures de la table des faits

N°	Mesures ou faits	Signification
1	MontantFinancement	Le montant total financé par projet à ce trimestre
2	MontantAutoAssemble	Le montant total autorisé à être consommé par le gouvernement
3	CoutInitial	Il renseigne sur le coût prévu pour le projet
4	CoutActuel	Il renseigne sur le coût réel du projet par trimestre
5	MontantEngagement	Montant total engagé au cours du trimestre par projet
6	MontantOrdonnancement	Montant total ordonné par trimestre.
7	MontantDecaissement	Le montant total décaissé par projet au cours du trimestre.
8	MontantAlloueCommune	Le montant alloué par commune bénéficiaire.
9	MoisExecutionFinanciere	Il renseigne sur le mois pour lequel le montant total décaissé a été enregistré.
10	AnneeExecutionFinanciere	Il renseigne sur l'année dans laquelle le montant total décaissé a été enregistré.

Source : Conçus par nous-mêmes

Un tableau récapitulatif présente toutes les dimensions avec leurs attributs et les mesures de la table de faits. Il est représenté à la figure 15.

Figure 15 : Tableau récapitulatif des dimensions du datamart sur l'Exécution des projets

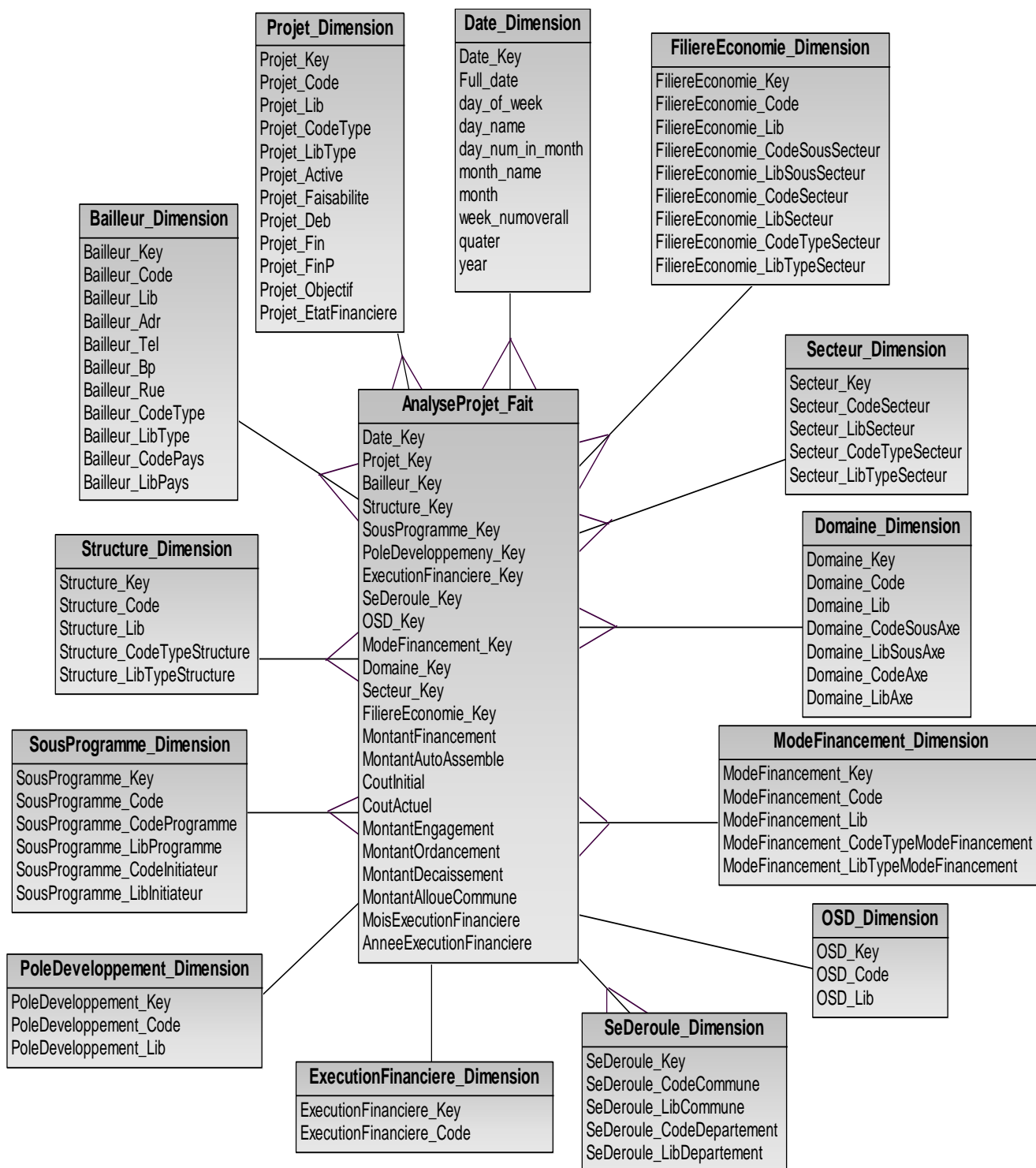
	Projet_Dimension	Bailleur_Dimension	Structure_Dimension	SousProgramme_Dimension	PoleDeveloppement_Dimension	ExecutionFinanciere_Dimension	SeDeroule_Dimension	OSD_Dimension	ModeFinanciere_Dimension	Domaine_Dimension	Secteur_Dimension	FiliereEconomie_Dimension	Date_Dimension
	Projet_Key	Bailleur_Key	Structure_key	SousProgramme_Key	PoleDeveloppement_Key	ExecutionFinanciere_Key	SeDeroule_Key	OSD_Key	ModeFinancement_Key	Domaine_Key	Secteur_Key	FiliereEconomie_Key	Date_Key
	Projet_Code	Bailleur_Code	Structure_Code	SousProgramme_Code	PoleDeveloppement_Code	ExecutionFinanciere_Code	SeDeroule_CodeCommune	OSD_Code	ModeFinancement_Code	Domaine_Code	Secteur_Code	FiliereEconomie_Code	Full_date
	Projet_Lib	Bailleur_Lib	Structure_Lib	SousProgramme_CodeProgramme	PoleDeveloppement_Lib		SeDeroule_LibCommune	OSD_Lib	ModeFinancement_Lib	Domaine_Lib	Secteur_LibSecteur	FiliereEconomie_Lib	day_of_week
	Projet_CodeType	Bailleur_Adr	Structure_CodeType	SousProgramme_LibProgramme			SeDeroule_CodeDepartement		ModeFinancement_CodeTypeModeFinancement	Domaine_CodeSousAxe	Secteur_CodeTypeSecteur	FiliereEconomie_CodeSousSecteur	day_name
	Projet_LibType	Bailleur_Tel	Structure_LibType	SousProgramme_CodeInitiateur			SeDeroule_LibDepartement		ModeFinancement_LibTypeModeFinancement	Domaine_LibSousAxe	Secteur_LibTypeSecteur	FiliereEconomie_LibSousSecteur	day_num_in_month
	Projet_Active	Bailleur_Bp		SousProgramme_LibInitiateur						Domaine_CodeAxe		FiliereEconomie_CodeSecteur	month_name
	Projet_Faisabilite	Bailleur_Rue								Domaine_LibAxe		FiliereEconomie_LibSecteur	month
	Projet_Deb	Bailleur_CodeType										FiliereEconomie_CodeTypeSecteur	week_numoverall
	Projet_Fin	Bailleur_LibType										FiliereEconomie_LibTypeSecteur	quater
	Projet_FinP	Bailleur_CodePays											year
	Projet_Objectif	Bailleur_LibPays											
	Projet_EtatFinanciere												
	Faits de AnalyseProjet_Fait : MontantFinancement, MontantAutoAssemble, CoutInitial, CoutActuel, MonatantEngagement, MontantOrdonnancement, MontantDecaissement, MontantAlloueCommune, MoisExecutionFinanciere, AnneeExecutionFinanciere												

Source : Conçus par nous-mêmes

2.4 Modèle dimensionnel

Le modèle de donnée du datamart est présenté sur la figure ci-dessous. Il présente les dimensions et la table des faits.

Figure 16 : Schéma en étoile sur l'Exécution des projets



Source : Conçus par nous-mêmes

3 Conception du datamart sur l'évolution des prévisions d'investissements publics

3.1 Présentation du datamart

Le datamart sur l'évolution des prévisions d'investissements publics a pour objectif de fournir toutes les informations sur le suivi des indicateurs de performances. Ce datamart permet ainsi de connaître le taux d'exécution des indicateurs de performances.

3.2 Grain du datamart

Le grain de ce datamart est le suivi individuel par projet des indicateurs de performance par trimestre.

3.1 Modèle logique

3.1.1 Tables de dimensions

Tableau 8 : Description des tables de dimension du datamart sur l'évolution des prévisions d'investissements publics

N°	Table de Dimension	Signification
1	Projet_Dimension	cette table regroupe tous les projets mis en œuvre par le gouvernement béninois.
2	Bailleur_Dimension	Cette table renseigne sur les bailleurs des projets.
3	Structure_Dimension	Elle permet d'avoir les ministères, institutions et autres organismes qui exécutent la mise en œuvre des différents projets.
4	SousProgramme_Dimension	Cette entité donne le sous-programme concerné de chaque projet.
5	PoleDeveloppement_Dimension	Elle donne le pôle de développement couvert par le projet.
6	ExécutionPhysique_Dimension	Cette table donne les informations sur l'exécution physique de chaque projet
7	OSD_Dimension	Cette entité permet de renseigner les

		Orientations Stratégiques de Développement que chaque projet met en œuvre.
8	ModeFinancement_Dimension	Elle précise le type de financement du projet.
9	Domaine_Dimension	Elle donne le domaine dans lequel s'exécute le projet.
10	Secteur_Dimension	Cette table donne le secteur dans lequel le projet est mis en œuvre.
11	Indicateurs_Dimension	Elle spécifie les filières économiques du pays dans lequel le projet est élaboré.
12	Date_Dimension	Cette table représente la table de date.
13	Beneficiaire_Dimension	Elle possède la liste des bénéficiaires de chaque projet.

Source : Conçus par nous-mêmes

3.1.2 Table des faits

Tableau 9 : Description des mesures du datamart sur l'évolution des prévisions d'investissements publics

N°	Mesures ou faits	Signification
1	MontantFinancement	Le montant total financé par projet à ce trimestre
2	MontantDecaissement	Le montant total décaissé par projet au cours du trimestre.
3	ValeurdeBasePrevu	La valeur de base que l'indicateur de performance devait atteindre.
4	TauxExecutionPhysique	Le taux d'exécution physique du projet.
5	MoisExecutionPhysique	Il renseigne sur le mois au cours duquel le montant total décaissé a été renseigné.
6	AnneeExecutionPhysique	Il renseigne sur l'année au cours de laquelle le montant total décaissé a été renseigné.

Source : Conçus par nous-mêmes

Un tableau récapitulatif présente toutes les dimensions avec leurs attributs et les mesures de la table de faits. Il est représenté à la figure 17.

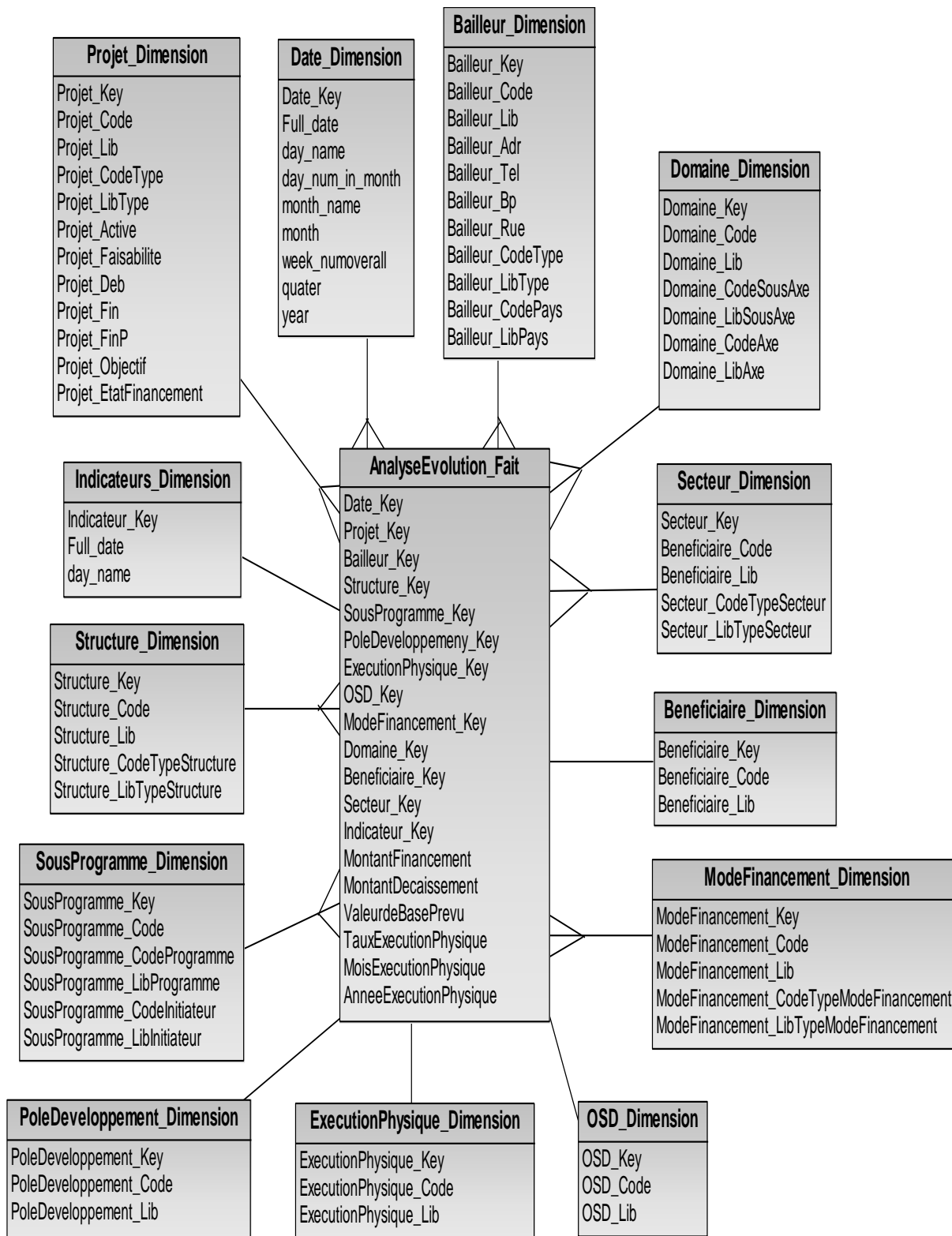
Figure 17 : Tableau des dimensions et fait du datamart Evolution des prévisions d'investissements publics

	Projet_Dimension	Bailleur_Dimension	Structure_Dimension	SousProgramme_Dimension	PoleDeveloppement_Dimension	ExecutionPhysique_Dimension	Indicateur_Dimension	OSD_Dimension	ModeFinanciere_Dimension	Domaine_Dimension	Secteur_Dimension	Beneficiaire_Dimension	Date_Dimension
	Projet_Key	Bailleur_Key	Structure_key	SousProgramme_Key	PoleDeveloppement_Key	ExecutionPhysique_Key	Indicateur_Key	OSD_Key	ModeFinancement_Key	Domaine_Key	Secteur_Key	Beneficiaire_Key	Date_Key
	Projet_Code	Bailleur_Code	Structure_Code	SousProgramme_Code	PoleDeveloppement_Code	ExecutionPhysique_Code	Indicateur_Code	OSD_Code	ModeFinancement_Code	Domaine_Code	Secteur_Code	Beneficiaire_Code	Full_date
	Projet_Lib	Bailleur_Lib	Structure_Lib	SousProgramme_CodeProgramme	PoleDeveloppement_Lib	ExecutionPhysique_Lib	Indicateur_Lib	OSD_Lib	ModeFinancement_Lib	Domaine_Lib	Secteur_LibSecteur	Beneficiaire_Lib	day_of_week
	Projet_CodeType	Bailleur_Adr	Structure_CodeType	SousProgramme_LibProgramme					ModeFinancement_CodeTypeModeFinancement	Domaine_CodeSousAxe	Secteur_CodeTypeSecteur		day_name
	Projet_LibType	Bailleur_Tel	Structure_LibType	SousProgramme_CodeInitiateur					ModeFinancement_LibTypeModeFinancement	Domaine_LibSousAxe	Secteur_LibTypeSecteur		day_num_in_month
	Projet_Active	Bailleur_Bp		SousProgramme_LibInitiateur						Domaine_CodeAxe			month_name
	Projet_Faisabilite	Bailleur_Rue								Domaine_LibAxe			month
	Projet_Deb	Bailleur_CodeType											week_numoverall
	Projet_Fin	Bailleur_LibType											quater
	Projet_FinP	Bailleur_CodePays											year
	Projet_Objectif	Bailleur_LibPays											
	Projet_EtatFinanciere												
Faits de													
AnalyseEvolution_Fait :													
MontantFinancement,													
MontantDecassement,													
ValeurdeBasePrevu,													
TauxExecutionPhysique,													
MoisExecutionPhysique,													
AnneeExecutionPhysique													

Source : Conçus par nous-mêmes

3.2 Modèle dimensionnel

Figure 18 : Schéma en étoile sur l'Evolution des prévisions d'investissements publics



Source : Conçus par nous-mêmes

Section 2 : Conception physique

Paragraphe 1 : Transformation et outil de développement

1 Règles de transformation Source / Destination

Après la modélisation dimensionnelle, nous avons procédé à la correspondance et aux transformations nécessaires pour la réalisation physique du modèle. A cet effet, un récapitulatif des transformations que subiront les données existantes au niveau des applications opérationnelles au niveau de l'ETL vers la base de datawarehouse est présenté en annexe 12.

2 Choix de l'outil de développement

L'outil utilisé pour la construction du datawarehouse est Microsoft SQL Server 2008. Les raisons de ce choix se situent à divers niveaux. Entre autres, nous pouvons citer :

- la facilité d'utilisation à cause de la documentation existant sur ce produit ;
- la gamme des outils utilisée par le MDAEP est la gamme Microsoft d'où, cette proposition ;
- la configuration standard (annexe 13) requis pour son installation est accessible par la plupart des petites structures ;
- la possibilité de construire un système datawarehouse complet en l'utilisant uniquement.

Paragraphe 2 : Modèle physique

Cette phase nous conduit à mieux appréhender l'architecture physique dans laquelle notre datawarehouse sera opérationnelle. L'objectif final sera d'allier performance et modèle physique de données. Plusieurs facteurs nous conduisent à déterminer l'infrastructure adéquate pour la construction de notre ED. Nous présenterons brièvement quelques-uns.

1 Le volume de données

Pendant la construction d'un entrepôt de données, il est nécessaire de le dimensionner en choisissant le Système de Gestion des Bases de Données (SGBD) de la machine cible. L'estimation porte le plus souvent sur la table des faits et bien entendu en tenant compte des dimensions et de la durée de stockage. A cet effet, nous proposons une approximation de l'espace disque requis pour nos bases de données. La méthode d'estimation consistera à évaluer le nombre de lignes moyen de projets mis en œuvre au cours des 12 derniers trimestres puis de multiplier ce nombre par la taille d'un enregistrement de la table des faits, ensuite multiplier par le nombre d'années de données historiques à garder. Le cas pratique se repose sur le modèle dimensionnel du Suivi financier des projets. En pratique, cela donne :

Temps d'historisation : 6 ans * 4 trimestres = 24 trimestres

Nombre de projets par trimestre = 1822 (Janvier 2010 à Décembre 2011)

Nombre de champs de clé = 13

Nombre de champs de fait = 10

Table des faits = $24 * 1822 * (20 * 8 + 3 * 2) = 31200 * (160 + 6) = 7\,258\,848$
Octets ~ 6,9225 GO

L'approximation de ce datamart se situe à 7 GO. En considérant le datamart sur les prévisions, nous pouvons proposer 1 TO pour disque dur. Cette capacité permettra de supporter les développements à court terme, et les extensions futures du système.

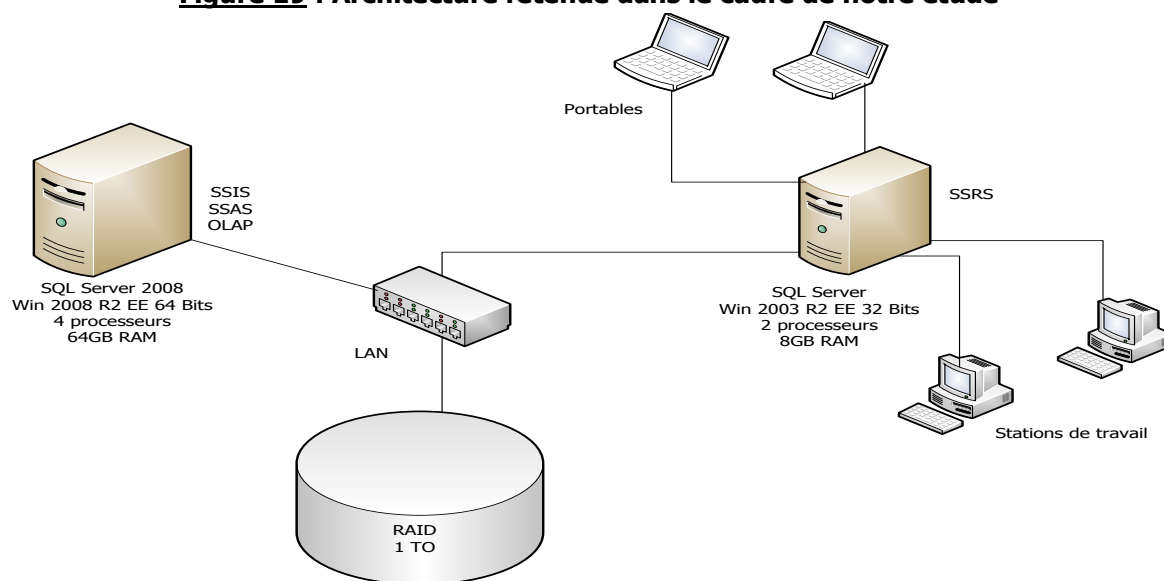
2 Performance

Les différents composants de SQL Server 2008 sont gourmands en mémoire. Dans le cadre du projet actuel, il est recommandé d'après le laboratoire MTI Microsoft¹⁷, un serveur 64 Bits avec 2 GB de mémoire au minimum ; et le processeur ayant 2 GHz. L'annexe 13 présente les spécifications plus en détail. L'architecture qui

¹⁷ SQL SERVER 2008 - Julien Stuyck

cadre pour la mise en place de l'ED est une base de données SQL Server et un serveur de Reporting séparés.

Figure 19 : Architecture retenue dans le cadre de notre étude



Source : Conçus par nous-mêmes

La conception du système à mettre en place est réalisée dans cette partie. Nous pouvons donc passer à la face pratique.

CHAPITRE IV : CREATION DU MODELE PHYSIQUE

Les différentes étapes de construction de l'entrepôt de données seront développées dans cette partie. Le prototype réalisé est basé sur le datamart du suivi financier des projets.

Section 1 : Mise en place des bases de données

La construction physique des différentes bases de données sera effectuée dans cette partie.

Paragraphe 1 : Base de donnée du datawarehouse

C'est la création de la base qui contient tous les informations extraites des différentes bases de production.

1 Script de création de la base de données

Un extrait du script de création de la base se présente comme suit :

```
USE [master]
GO
CREATE DATABASE [bd_projet] ON PRIMARY
( NAME = N'bd_projet', FILENAME = N'C:\Program Files\Microsoft SQL
Server\MSSQL10.MSSQLSERVER\MSSQL\DATA\bd_projet.mdf' , SIZE = 11264KB ,
MAXSIZE = UNLIMITED, FILEGROWTH = 1024KB )
LOG ON
( NAME = N'bd_projet_log', FILENAME = N'C:\Program Files\Microsoft SQL
Server\MSSQL10.MSSQLSERVER\MSSQL\DATA\bd_projet_log.ldf' , SIZE = 11200KB ,
MAXSIZE = 2048GB , FILEGROWTH = 10%)
GO
ALTER DATABASE [bd_projet] SET COMPATIBILITY_LEVEL = 100
GO
IF (1 = FULLTEXTSERVICEPROPERTY('IsFullTextInstalled'))
begin
EXEC [bd_projet].[dbo].[sp_fulltext_database] @action = 'enable'
end
GO
```

```
ALTER DATABASE [bd_projet] SET ANSI_NULL_DEFAULT OFF
GO
ALTER DATABASE [bd_projet] SET ANSI_NULLS OFF
GO
ALTER DATABASE [bd_projet] SET ANSI_PADDING OFF
GO
ALTER DATABASE [bd_projet] SET ANSI_WARNINGS OFF
```

2 Script de création destables

Le script de création de la table BAILLEUR sera présenté en guise d'exemple à cause du nombre important de tables existant.

```
USE [bd_projet]
GO
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
CREATE TABLE [dbo].[Bailleur](
    [CodeBailleur] [nchar](10) NOT NULL,
    [LibBailleur] [nchar](255) NULL,
    [AdrBailleur] [nchar](255) NULL,
    [TelBailleur] [nchar](15) NULL,
    [BPBailleur] [nchar](10) NULL,
    [RueBailleur] [nchar](50) NULL,
    [CodeTypeBailleur] [nchar](10) NULL,
    CONSTRAINT [PK_Bailleur] PRIMARY KEY CLUSTERED
(
    [CodeBailleur] ASC
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF,
IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS =
ON) ON [PRIMARY]
) ON [PRIMARY]
GO
```

Paragraphe 2 : Base de données dimensionnelle

La base dimensionnelle est créée à l'image du cube à réaliser. Cette base contient toutes les données que le schéma du cube aurait besoin pour les traitements.

1 Script de création de la base dimensionnelle

Elle se présente sous le même format que la base du datawarehouse à part le nom de la base.

2 Script de création de la table de dimension Bailleur

```
USE [bd_projet_dimensionnelle]
GO

SET ANSI_NULLS ON
GO

SET QUOTED_IDENTIFIER ON
GO

CREATE TABLE [dbo].[Bailleur_Dimension](
    [Bailleur_Key] [bigint] IDENTITY(1,1) NOT NULL,
    [Bailleur_Code] [nchar](10) NULL,
    [Bailleur_Lib] [nchar](255) NULL,
    [Bailleur_Adr] [nchar](255) NULL,
    [Bailleur_Tel] [nchar](15) NULL,
    [Bailleur_Bp] [nchar](10) NULL,
    [Bailleur_Rue] [nchar](50) NULL,
    [Bailleur_CodeType] [nchar](10) NULL,
    [Bailleur_LibType] [nchar](255) NULL,
    [Bailleur_CodePays] [nchar](10) NULL,
    [Bailleur_LibPays] [nchar](255) NULL,
    CONSTRAINT [PK_Bailleur_Dimension] PRIMARY KEY CLUSTERED
    (
        [Bailleur_Key] ASC
    )WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF,
    IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS =
    ON) ON [PRIMARY]
) ON [PRIMARY]

GO
```

Section 2 : Construction de l'entrepôt de données

Dans la première partie, nous avons proposé la réalisation d'un Entrepôt de Données (ED) pour un meilleur suivi des PIP au Bénin. Après les différentes étapes suivies depuis la conception à la construction physique de l'ED, nous sommes en mesure d'interroger notre base de données obtenue dans le but de réaliser des états de suivi. La deuxième étape présente la facilité d'accès aux données importées des bases de production dans la réalisation des différents états de suivi. Une brève présentation de la construction de l'ETL qui a permis l'extraction des données des diverses bases de production fera suite. Les différentes bases de données rencontrées sont conçues soit en Access, en Excel, en World et sous MySQL.

Paragraphe 1 : Mise en œuvre et déploiement des cubes

Des différentes situations de suivi présentées dans le rapport d'avancement sur le PIP, nous avons choisi d'en réaliser quelques-uns parmi lesquels : la répartition du PIP par source de financement. Cet état présente la contribution financière de chaque bailleur de fonds dans la réalisation du PIP. Cette contribution est évaluée par le poids et le taux d'accroissement.

Le prototype réalisé porte sur les données de 3 ans c'est-à-dire de 2010 à 2012. En effet, à cause de la difficulté rencontrée dans le traitement des données, trois années ont pu être réellement récupérées dans la base. De nombreuses données ont été laissées par l'ETL car ne respectant pas tous les critères. Ceci entraîne donc que les résultats obtenus ne seront pas forcément le même que les situations réelles, mais le résultat est fiable. L'utilisation du livre Learning SQL Server 2008 Reporting Services (Jayaram Krishnaswamy, 2009), nous a permis de suivre les différentes étapes pour la construction de rapports. .

1 Réalisation du rapport sous Microsoft Visual Studio

1.1 Réalisation du rapport

Après la réalisation du rapport sur la répartition du PIP par source de financement, nous voulons simuler un dirigeant qui désire obtenir tous les financements des projets par source de financement sur une période de 3 ans. Il a la possibilité d'obtenir l'état chaque année ou sur un seul état pour les trois années. Pour cela, il a besoin d'avoir un profil pour accéder aux divers rapports disponibles et de sélectionner ce qui l'intéresse. Le rapport étant paramétré sur l'année, il n'aura qu'à exprimer son besoin par l'entrée de cette dernière. Les informations disponibles s'afficheront suivant les critères de ces derniers. L'exemple de rapport paramétré permettant de renseigner l'année de recherche a été réalisé sous Microsoft Visual Studio 2008.

L'exemple suivant présentera l'obtention d'un état de répartition du PIP par source de financement sur trois ans.

1.2 Publication de l'état sur une page web

Une fois le rapport construit, nous allons le publier sur un serveur afin de le rendre accessible aux utilisateurs via un site web. Ces derniers suivant leurs profils pourront y accéder par un browser en tapant l'adresse du site. Cette action consiste à déployer l'état. L'outil utilisé à cet effet est le Report Manager de Microsoft.

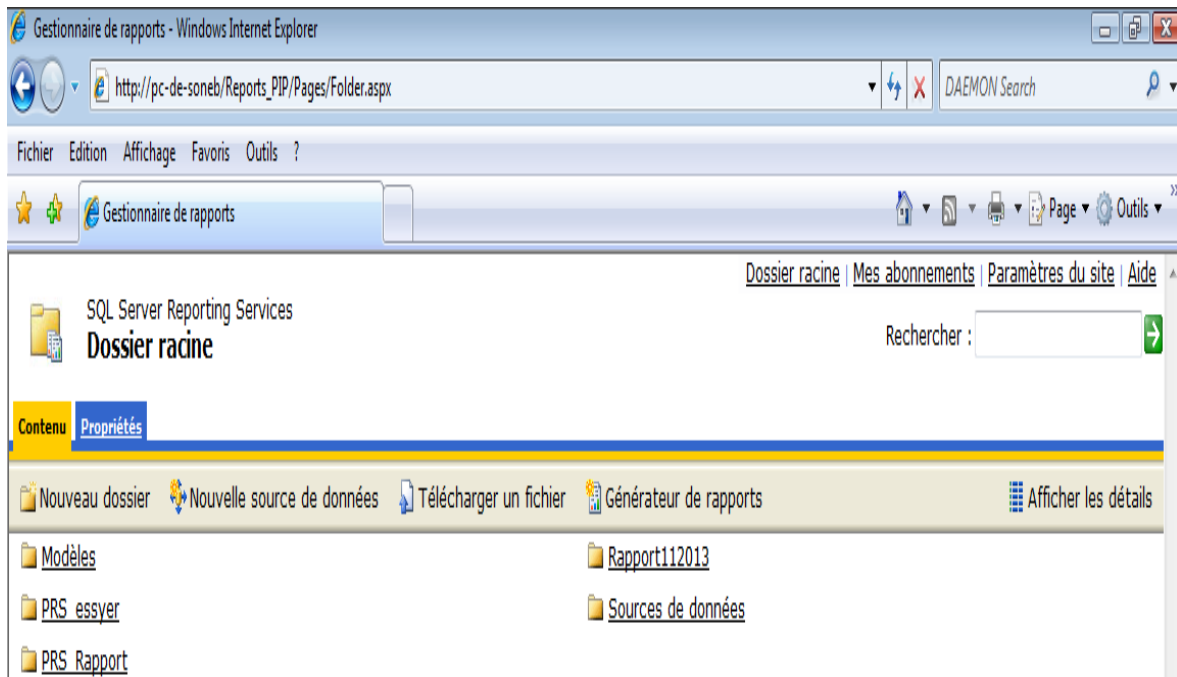
➤ Pré-requis

Il faut à cet effet :

- activer l'outil OverWriteDataSources en le mettant true et spécifier le chemin pour accéder au serveur ;
- Déployer le projet

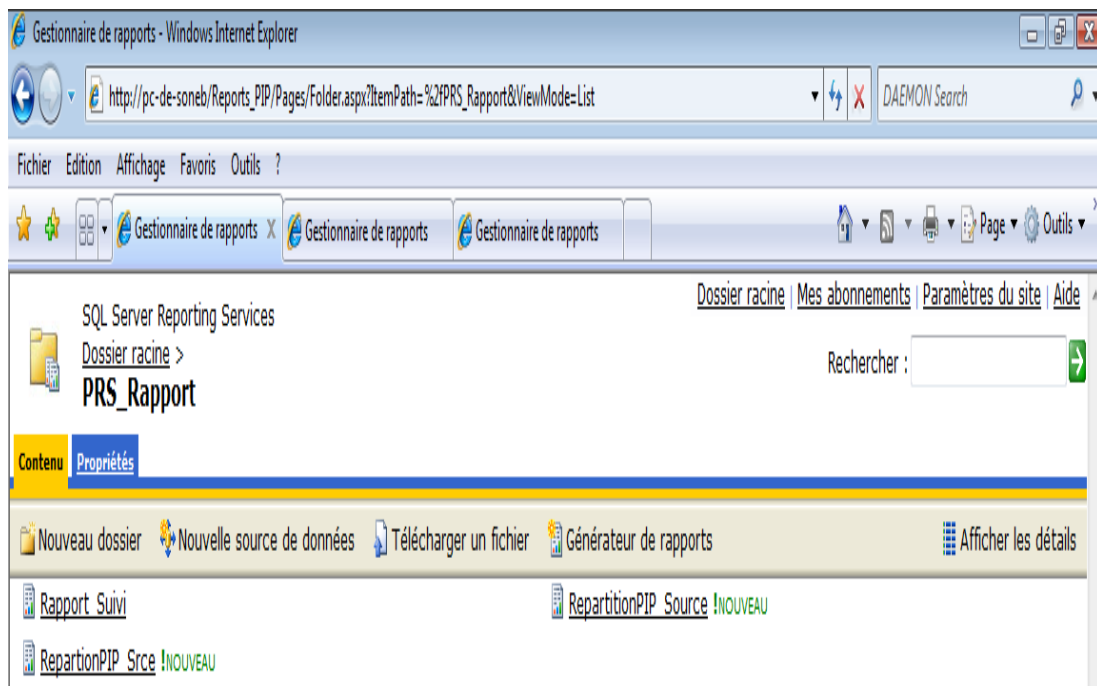
➤ Affichage du rapport

Pour accéder au serveur, il taper le lien suivant : http://pc-de-soneb/Reports_PIP/. En tapant cette adresse, nous obtenons :

Figure 20 : Page d'accueil du Gestionnaire de rapport

Source : Conçus par nous-mêmes

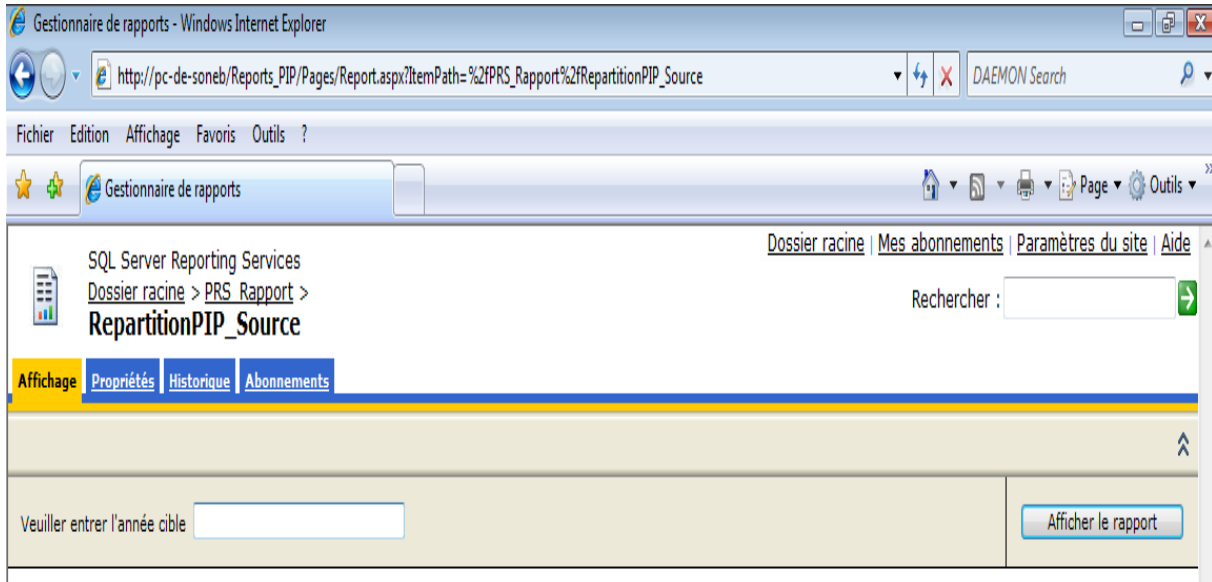
Cette page présente tous les différents rapports réalisés et rendus disponibles pour les décideurs. Dans notre exemple, nous cliquons sur le dossier PRS_Rapport car c'est ce dossier qui contient l'état recherché, nous obtenons :

Figure 21 : Affichage des rapports disponibles dans le rapport PRS_Rapport

Source : Conçus par nous-mêmes

Lorsque les rapports sont publiés pour la première fois, le serveur le signale. C'est à cet effet que nous remarquons "!Nouveau" devant des noms de rapport. En cliquant sur le nouveau rapport RepartitionPIP_Srce, on obtient :

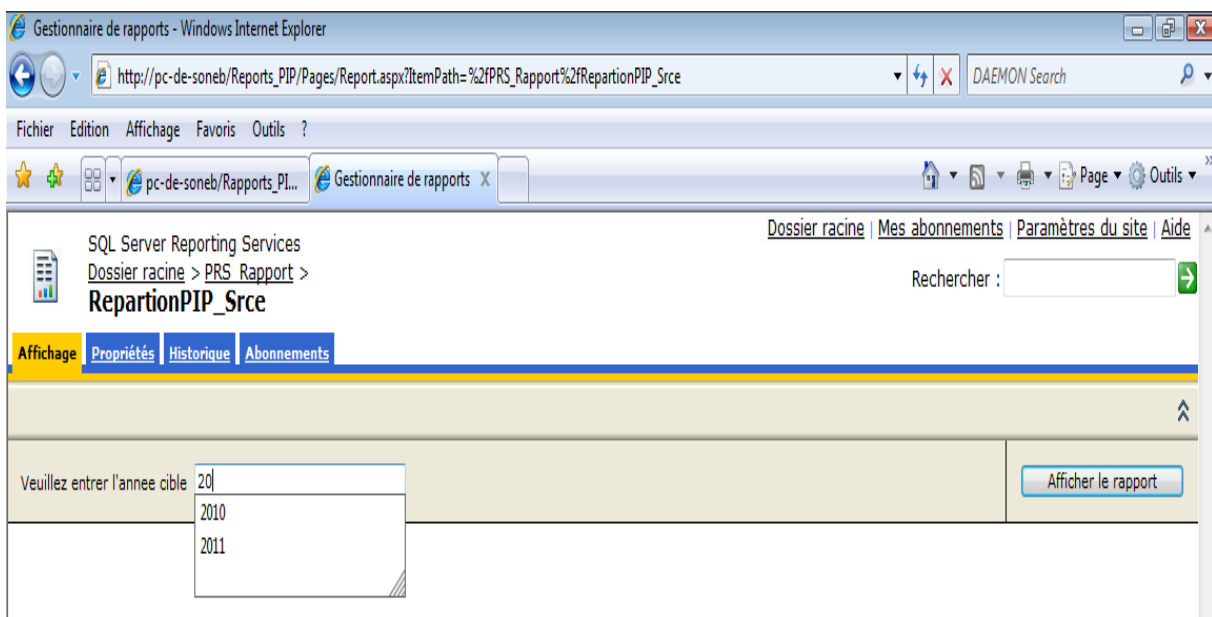
Figure 22 : Affichage du rapport dans le Gestionnaire de rapports



Source : Conçus par nous-mêmes

Le rapport étant paramétré, nous avons une invite de commande demandant l'année cible, ici nous tapons 2010 :

Figure 23 : Saisie du paramètre de rapport dans le Gestionnaire de rapport



Source : Conçus par nous-mêmes

Le rapport suivant s'affiche en validant :

Figure 24 : Affichage du rapport Répartition par source de financement

Source de financement	Prevision 2010 (en millions de FCFA)	Poids Prévision 2010 par rapport au Total Général (%)	Prevision 2011 (en millions de FCFA)	Poids Prévision 2011 par rapport au Total Général (en%)	Taux d'accroissement de 2010 par rapport à 2011 (en %)	Prevision 2012 (en millions de FCFA)	Poids Prévision 2012 par rapport au Total Général (%)	Taux d'accroissement de 2011 par rapport à 2012 (en %)
GIZ	0	0,00	221782	2,53	0,00	0	0,00	100,00
Bailleurs multilatéraux	0	0,00	3404	0,04	0,00	0	0,00	100,00
BAD	0	0,00	35056	0,40	0,00	0	0,00	100,00
FED	0	0,00	595734	6,79	0,00	0	0,00	100,00
PNUD	0	0,00	11500	0,13	0,00	0	0,00	100,00
Budget National	30008	8,63	6747451	76,87	0,00	0	0,00	99,56
BOAD	106582	30,65	0	0,00	0,00	0	0,00	0,00
OPEP	0	0,00	7015	0,08	0,00	0	0,00	100,00
BID	0	0,00	446119	5,08	0,00	0	0,00	100,00
Autres	0	0,00	49698	0,57	0,00	0	0,00	100,00
KFW	0	0,00	5265	0,06	0,00	0	0,00	100,00
Belgique	0	0,00	90900	1,04	0,00	0	0,00	100,00

Source : Conçus par nous-mêmes

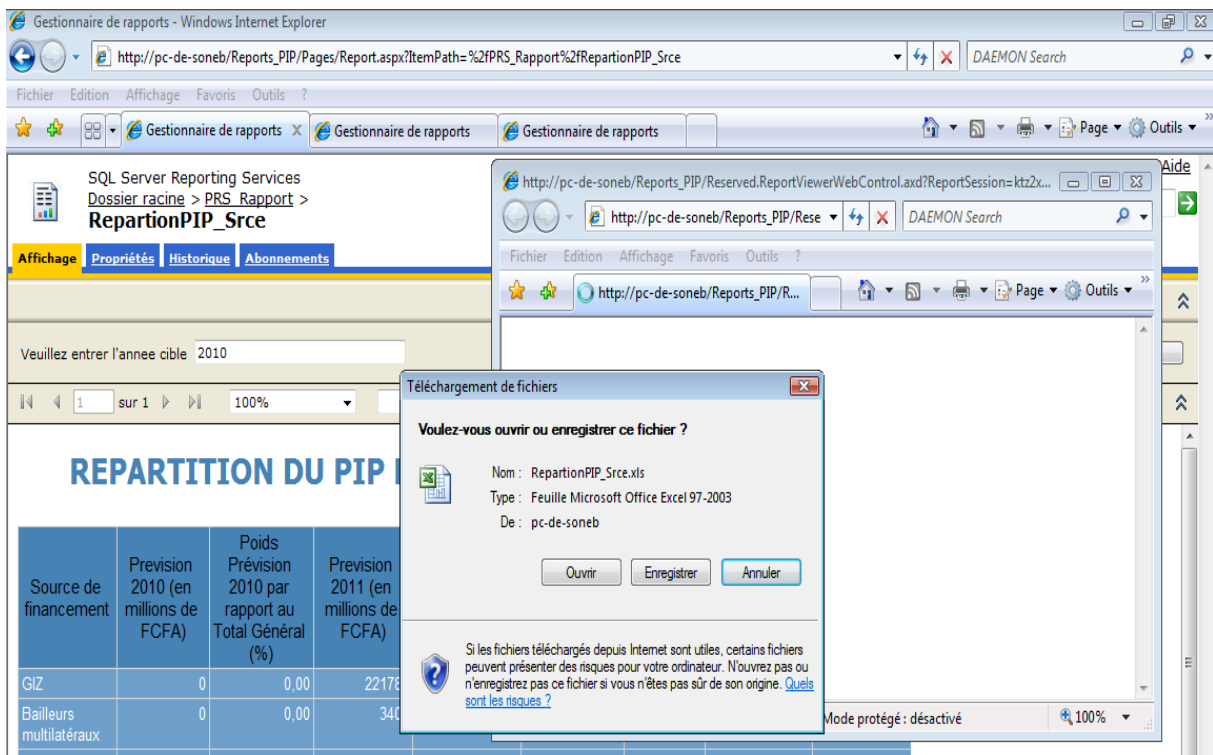
Le tableau affiché présente toutes les prévisions par année pour chaque bailleur, calcule le poids de ce montant et le taux d'accroissement de la prévision.

Ainsi le décideur, peut suivant ses critères dans la zone de paramètre, voir son état s'afficher. Arriver à exporter un rapport sous un format désiré devient aisé. Le gestionnaire de rapports offre de nombreuses possibilités d'exporter le rapport. En sélectionnant le format d'export, dans notre cas Excel, nous obtenons :

Figure 25 : Sélection du format d'export

Source : Conçus par nous-mêmes

Figure 26 : Exécution de la commande d'export sous Excel



Source : Conçus par nous-mêmes

Figure 27 : Affichage du rapport exporté sous Excel

The screenshot shows the Excel spreadsheet with the following data:

Source de financement	Prevision 2010 (en millions de FCFA)	Poids Prévision 2010 par rapport au Total Général (%)	Prevision 2011 (en millions de FCFA)	Poids Prévision 2011 par rapport au Total Général (en%)	Taux d'accroissement de 2010 par rapport à 2011 (en %)	Prevision 2012 (en millions de FCFA)	Poids Prévision 2012 par rapport au Total Général (%)	Taux d'accroissement de 2011 par rapport à 2012 (en %)
GIZ	0	0,00	221782	2,53	0,00	0	0,00	100,00
Baillleurs	0	0,00	3404	0,04	0,00	0	0,00	100,00
BAD	0	0,00	35056	0,40	0,00	0	0,00	100,00
FED	0	0,00	595734	6,79	0,00	0	0,00	100,00
PNUD	0	0,00	11500	0,13	0,00	0	0,00	100,00
Budget	30008	8,63	6747451	76,87	0,00	0	0,00	99,56
BOAD	106582	30,65	0	0,00	0,00	0	0,00	0,00
OPEP	0	0,00	7015	0,08	0,00	0	0,00	100,00
BID	0	0,00	446119	5,08	0,00	0	0,00	100,00
Autres	0	0,00	49698	0,57	0,00	0	0,00	100,00
KFW	0	0,00	5265	0,06	0,00	0	0,00	100,00
Belgique	0	0,00	90900	1,04	0,00	0	0,00	100,00
Pays-Bas	0	0,00	113900	1,30	0,00	0	0,00	100,00
AFD	0	0,00	426690	4,86	0,00	0	0,00	100,00
Danemark	211152	60,72	0	0,00	0,00	0	0,00	0,00
GEF	0	0,00	23000	0,26	0,00	0	0,00	100,00

Source : Conçus par nous-mêmes

Nous obtenons de ce fait le tableau dans un outil que de nombreux utilisateurs connaissent et peuvent facilement traiter l'information. Le décideur peut à son aise obtenir les données sous des diagrammes, les imprimer, etc.

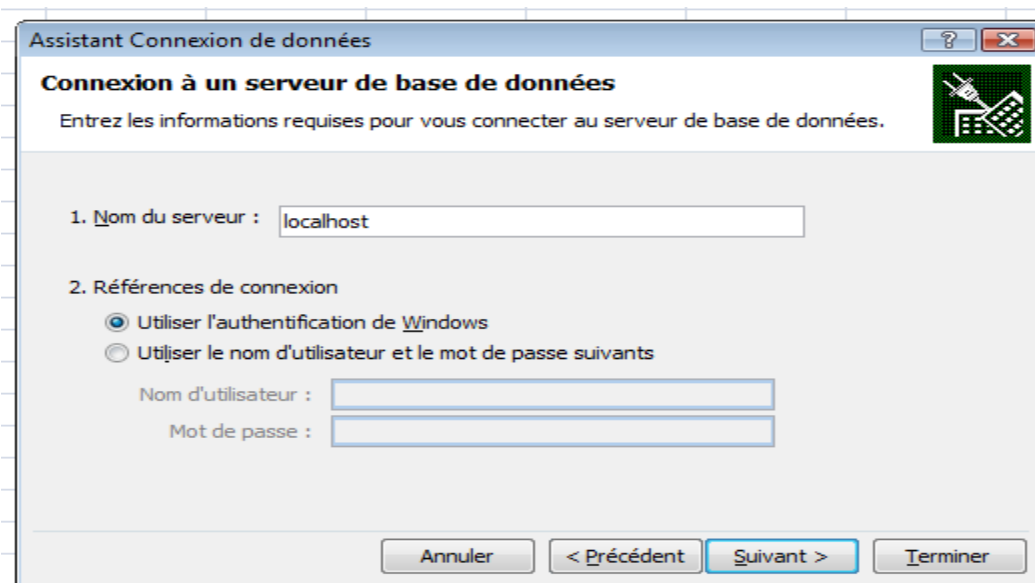
Nous pouvons aussi accéder aux données multidimensionnelles sous Microsoft Excel. Et comme précédemment, l'utilisateur peut traiter l'information comme il le désire (graphes, impression, etc.). Dans le cas d'espèce nous utilisons Excel 2007. Les cubes sont accessibles via un outil OLAP, un explorateur de cube ou une application framework.

2 Restitution des informations par Excel/OLAP

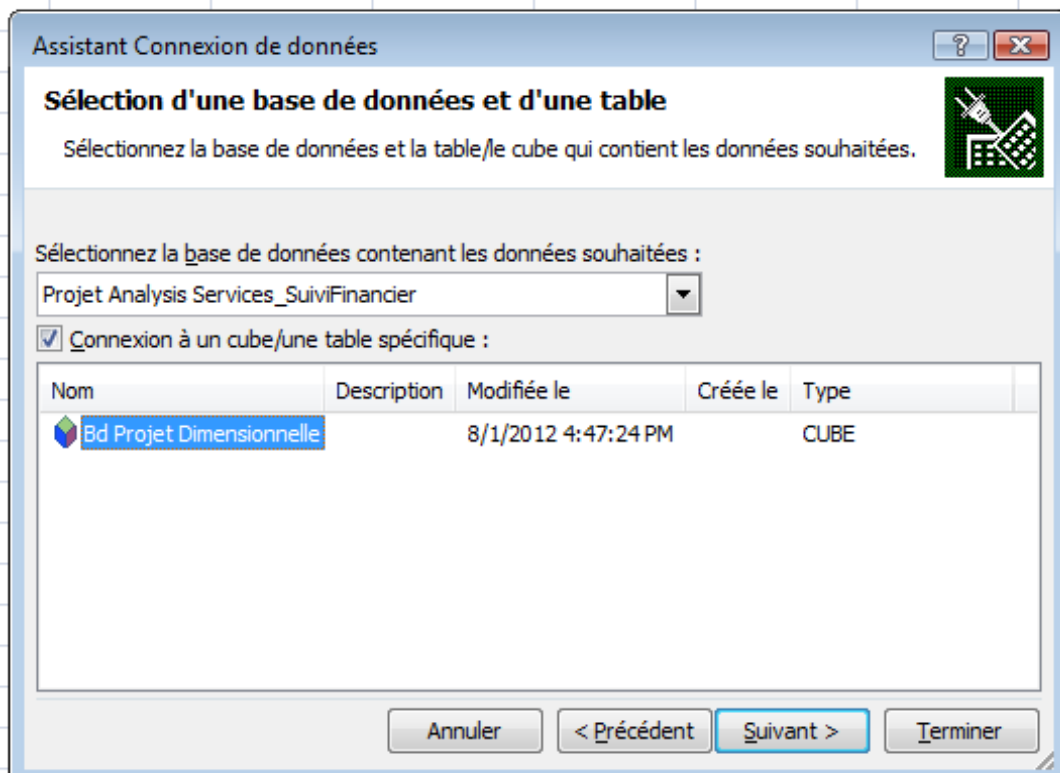
L'avantage de cette méthode est que le décideur a un accès sur des cubes conçus pour effectuer des analyses. Ce dernier maîtrisant mieux les données pourra lancer des requêtes suivant les axes qu'il désire et obtenir des résultats rapidement car les informations existent et sont disponibles. Pour ce fait, la procédure d'accès à ces données est décrite comme suit :

- Ouvrir Excel 2007 ;
- Se connecter à la base de données multidimensionnelle ;
- Procéder à une requête d'Analyse croisée dynamique comme suit:

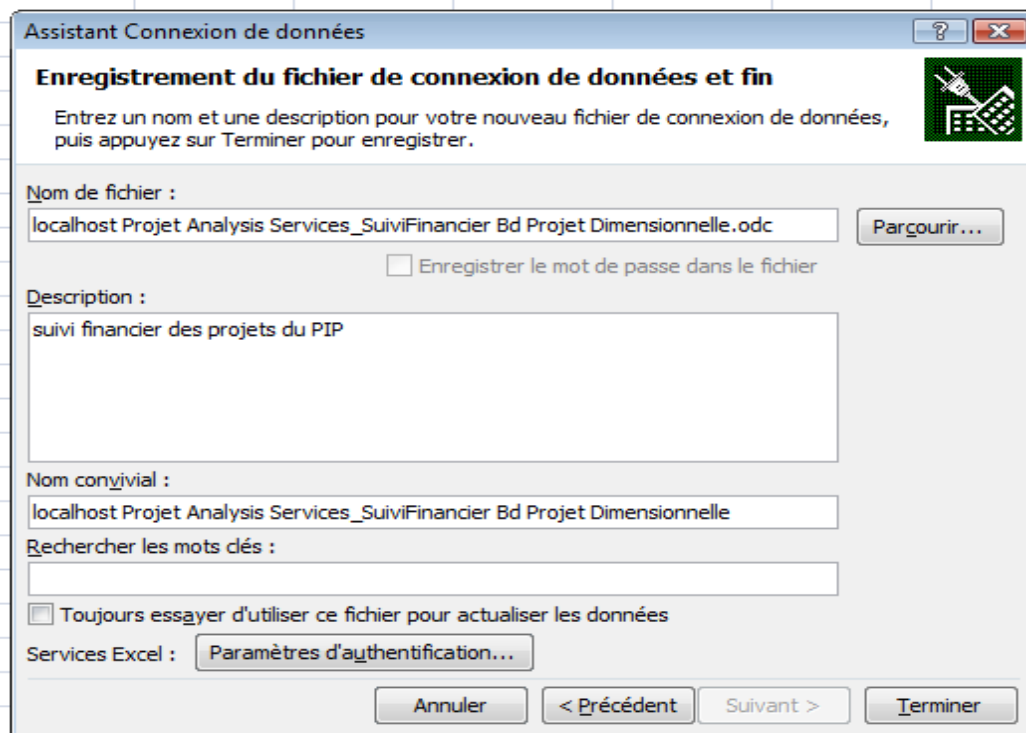
- **Figure 28: Connexion à la base de données dimensionnelle**



Source : Conçus par nous-mêmes

Figure 29 : Choix du cube à utiliser comme source de données

Source : Conçus par nous-mêmes

Figure 30 : Enregistrement du fichier de connexion

Source : Conçus par nous-mêmes

Figure 31 : Création d'un tableau de données montrant quelques indicateurs par structure

	Valeurs			
Étiquettes de lignes	Montant Decaissement	Montant Engagement	Montant Ordonnancement	Montant Alloue Commune
Ministère de la Justice, de la Législation et des Droits de l'Homme (MJLDH)	0	0	0	600
Ministère de la Santé (MS)	0	457968	122080	92
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique (M)	135000	177570	135000	30
Ministère de l'Industrie, du Commerce, des Petites et Moyennes Entrepris	0	0	0	145
Ministère des Affaires Etrangères, de l'Intégration Africaine, de la Francop	10000	20818	20818	840
Total général	145000	656356	277898	1707

Source : Conçus par nous-mêmes

Les possibilités d'accès aux données disponibles dans l'ED sont multiples. Une formation devrait être donnée aux décideurs et utilisateurs du système pour accroître leurs connaissances.

Nous voilà capable de réaliser divers rapports suivant des critères que nous désirons. Notre tâche est d'expliquer comment nous y sommes arrivés.

Paragraphe 2 : Mise en œuvre de l'ETL

La réalisation des résultats attendus par les dirigeants demande un travail préalable. Il s'agit d'extraire les données des bases de production et de les rendre disponibles dans l'entrepôt de données. Après cette étape, la construction des différents cubes permettant diverses analyses pourra être effectuée.

1 Extraction des données des différentes bases de production

L'extraction des différentes données ayant les SGBD de la gamme Microsoft s'est déroulée sans difficultés majeures car les outils étaient intégrés à SQL Server 2008. Cependant, l'extraction des données au niveau de la base MySQL a nécessité l'installation d'un connecteur ODBC. Le connecteur utilisé dans notre cas est : MySQL ODBC 3.51, téléchargeable à l'adresse <http://dev.mysql.com/downloads/connector/odbc/>.

1.1 Extraction des données sous MySql

Les données doivent être extraites de la base pipgestion avant les données des autres sources.

➤ Création de la source ODBC

La procédure d'installation et de configuration du connecteur est disponible sur le site dbperf.wordpress.com/2010/07/22/link-mysql-to-ms-sql-server2008/.

➤ Extraction de la base de données « pipgestion »

Cette tâche demande la création d'un projet SSIS. Les transferts de données se sont effectués suivant les liaisons des tables. C'est-à-dire que si une table A a une clé étrangère avec une table B, il faut charger la table B avant la table A. C'est sur cette base que les travaux se sont déroulés.

Le processus de transfert de données de MySQL vers SQL Server est décrit dans l'article [belle'ssqlmusing](http://www.sqlmusings.com/2011/03/23/transfer-data-from-mysql-to-sql-server-using-ssis/) à l'adresse

<http://www.sqlmusings.com/2011/03/23/transfer-data-from-mysql-to-sql-server-using-ssis/>.

➤ Transformation des données

Après cette étape, les données extraites sont transformées. L'outil Conversion de données est utilisé pour effectuer ce traitement.

➤ **Transfert des données vers la base datawarehouse**

Le chargement de la base datawarehouse s'effectue en choisissant l'outil Dimension à variation lente. Cet outil permet de charger la source de données et en cas de redondance de données, n'exécute pas le script.

1.2 Extraction des données sous fichier Excel

Le fichier Excel reçu a subi différentes transformations du fait de la non concordance de ses informations avec celles de SIAPIP. En effet, les projets recensés sur le fichier Excel et retraçant l'état d'avancement de chaque projet ne pouvaient être comparés à ceux existant dans la Base de données PIPGestion. Nous avons alors procédé à une correspondance en attribuant les numéros réels existant dans la base SIAPIP, aux projets du fichier Excel. Ainsi nous obtenons le fichierEXECUTION_FINANCIERE.Xcl présenté à la figure 45. Nous remarquons la colonne ID_PROJET qui renseigne l'identifiant du projet. Avec cet identifiant, nous arriverons à lier les projets existant dans pipgestion pour avoir une seule base de données. L'extraction de ce fichier dans l'entrepôt de données suit pratiquement les mêmes principes qu'au niveau de la Source ADO NET. Cependant la source à utiliser est Source Excel. L'article SQL Server 2005 Integration Services : Import/export Excel écrit par Fleur-Anne BLAIN (Fleur-Anne BLAIN, 2008), présente la procédure à mettre en application pour réussir l'importation des données de la source Excel à la source SQL Server. Il est téléchargeable sur www.fabian.developpez.com.

1.3 Déploiement et Vérification du package SSIS

L'exécution du package est possible via Visual Studio comme nous l'avons fait précédemment. Cependant, un déploiement du package sur le serveur est plus avantageux, en particulier pour automatiser l'exécution. L'article de Fleur-Anne Blain intitulé Présentation de SQL Server Integration Services : SSIS traite de cette question et montre la procédure pour déployer et de procéder à la vérification un package SSIS.

1.4 Création et déploiement des cubes

- Création d'un nouveau projet SASS ;
- Connexion d'une source de données ;
- Création d'une vue de données source pour le cube ;
- Création d'un nouveau cube.

Le cube est ensuite déployé pour son utilisation. C'est après cette étape que les rapports présentés en début de cette section sont réalisés grâce au SQL Server Reporting Services (SSRS). SSRS offre de nombreuses autres possibilités comme l'abonnement à un état, la publication régulière d'un état via un processus automatisé, etc.

2 Sécurité et conditions de mise en œuvre

2.1 Sécurité

2.1.1 Sécurité d'accès à la base de données

Les utilisateurs accèderont à la base suivant des profils définis.

- Les sauvegardes et les restaurations du serveur d'informations centralisées devront se dérouler selon des formes assurant une totale sécurité physique et logique.
- Au moins deux niveaux de sensibilité à la sécurité devront être définis pour implémenter un nouveau standard de confidentialité dans l'organisation du ministère. Les informations géographiques de type général se verront attribuer un moindre niveau de sécurité. Les noms, les numéros de compte et des informations privées de type financier ou médical se verront affecter un niveau de sécurité plus élevé.
- Une base de données d'audit retraçant chaque utilisation d'informations personnelles doit accompagner la base de données centrale. Cette base de données d'audit doit permettre de satisfaire l'exigence de notification à chaque individu de toutes les utilisations de ses informations personnelles, en précisant qui est le demandeur et quelle est l'utilisation. La base de données

d'audit comporte le moins possible de contenu confidentiel, se bornant à indiquer l'utilisation finale des informations.

- Une interface doit être prévue pour établir l'authenticité des demandes individuelles et fournir ensuite au demandeur une copie de toutes ses informations personnelles présentes dans la base de données. Une seconde interface doit permettre à l'individu de contester, de commenter ou de corriger les informations.
- Un mécanisme doit exister pour détruire effectivement les informations qui s'avèrent incorrectes, détenues illégalement ou périmées. Les informations détruites doivent l'être de manière à ne pas pouvoir réapparaître ultérieurement.

2.1.2 Sécurité au niveau des applications web

Une étape d'identification des besoins de sécurité et des menaces doit être suivie. Après cette étape, il convient de concevoir la sécurité de l'application Web, c'est-à-dire de définir précisément les mécanismes de sécurité qui permettront d'assurer l'atteinte des objectifs de sécurité et de répondre aux menaces (authentification, contrôle d'accès, protection contre les injections, etc.). Un document de conception doit décrire formellement ces mécanismes, en fournissant une bonne visibilité sur la manière dont la sécurité sera organisée et les menaces seront gérées. Les mécanismes de sécurité ainsi conçus doivent en outre respecter le principe de la défense en profondeur, qui veut que si une ligne de défense est défaillante ou franchie par l'attaquant, une deuxième voire une troisième ligne soient en place pour protéger les ressources.

Une fois la conception de l'application et de sa sécurité réalisée, vient la phase d'implémentation qui voit les développeurs générer le code source et assembler les composants. Il est alors primordial que les personnes en charge de l'implémentation aient été spécialement formées au développement sécurisé d'applications Web. En outre, des outils doivent être fournis aux équipes :

- un référentiel documentaire qui présente les bonnes pratiques de développement sécurisé ;

- une check-list qui permet de s'autocontrôler et de s'assurer que rien n'a été oublié ;
- des API ou un framework « sécurité » qui évitent d'avoir à recréer des mécanismes de sécurité de base (authentification, filtrage des données, etc.). Une attention particulière doit être portée sur l'utilisation d'API ou de framework internes ou externes qui doivent être validés ou audités afin de s'assurer qu'ils ne portent pas de vulnérabilités ou de code malveillant.

Dans cette phase également, la possibilité pour les équipes de consulter un expert sécurité afin de valider les mécanismes de sécurité développés est un plus. Les équipes peuvent également se référer au Guide de conception et d'implémentation d'applications Web sécurisées de l'OWASP¹⁸.

A l'issue de l'implémentation, de même qu'une recette est réalisée par les utilisateurs, la sécurité de l'application doit être vérifiée, afin de valider de manière pragmatique que l'application se comporte de manière sécurisée face aux attaques.

2.2 Conditions de mise en œuvre

2.2.1 Mise en service

La mise en route de l'entrepôt de données demande nécessairement que la plupart des données qui doivent être utilisées, soient déjà intégrées. De plus, les applications permettant d'accéder aux données doivent être créées.

La formation des utilisateurs et l'assistance associées à la mise en service sont envisagées. En effet, l'acceptation de l'entrepôt de données étant indispensable à son succès, la formation est incontournable. Le programme de formation doit porter sur tout ce que fournissent l'entrepôt de données, les données, et des applications analytiques. La stratégie à adopter en matière de support dépend des attentes du management ainsi que de la réalité des produits offerts par l'entrepôt de données. Le support est souvent structuré en deux niveaux – la première ligne

¹⁸ SQL Injection from OWASP, https://www.owasp.org/index.php/SQL_Injection

d'expertise est répartie au sein des activités du ministère et une deuxième ligne de défense est représentée par un support centralisé. Nous devons trouver les ressources nécessaires au support et définir les modalités de son utilisation. Il convient également de préparer un plan pour la maintenance et la mise en service des modifications et une stratégie de diffusion d'informations sur l'entrepôt de données.

La diffusion d'un produit logiciel passe par une série de phases avant sa disponibilité définitive et il en va de même pour l'entrepôt de données¹⁹. La phase de test alpha est représentée par un test de toutes les applications d'un bout à l'autre par le noyau dur de l'équipe du projet. Comme dans tout test d'intégration, nous rencontrerons forcément des problèmes et nous avons prévu du temps dans la planning pour les inévitables corrections. Le test beta fait appel à un nombre limité d'utilisateurs qui font un test d'acceptation, concernant plus spécialement l'adéquation des produits aux besoins de l'entreprise et leur qualité intrinsèque. Après ces différents tests, nous mettrons l'entrepôt de données à disposition de tous ses utilisateurs.

2.2.2 Maintenance et expansion

Après la mise en service de l'entrepôt de données, nous devons continuer d'investir des ressources dans les domaines suivants :

Support . Aussitôt après la mise en service, le support aux utilisateurs est un élément crucial pour leur faire apprécier l'utilisation de l'entrepôt de données. Au cours des premières semaines consécutives à la formation des utilisateurs, l'équipe de support doit travailler avec eux en anticipant leurs besoins. Une équipe sera dégagée temporairement et installée près des utilisateurs, pour qu'ils accèdent facilement au support et aussi pour des corrections éventuelles.

¹⁹ Ralph kimball et Margy Ross (2002) : « Entrepôts de données, Guide pratique de modélisation dimensionnelle » (2^e édition),pp 373

Formation. Nous devons offrir un programme de formation permanente aux utilisateurs de l'entrepôt de données. Le programme officiel doit comporter des cours de rappels et des cours avancés, ainsi qu'une répétition des cours d'initiation. Une formation informelle peut être aussi proposée aux développeurs ainsi qu'aux utilisateurs intensifs pour promouvoir les échanges d'idées.

Support technique. 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. Les décisions relatives à la croissance de l'entrepôt de données ne doivent pas être prises au hasard. Les priorités doivent être fixées en fonction des besoins de l'entreprise. A cet effet, il est souhaitable, qu'un comité de soutien de l'entrepôt de données, composé de responsables représentant les utilisateurs et qui participe à la fixation des priorités, soit mis sur pied.

Ce chapitre a permis de présenter quelques fonctionnalités disponibles sur le prototype réalisé. Il présente surtout les diverses possibilités offertes aux dirigeants pour améliorer la recherche des différentes données indispensables à leurs études. Ces données étant pour la plupart disponibles dans l'Entrepôt de Données. Les conditions à mettre en place pour assurer la sécurité au niveau de l'entrepôt de données ont été abordées ainsi que les conditions de mise en œuvre.

La deuxième partie présente la phase pratique de notre travail de recherche. Elle a permis de présenter d'une part la conception du modèle de l'entrepôt à mettre en place et d'autre part notre prototype.

CONCLUSION GENERALE

Le gouvernement béninois se base sur les résultats de suivi des Programmes d'Investissements Publics (PIP) pour étudier l'évolution de son développement. Les tâches effectuées pour élaborer les différents rapports de suivi et d'analyse rencontrent aujourd'hui des difficultés. En effet, les Technologies de l'Information et de la communication permettent de concevoir des systèmes d'informations particulièrement performants qui fournissent d'importantes informations. Cependant, les systèmes opérationnels ne sont pas conçus pour permettre l'utilisation des données produites dans un processus d'aide à la décision.

Le datawarehouse a pour but de fournir des observations qui alimenteront, après recoupement avec d'autres sources (économiques, démographiques, marketing, etc.) la compréhension du système de suivi et permettront de suivre l'activité, d'analyser son efficacité, d'optimiser les moyens, le tout en vue de faciliter l'orientation de l'action. Le datawarehouse permet au décideur de travailler dans un environnement informationnel, référencé, homogène et historisé. Les décideurs disposent ainsi de toutes les informations nécessaires pour des analyses avancées.

Par ailleurs, l'analyse multidimensionnelle permet d'extraire du datawarehouse, une connaissance quantitative de l'activité de l'organisation selon les dimensions ou les aspects qui préoccupent le décideur à un instant donné.

Néanmoins, le datawarehouse n'a pas vocation à fournir des indicateurs pour un pilotage opérationnel au jour le jour ou pour le suivi de dossiers individuels : il faut que chaque processus de production fournisse les indicateurs et alarmes qui permettent aux responsables opérationnels de piloter les travaux au plus près de leur réalisation.

Le projet de mise en place d'un datawarehouse pour un meilleur suivi des Programmes d'Investissements au Bénin, a été initié pour rendre l'information décisionnelle à la portée des décideurs. Cet outil facilitera les traitements dans l'élaboration des différentes statistiques.

Cependant si cet outil rend le ministère performant, cela pose aussi de nouveaux problèmes (problèmes de confidentialité, de compétences pour l'analyse, etc.). Il est urgent pour le ministère de résoudre ces différents problèmes, dans le but de sécuriser les tâches exécutées sur le datawarehouse.

Notre étude nous a permis de comprendre les différents facteurs qui interviennent dans l'analyse de l'évolution du développement. L'application de réalisation d'un entrepôt de données avec SQL Server 2008, nous a permis de découvrir les multiples outils qu'offre ce logiciel.

Le projet de mise en place d'un entrepôt de données pour le gouvernement béninois est immense. Nous n'avons exploré qu'une infime partie de ce vaste projet. Les travaux de recherche continueront pour couvrir tous domaines ciblés dans le but de construire un entrepôt exploitable par tous les acteurs de décision. Le périmètre d'application de ce projet pourrait s'étendre aux techniques de simulation et de prévisions à travers la mise à contribution d'autres outils tiers comme le datamining.

BIBLIOGRAPHIE

RÉFÉRENCES BIBLIOGRAPHIQUES

(B. Inmon, 1996) W.H Inmon, "Building the Data Warehouse", John Wiley & Sons, 1996.

(Bouquin, 2003) Bouquin Henry ; "Le contrôle de gestion" ; P.U.F ; 2003.

(Jean-Marie Gouarné, 1997) Jean-Marie Gouarné, "Le Projet Décisionnel Enjeux, Modèles, Architectures du Data Warehouse", Éditions Eyrolles ,1997

(Kimball, 1996) Ralph Kimball, "The data warehouse toolkit: practical techniques for building dimensional data warehouses", John Wiley & Sons, ISBN 0-471-15337-0, 1996.

(Kimball, 2000) Ralph Kimball, "Concevoir et déployer un datawarehouse", Editions Eyrolles, ISBN: 2-212-09165-6, 2000.

(Kimball et Caserta, 2004) Ralph. Kimball et Joe. Caserta, "The Data warehouse ETL Toolkit", Wiley Publishing, 2004

(Kortink et al, 1999) Mark A. R. Kortink, Daniel L. Moody, "From entities to Stars, Snowflakes, Clusters, Constellations and Galaxies: A Methodology for Data Warehouse design", Proceedings 18th international conference on conceptual Modeling, 15-18 novembre 1999.

(Shanks, 1996) G.G Shanks, "Enterprise Data Architectures: A study or practice", Proceedings First Australian Data Management Conference, Melbourne, Australia, Déc. 2-3 1996

OUVRAGES GENERAUX

Joy Mundy, Warren Thornthwaite and Ralph Kimball (2006): "The Microsoft Data Warehouse Toolkit: With SQL Server 2005 and the Microsoft Business Intelligence Toolset", John Wiley & Sons.

Vincent Rainardi (2008): "Building a Data Warehouse: With Examples in SQL Server", Apress.

Paul Turley et al (2007): "Microsoft SQL Server 2005 Integration Services Step by Step", Microsoft Press.

Lynn Langit (2007), "Foundations of SQL Server 2005 Business Intelligence", Apress.

Ralph Kimball et Margy Ross, "Entrepôts de Données, Guide pratique de modélisation dimensionnelle", 2^e édition, Vuibert.

Adullact (2008), "Solutions Open Source de Business Intelligence", Copyright © 2008 - ADULLACT

MEMOIRES

Antonio BOKO (2008), "Etude et mise en œuvre d'un datawarehouse pour l'analyse du trafic de communication GSM de Moov Bénin", Mémoire pour l'obtention du diplôme de master professionnel en Informatique de Gestion, Ecole Nationale d'Economie Appliquée et de Management (ENEAM).

Msougar Younès (2003), "Mise en place d'un datawarehouse : Data Mart finance Lafarge Maroc", Institut Supérieur de Commerce et d'Administration des Entreprises (I S C A E), CASABLANCA

Fleur-Anne Blain (2008), "Automatisation de l'exécution d'un package SSIS sur Microsoft SQL Server 2005", Date de mise à jour le 01/04/2008.

Fleur-Anne Blain (2008), "Sql Server 2005 Integration Services : Import/export Excel", Date de mise à jour le 29/04/2008.

Olivier DELMOTTE (2007), "Introduction à Sql Server Integration Services (SSIS)", Date de mise à jour le 07/06/2007.

Thibault Denizet, "SSIS Implémenter un flux, version 1.0".

Belle (2011), "Tranfer data from MySQL to SQL Server Using SSIS", posted 23/03/2011.

Mohamed Taslimanka Sylla (2008), "Conception et déploiement d'un cube Analysis Services 2005", 07/04/2008.

Jayaram Krishnaswamy (2009), "Learning SQL Server 2008, Reporting Services 2008, Packt Publishing, 2009.

Valentin Pauline (2008), "Introduction à la BI avec SQL Server 2008 version 1.0".

Sikha Saha Bagui et Richard Walsh Earp (2008), "Essential SQL on SQL Server 2008", Jones and Bartlett Publishers

ANNEXES

TABLE DES ANNEXES

ANNEXE 1 : RESUME DE LA DECLARATION DE PARIS.....	103
ANNEXE 2 : ORGANIGRAMME DU MINISTERE DU DEVELOPPEMENT, DE L'ANALYSE ECONOMIQUE ET DE LA PROSPECTIVE (MDAEP).....	104
ANNEXE 3 : REPARTITION DU PIP PAR SOURCE DE FINANCEMENT	105
ANNEXE 4 : REPARTITION DU PIP PAR SECTEUR DE DEVELOPPEMENT EN MILLIONS DE FCFA ...	106
ANNEXE 5 : REPARTITION ET CLASSIFICATION PAR ORDRE CROISSANT DU PIP GESTION 2010 PAR MINISTERE ET INSTITUTION DE L'ÉTAT.....	107
ANNEXE 6 : POINT D'EXECUTION FINANCIERE PAR SOURCE DE FINANCEMENT AU 31 MARS 2010(EN MILLIONS DE FRANCS CFA)	110
ANNEXE 7 : LES 12 REGLES D'OLAP	111
ANNEXE 8 : SCHEMA DE L'INFRASTRUCTURE D'UN DATAWAREHOUSE.....	112
ANNEXE 9 : QUESTIONNAIRE.....	113
ANNEXE 10 : MODELE CONCEPTUEL DES DONNEES DU SIAPIP	117
ANNEXE 11 : MODELE CONCEPTUEL DU DATAWAREHOUSE	118
ANNEXE 12 : EXTRAIT DU TABLEAU DE TRANSFORMATION DES DONNEES SOURCES VERS LE DATAWAREHOUSE.....	119
ANNEXE 13 : CONFIGURATION MATERIELLE NECESSAIRE POUR L'INSTALLATION DE SQL SERVER 2008.....	121

ANNEXE 1 : Résumé de la déclaration de Paris

La Déclaration de Paris a été signée en mars 2005 par des Partenaires Techniques et Financiers (PTF) et des pays partenaires y compris le Bénin. Elle est basée sur un principe d'engagement mutuel pour améliorer l'efficacité de l'aide, et comprend un cadre pratique et concret visant à améliorer la qualité de l'aide et son impact sur le développement. Les engagements s'articulent autour de cinq (05) grands principes.

Appropriation : les pays partenaires exercent une réelle maîtrise sur leurs politiques et stratégies de développement et assurent la coordination de l'action à l'appui au développement.

Alignement : les donateurs font reposer l'ensemble de leur soutien sur les stratégies nationales de développement, les institutions et les procédures des pays partenaires.

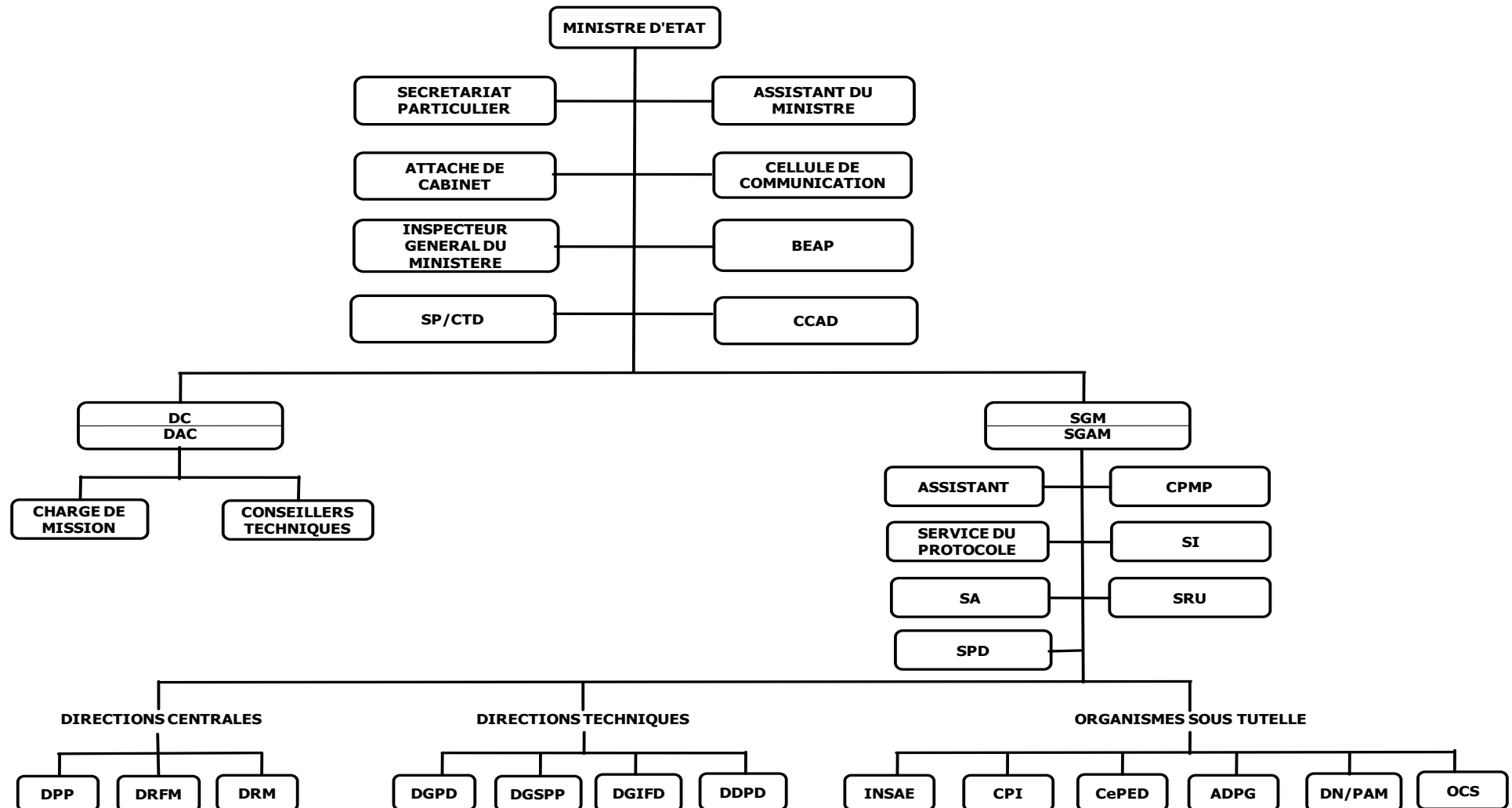
Harmonisation : les donateurs s'emploient à mieux harmoniser leurs actions et à leur conférer une plus grande efficacité collective.

Gestion axée sur les résultats : la gestion et les processus de décision relatifs aux ressources affectées au développement sont améliorés en vue d'obtenir des résultats.

Responsabilité mutuelle : les donateurs et les pays partenaires se tiennent mutuellement responsables des résultats obtenus en matière de développement.

Source : OCDE ; Déclaration de Paris

ANNEXE 2 : Organigramme du Ministère du Développement, de l'Analyse Economique et de la Prospective (MDAEP)



Source : DECRET N° 2008-107 DU 10 MARS 2008

ANNEXE 3 : Répartition du PIP par source de financement

Source de financement	Prévision 2010 (en millions de FCFA)	Prévision 2009 (en millions de FCFA)	Poids prévision 2010 par rapport au Total Général(en %)	taux d'accroissement de 2009 à 2010 (en %)
BUDGET NATIONAL	279 350,37	272 798,19	59,54	2,40
- <i>CPT BN</i>	<i>74 206,76</i>	<i>56 541,71</i>	<i>15,82</i>	<i>31,24</i>
- <i>BN EXC</i>	<i>199 493,03</i>	<i>212 544,48</i>	<i>42,52</i>	<i>-6,14</i>
- <i>TEED/TTE</i>	<i>800,00</i>	<i>800,00</i>	<i>0,17</i>	<i>0,00</i>
- <i>Portefeuille</i>	<i>4 850,57</i>	<i>2 912,00</i>	<i>1,03</i>	<i>66,57</i>
- <i>Aide Budgétaire</i>	<i>0,00</i>	<i>0,00</i>	<i>0,00</i>	<i>0,00</i>
Autofinancement	0,00	380,00	0,00	-100,00
Collectivités Locales	0,00	0,00	0,00	0,00
Autres Sources Intérieures	0,00	0,00	0,00	0,00
SOURCES INTERIEURES	279 350,37	27 3178,19	59,54	2,26
Prêts	105 594,68	103 662,43	22,51	1,86
Dons	84 246,32	114 906,54	17,96	-26,68
SOURCES EXTERIEURES	189 841,000	218 568,98	40,46	-13,14
TOTAL GENERAL	469 191,37	491 747,167	100,00	-4,59

Source : DGSP/MPDEPP-CAG

ANNEXE 4 : Répartition du PIP par secteur de développement en millions de FCFA

Secteurs	Prévisions 2010	Prévisions 2009	Poids prévision 2010 par rapport au Total Général (en %)	Evolution de 2009 à 2010 (en %)
SECTEUR RURAL	36 012,93	46 450,00	7,68	-22,47
INDUSTRIE, COMMERCE ET SERVICES	15 907,42	12 685,92	3,39	25,39
EAU ET ELECTRICITE	57 068,08	61 845,50	12,16	-7,72
INFRASTRUCTURES	151 712,89	160 264,36	32,33	-5,34
ARTISANAT ET TOURISME	5 783,70	6 062,99	1,23	-4,61
TOTAL SECTEURS PRODUCTIFS	266 485,02	287 308,77	56,80	-7,25
SANTE	30 555,43	52 657,62	6,51	-41,97
EDUCATION	34 833,31	39 124,68	7,42	-10,97
HABITAT	51 441,48	37 444,02	10,96	37,38
ENVIRONNEMENT	10 588,67	14 398,65	2,26	-26,46
AUTRES SECTEURS SOCIAUX	22 824,62	9 496,00	4,86	140,36
TOTAL SECTEURS SOCIAUX	150 243,50	153 120,97	32,02	-1,88
ADMINISTRATION (Souveraineté)	52 462,84	51 317,42	11,18	2,23
TOTAL GENERAL	469 191,37	491 747,17	100,00%	-4,59

Source : DGSP/MPDEPP-CAG

ANNEXE 5 : Répartition et classification par ordre croissant du PIP gestion 2010 par Ministère et Institution de l'Etat

N° D'ORDRE	INSTITUTIONS / MINISTERES	PREVISIONS 2010(EN MILLIONS DE FCFA)	POIDS (en %)
01	HAUTE AUTORITE DE L'AUDIOVISUEL ET DE LA COMMUNICATION	135,46	0,03
02	COUR SUPREME	408,44	0,09
03	MINISTERE DES RECHERCHES PETROLIERES ET MINIERES	476,87	0,10
04	MINISTERE DELEGUE AUPRES DU PRESIDENT DE LA REPUBLIQUE CHARGE DE L'ECONOMIE MARITIME, DES TRANSPORTS MARITIMES ET DES INFRASTRUCTURES PORTUAIRES	902,14	0,19
05	PRESIDENCE DE LA REPUBLIQUE	1 521,17	0,32
06	MINISTERE DE LA CULTURE, DE L'ALPHABETISATION ET DE LA PROMOTION DES LANGUES NATIONALES	1 663,95	0,35
07	MINISTERE CHARGE DES RELATIONS AVEC LES INSTITUTIONS	1 718,04	0,37
08	MINISTERE DE LA REFORME ADMINISTRATIVE ET INSTITUTIONNELLE	1 854,90	0,40
09	MINISTERE DU COMMERCE	2 031,12	0,43
10	MINISTERE CHARGE DES PME ET DE LA PROMOTION DU SECTEUR PRIVE	2 075,48	0,44
11	MINISTERE DES AFFAIRES ETRANGERES, DE L'INTEGRATION AFRICAINE, DE LA FRANCOPHONIE ET DES BENINOIS DE L'EXTERIEUR	2 161,32	0,46
12	MINISTERE DE LA FAMILLE ET DE LA SOLIDARITE NATIONALE	2.407,28	0,51
13	MINISTERE DE LA JUSTICE, DE LA LEGISLATION ET DES DROITS DE L'HOMME	2 648,66	0,56
14	MINISTERE DELEGUE AUPRES DU PRESIDENT DE LA REPUBLIQUE CHARGE DE LA COMMUNICATION ET DES TECHNOLOGIES DE L'INFORMATION ET DE LA COMMUNICATION	3 097,21	0,66

15	MINISTERE DU TRAVAIL ET DE LA FONCTION PUBLIQUE	3 189,96	0,68
16	MINISTERE DE LA DEFENSE NATIONALE	3 547,08	0,76
17	MINISTERE DE L'INDUSTRIE	3 927,44	0,84
18	MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE	4 251,70	0,91
19	MINISTERE DE LA JEUNESSE, DES SPORTS ET LOISIRS	4 277,26	0,91
20	MINISTERE DE L'ARTISANAT ET DU TOURISME	5 026,50	1,07
21	MINISTERE DE LA MICRO-FINANCE, DE L'EMPLOI DES JEUNES ET DES FEMMES	6 591,30	1,40
22	MINISTERE DE LA PROSPECTIVE, DU DEVELOPPEMENT, DE L'EVALUATION DES POLITIQUES PUBLIQUES ET DE LA COORDINATION DE L'ACTION GOUVERNEMENTALE	7 067,67	1,51
23	MINISTERE DE L'ENVIRONNEMENT ET DE LA PROTECTION DE LA NATURE	9 631,74	2,05
24	MINISTERE DE L'ENSEIGNEMENT SECONDAIRE, DE FORMATION TECHNIQUE ET PROFESSIONNELLE	11 406,32	2,43
25	MINISTERE DE L'INTERIEUR ET DE LA SECURITE PUBLIQUE	12 084,18	2,58
26	MINISTERE DE L'ECONOMIE ET DES FINANCES	19 064,04	4,06
27	MINISTERE DES ENSEIGNEMENTS MATERNEL ET PRIMAIRE	19 175,78	4,09
28	MINISTERE DE LA DECENTRALISATION, DE LA GOUVERNANCE LOCALE, DE L'ADMINISTRATION ET DE L'AMENAGEMENT DU TERRITOIRE	21 688,85	4,62
29	MINISTERE DE LA SANTE	39 376,54	8,39
30	MINISTERE DE L'AGRICULTURE, DE L'ELEVAGE ET DE LA PECHE	41 295,62	8,80
31	MINISTERE DE L'URBANISME, DE L'HABITAT, DE LA REFORME FONCIERE ET DE LA LUTTE CONTRE L'EROSION COTIERE	43 866,04	9,35
32	MINISTERE DE L'ENERGIE ET DE L'EAU	61 654,64	13,14

33	MINISTERE DELEGUE AUPRES DU PRESIDENT DE LA REPUBLIQUE CHARGE DES TRANSPORTS TERRESTRES, DES TRANSPORTS AERIENS ET DES TRAVAUX PUBLICS	128 966,67	27,49
	TOTAL	469 191,37	100,00

Source : Réalisée par la DGSPP du MPDEPP-CAG sur la base du PIP gestion 2010 et de la loi des Finances exercice 2010

ANNEXE 6 : Point d'exécution financière par source de financement au 31 mars 2010(en millions de francs CFA)

Source de financement	Prévision 2010	Réalisation	Taux d'exécution (en %)	Poids du montant d'exécution par rapport au Total Général de réalisation (en %)
BUDGET NATIONAL	279 350,37	14 953,14	5,35	56,44
- <i>BN CPT</i>	74 206,76	5 509,82	7,42	20,80
- <i>BN EXC</i>	199 493,03	9 443,30	4,73	35,64
- <i>TEED/TTE</i>	800,00	0,00	0,00	0,00
- <i>Portefeuille</i>	4 850,57	0,00	0,00	0,00
- <i>Aide Budgétaire</i>	0,00	0,00	0,00	0,00
Autofinancement	0,00	0,00	0,00	0,00
Collectivités Locales	0,00	0,00	0,00	0,00
Autres Sources Intérieures	0,00	0,00	0,00	0,00
SOURCES INTERIEURES	279 350,37	14 953,14	5,35	50,39
Prêts	105 594,683	12 065,15	11,43	40,66
Dons	84 246,317	2655,16	3,15	8,95
SOURCES EXTERIEURES	189 841,00	14 720,31	7,75	49,61
TOTAL GENERAL	469 191,37	29 673,442	6,32	100,00

Source : MPDEPP-CAG/DGSPP

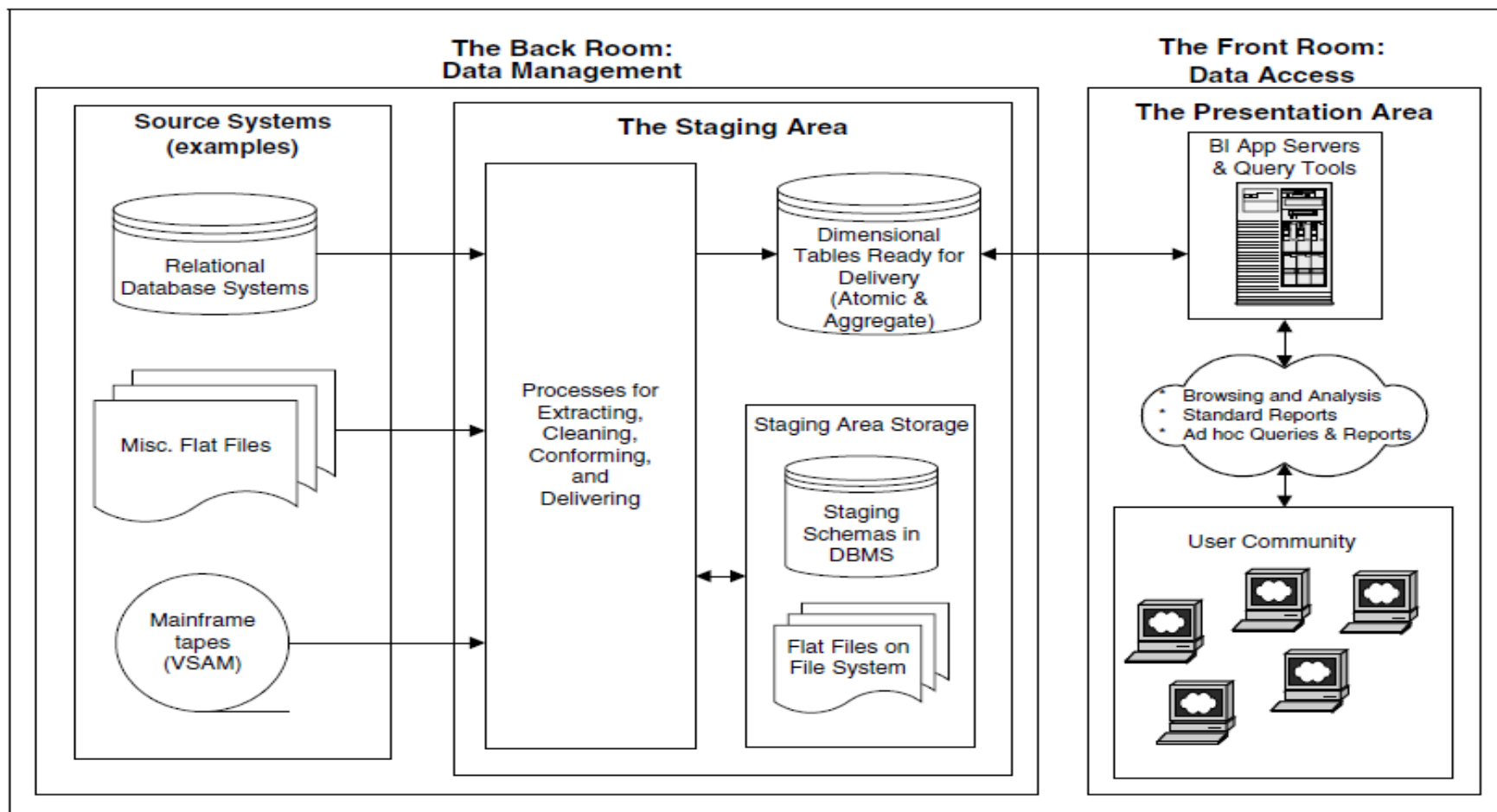
ANNEXE 7 : Les 12 règles d'OLAP

- 1) une vue multidimensionnelle des données.
- 2) La transparence vis à vis de l'utilisateur qui doit accéder à la BD par l'intermédiaire d'outils simples (tableur, par ex).
- 3) La BD doit disposer d'un modèle et d'outils permettant d'accéder à de multiples sources, d'effectuer les conversions et extractions nécessaires pour alimenter la Base OLAP.
- 4) Le modèle de données, le nombre de dimensions ou le nombre de niveaux d'agrégation doivent pouvoir changer, sans remettre en cause le fonctionnement de la base.
- 5) Architecture Client/Serveur.
- 6) Toutes les dimensions définies dans le modèle de données doivent être accessibles pour chacune des données.
- 7) Gestion des matrices creuses. Les parties vides du cube multidimensionnel doivent être stockées de manière à ne pas détériorer les temps d'accès.
- 8) Accessibilité simultanément par plusieurs utilisateurs.
- 9) Toutes les données stockées ou calculées dans le cube doivent être accessibles et les règles de gestion doivent toujours s'y appliquer. Toutes les tranches de cube doivent être visualisées.
- 10) Navigation aisée dans les données pour les utilisateurs, de manière intuitive.
- 11) Outil de présentation des données.
- 12) Nombre illimité de dimensions et de niveaux d'agrégation

Source : Extrait de O. Boussaid, 2011²⁰,

²⁰ O. Boussaid, 2011, « Introduction aux systèmes d'information décisionnels », Université de Lyon

ANNEXE 8 : Schéma de l'infrastructure d'un datawarehouse



Source : Outils Back Room et Front Room (Kimball et Caserta, 2004)

ANNEXE 9 : Questionnaire

Dans la mise en place d'un Entrepôt de Données (ED), il est indispensable de bien appréhender le Système d'Information Opérationnel. En effet, la structure désireuse d'une telle solution a déjà en son sein une organisation, des applications opérationnelles fonctionnelles et une infrastructure technique. Une bonne connaissance de ces différents éléments permettra de mieux gérer les ressources acquises dans l'élaboration de ce projet.

Votre contribution nous sera d'une grande aide, en répondant à ce questionnaire.

- **Présentation**

Mr/Mme _____

Fonction : _____

Niveau d'étude : _____

Ancienneté : _____

- **Questions**

1. Quelles sont les différentes tâches que vous exécutez à votre niveau ?
2. Comment travaillez-vous habituellement c'est-à-dire manuellement, automatiquement (donc utilisation d'une application), ou semi automatiquement (utilisation partielle d'une application et continuation manuelle) ?
3. Si vous utilisez un mode semi-automatique, en quoi l'application ou le tableur utilisé ne répond pas à vos prérogatives vous obligeant à continuer manuellement ?

4. Si vous utilisez une application, pouvez-vous nous la présenter ? Du point de vue fonctionnalités : menus ou fonctions, états ou situations sorti(e)s, du point de vue de la présentation : nom de l'application, sa description, et du point de vue technique : type d'application (client/serveur ou monoposte), base de données, logiciel de conception etc.). Le tableau N°1 vous est proposé pour vous aider à nous donner les caractéristiques.
5. D'où proviennent les données qui alimentent votre application ? Quel est le mode de transfert utilisé (direct si il y a communication entre les différentes applications donc par le biais de fichier ; indirect si il y a reprise de saisie des informations dans l'application concerné donc par le biais d'états imprimés, de fichier Word ou Excel copié sur des supports physiques,...).
6. Pour la/les application(s) dont vous disposez, des informations interviennent pour assurer la maintenance, la programmation etc..., pourrions-nous avoir une liste de ces derniers ? les informations désirées sont nom et prénoms, âge ou tranche d'âge, derniers diplômes obtenus, formations récentes, temps passés à ce poste, ancienneté dans la vie professionnelle, profil de formation (administrateur, programmeur ou maintenancier), profil exercé dans l'entreprise, description des tâches exécutées sur les différentes applications, autres informations nécessaires. Le tableau N°2 est élaboré pour vous faciliter dans votre réponse.
7. Présentez-nous l'infrastructure du système informatique pour le ou les applications de votre structure ? pour les postes clients (Système d'Exploitation (SE) installé, capacité disque dur, capacité mémoire, date d'acquisition, caractéristiques du processeur : vitesse, type de processeur ; état du poste, etc.) ; pour les postes serveur (idem, nombre de postes desservis, logiciels et SGBD installés : nom, version, etc.). Le tableau N°3 est élaboré pour recueillir ces informations.

8. Quels sont vos besoins ou les problèmes rencontrés que vous désirez résoudre pour une meilleure exécution de vos tâches ?

9. Est-ce possible d'avoir accès à votre application dans le but de mieux l'étudier pour la définition des données nécessaires dont vous aurez besoins dans l'ED ?

10. Avez-vous un manuel d'utilisation et un document technique présentant la description technique de l'application (Dictionnaire de Données (DD), Modèle Conceptuel de Données (MCD), etc.) ? Si ou pourrions-nous obtenir un exemplaire ?

Tableau N° 1 : Présentation générale de l'application -----

NOM	DESCRIPTION	FONCTIONNALITES	ETATS SORTIS	TYPE	SGBD	LOGICIEL DE CONCEPTION	SOURCES D'ALIMENTATION	
							NOM	MODE DE TRANSFERT

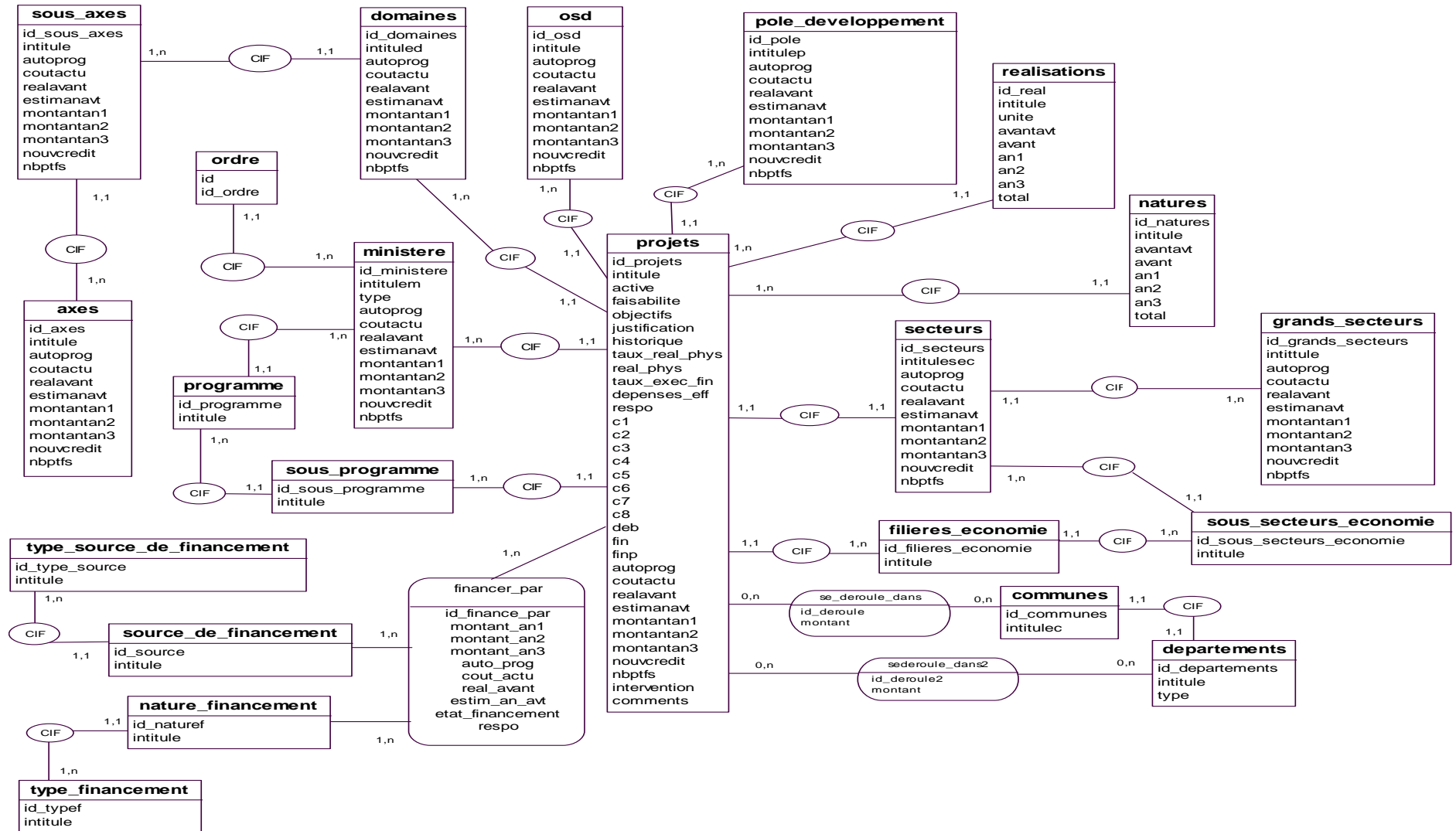
Tableau N° 1 : Liste des informaticiens de la structure -----

N°	RENSEIGNEMENTS							POFIL OU EMPLOI EXERCE	DESCRIPTION DES TÂCHES EXECUTEES
	NOM & PRENOMS	AGE OU TRANCHE D'AGE	NIVEAU D'ETUDE	FORMATIONS RECENTES	PROFIL DE FORMATION	ANCIENNETE A CE POSTE	ANCIENNETE DANS LA VIE PROFESSIONNELLE		

Tableau N° 2 : Infrastructure du système informatique de la structure -----

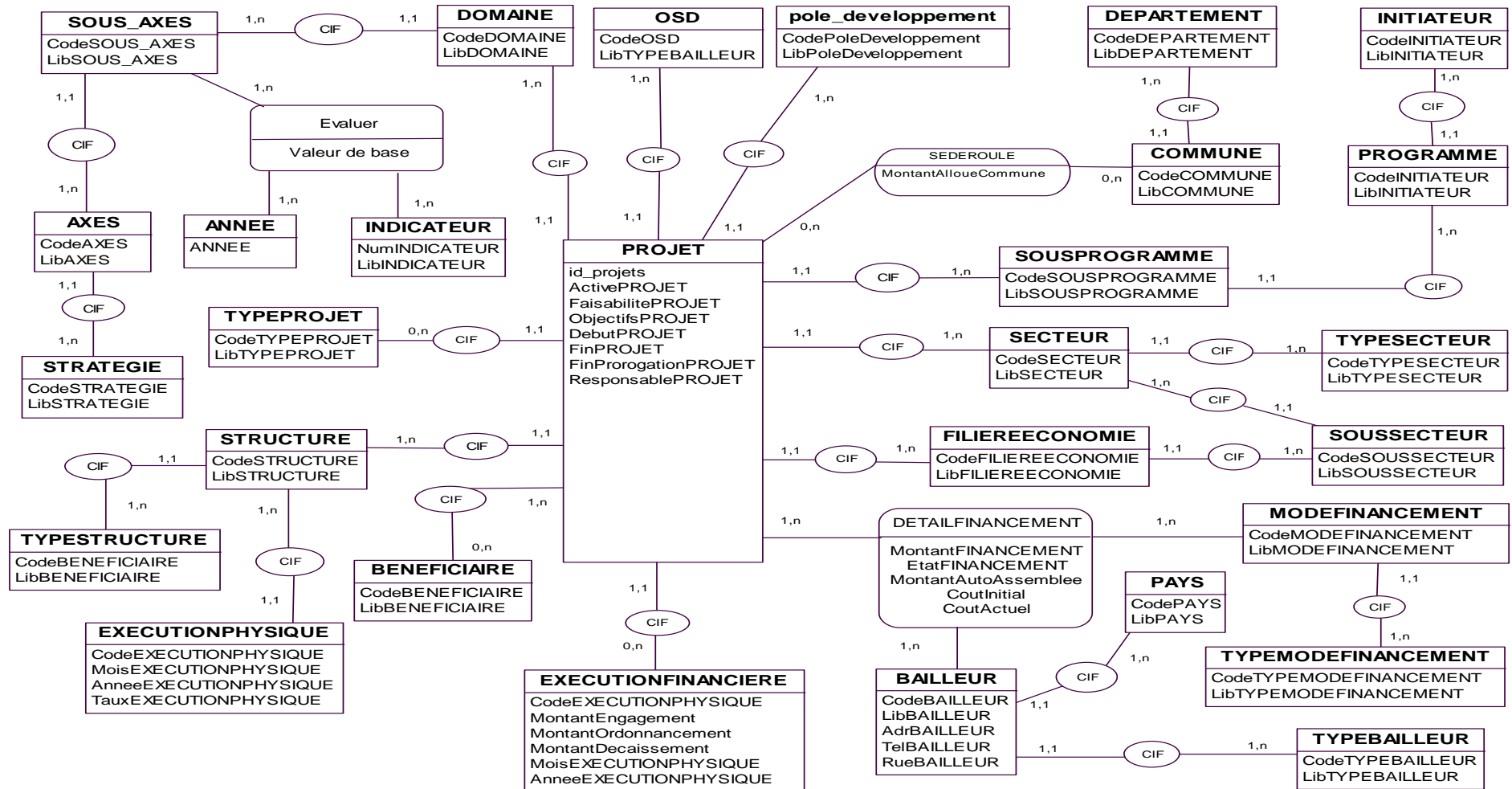
POSTES	CARACTERISTIQUES			SE INSTALLES	ANNEE D'ACQUISITION	ETAT DU POSTE	LOGICIELS ET SGBD INSTALLE(S)		OBSERVATIONS
	DISQUE DUR	MEMOIRE	PROCESSEUR				NOM	VERSIONS	

Annexe 10 : Modèle Conceptuel des Données du SIAPIP



Source : SIAPIP

Annexe 11 : Modèle Conceptuel du Datawarehouse



Source : Nos propres investigations

Annexe 12 : Extrait du tableau de transformation des données sources vers le Datawarehouse

BASE SIAPIP ET CASPPP							Datawarehouse (Stage)			BASE DIMENSIONNELLE			
N°	Propriété	Signification	Type	Longueur	Nature	Observation	Propriété	Type	Longueur	Propriété	Type	Longueur	Dimension
1	id_axes	Renseigne sur l'identifiant de l'axe	N	3	E	Automatique	CodeAxe	nchar	10	Domaine_ Code Axe	Nchar	10	Domaine_dimension
2	intitule	Intitulé de l'axe	A	255	E		LibAxe	nchar	255	Domaine_ Lib Axe	Nchar	255	
3	id_communes	Renseigne sur l'identifiant de la commune	N	3	E	Automatique	CodeCommune	nchar	10	Commune_Code	Nchar	10	Commune_dimension
4	intitulec	Intitulé de la commune	A	255	E		LibCommune	nchar	255	Commune_Lib	nchar	255	
5	id_departements	Identifiant du département	N	3	E		CodeDepartement	nchar	10	Commune _Code_Departement	nchar	10	
6	intitule	Intitulé du département	A	255	E		Lib_Departement	nchar	255	Commune _Lib_Departement	nchar	255	
7	id_domaines	Identifiant du domaine	N	3	E	Automatique	CodeDomaine	nchar	10	Domaine_Code	Nchar	10	Domaine_dimension
8	intituled	Intitulé du domaine	A	255	E		LibDomaine	nchar	255	Domaine_Lib	Nchar	255	
9	id_filières_economie	Identifiant du département	N	10	E	Automatique	CodeFiliereEconomie	nchar	10	FiliereEconomie_Code	Nchar	10	FiliereEconomie
10	intitule	Intitulé de la filière	A	255	E		LibFiliereEconomie	nchar	255	FiliereEconomie_Lib	Nchar	255	

MISE EN PLACE D'UN DATAWAREHOUSE POUR UN MEILLEUR SUIVI DES PROGRAMMES D'INVESTISSEMENTS PUBLICS AU BENIN

		économique											
11	montant_an1	Montant allouer la 1ière année	N	7	E		MontFinancement	bigint	8	Montant_Financement	bigint	8	AnalyseProjet_Fait
12	etat_financement	Etat du financement	A	255	E		EtatFinancement	nchar	255				AnalyseProjet_Fait
13	autoprog	Montant autorisé pour le programme	N	10	E		MontantAutoAssemble	bigint	8	MontantAutoAssemble	bigint	8	
14	coutactu	Cout actuel du projet	N	10	E		CoutActuel	bigint	8	CoutActuel	bigint	8	
15	id_grands_secteurs	Identifiant de grand secteur	N	3	E	Automatique	CodeTypeSecteur	nchar	10	FiliereEconomie_CodeTypeSecteur	Nchar	10	FiliereEconomie_dimension
16	intitule	Intitulé de grand secteur	A	255	E		LibTypeSecteur	nchar	255	FiliereEconomie_LibTypeSecteur	Nchar	255	
17	id_ministeres	Identifiant du ministère	N	3	E	Automatique	CodeStructure	nchar	10	Structure_Code	Nchar	10	Structure_Dimension
18	intitlem	Intitulé du ministère	A	255	E		LibStructure	nchar	255	Structure_Lib	Nchar	255	
19	type	Type du ministère	A	50	E		CodeTypeStructure	nchar	10	Structure_CodeType	nchar	10	
20	id_osd	Identifiant de l'OSD	N	3	E	Automatique	CodeOSD	nchar	10	OSD_Code	Nchar	10	OSD_Dimension
21	intitule	Intitulé de l'OSD	A	255	E		LibOSD	nchar	255	OSD_Lib	Nchar	255	

Source : Conçu par nous-mêmes

Annexe 13 : Configuration matérielle nécessaire pour l'installation de SQL Server 2008

Composant	Requis
Processeur	<p>Type du processeur:</p> <ul style="list-style-type: none"> - IA64 minimum: processeur Itanium ou plus - x64 minimum: AMD Opteron, AMD Athlon 64, Intel Xeon avec Intel EM64T support, Intel Pentium IV avec support EM64T <p>Vitesse du processeur:</p> <ul style="list-style-type: none"> - IA64 minimum: 1.6 GHz - IA64 recommandé: 1.6 GHz ou plus - x64 minimum: 1.6 GHz - x64 recommandé: 2.0 GHz ou plus
Framework	<ul style="list-style-type: none"> - .NET Framework 2.0 - SQL Server Native Client - SQL Server Setup support files
Mémoire	<p>RAM:</p> <ul style="list-style-type: none"> - IA64 minimum: 512 Mo - IA64 recommandé: 1 Go ou plus - IA64 maximum: Limite fixée par le système d'exploitation - x64 minimum: 512 Mo - x64 recommandé: 1 Go ou plus - x64 maximum: Limite fixée par le système d'exploitation

Source : Julien Stuyck, 2008

TABLE DES MATIERES

IDENTIFICATION DU JURY	3
DECLARATION D'ENGAGEMENT	4
DEDICACE.....	I
REMERCIEMENTS.....	II
LISTE DES SIGLES ET ABREVIATIONS	III
LISTE DES TABLEAUX	V
LISTE DES FIGURES.....	VI
RÉSUMÉ	VIII
ABSTRACT	IX
SOMMAIRE	X
INTRODUCTION GENERALE	1
PARTIE I : CONTEXTE DE L'ETUDE	4
CHAPITRE I : INVESTISSEMENT PUBLIC	5
<i>Section 1 : Généralités sur l'investissement public.....</i>	<i>5</i>
Paragraphe 1 : Définition de l'investissement public.....	5
1 Définition général de l'Investissement public.....	5
2 Investissement Public au Bénin	5
2.1 Planification du développement au Bénin	5
2.1.1 Etudes prospectives	6
2.1.2 Planification	6
2.1.3 Programmation	7
1.1.1 Budgétisation.....	8
2.2 Acteurs de l'élaboration de la planification du développement	8
Paragraphe 2 : Base de l'étude sur l'investissement public	9
1 Présentation générale du rapport d'avancement	9
2 Présentation technique du rapport d'avancement	10
2.1 Présentation du portefeuille actif.....	10
2.2 Répartition du PIP par gestion	10
2.2.1 Par source de financement	11
2.2.2 Par secteur	11
2.2.3 Par OSD	11

2.3 Point d'exécution financière et réalisations physiques du portefeuille actif	12
<i>Section 2 : Etude du système</i>	<i>12</i>
Paragraphe 1 : Difficultés rencontrées et approches de solution	13
1 Difficultés rencontrées dans le suivi du PIP	13
1.1 Problèmes au niveau du traitement des données du PIP	13
1.1.1 Planification et programmation du PIP	13
1.1.2 Suivi du PIP	14
1.2 Problèmes au niveau du matériel	15
1.3 Problèmes au niveau du personnel	15
2 Analyse du système et approche de solution	15
2.1 Analyse du système.....	15
2.2 Approche de solution.....	16
2.2.1 Améliorer le suivi du PIP par renouvellement du parc et formation continue du personnel informatique.	17
2.2.2 Améliorer le suivi du PIP par intégration d'applications d'entreprise.....	17
2.2.3 Améliorer le suivi du PIP par la mise en place d'un datawarehouse	18
Paragraphe 2 : Problématique et objectifs de l'étude	20
1 Problématique de l'étude	20
2 Objectifs de l'étude	22
CHAPITRE II : ETAT DE L'ART	23
<i>Section 1 : Présentation générale du datawarehouse</i>	<i>23</i>
Paragraphe 1 : Généralités sur le datawarehouse	23
1 Datawarehouse (Entrepôt de Données ED)	23
1.1 Définition.....	23
1.2 Architecture d'un datawarehouse	25
1.2.1 Les applications opérationnelles sources.....	26
1.2.2 La préparation des données.....	26
1.2.3 La présentation des données	26
1.2.4 Les outils d'accès aux données	26
2 Autres concepts	27
2.1 Analyse multidimensionnelle	27
2.2 Les différentes approches d'OLAP.....	27

2.2.1 Implémentation	27
2.2.2 Utilisation	28
2.3 Autres outils du décisionnel	28
Paragraphe 2 : Considérations pratiques pour la mise en place d'un datawarehouse	29
1 Les grandes approches de mise en place d'un système datawarehouse	29
1.1.1 Datamarts Indépendants	29
1.1.2 Datamarts dépendants ODS.....	30
1.1.3 Hybride	30
2 Infrastructure du datawarehouse	30
2.1 Infrastructure des outils d'arrière-plan (back room).....	31
2.1.1 Plates-formes matérielles et systèmes d'exploitation	31
2.1.2 Architectures de traitement en parallèle	32
2.1.2.1 Architecture SMP (Fonctionnement en multiprocesseurs symétrique)	32
2.1.2.2 Architecture MPP (Traitement massivement parallèle)	33
2.1.2.3 Architecture NUMA (Architecture de mémoire non conforme)	33
2.1.3 Facteurs stimulant les performances matérielles	34
2.1.4 Niveau de service attendu	34
2.1.5 Stockage secondaire	35
2.1.6 Autres facteurs	35
2.2 Infrastructure des outils frontaux (front room)	35
2.2.1 Serveur d'application.....	35
2.2.2 Poste de travail.....	35
2.3 Connectivité et réseau	36
2.3.1 Bande passante	36
2.3.2 Accès à distance	36
2.3.3 Passerelles	36
2.3.4 Transfert de fichiers	36
2.3.5 Connectivité des bases de données	37
2.3.6 Service d'annuaire.....	37
<i>Section 2 : Etudes antérieures et modélisation dimensionnelle.....</i>	<i>37</i>

Paragraphe 1 : Etudes de recherches antérieures effectuées sur le thème de datawarehouse	38
1 Travaux effectués en général	38
1.1 Au Maroc.....	38
1.2 Au Burkina-faso,	39
1.3 En Algérie	39
2 Au Bénin.....	40
Paragraphe 2 : Modélisation dimensionnelle	41
1 Généralités sur les modèles dimensionnels.....	41
1.1 Problématiques des Entrepôts de Données	41
1.1.1 Conception	42
1.1.2 Alimentation	42
1.1.3 Interrogation	43
1.1.4 Présentation	44
1.2 Concepts des modèles dimensionnels	44
1.2.1 Objectifs.....	44
1.2.2 Concepts fondamentaux	44
1.2.2.1 Dimension	44
1.2.2.2 Fait	45
1.2.2.3 Indicateurs	45
1.2.2.4 Structure de la base de données.....	45
1.2.3 Modèles dimensionnels.....	46
1.2.3.1 Schéma en étoile	46
1.2.3.2 Schéma en constellation.....	47
1.2.3.3 Le schéma en flocon de neige	47
1.2.3.4 Le schéma en grappe :.....	47
2 Méthodologie de conception.....	48
2.1 Définition du modèle de datawarehouse	48
2.1.1 Classification des entités.....	48
2.1.1.1 Entité de transaction	48
2.1.1.2 Entité Composante.....	49
2.1.1.3 Entité de classification	49
2.1.1.4 Ambiguïtés	50
2.1.2 Hiérarchies	50

2.1.2.1	Identification des hiérarchies	50
2.1.2.2	Réduction de la hiérarchie	50
2.1.2.3	L'agrégation : une fonction des entités de transaction	50
PARTIE II	: ETUDE CONCEPTUELLE	52
CHAPITRE III	: TRAVAUX REALISES.....	53
<i>Section 1</i>	: <i>Conception dimensionnelle</i>	<i>53</i>
Paragraphe 1	: Méthodologie adoptée.....	53
1	Présentation	53
2	BESOINS DE L'ENTREPRISE	54
2.1	Besoins fonctionnels.....	54
2.1.1	DGSP : Direction Générale du Suivi des Projets et Programmes	54
2.1.2	DGIFD : DIRECTION GENERALE DES INVESTISSEMENTS ET DU FINANCEMENT DU DEVELOPPEMENT	55
2.1.3	Ensemble des directions :.....	56
2.2	Besoins non fonctionnels	56
3	Etude de faisabilité et risques associés	56
3.1	Cartographie des sources	57
3.2	Documentation des sources	58
3.3	Identification des problèmes potentiels et filtrage des données	59
3.4	Risques liés à l'étude	59
4	Thèmes analytiques et processus métiers	60
Paragraphe 2	: Modélisation dimensionnelle	61
1	Bus Matrix	62
2	Conception du datamart sur le suivi financier des projets.....	63
2.1	Présentation du datamart	63
2.2	Grain du datamart.....	63
2.3	Modèle logique	63
2.3.1	Tables de dimensions de l'Exécution des projets	63
2.3.2	Table des faits	64
2.4	Modèle dimensionnel	67
3	Conception du datamart sur l'évolution des prévisions d'investissements publics.....	68

3.1 Présentation du datamart	68
3.2 Grain du datamart.....	68
3.1 Modèle logique	68
3.1.1 Tables de dimensions	68
3.1.2 Table des faits	69
3.2 Modèle dimensionnel	71
<i>Section 2 : Conception physique</i>	<i>72</i>
Paragraphe 1 : Transformation et outil de développement	72
1 Règles de transformation Source /Destination	72
2 Choix de l'outil de développement	72
Paragraphe 2 : Modèle physique.....	72
1 Le volume de données.....	73
2 Performance	73
CHAPITRE IV : CREATION DU MODELE PHYSIQUE	75
<i>Section 1 : Mise en place des bases de données</i>	<i>75</i>
Paragraphe 1 : Base de donnée du datawarehouse	75
1 Script de création de la base de données	75
2 Script de création destables	76
Paragraphe 2 : Base de données dimensionnelle	77
1 Script de création de la base dimensionnelle	77
2 Script de création de la table de dimension Bailleur	77
<i>Section 2 : Construction de l'entrepôt de données</i>	<i>78</i>
Paragraphe 1 : Mise en œuvre et déploiement des cubes	78
1 Réalisation du rapport sous Microsoft Visual Studio	79
1.1 Réalisation du rapport	79
1.2 Publication de l'état sur une page web	79
2 Restitution des informations par Excel/OLAP	84
Paragraphe 2 : Mise en œuvre de l'ETL.....	86
1 Extraction des données des différentes bases de production	87
1.1 Extraction des données sous MySql	87
1.2 Extraction des données sous fichier Excel	88
1.3 Déploiement et Vérification du package SSIS	88
1.4 Création et déploiement des cubes	89
2 Sécurité et conditions de mise en œuvre	89

2.1 Sécurité.....	89
2.1.1 Sécurité d'accès à la base de données.....	89
2.1.2 Sécurité au niveau des applications web.....	90
2.2 Conditions de mise en œuvre.....	91
2.2.1 Mise en service.....	91
2.2.2 Maintenance et expansion.....	92
CONCLUSION GENERALE	94
BIBLIOGRAPHIE	97