
TABLE DES MATIÈRES

Chapitre 1 : Présentation générale.....	3
1. Présentation de la structure d'accueil.....	3
1.1 Historique.....	3
1.2 Mission.....	3
1.3 Organisation	4
2. Présentation du projet	6
2.1 Contexte	6
2.2 Problématique	7
2.3 Objectifs.....	8
Chapitre 2 : État de l'art.....	9
1. Etude de l'existant.....	9
1.1 Description de l'existant	9
1.1.1 La comptabilité Budgétaire.....	9
1.1.2 La gestion de la trésorerie	9
1.2 Bilan de l'existant	10
2. Concepts fondamentaux	10
2.1 Modélisation multidimensionnelle.....	10
2.1.1 Concept de fait.....	10
2.1.2 Concept de dimension.....	11
2.1.3 Concept d'hierarchie.....	11
2.1.4 Concept de mesure.....	11
2.2 Stockage des données.....	12
2.2.1 DataWarehouse.....	12
2.2.2 Cube.....	12
2.3 Le modèle en étoile.....	14
2.4 Architecture OLAP	14
2.4.1 Règles de base.....	14
2.4.2 Caractéristique majeure des produits OLAP	15



2.4.3	Différents types de modèle OLAP	16
Chapitre 3 : Conception de notre solution.....		19
1.	Conception architecturale.....	19
1.1	Architecture	19
1.2	Schéma général de l'application.....	19
1.3	Schéma actuel de l'application.....	20
1.3.1	Architecture du système existant.....	20
1.3.2	Architecture de notre solution.....	21
2.	Modélisation	21
2.1	Reverse Engineering	21
2.2	Modélisation du DataWarehouse.....	23
Chapitre 4 : Implémentation de la solution.....		40
1.	Solution proposée	40
2.	Présentation des outils de développement	40
2.1	Talend Open Studio	40
2.2	Suite Microsoft BI	41
3.	Mise en Œuvre	42
3.1	Alimentation de l'entrepôt de données	42
3.1.1	Extraction	42
3.1.2	Transformation	47
3.1.3	Chargement	50
3.1.4	Description des flux d'alimentation	53
1.	Mode d'alimentation du DataWarehouse.....	53
3.2	Mise en place du Cube de données.....	53
3.3	Spécification détaillée des restitutions	56
3.3.1	Paramètres	56
3.3.2	Filtres.....	57
3.3.3	Navigation inter rapport.....	58
Chapitre 5 : Présentation et exploitation des résultats		59
1.	Rapports	59

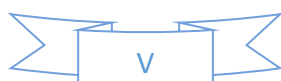
Chapitre 6 : Conclusion et perspectives	64
1. Conclusion.....	64
2. Perspectives	64



TABLES DES FIGURES

1.1 Organigramme de le DISI.....	18
2.1 Architecture ROLAP.....	31
2.2 Architecture MOLAP.....	32
2.3 Architecture HOLAP.....	33
2.4 Architecture générale du SID de l'UCAD.....	35
2.5 Architecture actuelle du DWH existant.....	35
2.6 Architecture actuelle de notre solution.....	36
2.7 Modèle relationnel des finances après reverse engineering(1)	37
2.7 Modèle relationnel des finances après reverse engineering (2)	38
2.9 – Modèle du mouvement trésorerie des finances.....	39
3.6 – Liquidations.....	40
3.7 – Crédits.....	42
3.8 – Engagements.....	43
3.9 – DBM.....	45
3.10 – Mandats.....	46
3.11 - Ordres_Mandats.....	48
3.12 : Ordres_Recettes.....	50
3.13 - Prévision_Période.....	51
3.14 - Ressources_Générales.....	53
3.15 - Ordres_Depenses.....	54
4.1-Connexion aux bases de données.....	58
4.2-Récupérer le schéma des finances (1)	59
4.3-Récupérer le schéma des finances (2)	59
4.4-Récupérer le schéma des finances (3)	60
4.5-Récupérer le schéma des finances (4)	60

4.6- Schéma des synonymes des finances.....	61
4.7-Collecte des données (1)	62
4.8-Collecte des données (2)	62
4.9-Collecte des données (3)	63
4.10 - nettoyage des données (1)	64
4.10 - nettoyage des données (2)	65
4.10 - nettoyage des données (3)	66
4.14 – Chargement des Dimensions (1)	68
4.14 – Chargement des Dimensions (2)	68
4.14 – Chargement des Dimensions (3)	69
4.17 – Alimentation du DWH.....	70
4.18 – Utilisation des paramètres.....	74
4.19 – Utilisation des filtres	75
4.20 – Navigation inter rapport.....	75
4.21 – Situation d’exécution de trésorerie par banque.....	77
4.22 – Situation d’exécution du budget par rubrique.....	78
4.23 – Répartition du budget par établissement et par année.....	79
4.24 – Somme des dépenses budgétaires par structure.....	80
4.25 – Exemple de rapport sur Excel.....	79
4.26 – Total mandat par établissement.....	79



Liste des Tableaux

1.1 – Les applications de l’UCAD.....	20
3.1-Les dimensions des mouvements de trésorerie des finances.....	39
3.2 – Liquidations.....	40
3. 3 – Crédits	42
3.4 – Engagements.....	44
3.5 – DBM.....	45
3.6 – Mandats.....	47
3.7 - Ordres_Mandats.....	49
3.8 : Ordres_Recettes.....	50
3.9 - Prévision_Période.....	52
3.10- Ressources_Générales.....	53
3.11 - Ordres_Depenses.....	55
4.1– Cubes des Finances.....	71
4.2– Dimensions du cube Budget.....	71
4.3 – Dimensions du cube trésorerie.....	72

INTRODUCTION GÉNÉRALE

Les possibilités techniques et les enjeux économiques actuels incitent toute entreprise à réaliser des systèmes de plus en plus complexes. La concurrence accrue impose aux entreprises une réaction rapide et une prise de décisions pertinente suite à l'évolution des marchés.

De ce fait, les entreprises cherchent de plus en plus à informatiser leurs systèmes afin de proposer des services automatisés. Beaucoup d'entre elles, ont à leur disposition des applications informatiques prenant en charge les services métiers et supports. Ces applications, si elles ne sont pas intégrées, elles fonctionnent généralement en silo produisant chaque jour une grande masse de données. C'est le cas de l'UCAD avec ses applications (Scolarité, GRH, Finances, Paie, Gaps-edu etc.).

Ainsi, l'accroissement des quantités d'informations requises rend leur exploitation de plus en plus difficile. Sachant que l'information est devenue une ressource stratégique pour l'institution, les décideurs doivent être bien outillés pour mener à bien leurs missions. Ils doivent prendre régulièrement, dans des délais très courts, des décisions qui se doivent d'être les plus opportunes et qui influent grandement sur la stratégie de l'UCAD et donc sur sa croissance. Elles doivent impérativement être prises sur la base d'informations factuelles qui traduisent fidèlement l'activité de l'institution. En effet, il s'agit de prendre des décisions fondées, basées sur des informations claires, fiables et pertinentes afin d'être en mesure d'anticiper et d'agir avec agilité.

C'est dans ce contexte que le paradigme des SID a émergé afin d'apporter une réponse aux besoins de prise de décisions et de fournir aux décideurs des moyens d'accéder aux données pour un meilleur pilotage des activités. En effet, ils offrent aux décideurs des informations croisées et multi sectorielles de qualité sur lesquelles ils pourront s'appuyer pour arrêter leurs choix décisionnels. Le SID s'avère très efficace, dès lors qu'il s'agit de faire du reporting ou de l'analyse de données pour arriver à fournir des tableaux de synthèse.

Avec plus de 80.000 étudiants, 2.085 enseignants, 2.738 personnels administratifs, techniques et de services et un budget de 30.000.000.000F CFA en 2019, l'UCAD possède un volume significatif de données complexes et a besoin d'un système efficace pour la gestion durable de sa gouvernance. A ce propos, l'UCAD a mis en place ses premières briques décisionnelles avec l'intégration des données Scolarité et RH. Dans le cadre de ce travail, il s'agit d'intégrer la brique des finances à l'entrepôt existant pour étayer les indicateurs de performance et obtenir une vision globale et transversale sur l'ensemble de ses activités. C'est dans ce sens que nous traitons, dans le cas de notre mémoire le thème « **Mise en place d'un Système d'Information Décisionnel pour la gouvernance à l'UCAD : Intégration des données budgétaires et financières** » nous a été soumis.

Nous envisageons de répondre à la problématique en adoptant la démarche suivante. D'abord nous présentons la structure d'accueil et le sujet de notre mémoire dans le **chapitre 1**. Puis, dans le **chapitre 2**, nous exposons un état de l'existant et des concepts

fondamentaux nécessaires à la mise en place de la solution. Ensuite, dans le **chapitre 3**, nous faisons une description de l'architecture de la solution, du travail de reverse engineering qui nous a permis de faire la modélisation multidimensionnelle. Puis dans le **chapitre 4**, nous parlerons de la solution proposée et des outils de développement ayant permis la mise en œuvre de la solution. Dans le **chapitre 5** nous faisons une présentation et une exploitation des résultats obtenus. Enfin, **en chapitre 6**, nous faisons un bilan de notre travail et dégageons quelques perspectives.

Chapitre 1 : Présentation générale

1. Présentation de la structure d'accueil

1.1 Historique

Dans le cadre de son soutien aux Universités Francophones, la Coopération Française lance, en 1996, une action de coopération avec l'Université Cheikh Anta Diop concrétisée à travers le projet d'Appui à la Réforme Universitaire (PARU) pour aider l'université à se doter d'outils de gestion (gestion de la scolarité, gestion de la paie etc.).

Suite aux succès dans le développement et la mise en place des applications de gestion et conformément aux dispositions de l'Arrêté rectoral N ° 299/U du 11 Mars 1999, l'Université Cheikh Anta Diop de Dakar (UCAD) crée la Direction de l'Informatique de Gestion (DIG).

À compter du 21 février 2012, l'arrêté 0503/U change la dénomination de la Direction de l'informatique de Gestion (DIG) du rectorat en Direction des Systèmes d'Information (DSI) qui, fusionnée avec le centre de Calcul, devient la Direction de l'informatique et des Systèmes d'information (DISI).

1.2 Mission

Ayant pour mission la mise en œuvre de la politique des systèmes d'information et des technologies de l'information et de la communication définie par les autorités dans les domaines de l'enseignement, de la recherche, de la gestion ainsi que de la documentation, la DISI est, à ce titre, chargée :

- de mettre en œuvre la politique informatique de l'UCAD ;
- d'élaborer les plans directeurs informatiques de l'UCAD et en assurer la mise en œuvre et le suivi ;
- de participer et d'assister à la réalisation des tableaux de bord et de pilotage pour l'Institution et ses différentes composantes ;
- d'assurer une veille technologique permanente et la traduire dans des projets d'avenir au service des missions de l'Université, particulièrement l'enseignement et la recherche, et au service des utilisateurs des TIC dans l'Institution ;
- de mettre à jour et de renforcer les infrastructures informatiques et de télécommunication de même que les ressources humaines spécialisées, pour les mettre au service de l'administration, de l'enseignement et de la recherche ;
- de développer et de constituer un système d'information intégré et global ainsi que son référentiel et en assurer une cohérence globale;
- de proposer et de concourir à des actions de formations des acteurs (personnel, étudiants) en matière de technologie de l'information et de la communication et en enseignement à distance (E-learning);
- de développer les formations ouvertes et à distance;



- d'assurer la maîtrise d'œuvre de la politique informatique de l'UCAD;
- d'apporter son concours à la promotion de nouvelles formes d'enseignement et au développement de la recherche scientifique : FOAD, e-learning;
d'assurer l'accès sécurisé à l'information et aux applications de gestion de l'UCAD et en garantir leur intégrité et leur fiabilité ;
- de constituer un guichet unique permettant d'offrir à ses usagers des infrastructures de pointe, une assistance et une gestion de la demande homogène et unifiée;
- de fournir des tableaux de bords et de statistiques pour le pilotage de l'institution.

1.3 Organisation

Comme toute structure, la Direction de l'Informatique et des Systèmes d'Information (DISI) est organisée sous forme de directions et de divisions. La DISI compte deux départements et deux divisions. Ses deux départements sont le Département des Infrastructures et des SIG (DIG) et le Département des Ressources Technologiques et pédagogiques (DRTP). Ses deux divisions sont la

Division Administration et Organisation (DAO) et la Division des Systèmes d'Information Décisionnel (DSID). Les deux divisions sont dans des projets en cours. Chaque département ou division a un ensemble de tâches qui lui est spécifique.

- Département des Infrastructures et des SIG (DIG) : il est chargé de la mise en œuvre du système d'information et est responsable des infrastructures informatiques et des réseaux. Il gère l'annuaire électronique et la messagerie et veille aussi au développement cohérent des applications de gestion. Le DIG est composé de quatre (04) divisions :
 - DSIG : Division des Systèmes d'Information de Gestion ;
 - DIC : Division de l'Information et de la communication ;
 - DSRS : Division Réseaux Systèmes et sécurité ;
 - DEL : Division Exploitation et Logistique.
- Département des Ressources technologiques et pédagogiques (DRTP) : c'est un service d'appui aux établissements, chargé d'assurer l'intégration et le développement des TIC dans les activités d'enseignement / apprentissage et la recherche de l'université. Le DRTP est organisé en trois divisions :
 - DFC : division formation et certification ;
 - DED : division enseignement à distance ;
 - DAM : division audiovisuelle et multimédia.
- Division Administration et Organisation (DAO) : elle est chargée d'assurer la gestion des affaires administratives et l'organisation de la direction.
- Division des Systèmes d'Information Décisionnel (DSID) : elle est chargée de coordonner le développement des tableaux de bord et du système d'information des indicateurs qui permettent d'éclairer les prises de décisions relatives à la gestion de l'université.

Ci-dessous, nous avons l'organigramme de la DISI.

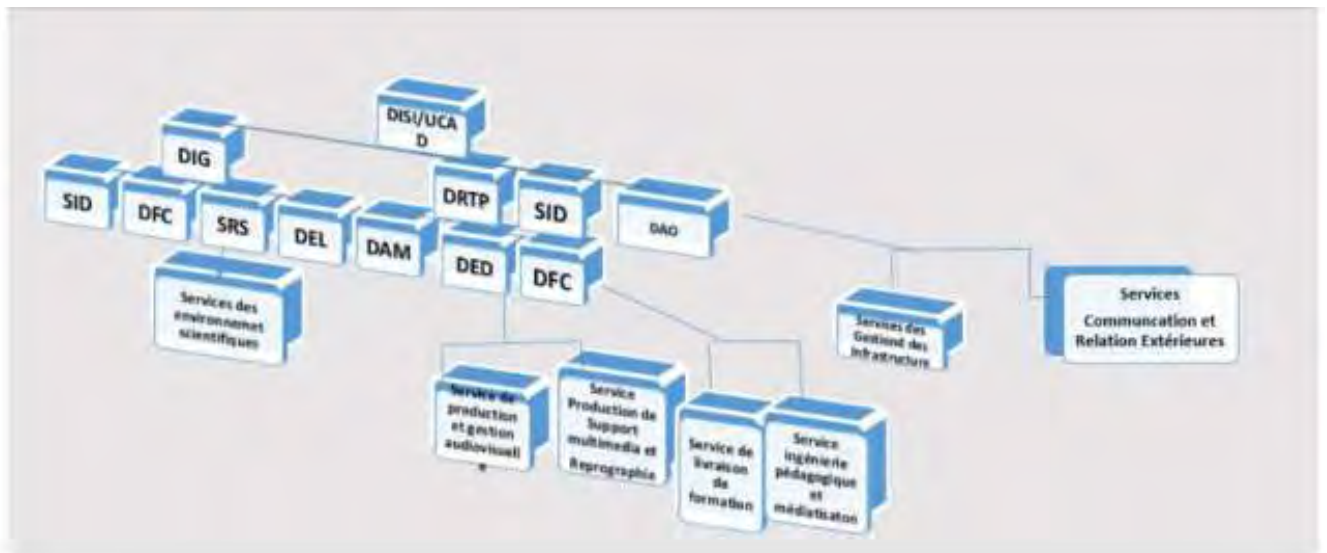


FIGURE 1.1- Organigramme de le DISI

2. Présentation du projet

Dans la perspective d'avoir un suivi quotidien des activités, d'obtenir le plus rapidement un portrait global, de procurer une vision transversale aux décideurs afin d'assurer la pérennité de la bonne gouvernance à l'UCAD, un projet de mise en place d'un système d'information décisionnelle a été mis en place. Dans ce cadre il a été intégré :

- tout d'abord les données de la scolarité qui permettent de mesurer les effets de massification et de réussite ;
- ensuite, les données de ressources humaines qui ont permis d'évaluer les effectifs et de mesurer le niveau d'encadrement par croisement avec les données de la scolarité ;

Et aujourd'hui, il s'agit d'intégrer les données financières d'ordre budgétaire et de trésorerie dans le SID pour mesurer les recouvrements des recettes, consommations et exécutions des dépenses. Cette partie vient en appoint aux deux premières activités et contribuent à donner aux décideurs les meilleurs outils de gouvernance institutionnelle.

Nous allons, dans cette partie présenter le contexte puis la problématique et terminer par les objectifs.

2.1 Contexte

La massification est l'une des causes profondes de toutes les grandes difficultés que l'UCAD rencontre de nos jours. En rappel, elle compte aujourd'hui plus de 80.000 étudiants pour un personnel qui compte 2.085 enseignants et 2.738 personnels administratifs et un budget de 30.000.000.000F alors qu'en 2001 elle en comptait moins de 11.000 étudiants pour 1000 enseignants et 1200 personnels administratifs pour un budget de 11.000.000.000F.

L'UCAD dispose de plusieurs applications permettant d'assurer sa gestion quotidienne :

Applications	Intégrées dans DWH
Application de gestion des inscriptions administratives et pédagogiques	OUI
Application de gestion des ressources humaines	OUI
Application de gestion de la paie	NON

Application pour la Recherche (voyages d'études).	NON
Applications IRIS et RADIS pour les statistiques	NON CONCERNEES
Application de gestion budgétaire et financière	NON

TABLE 1.1 – Les applications de l'UCAD

L'UCAD possède ainsi un très grand volume d'informations très riche, qu'elle collecte tous les jours à partir des bases de données de production de ces différentes applications.

Depuis quatre ans, la DISI est en train de mettre en œuvre un SID dont le but est de permettre l'analyse de tout le volume de données afin d'en tirer profit et d'améliorer la gouvernance.

Un premier projet financé par la Banque Mondiale dans le cadre du contrat de performance (CDP) a permis de développer une première application IRIS qui permettait la mise en place des indicateurs de performance.

Par la suite, une application nommée RADIS a été développée pour la mise en place de tableaux de bord et d'autres indicateurs non prévus dans IRIS avec un modèle de base de données relationnelle qui ne prend pas en compte l'historisation des données.

Trois ans auparavant, un premier groupe d'étudiants de la section informatique avait démarré le projet de mise en place du Datawarehouse en commençant par les données administratives et pédagogiques.

L'année dernière, une deuxième génération d'étudiants avait étendu le Data Warehouse en y intégrant les données des Ressources humaines.

2.2 Problématique

La DISI depuis quelques années est dans une perspective de mise en place d'un SID. En effet, comme annoncé précédemment, l'institution dispose d'un ensemble d'applications dédiées à la production qui assurent la bonne gouvernance à l'UCAD. Ces applications étant opérationnelles, permettent chacune de son côté de produire des rapports spécifiques aux métiers. Cela veut dire que, si nous prenons l'exemple de l'application chargée de gérer la scolarité, elle offre des rapports qui permettent de suivre l'évolution de tout ce qui a trait à la scolarité. Compte tenu de cette situation, les analyses effectuées ne sont spécifiques qu'à un métier bien défini. Or les décideurs ont besoins d'avoir une vision transversale de l'information car une décision à prendre au niveau pédagogique peut être impactée par les ressources humaines ou les finances.

Donc pour pallier cela, la DISI voudrait qu'à partir d'une seule application, qu'il soit possible de croiser les données de la scolarité, de la RH, des finances etc. pour fournir

aux décideurs des rapports leur permettant d'avoir une vision transversale des activités de l'UCAD, d'un établissement, d'un service ou même d'une formation. Par exemple, il serait intéressant de connaître les dépenses ou des recettes faites par l'institution ou par l'établissement suivant les effectifs des étudiants ou des enseignants.

Les données des ressources humaines et de la scolarité son déjà dans le Datawarehouse, le problème de la DISI, aujourd'hui est :

- D'intégrer les données des finances.
- D'assurer l'historisation des données, pour permettre aux décideurs d'analyser et de pouvoir anticiper sur les besoins futurs.
- De mettre des cubes multidimensionnels.
- D'exporter des données sous Excel.
- De produire des rapports sur lesquels, les décideurs pourront se baser pour arrêter leurs choix décisionnels.

2.3 Objectifs

Pour apporter des réponses aux questions soulevées dans la problématique, un ensemble d'objectifs a été assigné au présent projet. Le but principal de ce dernier est d'intégrer les données qui se trouvent dans l'application des finances dans le Datawarehouse et de prendre en compte l'historisation des données.

Les autres objectifs, se déclinent ainsi :

- ✓ Mettre en place des cubes pour avoir le taux d'exécution pour une activité donnée.
- ✓ Faire des analyses et exploiter les résultats de ces analyses pour une meilleure prise de décision.
- ✓ Obtenir des statistiques pour avoir des informations.
- ✓ Offrir une possibilité d'extraire les données vers Excel.
- ✓ Permettre d'avoir une vue globale sur l'ensemble des activités des finances et des autres services.
- ✓ Donner aux utilisateurs la possibilité d'avoir une vue multidimensionnelle des données, ce qui diminue le temps alloué à la prise de décision,
- ✓ Permettre aux décideurs de prendre des décisions dans un court délai en ayant accès aux informations pertinentes.

Chapitre 2 : État de l'art

1. Etude de l'existant

1.1 Description de l'existant

La gestion des finances à l'université Cheikh Anta Diop de Dakar est essentiellement composée de trois parties :

- La comptabilité budgétaire ;
- La gestion de la trésorerie ;
- La comptabilité générale.

Cependant, seules la comptabilité budgétaire et la gestion de la trésorerie sont opérationnelles dans XAALIS.

1.1.1 La comptabilité Budgétaire

Toutes les opérations comptables à l'UCAD démarrent avec la comptabilité budgétaire. Il s'agit d'une comptabilité de dépense qui gère les crédits mis à la disposition par l'état. L'ordonnateur (Recteur ou tout Chef d'établissement à autonomie financière) prescrit l'exécution des recettes, engage les dépenses et en ordonne le paiement. L'Agent Comptable est responsable, pour sa part, de la gestion de la trésorerie. Il n'est pas sous l'autorité du Recteur. C'est un agent du ministère des finances.

Dans la comptabilité budgétaire, on effectue des prévisions de dépenses (subvention de l'état, ressources propres, matériels informatique, voyage d'étude) et de ressources. Il convient de préciser que les bases de la comptabilité budgétaire exigent que l'équilibre recettes-dépenses (fonctionnement comme investissement) soit respecté. Ces prévisions sont soumises à approbation en Assemblée d'Université puis à l'Assemblée Nationale (Ministère des finances).

1.1.2 La gestion de la trésorerie

La trésorerie est une continuité du budget. Elle englobe tout ce qui est virement et chèque. Elle permet de procéder à la paie des mandats, de recevoir des ordres de recettes, des ordres de dépenses et des bordereaux de recettes, à l'intérieur desquels, on trouvera des opérations de trésorerie, qui symbolisent les bons de caisses et les avis de crédits.

Les opérateurs qui interviennent au niveau de la trésorerie sont :

- Les caissiers ;
- Le comptable trésorerie ;

- Le comptable budget.

Le comptable budget est chargé de constater l'arrivée de toute pièce d'établissement de l'ordonnateur et d'en effectuer la supervision afin de voir si tout est conforme. Lorsqu'il appose son visa, il bloque ainsi l'accès en écriture à toutes les opérations qui ont été produites au niveau du budget. Le bon d'engagement qu'il a reçu ne peut plus être modifié par l'ordonnateur de même que la liquidation.

1.2 Bilan de l'existant

La solution XAALIS en charge de la comptabilité budgétaire et de la gestion de la trésorerie dispose d'un système de reporting. Cependant le système de reporting existant est spécifique au métier et il est impossible d'ajouter d'autres reporting.

A ce jour, la brique « Finances » n'est pas intégrée dans RADIS. Ainsi, le besoin identifié dans le cadre de cette étude est d'intégrer les données des finances dans le Datawarehouse, de produire des rapports et de pouvoir croiser ces données aux données des autres métiers intégrées dans l'entrepôt décisionnel à des fins d'aide à la décision. Par exemple, calculer le coût de formation d'un étudiant ou le cout d'une formation.

2. Concepts fondamentaux

2.1 Modélisation multidimensionnelle

L'intérêt pour l'analyse des données s'est développé énormément ces dernières années. Les entreprises se sont rendu compte de l'efficacité de la technologie OLAP dans l'analyse et l'exploration des données. La manière la plus appropriée pour faciliter cette analyse OLAP, est la modélisation multidimensionnelle des données. Cette dernière est une méthode de conception logique qui vise à présenter les données sous une forme standardisée, intuitive et qui permet des accès hautement performants.

Les données de la modélisation multidimensionnelle sont organisées de manière à mettre en évidence le sujet de l'analyse (la table des faits) et les différentes perspectives de l'analyse (les tables des dimensions) qui symbolisent les différentes valeurs de l'activité analysée. En outre, la modélisation multidimensionnelle permet de modéliser un sujet analysé comme un point dans un espace à plusieurs dimensions. Elle repose sur le concept de faits et de dimensions.

Nous allons définir quelques concepts clés de l'informatique décisionnelle nécessaires pour la mise en place de notre solution.

2.1.1 Concept de fait

Le fait : il représente le sujet analysé qui est formé de mesures d'activité analysée. Les mesures sont généralement numériques pour permettre de résumer un grand nombre d'enregistrements (les additionner, calculer le minimum, le maximum,). On distingue trois types de faits : le fait additif, le fait semi additif et le fait non additif.

- Fait additif : un fait est additif s'il est additionnable suivant toutes les dimensions,
- Fait semi-additif : un fait est semi-additif s'il n'est additionnable que suivant certaines dimensions,
- Fait non additif : il s'agit d'un fait non additionnable quel que soit la dimension.

Table de fait : La table de faits est la clef de voûte du modèle dimensionnel où sont stockés les indicateurs de performances. Elle rassemble les données observables (les faits) que l'on possède sur un sujet et que l'on veut étudier, selon divers axes d'analyse (les dimensions). Elle contient toutes les mesures de performances (par exemple : montant de ventes, quantité vendue) et les références (clés étrangères) vers des tables de dimension.

2.1.2 Concept de dimension

La dimension : Une dimension est un axe qui permet d'analyser une mesure selon différentes perspectives. Les dimensions fournissent le contexte (qui, quoi, quand, où, pourquoi et comment) des faits. Une dimension peut être discrète, ce qui implique que ses membres n'ont pas d'ordre particulier (par exemple les départements du Sénégal). Elle peut également être continue, dans ce cas l'ordre des membres a une importance : la dimension temps.

Table de dimension : Une table dimension est une table qui contient les axes d'analyse (les dimensions) selon lesquels on veut étudier des données observables (les faits) qui, soumises à une analyse multidimensionnelle, donnent aux utilisateurs des renseignements nécessaires à la prise de décision. Les tables de dimension servent à enregistrer les descriptions textuelles des dimensions de l'activité. Chacune de ces descriptions textuelles concerne un membre de la dimension concernée.

2.1.3 Concept d'hierarchie

Une hiérarchie est une organisation logique des membres d'une dimension de manière hiérarchique. Elle représente pour l'utilisateur des chemins de consolidation d'indicateurs (faits). Elle est constituée de niveaux et chaque niveau est représenté par une entité. Certaines entités sont rattachées à d'autres par des liens d'appartenance ou de regroupement hiérarchique.

2.1.4 Concept de mesure

Une mesure est un indicateur numérique représentant une grandeur (comme par exemple le chiffre d'affaire ou le prix d'un produit). Elle permet d'évaluer une quantité aux croisements de dimensions d'analyse. Ces mesures se trouvent dans la table de fait. Une mesure est accompagnée d'un ensemble de fonctions d'agrégation qui permettent de l'agréger en fonction des axes d'analyse.

2.2 Stockage des données

Les données informatisées au sein de l'UCAD ont pris une importance qui ne cesse de croître. L'UCAD se voit ainsi confrontée à un problème de volume de données, rendant leur exploitation difficile avec des bases de données relationnelles. Afin de rendre ces données exploitables facilement, elles peuvent être stockées d'abord dans des entrepôts de données puis dans des cubes et ou éventuellement des datamarts selon les spécifications des besoins métiers.

2.2.1 DataWarehouse

Un DataWarehouse ou un entrepôt de données, constitue l'espace de stockage sur lequel repose un système décisionnel. Ce terme, est apparu dans les années 90 avec Bill Inman qui fut le premier à l'avoir employé. Ce dernier définit un entrepôt de données comme : « une collection de données orientées sujet, intégrées, variant selon le temps et non volatiles, qui sert de support au processus de prise de décision des acteurs du management [(les décideurs)] ». Plus précisément la collection de données est [Inman, 1996] :

- Orientée sujet : les données constituent des granules d'information concernant des sujets d'analyse plutôt que les opérations de gestion se déroulant au sein de l'entreprise.
- Intégrée : les données sont centralisées dans un entrepôt à partir d'un ensemble de sources de données variées. Les données sont fusionnées et agencées au sein d'une vision cohérente.
- Variant selon le temps : Toutes les données d'un entrepôt sont identifiées par des périodes temporelles spécifiques. On parle d'historisation des données [Kimball, 1996], [Inmon, 1996], [Teste, 2000].
- Non volatile : Les données d'un entrepôt sont stables. Il est possible d'ajouter des données, mais on ne modifie pas les données déjà intégrées. Il est toutefois possible de les archiver.

2.2.2 Cube

Dans les entreprises, les outils d'analyse sont très limités. Il s'agit souvent des rapports PDF ou autres fichiers Excel. Si ces derniers permettent une vue d'ensemble de l'activité, ils trouvent rapidement leurs limites lorsqu'il s'agit de rentrer dans des analyses plus fines. Afin d'en améliorer la performance et obtenir une vue transversale de l'activité, il est alors possible de se tourner vers les cubes OLAP.

Une Cube également appelé hypercube, est une représentation abstraite d'informations multidimensionnelles exclusivement numériques, utilisée par l'approche OLAP et prévue à des fins d'analyse. Le cube OLAP est constitué de faits numériques appelés mesures et de dimensions.

➤ Navigation dans les données

Plusieurs opérations sont possibles sur un cube, offrant ainsi la possibilité de naviguer dans les données qui le constituent.

❖ Slice & Dice

Le « Slicing » et le « Dicing » sont des techniques qui offrent la possibilité de faire des tranches dans les données par rapport à des filtres de dimension bien précis. Ce sont des opérations liées à la structure, c'est-à-dire qu'elles se font sur les dimensions. Cependant, il existe une petite nuance entre ces deux opérations : le Slicing consiste à faire une sélection de tranches du cube selon des prédicats et selon une dimension. Tandis que le Dicing, quant à lui, peut être vu comme une extraction d'un sous cube.

❖ Drill-down & Roll-up

Ces méthodes, appelées aussi « forage vers le bas/vers le haut », sont les méthodes les plus répandues pour une navigation dans un entrepôt de données. Elles consistent à représenter les données du cube à un niveau de granularité inférieur, dans le cas du « Drill down », ou un niveau supérieur, c'est le « Roll-up ». En somme, ces deux opérations permettent de contrôler le niveau de détail des données du cube.

Elles se basent sur la notion de hiérarchisation des attributs d'une dimension et la différenciation entre tous les niveaux de hiérarchie disponibles dans les différentes dimensions.

➤ Langage de requêtage sur un cube

Il est important de savoir interroger les données multidimensionnelles provenant de cubes OLAP. Pour cela, on utilise un langage spécialement dédié en ce sens : il s'agit du MDX.

MDX est l'acronyme de Multi Dimensional eXpression. C'est un langage de requêtes OLAP pour les bases de données multidimensionnelles. Il existe depuis 1997, et est fait pour naviguer dans les bases multidimensionnelles, et définir des requêtes sur tous leurs objets (dimensions, hiérarchies, niveaux, membres et cellules), afin d'obtenir une représentation sous forme de tableaux. MDX est proche du SQL avec ses instructions select et where, mais la similarité ne va pas plus loin. MDX fournit la syntaxe DDL (Data Definition Language) pour gérer les structures de données. Cependant, ses fonctionnalités peuvent être plus complexes et robustes que les fonctionnalités de SQL.

Avec MDX, des parties spécifiques des données d'un cube peuvent être extraites puis manipulées pour analyse. Cela permet un examen approfondi et flexible des données du cube OLAP. Les utilisateurs de MDX peuvent profiter de fonctionnalités telles que les mesures calculées, les opérations numériques et les dimensions des axes et des tranches.

2.3 Le modèle en étoile

Pour intégrer les données des Finances dans le Data Warehouse de la DISI, nous allons utiliser un modèle en étoile.

Le modèle en étoile consiste en une grande table de fait encerclée par d'autres tables qui contiennent les éléments descriptifs du fait. C'est la structure de données la plus utilisée et la plus appropriée aux requêtes et analyses des utilisateurs d'entrepôts de données. En effet, elle permet une économie de jointures à l'interrogation, ce qui le rend optimisé pour les requêtes d'analyse.

C'est un modèle simple à créer, stable et intuitivement compréhensible par les utilisateurs finaux.

- Avantages
 - Nombre de jointures limité ;
 - Facilité de navigation.
- Inconvénient
 - Redondance dans les dimensions.

2.4 Architecture OLAP

OLAP, acronyme de Online Analytical Processing, est l'ensemble des technologies qui, se basant sur une représentation multidimensionnelle des données, permet aux analystes et décideurs de traiter leurs données de façon analytique, interactive et rapide. Il donne la possibilité à ces derniers de voir les données de l'entreprise sous plusieurs angles.

C'est un outil de reporting dont la couche d'analyse permet de générer les indicateurs de résultats en fonction du contenu d'un entrepôt de données. OLAP est une méthode d'analyse apparue pour la première fois en 1993, dans le livre blanc intitulé *Providing OLAP to User-Analysts : An IT Mandate*, réalisé par E.F. Codd qui est l'un des concepteurs des bases de données relationnelles.

2.4.1 Règles de base

E.F.Codd dans son ouvrage cité précédemment, a énuméré douze règles qui permettaient à une technologie de faire de l'analyse sur les données. Ainsi, pour qu'un outil puisse être qualifié d'outil OLAP, il doit respecter les caractéristiques énoncées par ces douze règles. Celles-ci sont :

- Vue conceptuelle multidimensionnelle : elle constitue le cœur des produits OLAP. Le Modèle OLAP est multidimensionnel par nature. Une base OLAP offre une vue multidimensionnelle des données ;
- Transparence : que l'OLAP fasse ou non partie du produit frontal habituel de l'utilisateur (par exemple, feuille de calcul ou package graphique), il doit être transparent ;

- Accessibilité : l'utilisateur doit pouvoir accéder sans ambiguïté à toutes les données nécessaires à ses analyses. La complexité et l'hétérogénéité des données sont masquées par les outils OLAP ;
- Performances de rapport cohérentes : l'analyste utilisateur OLAP ne doit pas percevoir de dégradation significative des performances de génération de rapports. En d'autres termes, la performance des reporting reste stable indépendamment du contexte d'analyse ;
- Architecture client-serveur : les produits OLAP doivent être capables de fonctionner dans un environnement client-serveur. À cette fin, il doit être possible de rattacher les différentes sources avec un peu de programmation ;
- Dimensionnalité générique : Chaque dimension doit être équivalente dans sa structure et ses capacités opérationnelles. Toute fonction appliquée à une dimension s'applique aussi aux autres. Il ne doit exister qu'une seule structure logique pour toutes les dimensions ;
- Manipulation de matrice dynamique clairsemée : Le schéma physique des outils OLAP doit s'adapter pleinement au modèle analytique spécifique en cours de création pour assurer une gestion optimale de la matrice creuse. C'est-à-dire qu'il doit assurer la gestion dynamique de la mémoire physique nécessaire pour stocker les données non nulles ;
- Supports multi-utilisateurs : les outils OLAP doivent fournir un accès simultané (récupération et mise à jour, intégrité, sécurité, gestion des accès concurrents aux données) ;
- Opérations multidimensionnelles sans restriction : le serveur OLAP permet la réalisation d'opérations inter dimensions sans restriction ;
- Manipulation intuitive des données : le serveur OLAP permet une manipulation aisée des données ;
- Rapports flexibles : la simplicité (ou souplesse) de l'édition des rapports est intrinsèque au modèle ;
- Dimensions et niveaux d'agrégation illimités : le nombre de dimensions et de niveaux d'agrégation possibles est suffisant pour autoriser les analyses les plus poussées. Tout outil OLAP doit gérer au moins 15 à 20 dimensions dans le même modèle. Le nombre de niveaux d'agrégation est illimité.

2.4.2 Caractéristique majeure des produits OLAP

L'acronyme FASMI (Fast Analysis of Shared Multidimensional Information) permet de résumer la définition des produits OLAP.

- Fast : le temps de réponse aux demandes des utilisateurs doit être entre 1 et 20 secondes ;
- Analysis : faire face à toutes les logiques d'affaire et de statistiques, ainsi que fournir la possibilité aux utilisateurs de construire leurs calculs et leurs analyses sans avoir à programmer ;
- Shared : le système doit créer un contexte où la confidentialité est préservée et doit gérer les cas où plusieurs utilisateurs ont des droits en écritures ;

- Multidimensional : caractéristique majeure, les produits OLAP doivent fournir des vues conceptuelles multidimensionnelles des données et supporter des hiérarchies de dimensions ; c'est l'intérêt de OLAP, les données doivent pouvoir être consultées sous différents angles de vue ;
- Information : Accès à toutes les données et informations, où qu'elles résident et de façon illimitée en termes de volumétrie.

2.4.3 Différents types de modèle OLAP

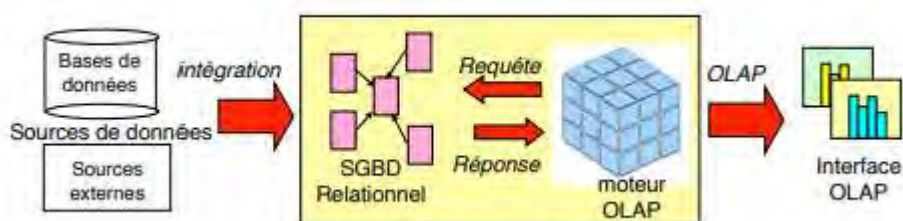
Les systèmes décisionnels s'appuient sur une architecture trois tiers :

- ✓ Le premier tiers est un SGBD : les données pertinentes pour l'analyse sont extraites des bases de données transactionnelles, nettoyées et transformées avec les outils ETL et intégrées dans un entrepôt de données maintenu par le SGBD ;
- ✓ Le deuxième niveau de l'architecture est un serveur OLAP, tels que par exemple Mondrian, DB2 Olap Server, Oracle OLAP ou Microsoft Analysis Services. Le serveur OLAP génère et maintient l'hypercube. Il fournit une vue multidimensionnelle des données et implémente l'ensemble des opérateurs OLAP permettant de les analyser ;
- ✓ Le dernier niveau est un client OLAP, par exemple Microstrategy, Cognos, JPivot etc. Il offre une interface utilisateur composée d'outils de reporting, d'analyse interactive, et parfois de fouille de données. Son rôle est de permettre la visualisation de l'information multidimensionnelle et la découverte des connaissances.

Ainsi, depuis 1993, il s'est créé à peu près autant de technologies OLAP que de fournisseurs. Progressivement, le marché s'est structuré autour de trois approches :

a. ROLAP

Relational OLAP ou ROLAP est un mode de stockage qui utilise un SGBD relationnel et qui organise les données au sein des relations correspondant aux tables du schéma de stockage. Il est capable de simuler le comportement d'un SGBD multidimensionnel en exploitant un SGBDR classique. Le serveur extrait les données par les requêtes SQL et les interprète selon une vue multidimensionnelle pour les présenter au module client.



FIGUTRE 2.1- Architecture ROLAP

- Avantages des systèmes ROLAP :
 - souplesse ;
 - évolution facile ;
 - permettent de stocker de gros volumes de données ;
 - peu coûteux.
- Inconvénients des systèmes ROLAP :
 - peu efficaces pour les calculs complexes ;
 - temps de réponse plus lents.

b. MOLAP

Ce système est basé sur un stockage par tableaux (techniques des matrices creuses) et une indexation rapide des données calculées. Il stocke les données basiques et leurs agrégations (toutes les intersections possibles entre les axes) dans une base multidimensionnelle sur un serveur spécialisé OLAP. Les langages de requêtes MOLAP sont optimisés pour les applications d'analyse et la meilleure performance en termes de temps de réponse pour des requêtes complexes.

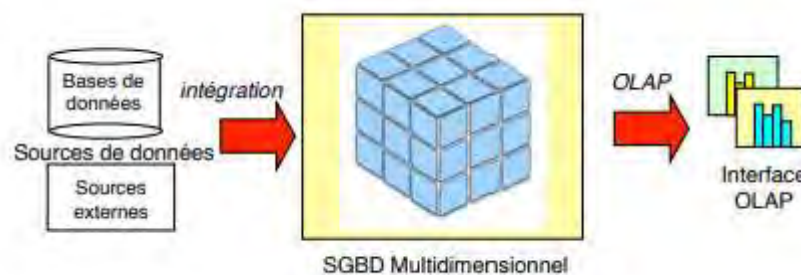


FIGURE 2.2- Architecture MOLAP

- Avantages des systèmes MOLAP :
 - temps de réponse satisfaisant ;
 - rapide et performante ;
 - conçus exclusivement pour l'analyse multidimensionnelle.
- Inconvénients des systèmes MOLAP :
 - difficiles à mettre en place ;
 - formats souvent propriétaires ;
 - peu adaptés pour les gros volumes de données.

c. HOLAP

HOLAP est un modèle hybride, qui découle de la combinaison de ROLAP avec MOLAP. Selon (Moorman, 1999), il s'agit de garder les grands volumes de données détaillées dans la partie ROLAP du système, car celle-ci présente de meilleures performances pour la gestion des grands volumes et intègre plus facilement les mises à jour incrémentales de l'entrepôt. La partie MOLAP du système permet un accès plus rapide aux données fréquemment consultées et stocke donc les cuboïdes contenant les agrégats matérialisés. Les systèmes HOLAP cherchent à apporter une convivialité d'utilisation tout en s'appuyant sur des modèles spécifiques permettant de passer d'un modèle relationnel à un modèle multidimensionnel avec des notions de faits, de hiérarchies, de dimensions.

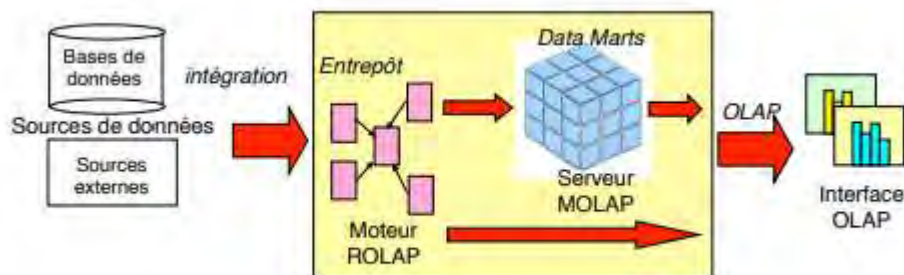


FIGURE 2.3-Architecture HOLAP

- Avantages des systèmes HOLAP :
 - bon compromis au niveau des coûts et des performances ;
 - permet de gérer de très grandes quantités de données et d'avoir des temps de réponses acceptables en analyse OLAP.
- Inconvénient des systèmes HOLAP :
 - difficile de les mettre en œuvre.

Chapitre 3 : Conception de notre solution

1. Conception architecturale

L'architecture de notre système décisionnel peut être représentée en quatre niveaux.

- Le premier niveau est celui des sources de données à savoir le système de gestion de l'entreprise contenant les bases de données opérationnelles ainsi que les sources externes ;
- Le deuxième niveau concerne la récupération, la transformation des données, puis l'alimentation de l'entrepôt de données ;
- Le troisième niveau est celui de l'exploitation de l'entrepôt en question ;
- Le quatrième niveau fournit à l'utilisateur final les moyens de composer ses propres analyses et restitutions des données.

1.1 Architecture

L'architecture mise en place est constituée d'une base de données pour le stockage, d'un outil de type ETL (Extraction Transformation Loading) pour alimenter la base à partir des systèmes sources et de différents outils pour restituer les informations aux utilisateurs (reporting, analyse, outil statistique, etc.). Dans la suite, sont détaillées les différentes étapes nécessaires à la mise en place de notre SID

- **Acquisition de l'information** : Cette phase va faire intervenir les processus ETL qui se chargeront de récupérer toutes les données nécessaires depuis les différentes sources de données. Il convient alors de localiser ces données. Ces données peuvent provenir de différentes sources (fichiers CSV, ORACLE). Dans l'acquisition des données, nous allons procéder aux étapes suivantes :
 - Extraction : collection de données utiles ;
 - Préparation : transformation des caractéristiques des données dans le modèle de l'entrepôt (filtrer, trier, homogénéiser, nettoyer, etc.)
 - Chargement : insertion des données dans l'entrepôt après nettoyage.
- **Stockage** : Les données sont chargées dans l'entrepôt pour traiter des applications décisionnelles.
- **Restitution** : Mise à disposition des données pour les utilisateurs finaux afin de pouvoir les interroger, les visualiser ou les analyser. Il existe plusieurs outils de restitution (tableaux de bord, requêtes SQL, analyse multidimensionnelle, data mining ...)

1.2 Schéma général de l'application

A l'UCAD, nous avons plusieurs sources de données. Ces données proviennent de différentes bases données de production à savoir les bases de la scolarité, des ressources

humaines, de la paie, de la comptabilité, des finances, etc. Ces différentes bases de production vont suivre l'ensemble des traitements nécessaires cités ci-haut pour leur mise en cohérence. Une fois que ces données sont traitées pour être cohérentes, elles seront stockées dans le DataWarehouse. Suite à leur stockage, on les met dans des datamarts et dans des cubes pour des besoins métiers. A partir des cubes ou datamarts, les données sont mises à disposition dans des tableaux de bord pour les utilisateurs finaux. Ci-dessous, nous avons le schéma général du SID avec les étapes citées en amont.

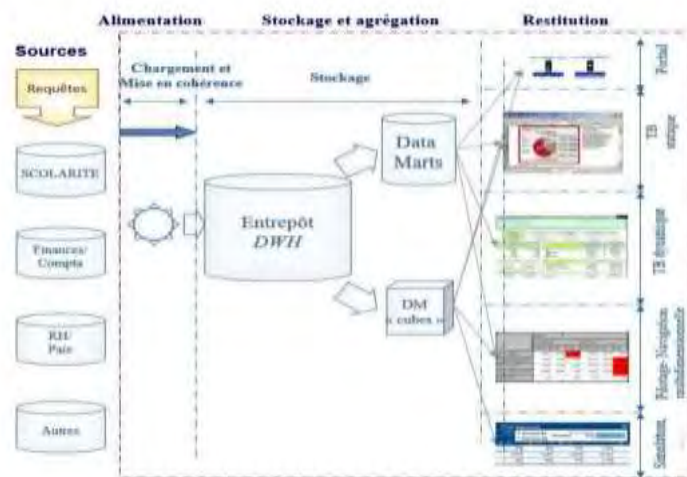


FIGURE 3.1 - Architecture générale du SID de l'UCAD

1.3 Schéma actuel de l'application

1.3.1 Architecture du système existant

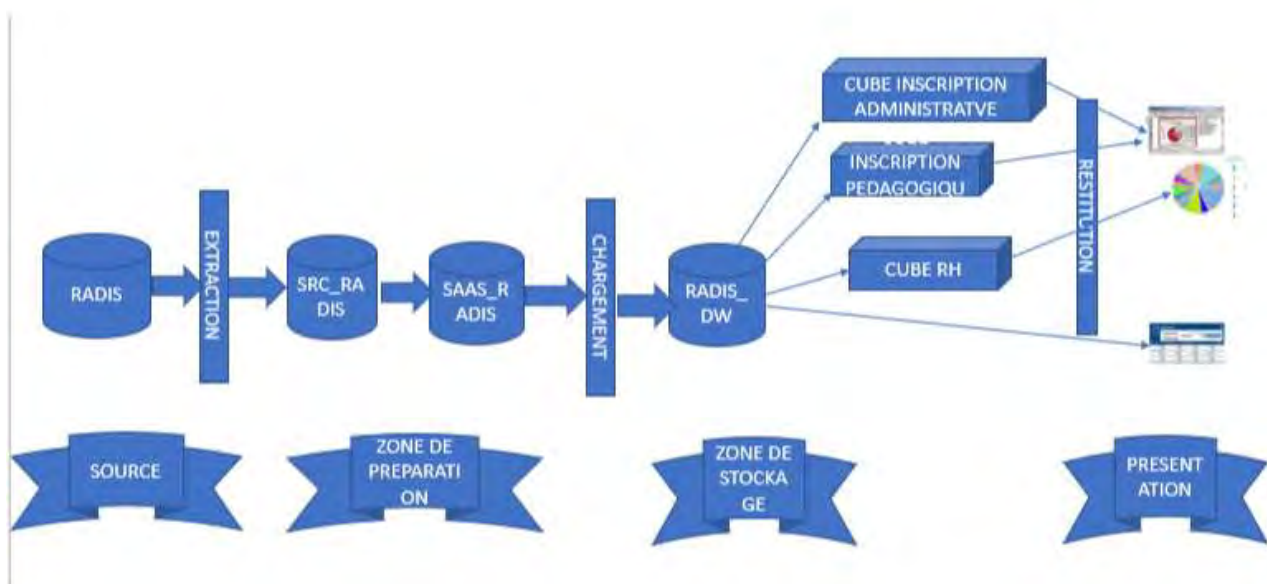


FIGURE 3.2 - Architecture actuelle du DWH existant

1.3.2 Architecture de notre solution

Les données que nous utilisons actuellement proviennent de la base de production oracle des finances. Ces données sont d'abord stockées dans une base de données oracle que l'on nomme SRC_RADIS. A partir de cette dernière, les données sont transférées dans une base de données nommée SAAS_RADIS. Les données de celle-ci suivront un certain nombre de traitements pour leur mise en cohérence. Après leur traitement, ces données sont stockées dans une base de données DW_FIN_RADIS, et c'est ainsi qu'est alimenté notre Datawarehouse.

Après le stockage de ces données propres dans le DataWarehouse, elles sont mises dans un cube OLAP pour des besoins métiers de la finance. A partir de ce cube ou même du DataWarehouse, les données sont mises à disposition dans des rapports ou tableaux de bord pour les utilisateurs finaux.

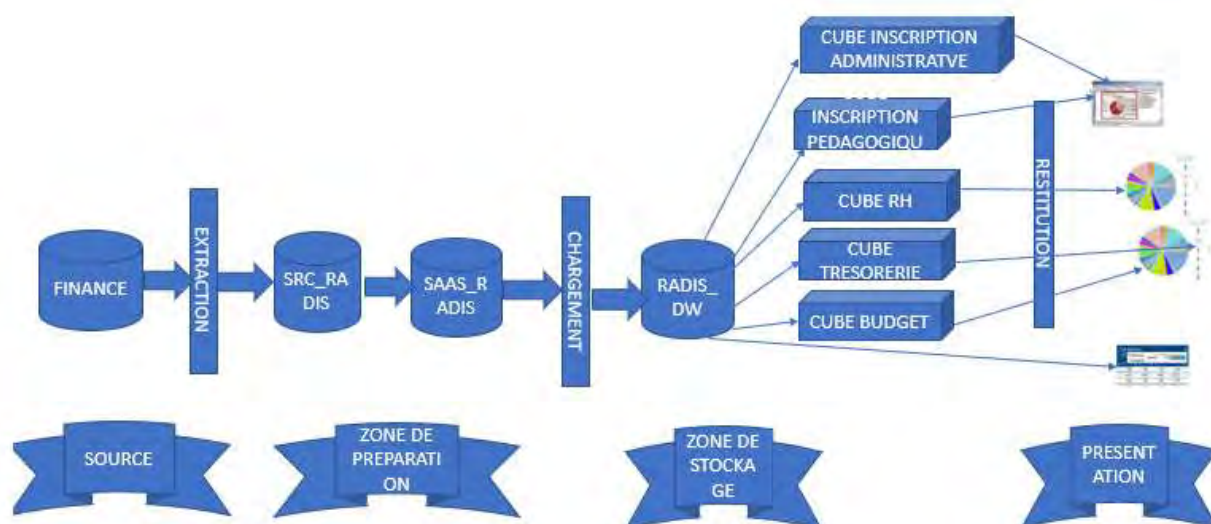


FIGURE 3.3 - Architecture actuelle de notre solution

2. Modélisation

2.1 Reverse Engineering

Le reverse engineering, plus communément appelé ingénierie inversée ou rétro-ingénierie représente l'étude d'un objet physique afin de découvrir la façon dont il fonctionne et dont il a été fabriqué. On peut aussi parler de concept inversé : c'est à dire partir d'un objet tout fait pour arriver à son implémentation étape par étape.

Réalisée généralement en deux phases, le reverse engineering consiste à l'identification des composants du système et de leurs relations, puis leurs représentations à un niveau d'abstraction plus élevé et plus compréhensible par les humains (spécifications, documentations, plans, schémas).

L'application XAALIS des finances est une solution pour laquelle il n'existe aucune documentation sur place. Pour parvenir à l'élaboration de nos différents modèles décisionnels, nous avons dû faire recours au reverse engineering. En effet, XAALIS est une application dont les états sont bâtis sur un ensemble de vues. Donc le travail consistait à recenser l'ensemble des tables référencées dans ces vues et des relations qui existaient entre ces différentes tables afin de pouvoir ressortir le modèle relationnel. Les images ci-dessous illustrent une partie du modèle qui découle de ce travail de reverse engineering :

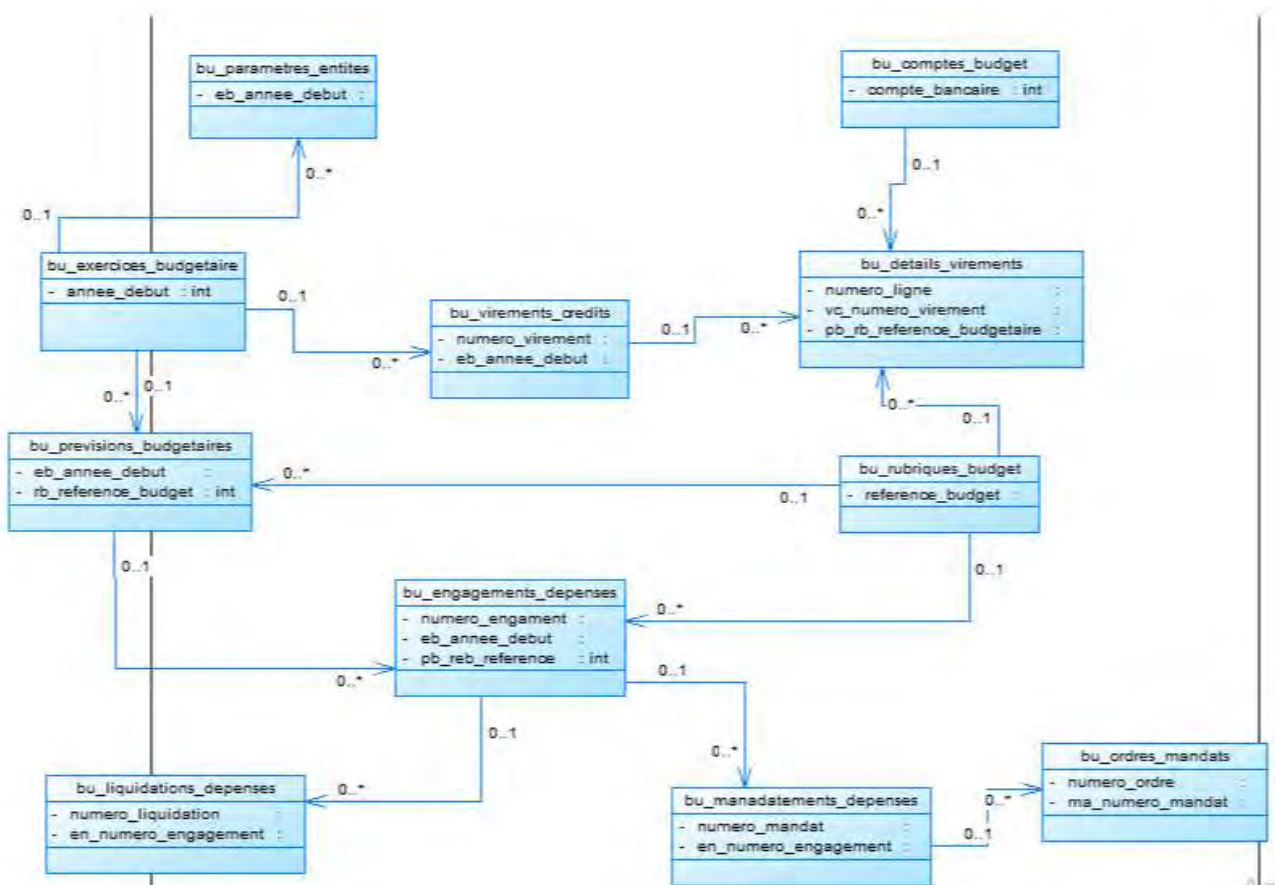


FIGURE 3 .4 – Modèle relationnel des finances après reverse engineering (1)

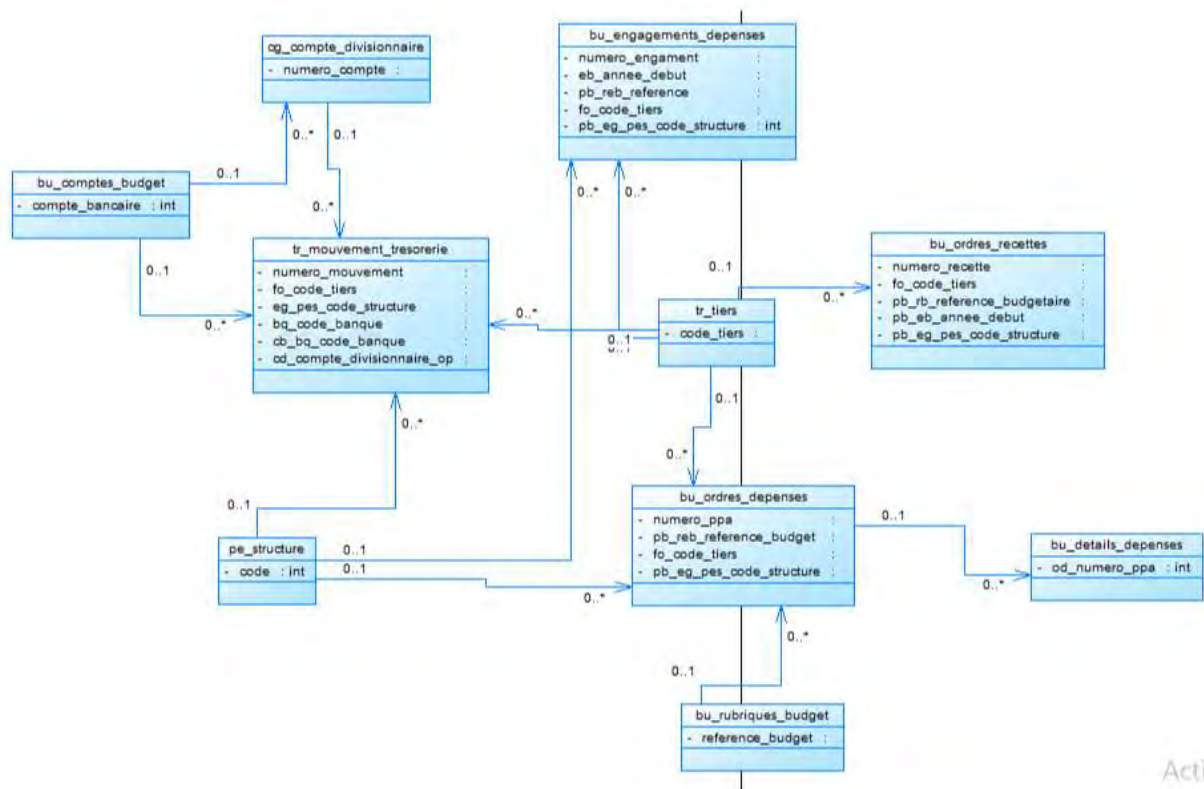


FIGURE 3.4 – Modèle relationnel des finances après reverse engineering (2)

2.2 Modélisation du DataWarehouse

Compte tenu du modèle relationnel élaboré grâce au reverse engineering, nous sommes parvenus à mettre en place des modèles multidimensionnels.

Dans la modélisation, nous allons utiliser le schéma en étoile qui nous a été imposé par l'existant et qui répond aussi à nos besoins. Nous avons des modèles : un pour la gestion de la trésorerie et les autres pour la gestion budgétaire. Chacun de ses modèles est spécifique à un sujet d'analyse et l'ensemble de ces derniers est spécifique à la finance.

a. Mouvements de trésorerie des finances

Ce modèle comprend 13 tables de dimensions et une table de fait centrale.

Les mouvements de trésorerie retracent toutes les opérations de trésorerie. En effet ils permettent de procéder au paiement des mandats, de recevoir des ordres de recettes et des ordres de dépenses, d'émettre des bordereaux dans lesquels se trouvent des opérations de trésorerie qui symbolisent des bons de caisse en fonction de la structure, du compte divisionnaire et du tiers.

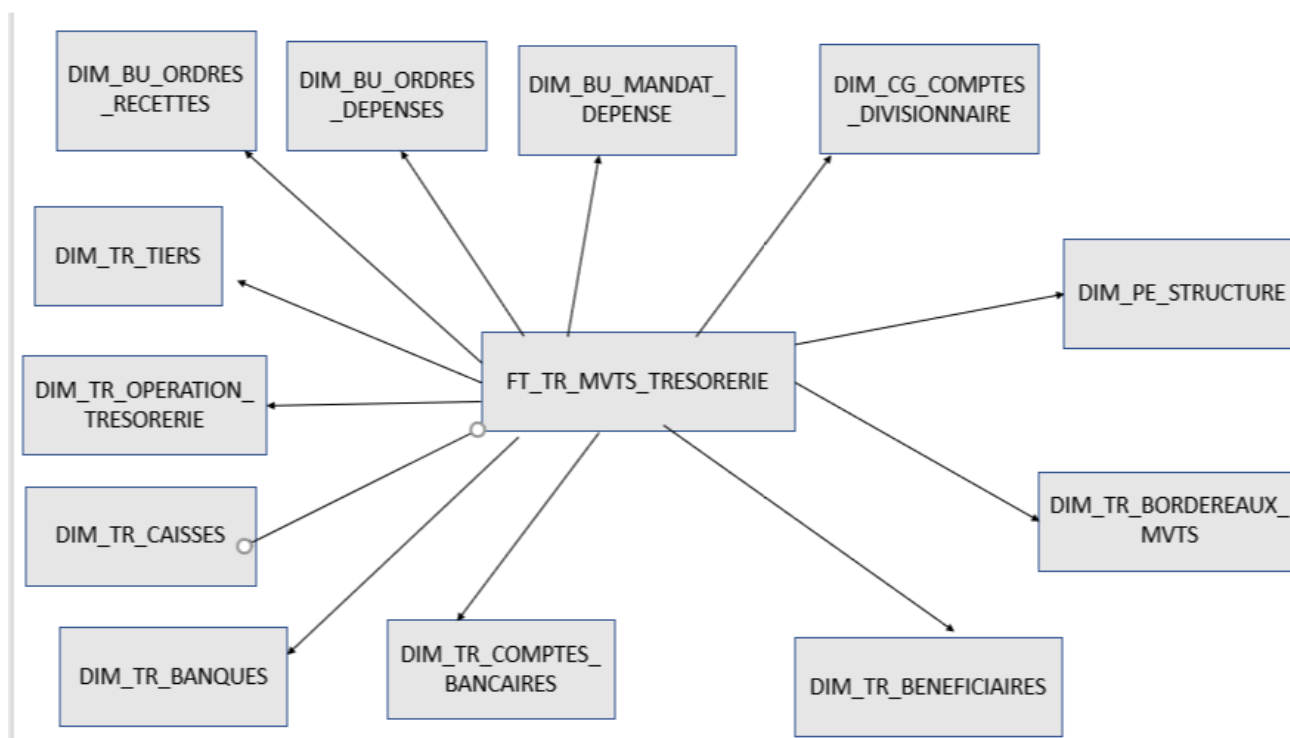


FIGURE 3.5 – Modèle des mouvements de trésorerie des finances

NOM DE DIMENSION	DESCRIPTION DES DIMENSIONS
DIM_BU_ORDRES_RECETTES	Permet de constater une ressource /recette.
DIM_BU_ORDRES_DEPENSES	Représente le paiement par anticipation et généralement effectuer en début de gestion lorsque le budget n'est pas encore approuvé.
DIM_BU_MANDATEMENT_DEPENSES	Autorise l'agent comptable a exécuter la dépense.
DIM_BU_CG_COMPTES_DIVISIONNAIRES	Contient La nomenclature comptable hors comptes de classe et comptes principaux.
DIM_BU_PE_STRUCTURES	Toutes les structures comptables pouvant être associées à un établissement a budget autonome
DIM_TR_TIERS	Représente toutes personnes physiques ou morales pouvant être associés à un fournisseur, un client, un employé ou un étudiant .
DIM_TR_BANQUES	Etablissement financière.
DIM_TR_COMPTES_BANCAIRES	Compte bancaire appartenant à un tiers ou à un établissement.
DIM_TR_OPERATION_TRESORERIE	Contient toute la codification des opérations de trésorerie ainsi que les limites des montants.
DIM_TR_CAISSE	Réserver aux opérations en espèce
DIM_TR_BENEFICIAIRES	Toutes personnes physiques ou morales impliquées dans une opération de trésorerie.
DIM_TR_BORDERAUX_MOUVEMENTS	Contient tous les mouvements de trésorerie d'une période où d'un type donné
DIM_DATE	Dimension temporelle qui permet de daté les fait.

TABLE 3.1-Les dimensions des mouvements de trésorerie des finances

b. Liquidations

La liquidation est une procédure qui constate la livraison et la conformité du cahier de charge.

Ce modèle comprend 6 tables de dimensions et une table de faits centrale. En effet une liquidation fait intervenir un tiers qui désigne le fournisseur, un établissement, une prévision budgétaire pour couvrir une rubrique budget, et honorer les engagements de dépenses.

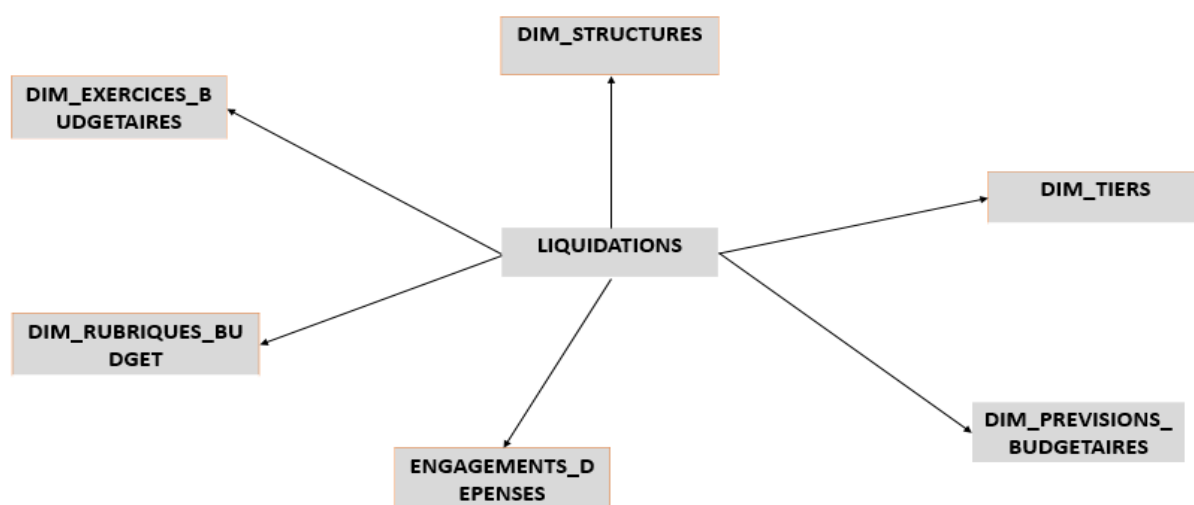


FIGURE 3.6 - Liquidations

Nom de la dimension	Description
DIM_STRUCTURES	Contient les informations sur les différentes structures telles que le code de la structure, l'agent comptable de la structure, l'adresse de cette dernière ...
DIM_RUBRIQUE_BUDGET	Contient l'ensemble des informations relatives au budget comme la référence budgétaire, le titre du budget, la section du budget, le nom de la rubrique, le type de la rubrique, l'état de la rubrique ...
DIM_EXERCICES_BUDGETAIRES	Dans cette dimension on trouve tout ce qui a trait à la gestion budgétaire notamment l'année de début, le

	libellé de l'exercice, le début de l'exercice, la fin de l'exercice ...
DIM_ENGAGEMENT_DEPENSES	Rassemble les informations sur les engagements de dépenses
DIM_TIERS	Renferme les informations sur les tiers à savoir le nom du tiers, la ville du tiers, l'adresse du tiers, le mail du tiers, le téléphone du tiers, etc.
DIM_PREVISIONS_BUDGETAIRES	Contient les informations sur les prévisions budgétaires: prévision annuelle, budget additionnelle, réalisation additionnelle ...

TABLE 3.2 - Liquidation

c. Crédits

Un crédit représente le cliché des crédits alloués.

Ce modèle comprend 5 tables de dimensions et une table de faits centrale. Il fait intervenir une structure, une rubrique budgétaire, un exercice budgétaire qui n'est rien d'autre que la gestion budgétaire, un engagement de dépense et un mandatement de dépenses.

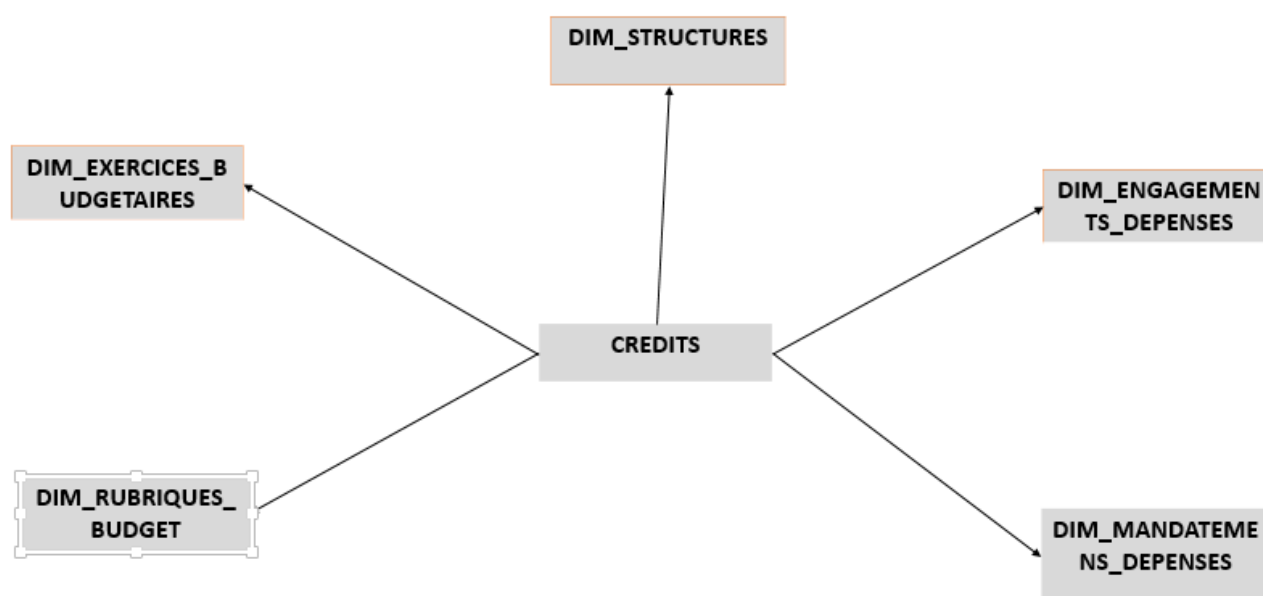


FIGURE 3.7 - Crédits

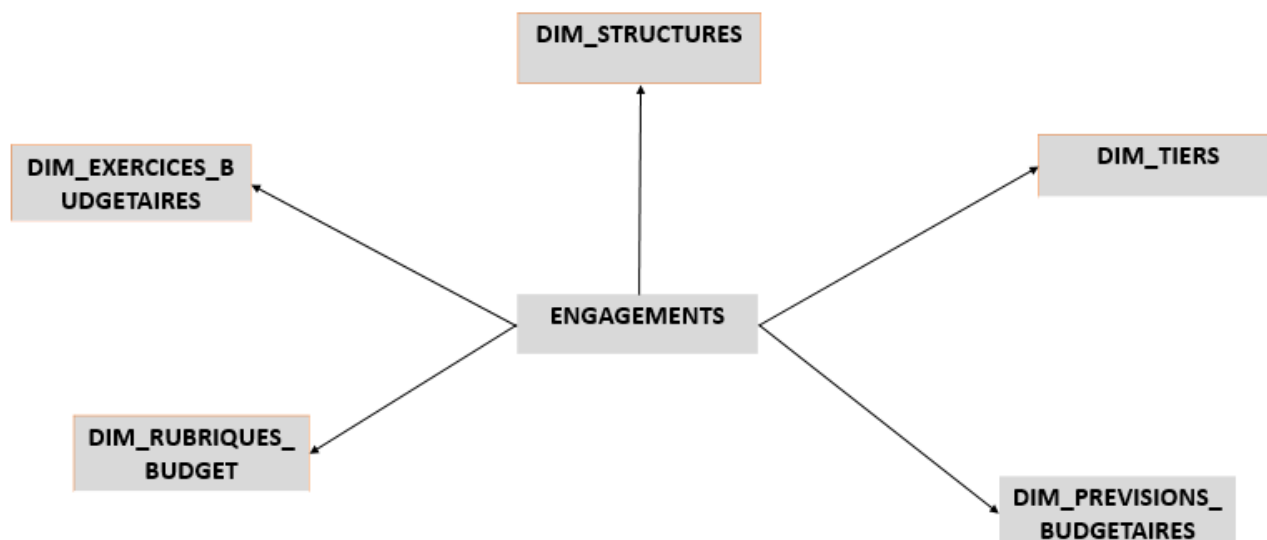
Nom de la dimension	Description
DIM_STRUCTURES	Contient les informations sur les différentes structures telles que le code de la structure, l'agent comptable de la structure, l'adresse de cette dernière ...
DIM_RUBRIQUE_BUDGET	Contient l'ensemble des informations relatives au budget comme la référence budgétaire, le titre du budget, la section du budget, le nom de la rubrique, le type de la rubrique, l'état de la rubrique ...
DIM_EXERCICES_BUDGETAIRES	Dans cette dimension on trouve tout ce qui a trait à la gestion budgétaire notamment l'année de début, le libellé de l'exercice, le début de l'exercice, la fin de l'exercice ...
DIM_ENGAGEMENT_DEPENSES	Rassemble les informations sur les engagements de dépenses
DIM_MANDATEMENTS_DEPENSES	Renferme les informations sur les mandatements de dépense à savoir l'objet de mandat, précompter, etc.

TABLE 3. 3 - Crédits

d. Engagements

Un engagement marque le début de la phase des dépenses. Il permet de mobiliser les crédits subvention de l'état, ressources informatiques, voyage d'étude et rémunération du personnel

Ce modèle est composé de 5 tables de dimensions et d'une table de faits centrale. Il fait intervenir une structure, une rubrique budgétaire, un exercice budgétaire, une prévision budgétaire et un tiers.



FIGURES 3.8 : Engagements

Nom de la dimension	Description
DIM_STRUCTURES	Contient les informations sur les différentes structures telles que le code de la structure, l'agent comptable de la structure, l'adresse de cette dernière ...
DIM_RUBRIQUE_BUDGET	Contient l'ensemble des informations relatives au budget comme la référence budgétaire, le titre du budget, la section du budget, le nom de la rubrique, le type de la rubrique, l'état de la rubrique ...
DIM_EXERCICES_BUDGETAIRES	Dans cette dimension on trouve tout ce qui a trait à la gestion budgétaire notamment l'année de début, le libellé de l'exercice, le début de l'exercice, la fin de l'exercice ...
DIM_TIERS	Renferme les informations sur les tiers à savoir le nom du tiers, la ville du tiers, l'adresse du tiers, le

	mail du tiers, le téléphone du tiers, etc.
DIM_PREVISIONS_BUDGETAIRES	Contient les informations sur les prévisions budgétaires: prévision annuelle, budget additionnelle, réalisation additionnelle ...

TABLE 3.4 – Engagements

d. DBM

DBM est l'acronyme de décision budgétaire modificative. Il permet de modifier les crédits. Certains l'appellent aussi virement.

Ce modèle comprend 4 tables de dimensions et une table de fait centrale. Dans ce modèle, interviennent une structure, une rubrique budgétaire, un exercice budgétaire, un détail de virements sachant qu'un virement peut en avoir un ou plusieurs détails.

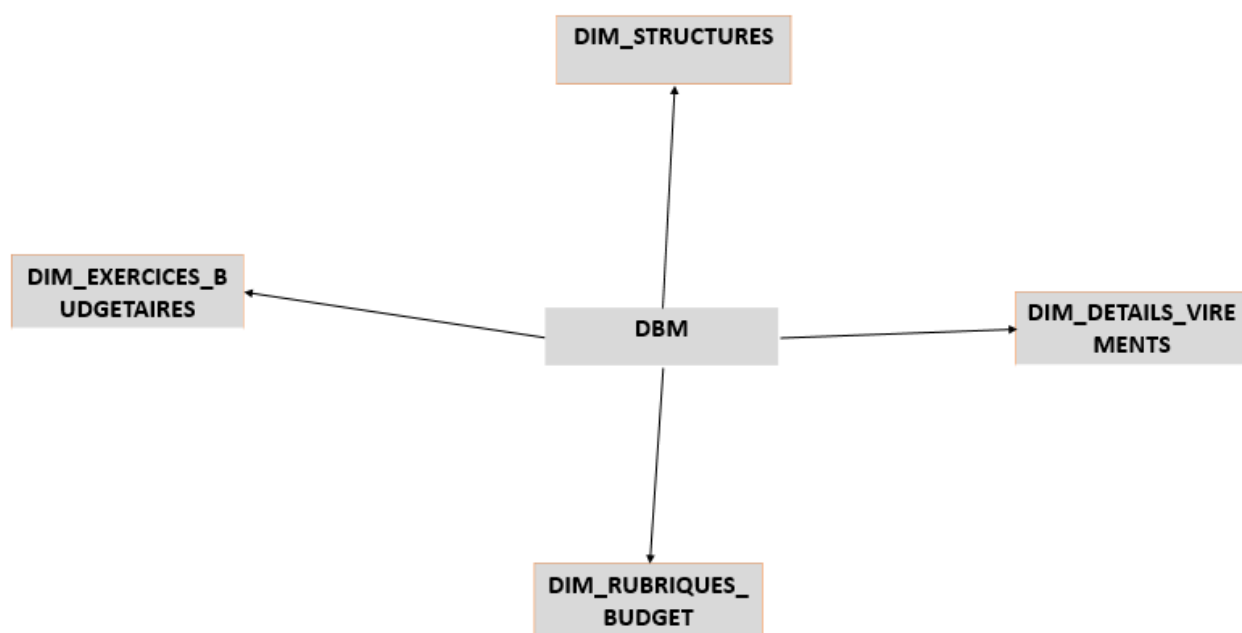


FIGURE 3.9 – DBM

Nom de la dimension	Description
DIM_STRUCTURES	Contient les informations sur les différentes structures telles que le code de la structure, l'agent comptable de la structure, l'adresse de cette dernière ...
DIM_RUBRIQUE_BUDGET	Contient l'ensemble des informations relatives au budget comme la référence budgétaire, le titre du budget, la section du budget, le nom de la rubrique, le type de la rubrique, l'état de la rubrique ...
DIM_EXERCICES_BUDGETAIRES	Dans cette dimension on trouve tout ce qui a trait à la gestion budgétaire notamment l'année de début, le libellé de l'exercice, le début de l'exercice, la fin de l'exercice ...
DIM_DETAILS_VIREMENT	Cette table contient les détails sur les virements

TABLE 3.5 - DBM

e. Mandats

Il permet de donner l'ordre à l'agent comptable de procéder à l'exécution des dépenses.

Ce modèle compte 9 tables de dimensions et une table de faits centrale. Il fait intervenir une structure, une rubrique budgétaire, un exercice budgétaire, une prévision budgétaire, une banque, un état de mandats, un engagement de dépenses, un mode de paiement et un tiers.

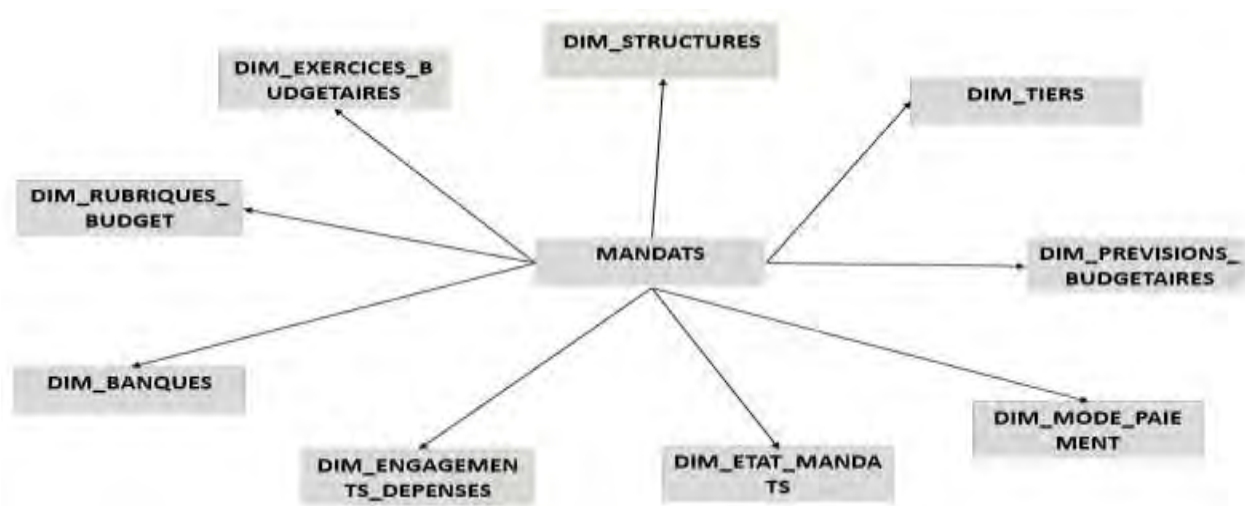


FIGURE 3.10 - Mandats

Nom de la dimension	Description
DIM_STRUCTURE	Contient les informations sur les différentes structures telles que le code de la structure, l'agent comptable de la structure, l'adresse de cette dernière ...
DIM_RUBRIQUE_BUDGET	Contient l'ensemble des informations relatives au budget comme la référence budgétaire, le titre du budget, la section du budget, le nom de la rubrique, le type de la rubrique, l'état de la rubrique ...
DIM_EXERCICES_BUDGETAIRES	Dans cette dimension on trouve tout ce qui a trait à la gestion budgétaire notamment l'année de début, le libellé de l'exercice, le début de l'exercice, la fin de l'exercice ...
DIM_ENGAGEMENT_DEPENSES	Rassemble les informations sur les engagements de dépenses
DIM_TIERS	Renferme les informations sur les tiers à savoir le nom du tiers, la ville du tiers, l'adresse du tiers, le mail du tiers, le téléphone du tiers, etc.

DIM_PREVISIONS_BUDGETAIRES	Contient les informations sur les prévisions budgétaires: prévision annuelle, budget additionnelle, réalisation additionnelle ...
DIM_ETAT_MANDATS	Contient les informations sur les états des mandats.
DIM_MODE_PAIEMENT	Contient les informations sur les modes de paiement.
DIMBANQUES	Renferme des informations sur les banques comme le nom, etc.

TABLE 3.6 - Mandats

f. Ordres_Mandats

Un ordre de mandat correspond à un détail de mandat.

Ce modèle compte 8 tables de dimensions et une table de faits centrale. Il fait intervenir une structure, une rubrique budgétaire, un exercice budgétaire, une prévision budgétaire, une banque, un bénéficiaire, un engagement de dépenses et un tiers.

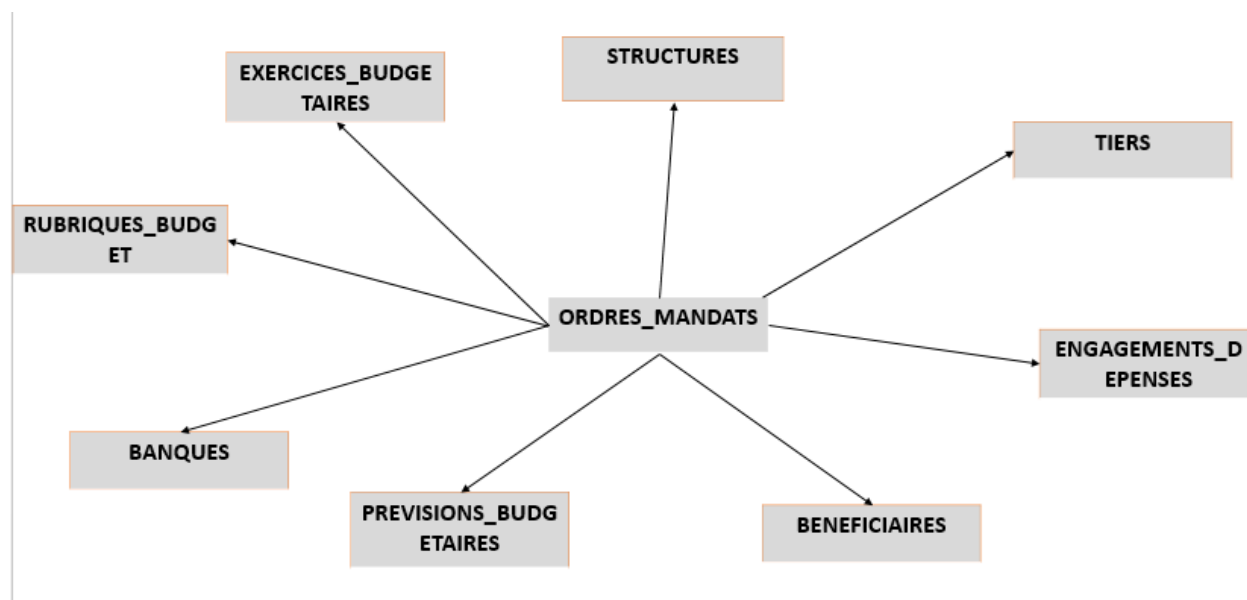


FIGURE 3.11 - Ordres_Mandats

Nom de la dimension	Description
DIM_STRUCTURES	Contient les informations sur les différentes structures telles que le code de la structure, l'agent comptable de la structure, l'adresse de cette dernière ...
DIM_RUBRIQUE_BUDGET	Contient l'ensemble des informations relatives au budget comme la référence budgétaire, le titre du budget, la section du budget, le nom de la rubrique, le type de la rubrique, l'état de la rubrique ...
DIM_EXERCICES_BUDGETAIRES	Dans cette dimension on trouve tout ce qui a trait à la gestion budgétaire notamment l'année de début, le libellé de l'exercice, le début de l'exercice, la fin de l'exercice ...
DIM_ENGAGEMENT_DEPENSES	Rassemble les informations sur les engagements de dépenses
DIM_TIERS	Renferme les informations sur les tiers à savoir le nom du tiers, la ville du tiers, l'adresse du tiers, le mail du tiers, le téléphone du tiers, etc.
DIM_PREVISIONS_BUDGETAIRES	Contient les informations sur les prévisions budgétaires : prévision annuelle, budget additionnelle, réalisation additionnelle ...
DIM_BENEFICIAIRES	Contient les informations sur les bénéficiaires.

TABLE 3.7 - Ordres_Mandats

g. Ordres_Recettes

Un ordre de recette, permet de constater toutes les ressources de l'institution.

Ce modèle compte 5 tables de dimensions et une table de faits centrale. Il fait intervenir une structure, une rubrique budgétaire, un exercice budgétaire, une prévision budgétaire et un tiers.

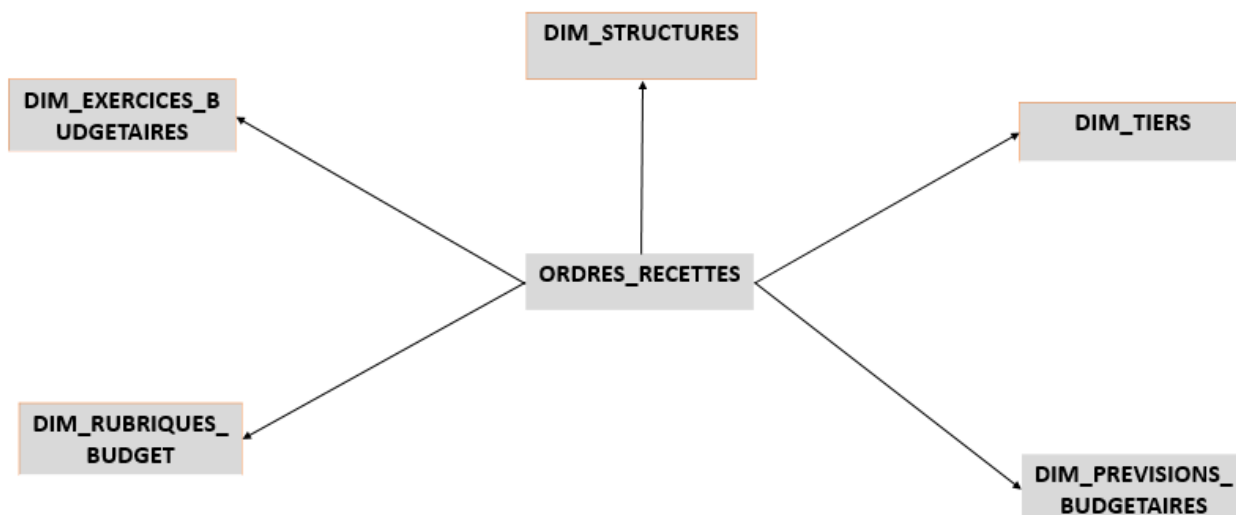


FIGURE 3.12 : Ordres_Recettes

Nom de la dimension	Description
DIM_STRUCTURES	Contient les informations sur les différentes structures telles que le code de la structure, l'agent comptable de la structure, l'adresse de cette dernière ...
DIM_RUBRIQUE_BUDGET	Contient l'ensemble des informations relatives au budget comme la référence budgétaire, le titre du budget, la section du budget, le nom de la rubrique, le type de la rubrique, l'état de la rubrique ...
DIM_EXERCICES_BUDGETAIRES	Dans cette dimension on trouve tout ce qui a trait à la gestion budgétaire notamment l'année de

	début, le libellé de l'exercice, le début de l'exercice, la fin de l'exercice ...
DIM_TIERS	Renferme les informations sur les tiers à savoir le nom du tiers, la ville du tiers, l'adresse du tiers, le mail du tiers, le téléphone du tiers, etc.
DIM_PREVISIONS_BUDGETAIRES	Contient les informations sur les prévisions budgétaires: prévision annuelle, budget additionnelle, réalisation additionnelle ...

TABLE 3.8 - Ordres_Recettes

h. Prévion_Période

Elle permet de faire une prévision du budget primitif pour une période donnée.

Ce modèle comprend 4 tables de dimensions et une table de faits centrale. Il fait intervenir une structure, une rubrique budgétaire, un exercice budgétaire et une prévision budgétaire.

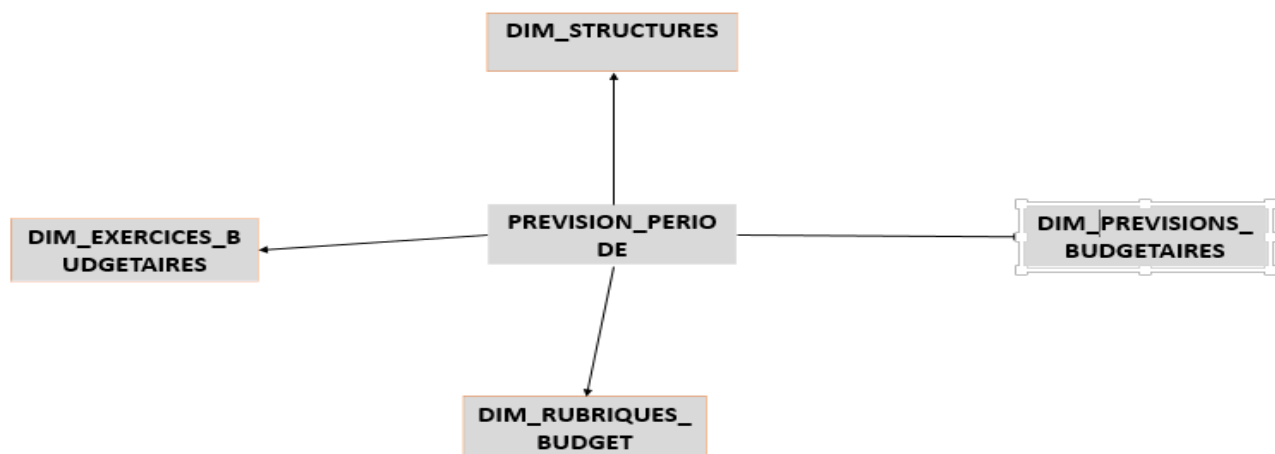


FIGURE 3.13 - Prévion_Période

Nom de la dimension	Description
DIM_STRUCTURES	Contient les informations sur les différentes structures telles que le code de la structure, l'agent comptable de la structure, l'adresse de cette dernière ...
DIM_RUBRIQUE_BUDGET	Contient l'ensemble des informations relatives au budget comme la référence budgétaire, le titre du budget, la section du budget, le nom de la rubrique, le type de la rubrique, l'état de la rubrique ...
DIM_EXERCICES_BUDGETAIRES	Dans cette dimension on trouve tout ce qui a trait à la gestion budgétaire notamment l'année de début, le libellé de l'exercice, le début de l'exercice, la fin de l'exercice ...
DIM_PREVISIONS_BUDGETAIRES	Contient les informations sur les prévisions budgétaires: prévision annuelle, budget additionnelle, réalisation additionnelle ...

TABLE 3.9 – Prévision_Période

i. ressources_générales

Les ressources générales permettent de faire un bilan des ressources de l'institution.

Ce modèle comprend 4 tables de dimensions et une table de faits centrale. Il fait intervenir une structure, une rubrique budgétaire, un exercice budgétaire et un paramètre entité.

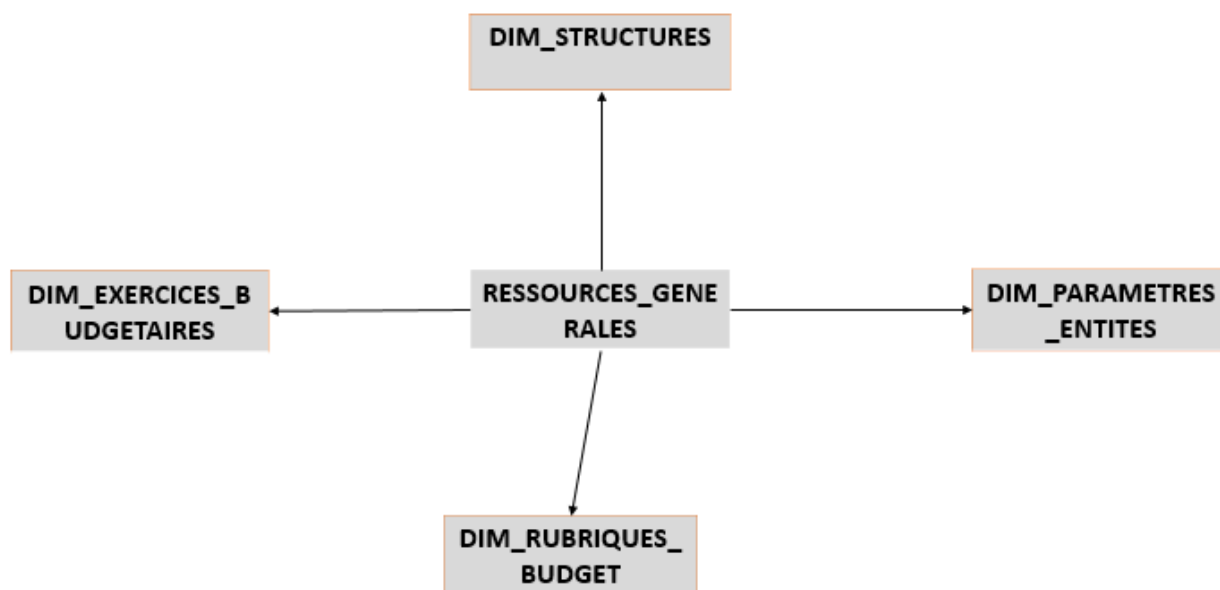


FIGURE 3.14 - Ressources_Générales

Nom de la dimension	Description
DIM_STRUCTURES	Contient les informations sur les différentes structures telles que le code de la structure, l'agent comptable de la structure, l'adresse de cette dernière ...
DIM_RUBRIQUE_BUDGET	Contient l'ensemble des informations relatives au budget comme la référence budgétaire, le titre du budget, la section du budget, le nom de la rubrique, le type de la rubrique, l'état de la rubrique ...
DIM_EXERCICES_BUDGETAIRES	Dans cette dimension on trouve tout ce qui a trait à la gestion budgétaire notamment l'année de début, le libellé de l'exercice, le début de l'exercice, la fin de l'exercice ...

TABLE 3.10 : Ressources_Générales

j. ordres_depenses

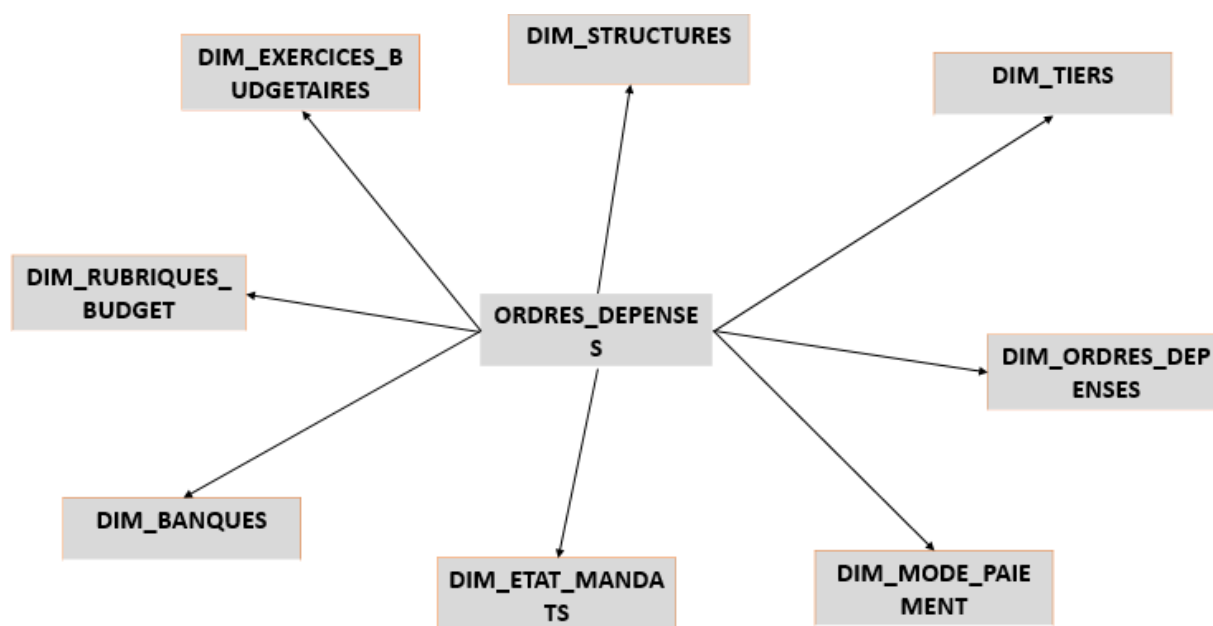


FIGURE 3.15 - Ordres_Depenses

Nom de la dimension	Description
DIM_STRUCTURE	Contient les informations sur les différentes structures telles que le code de la structure, l'agent comptable de la structure, l'adresse de cette dernière ...
DIM_RUBRIQUE_BUDGET	Contient l'ensemble des informations relatives au budget comme la référence budgétaire, le titre du budget, la section du budget, le nom de la rubrique, le type de la rubrique, l'état de la rubrique ...
DIM_EXERCICES_BUDGETAIRES	Dans cette dimension on trouve tout ce qui a trait à la gestion budgétaire notamment l'année de début, le libellé de l'exercice, le début de l'exercice, la fin de l'exercice ...
DIM_ORDRES_DEPENSES	Contient des informations telles que : numero_ppa, objet_ppa, montant_ppa, etc.

DIM_TIERS	Renferme les informations sur les tiers à savoir le nom du tiers, la ville du tiers, l'adresse du tiers, le mail du tiers, le téléphone du tiers, etc.
DIM_PREVISIONS_BUDGETAIRES	Contient les informations sur les prévisions budgétaires: prévision annuelle, budget additionnelle, réalisation additionnelle ...
DIM_ETAT_MANDATS	Contient les informations sur les états des mandats.
DIM_MODE_PAIEMENT	Contient les informations sur les modes de paiement.
DIM_BANQUES	Renferme des informations sur les banques comme le nom, etc.

TABLE 3.11 - Ordres_Depenses

Chapitre 4 : Implémentation de la solution

1. Solution proposée

Dans le projet global de mise en place du SID, pour les premières phases d'intégration des données les étudiants ont utilisé la suite Microsoft BI pour la mise en place de leurs solutions (intégration, analyse et reporting). Cette année, en plus de la suite Microsoft BI, il nous a été soumis l'outil Talend. En effet la DISI disposant des compétences en Talend et dans le besoin d'avoir des environnements diversifiés, voudrait que l'intégration des données des finances se fasse sous Talend.

L'outil Talend nous permettra donc d'extraire, de transformer et de charger les données des finances dans l'entrepôt existant et Microsoft BI, nous permettra de faire l'analyse et le reporting des différentes activités des finances.

La solution devrait permettre d'intégrer les données des finances, qui jusqu'à nos jours ne sont pas incluses dans l'entrepôt de données. Ainsi les décideurs seront en mesure d'avoir une vision transversale sur les données des finances, des RH et de la scolarité. Ce qui permettra d'avoir une vue unifiée des activités de l'UCAD.

2. Présentation des outils de développement

2.1 Talend Open Studio

C'est un ETL open source développé par la société française Talend. Il permet de créer graphiquement des processus de manipulation et de transformation de données puis de générer l'exécutable correspondant sous forme de programme Java ou Perl. Une longue liste de composants permet de se connecter à tout type de base de données ou d'applications. Talend a mis en place deux produits d'intégration de données :

- Talend Open Studio for Data Integration, outil de développement gratuit ;
- Talend Enterprise Data Integration, celui-ci intègre des fonctionnalités avancées de déploiement et de gestion distribuées sous licence commerciale. Source open-source-guide

Talend Open Studio est une solution orientée métadonnées, et de ce fait, peut vous aider à garantir l'homogénéité et la qualité des données traitées par votre job design. Sauvegardez vos métadonnées dans le référentiel pour pouvoir les réutiliser pour un autre composant. Ces métadonnées peuvent être : des connexions aux bases de données, des chemins d'accès et des schémas.

2.2 Suite Microsoft BI

SQL SERVER

C'est un SGBD (système de gestion de bases de données). Il peut être utilisé pour traiter des transactions, stocker et analyser des données. SQL Server est une suite de technologie qui permet d'une part le traitement de type OLTP (Online Transaction Processing) pour les données transactionnelles. D'autre part, SQL Serveur contient des modules qui permettent de faire un traitement de type OLAP (Online Analytical Processing) qui est un processus analytique permettant d'extraire et de visualiser de manière sélective des jeux de données.

SQL Server Management Studio (SSMS)

Il s'agit d'un environnement intégré pour la gestion de toute infrastructure SQL. SSMS est utilisé pour accéder, configurer, gérer, administrer et développer tous les composants de SQL Server. SSMS fournit un seul utilitaire complet qui combine un large groupe d'outils graphiques avec un certain nombre d'éditeurs de scripts riches pour fournir un accès à SQL Server aux développeurs et aux administrateurs de base de données de tous les niveaux de compétence.

SQL Server Integration Services (SSIS) :

SSIS un outil utilisé pour effectuer des opérations de processus d'ETL. À un niveau élevé, SSIS offre la possibilité de :

- récupérer des données à partir de n'importe quelle source ;
- effectuer diverses transformations sur les données par exemple : convertir d'un type en un autre, convertir en majuscules ou en minuscules, effectuer des calculs, etc.
- charger des données dans à peu près n'importe quelle source ;
- définir un flux de travail.

SQL Server Analysis Services (SSAS)

SSAS est une technologie de Microsoft Business Intelligence. Elle est utilisée pour créer des cubes à l'aide de données provenant de datamarts/ datawarehouse pour une analyse plus approfondie et plus rapide des données.

Microsoft SQL Server Reporting Services (SSRS) :

SSRS est une plate-forme de reporting prenant en charge les rapports traditionnels et interactifs diffusés sur le Web ou via des applications personnalisées. Il prend en charge diverses sources de données bidimensionnelles comme multidimensionnelles. Voici quelques caractéristiques SSRS :

- Récupérer des données de différentes sources ;
- Exporter les rapports en différent de formats ;

- Afficher les rapports de différentes manières : sous forme de tableaux, de graphes, etc.

3. Mise en Œuvre

3.1 Alimentation de l'entrepôt de données

3.1.1 Extraction

1. Création de connexion :

Pour accéder aux tables ou vues de la base de données XAALIS des finances, nous avons créé une connexion vers la cette base de données nommée SRC_FINANCE. Pour les connexions aux bases de données SRC_RADIS, SAAS_RADIS et l'entrepôt données RADIS elles existent déjà ; elles sont nommées respectivement :

- CONNEXION_SRC_RADIS,
- CONNEXION_SAAS_RADIS
- CONNEXION_RADIS

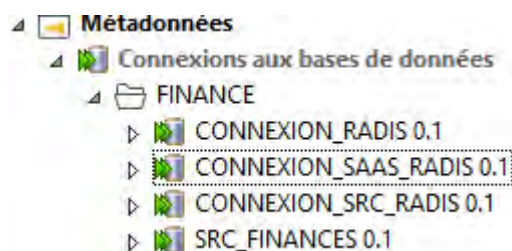


FIGURE 4.1- Connexion aux bases de données

2. Récupération des données :

Pour récupérer les données des finances nous utilisons la connexion SRC_FINANCE afin de pouvoir récupérer les schémas des synonymes car pour éviter de travailler directement sur les tables des synonymes ont été créés sur ces dernières.

Les images ci-dessous illustrent nos propos.

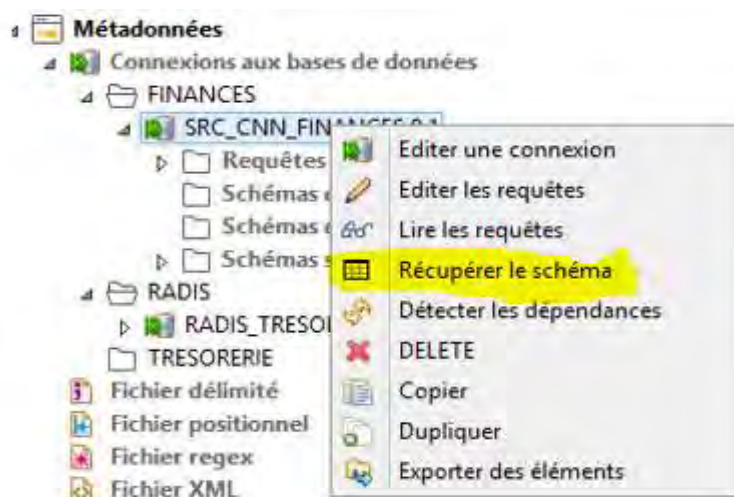


FIGURE 4.2- Récupérer le schéma des finances (1)

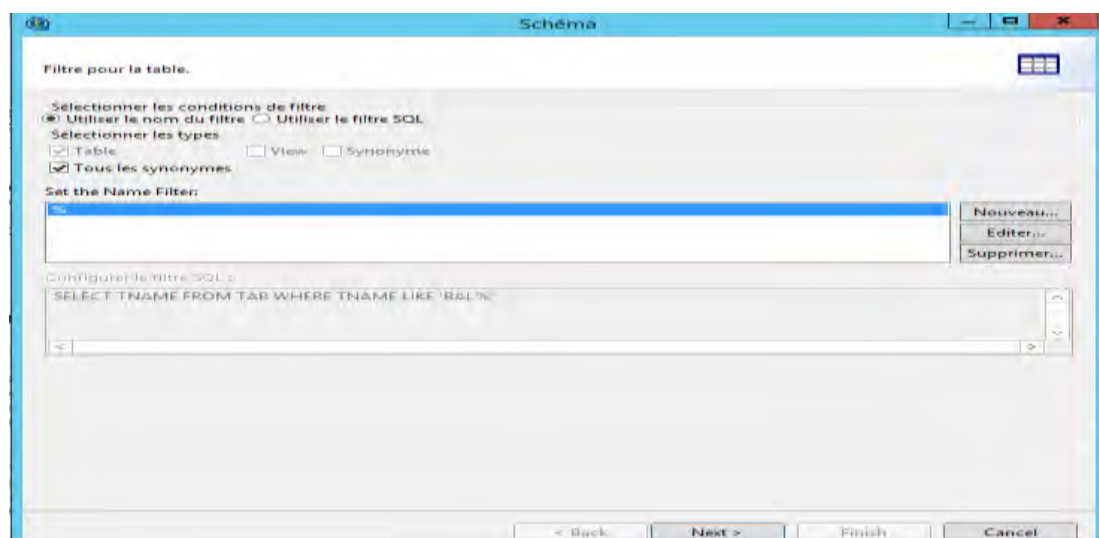


FIGURE 4.3-Récupérer le schéma des finances (2)

Ensuite nous avons sélectionné les tables des finances d'ordre budgétaire et de trésorerie.

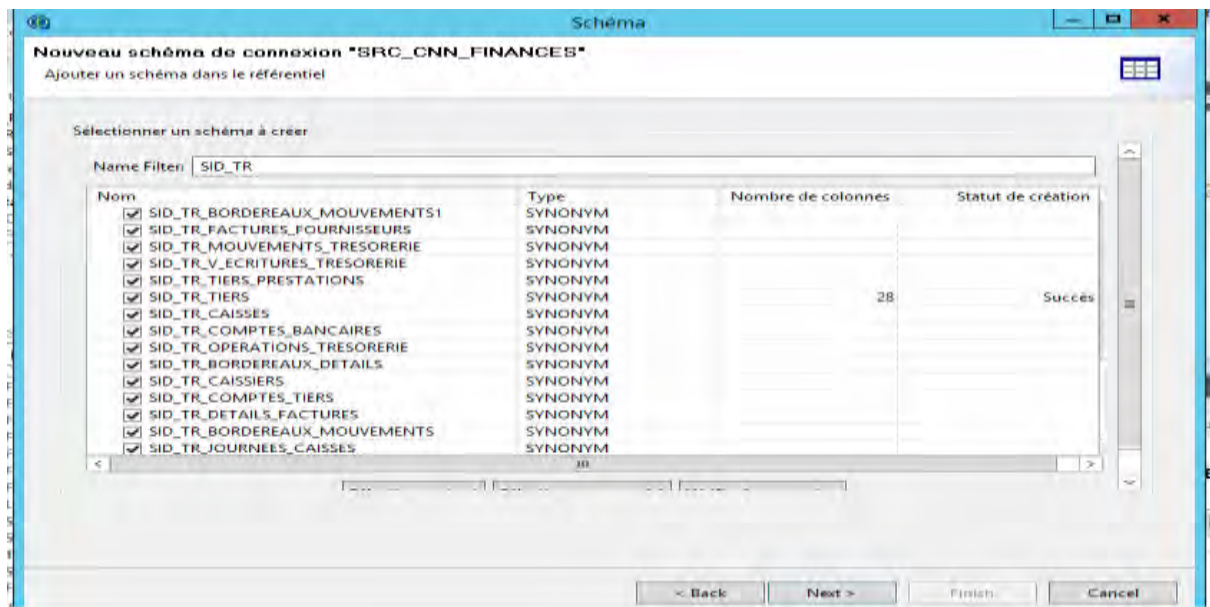
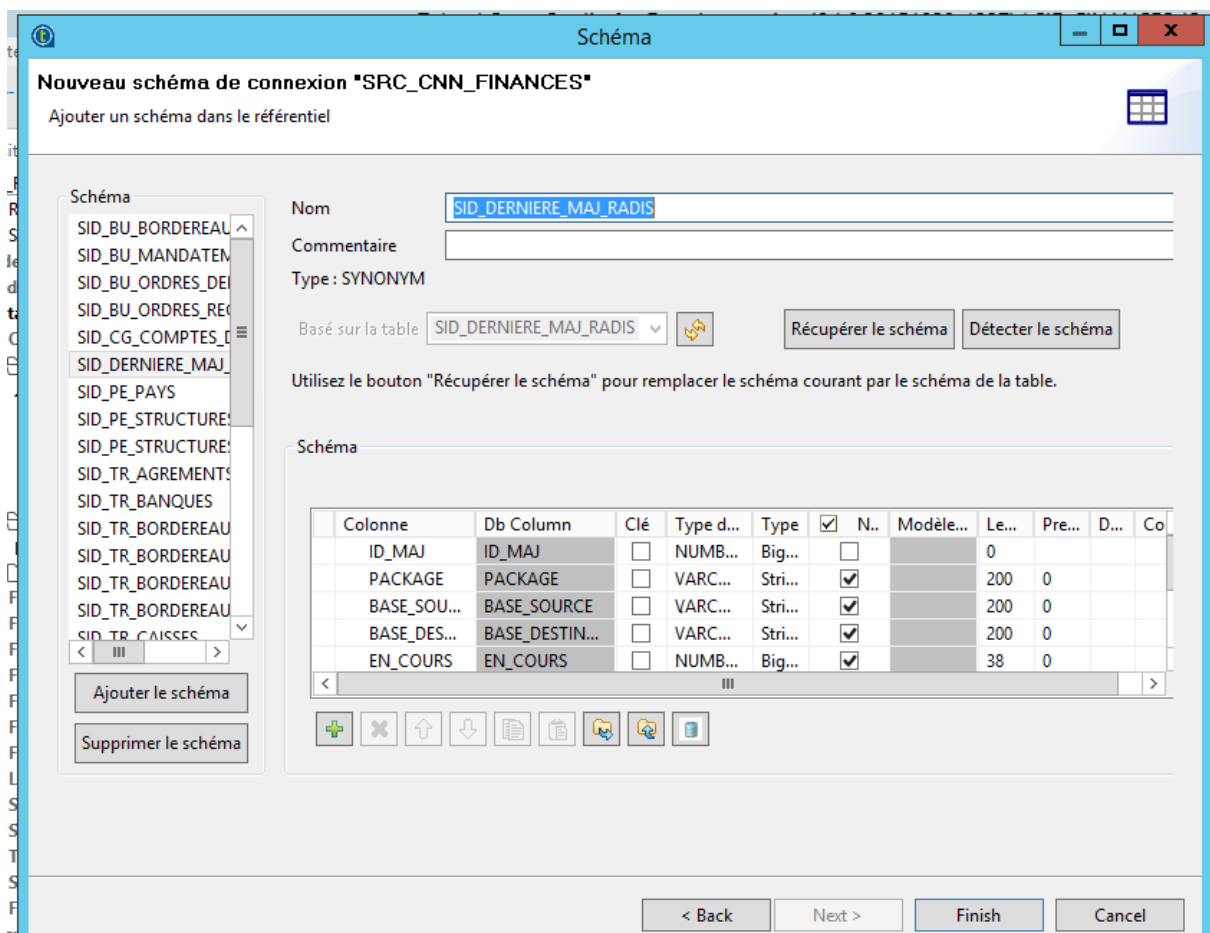
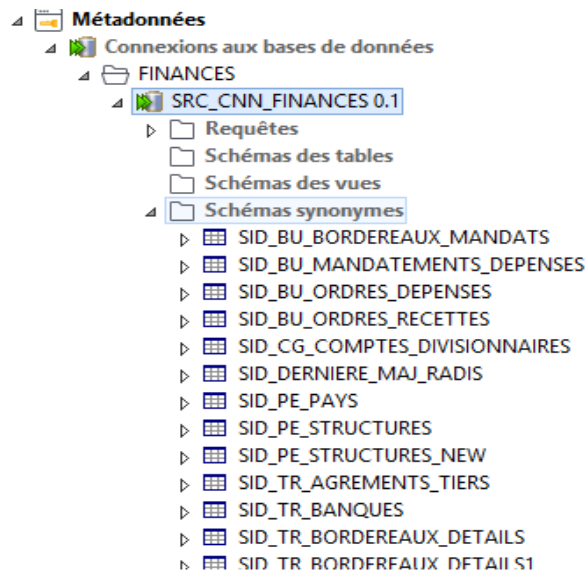


FIGURE 4.4-Récupérer le schéma des finances (3)

Enfin, nous obtenons toutes les tables sélectionnées avec leurs différentes caractéristiques.



FIGURES 4.5-Récupérer le schéma des finances (4)



FIGURES 4.6 - Schéma des synonymes des finances

Après avoir récupérer les données, on crée un job dans lequel, nous utilisons les composants :

- **tOracleInput** : qui permet de lire une base de données et d'extraire des champs à l'aide de requêtes ;
- **tMap** : il transforme et à partir d'une ou plusieurs sources vers une ou plusieurs destinations ;
- **tOracleOutput** : il exécute l'action définie sur la table et / ou sur les données d'une table, en fonction du flux entrant provenant du composant précédent.
- **tOracleRow** : il permet d'écrire rapidement et facilement des requêtes SQL,

et définissons des points de pause des connexions trigger OnSubjobOK et OnSubjobError, afin que l'exécution de notre job puisse être reprise, en cas d'échec de l'exécution du job, à partir du point de pause précédent l'erreur.

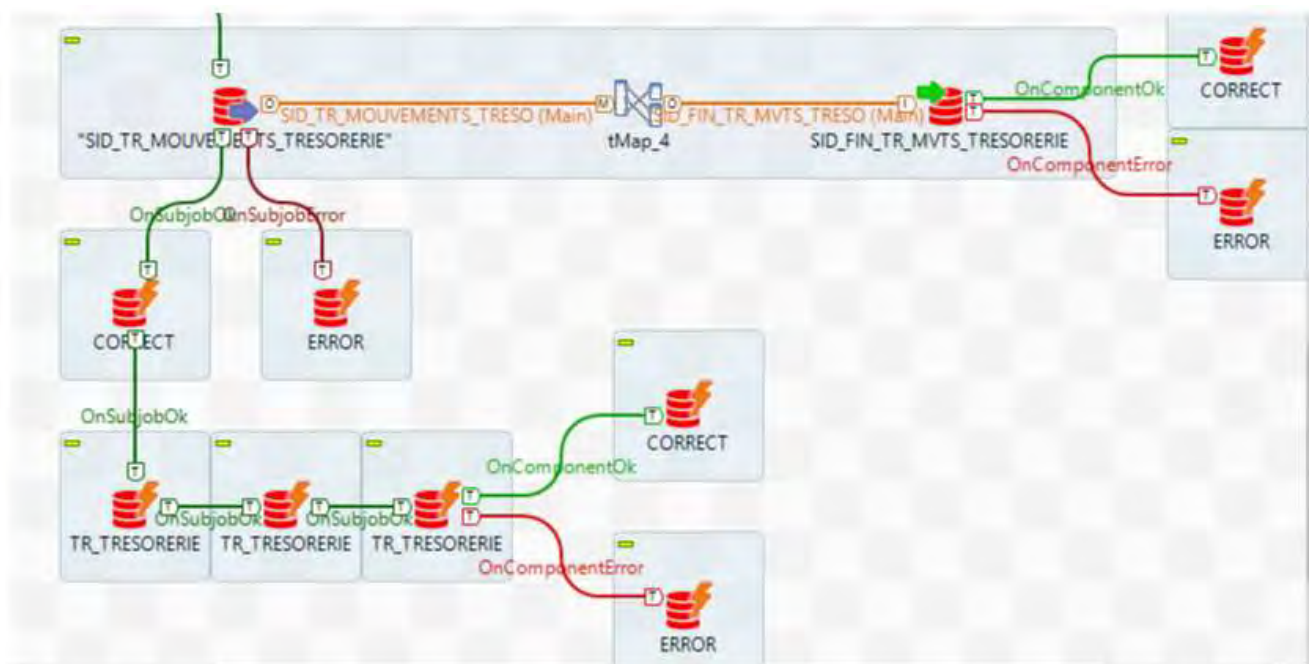


FIGURE 4.7- Collecte des données (1)

SID_TR_MOUVEMENTS_TRESO										SID_FIN_TR_MVTS_TRESO									
Colonne										Expression									
NUMERO_MOUEM...	<input checked="" type="checkbox"/>	String	<input type="checkbox"/>		15	0				SID_TR_MOUVEMENTS_TRESO.NUMER...	<input checked="" type="checkbox"/>	String	<input type="checkbox"/>		15	0			
TYPE_OPERATION	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		1	0				SID_TR_MOUVEMENTS_TRESO.TYPE_O...	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		1	0			
ETENDUE_MOUEM...	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		1	0				SID_TR_MOUVEMENTS_TRESO.ETENDU...	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		1	0			
REFERENCES_TITRE	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		20	0				SID_TR_MOUVEMENTS_TRESO.REFERE...	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		20	0			

FIGURES 4.8- Collecte des données (2)

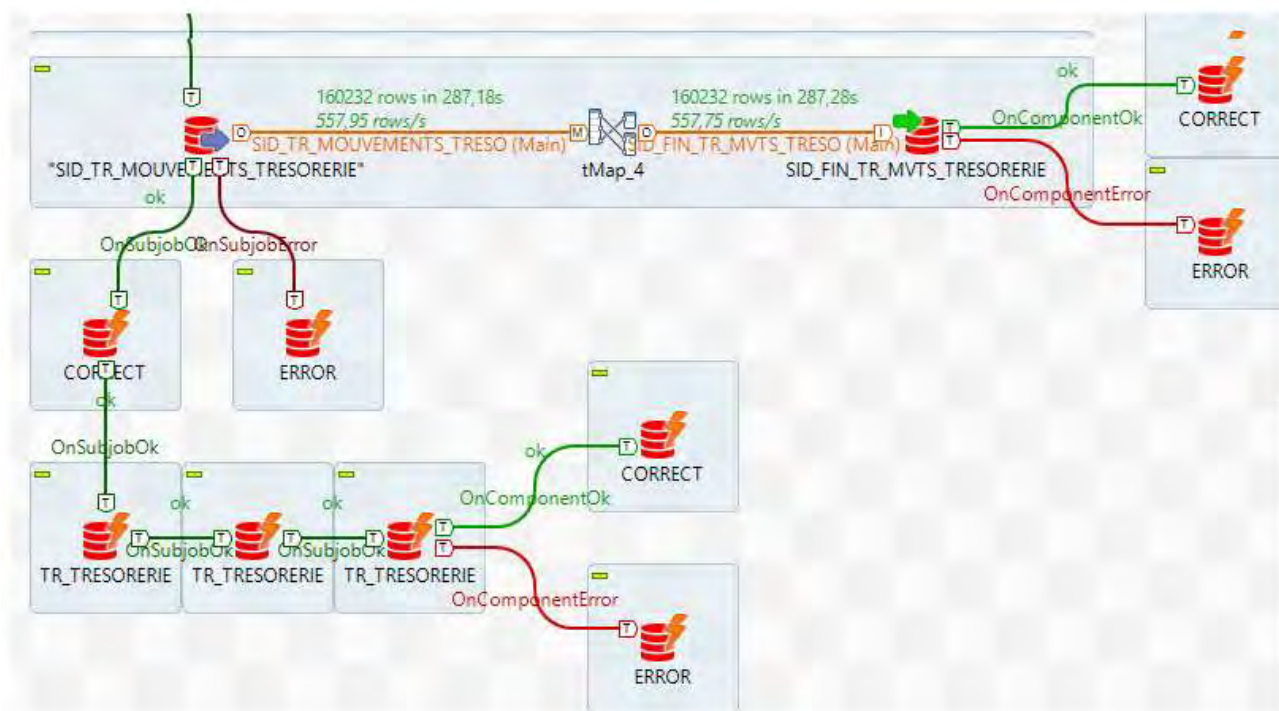


FIGURE 4.9- Collecte des données (3)

Ici nous collectons les données qui se trouvent dans la table mouvements trésorerie des finances tout en gérant l'historisation des données grâce à une table appelée SID_DERNIERE_MAJ_RADIS. Cette table permet de charger les données qui ont évolué ou qui ont été ajoutées récemment dans la base des finances.

SID_DERNIERE_MAJ_RADIS contient les champs suivants :

- Package : qui correspond au nom du job contenant les différentes tâches ;
- DATE_DERNIERE_MAJ_DEB : qui correspond à la date de début de MAJ
- DATE_DERNIERE_MAJ_FIN : qui correspond à la date de fin de MAJ
- EN_COURS : qui prend deux valeurs 0 et 1. 0 veut dire que les mises à jour ont été effectuées pour le package correspondant. Et 1 signifie que cette ligne de la table de MAJ est à mettre à jour.

3.1.2 Transformation

Après avoir récupéré l'ensemble des données nécessaires à la mise en place de l'entrepôt de données dans la base de données SRC_RADIS, on a la possibilité de récupérer le schéma des tables de cette dernière comme précédemment dans le cas de l'extraction. A ce stade nous allons passer maintenant à l'étape de nettoyage des données.

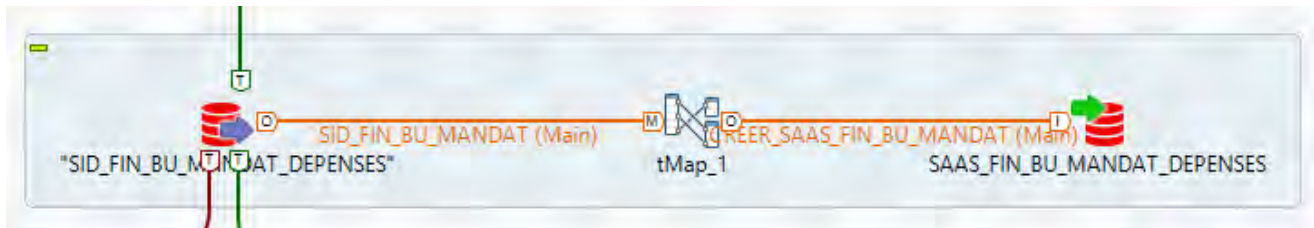
Étant donné que nous devons nettoyer plusieurs tables, nous allons prendre la table SID_BU_MANDATEMENTS_DEPENSES en guise d'exemple pour montrer la technique généralement utilisée pour le nettoyage des données.

Pour cela nous n'avons qu'à faire un glissé-déposé de la table en question en choisissant le composant tOracleInput pour nous connecter à la base de données SRC_RADIS afin de récupérer les données de cette table.

Ensuite nous utilisons un tOracleOutput où nous indiquons les actions à effectuer sur la table SAAS_BU_MANDATEMENTS_DEPENSES et sur les données.

Enfin, nous utilisons un tMap pour rediriger le flux d'entrée vers le flux de sortie.

L'image suivante illustre nos propos :



FIGURES 4.10 - nettoyage des données (1)

Une fois les connexions établies entre les composants, nous procédons au nettoyage des données. La méthode que nous adoptons est la suivante.

D'abord les données sont récupérées telles qu'elles sont stockées dans la table sources SID_BU_MANDATEMENTS_DEPENSES. Rappelons que les champs de cette table sont non typés. Cela permet de tout récupérer les données sans tenir compte des anomalies.

Ensuite nous créons une nouvelle table SAAS_BU_MANDATEMENTS_DEPENSES avec des champs aux formats souhaités. Par exemple, pour un champ de type chaîne de caractères représentant une date dans la table source SID_BU_MANDATEMENTS_DEPENSES, nous lui donnons son vrai type date dans la table SAAS_BU_MANDATEMENTS_DEPENSES comme l'illustre l'image suivante :

Commentaire

Type : TABLE

Basé sur la table SID_FIN_BU_MANDAT_DEP

Récupérer le schéma Détection le schéma

Utilisez le bouton "Récupérer le schéma" pour remplacer le schéma courant par le schéma de la table.

Schéma

Colonne	Db Column	Clé	Type de base ...	Type	<input checked="" type="checkbox"/>	N..	Modèle date (Ctrl...	Length	Precision
NUMERO_MANDAT	NUMERO_MANDAT	<input checked="" type="checkbox"/>	VARCHAR2	String	<input type="checkbox"/>			15	0
MODE_PAIEMENT	MODE_PAIEMENT	<input type="checkbox"/>	VARCHAR2	String	<input type="checkbox"/>			1	0
DATE_DEMANDE	DATE_DEMANDE	<input type="checkbox"/>	VARCHAR	Date	<input type="checkbox"/>		"dd-MM-yyyy"	7	0
DEMANDEUR	DEMANDEUR	<input type="checkbox"/>	VARCHAR2	String	<input checked="" type="checkbox"/>			30	0
MONTANT_MANDAT	MONTANT_MANDAT	<input type="checkbox"/>	DEC	BigDecimal	<input type="checkbox"/>			38	0
MONTANT_PRECOMPTE	MONTANT_PRECOMPTE	<input type="checkbox"/>	DEC	BigDecimal	<input checked="" type="checkbox"/>			38	0
DATE_ORDRE	DATE_ORDRE	<input checked="" type="checkbox"/>	DATE	Date	<input checked="" type="checkbox"/>		"dd-MM-yyyy"	7	0
AVIS_ORDONNATEUR	AVIS_ORDONNATEUR	<input type="checkbox"/>	VARCHAR2	String	<input type="checkbox"/>			1	0
REJET_ORDONNATEUR	REJET_ORDONNATEUR	<input type="checkbox"/>	VARCHAR2	String	<input checked="" type="checkbox"/>			50	0

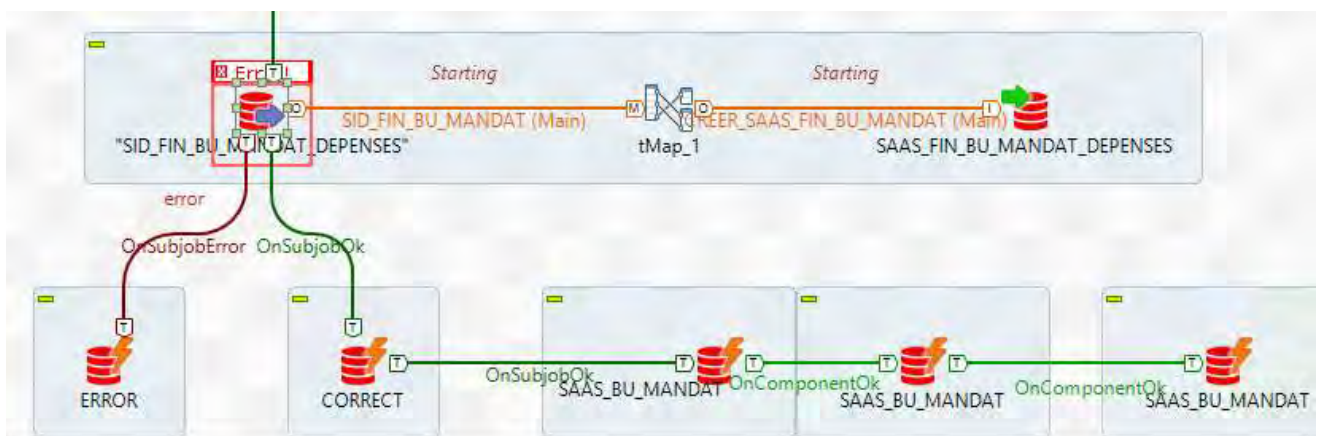
< III

< Back Next > Fini

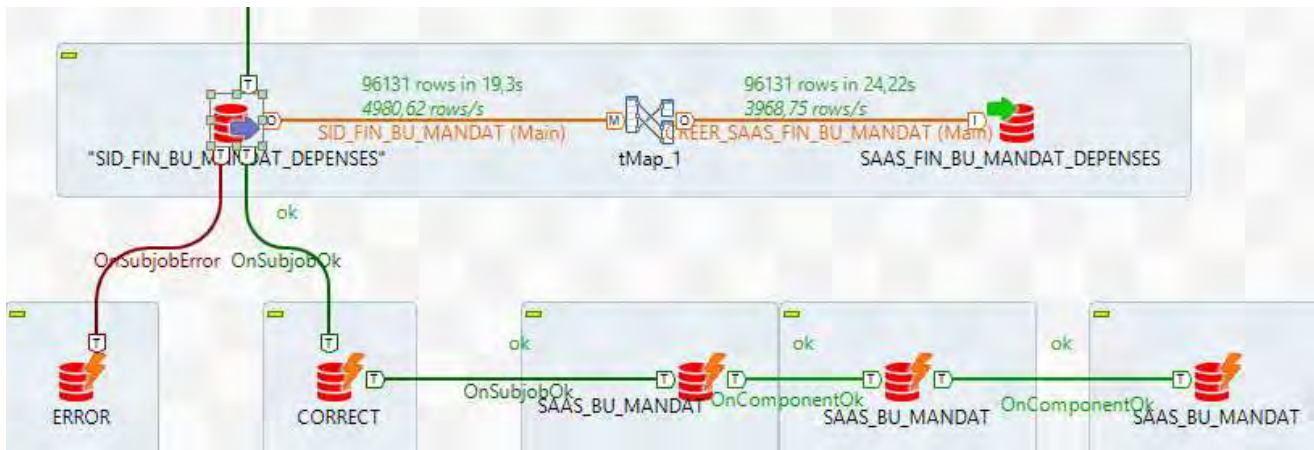
FIGURES 4.11 - nettoyage des données (2)

Enfin nous essayons de transférer les données de la table source vers la table saas. Lors du transfert si une donnée de la table source ne respecte pas le format de la table de destination, elle est redirigée dans la table de rejet ERREUR_EXECUTION.

Les images suivantes illustrent nos propos :



FIGURES 4.12 - nettoyage des données (3)



FIGURES 4.13 - nettoyage des données (4)

Comme dans le cas de la collecte des données, une table de MAJ a été utilisée pour transférer les données qui ont évolué ou qui ont été récemment chargées dans la base SRC_RADIS.

3.1.3 Chargement

Étant la dernière étape du processus l'ETL, il s'agit ici de charger les données déjà nettoyées dans l'entrepôt de données. Tout d'abord, nous créons la structure des tables de dimension et de la table de fait

1. Création de la structure des tables de dimension : Les dimensions ont à peu près la même structure que les tables de la base de données contenant les données propres (SAAS_RADIS). À la différence nous ajoutons sur chaque table de dimension, quatre colonnes. La première colonne va constituer la clé primaire de la table. Elle sera de type entier et auto-incrémente. Pour les trois autres colonnes, l'une est identifiant de la date de début d'insertion, l'autre est identifiant de la date de fin d'insertion qui ont permis d'historiser les dimensions et le troisième indique les dimensions qui sont actives ou non.

2. Chargement des dimensions : Comme nous l'avons expliqué dans le chapitre précédent Il existe plusieurs techniques pour charger les données dans les dimensions. Dans notre cas, nous avons adopté le chargement avec historisation en utilisant les SCD (Slowly Change Dimensions) à travers le composant tOracleSCD de Talend. En effet ce composant répond aux besoins de dimension à évolution lente, en lisant régulièrement une source de données et en enregistrant les modifications dans une table SCD dédiée.

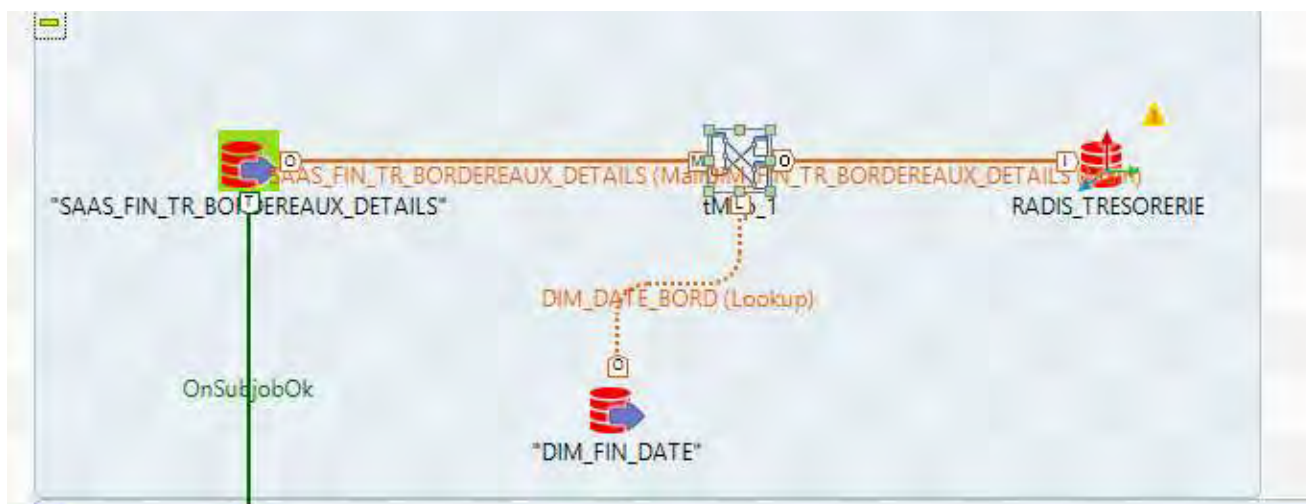
Le composant tOracleSCD comporte quatre types d'historisation :

- Type 0 - Ne rien faire ; les enregistrements de dimension ne sont pas mis à jour.
- Type 1 - Aucun historique n'est conservé. Les enregistrements de dimension sont mis à jour sur place.
- Type 2 - Un historique complet est conservé. Les enregistrements de dimension ont des dates de début et de fin pour déterminer leur date de validité.

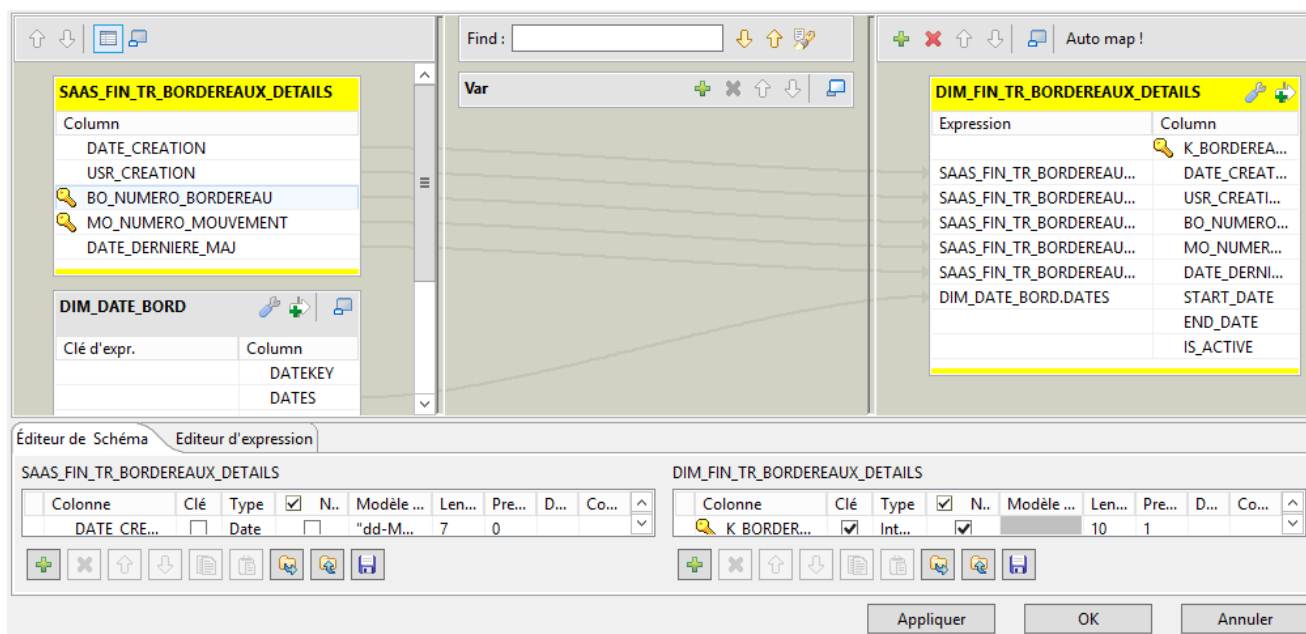
- Type 3 - Un historique limité est conservé. Généralement, l'historique le plus récent d'une colonne ou de deux est conservé dans la même ligne que les informations actuelles.
- Type 4 - Un historique complet est conservé dans une table d'historique ; la table de dimension a des enregistrements en cours.

Pour notre cas, nous avons utilisé le type 2 afin de gérer l'historisation des données de nos dimensions.

Les images illustrent la procédure de chargement des dimensions :



FIGURES 4.14 – Chargement des Dimensions (1)



FIGURES 4.15 – Chargement des Dimensions (2)

Filtre

Inutilisé

END_DATE

IS_ACTIVE

K_BORDEREUX_DETAILS

START_DATE

Clés source

BO_NUMERO_BORDEREAU

Clé de substitution

name

K_BORDEREUX_DETAILS

creation

Table max + 1

complement

Champs de type 0

Champs de type 1

Champs de type 2

DATE_CREATION

DATE_DERNIERE_MAJ

USR_CREATION

MO_NUMERO_MOUVEMENT

Versioning

type	name	creation	comple
start	START_DATE	Job start time	
end	END_DATE	Fixed year value	9999
version	scd_version		
active	IS_ACTIVE		

Champs de type 3

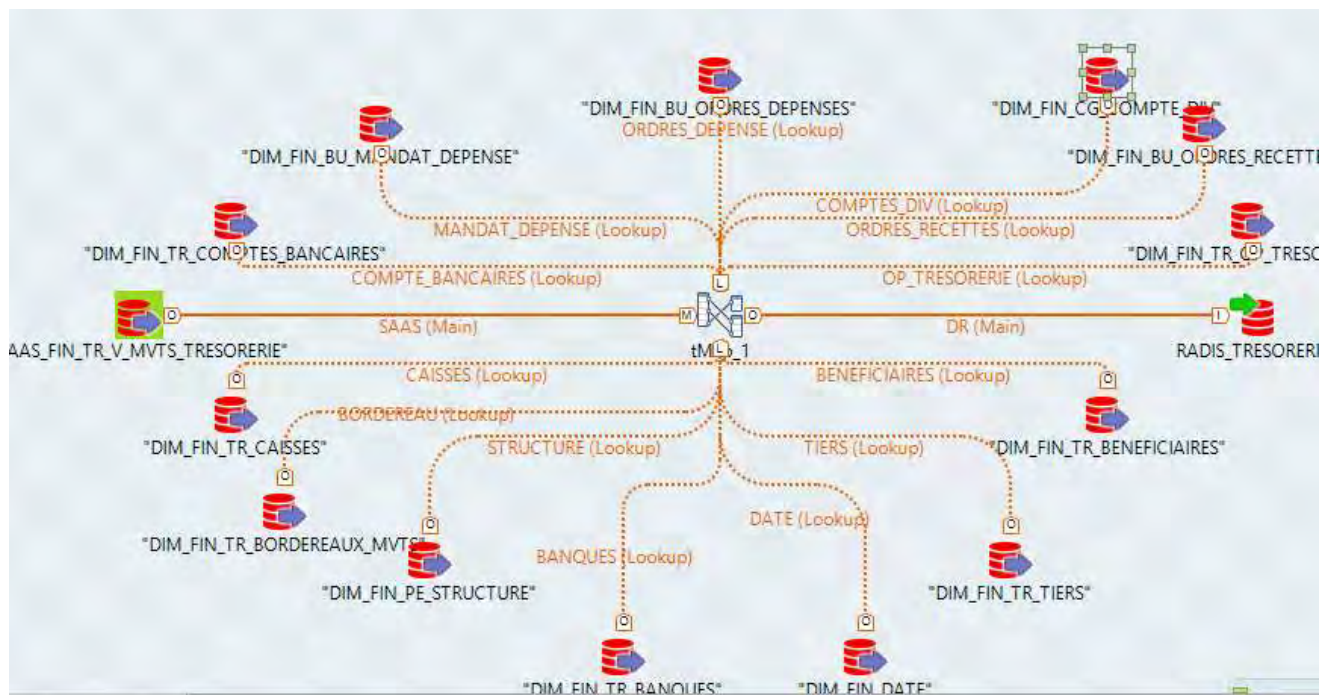
current value	previous value

OK

Cancel

FIGURES 4.16 – Chargement des Dimensions (3)

3. Chargement des tables de faits : Les tables de faits vont contenir l'ensemble des clés techniques des tables de dimension. Ainsi pour charger chaque table, nous faisons une série de jointures de gauche entre la table de fait et les dimensions



FIGURES 4.17 – Alimentation du DWH

Sur cette image la table de fait est alimentée avec les dimensions MAJ récemment.

3.1.4 Description des flux d'alimentation

Le Datawarehouse n'est pas statique, il est appelé à se mettre à jour afin de prendre en compte les nouvelles données qui sont ajoutées dans les sources depuis le dernier chargement. Donc il convient maintenant de décrire les fréquences de mise à jour du DataWarehouse mais aussi le mode d'alimentation qui est adopté. Donc, nous allons décrire le mode d'alimentation adopté.

1. Mode d'alimentation du DataWarehouse

Le mode d'alimentation désigne la manière dont le Datawarehouse est mis à jour. En effet il existe plusieurs façons de procéder pour la mise à jour du Datawarehouse comme nous l'avons tantôt expliqué. Dans le cas de ce projet nous avons utilisé la mise à jour avec historisation. C'est-à-dire que chaque nouvel enregistrement est d'abord vérifié dans la table de dimension destinatrice. S'il existe sa colonne IS_ACTIVE est mise à 0 et il est à nouveau ajouté avec les nouvelles valeurs de ses champs tout en mettant le IS_ACTIVE à 1 sinon il est tout simplement ajouté.

3.2 Mise en place du Cube de données

Nous avons créé deux cubes de données pour les finances : un pour la gestion de la trésorerie et un pour la gestion budgétaire

Voici les dimensions qui interviennent dans les cubes.

Domaine d'application	Sujet d'analyse	Cube	But
FINANCES	suivi des mouvements de trésorerie et des mandatements de dépenses budgétaire	CUBE_TRESORERIE CUBE_BUDGET	Produire des KPI qui permettent d'avoir une vue claire et simplifiée sur la gestion financière de l'UCAD.

TABLE 4.1– Cubes des Finances

Nom de la dimension
DIM_STRUCTURES
DIM_RUBRIQUE_BUDGET
DIM_EXERCICES_BUDGETAIRES
DIM_ORDRES_DEPENSES
DIM_TIERS
DIM_PREVISIONS_BUDGETAIRES
DIM_ETAT_MANDATS
DIM_MODE_PAIEMENT
DIM_BANQUES

TABLE 4.2– Dimensions du cube Budget

NOM DE DIMENSION
DIM_BU_ORDRES_RECETTES
DIM_BU_ORDRES_DEPENSES
DIM_BU_MANDATEMENT_DEPENSES

DIM_BU_CG_COMPTES_DIVISIONNAIRES
DIM_BU_PE_STRUCTURES
DIM_TR_TIERS
DIM_TR_BANQUES
DIM_TR_COMPTES_BANCAIRES
DIM_TR_OPERATION_TRESORERIE
DIM_TR_CAISSE
DIM_TR_BENEFICIAIRES
DIM_TR_BORDEREAUX_MOUVEMENTS
DIM_DATE

TABLE 4.3 – Dimensions du cube trésorerie

La hiérarchie détectée est : DIM_DATE (ANNEE, MOIS, JOUR)

Nous avons mise en place des KPI.

Un indicateur de performance clé (KPI), ou simplement indicateur, est un terme pour une mesure qui évalue la performance par rapport à un objectif utilisés par les organisations pour mesurer à la fois le succès et la qualité dans la réalisation des objectifs stratégiques, la mise en place de processus ou la fourniture de produits / services.

Par exemple les indicateurs “NOMINAL_DEPENSE “ et “NOMINAL_RECETTE“ peuvent être des indicateurs appropriés pour l’évaluation de la qualité des mouvements de trésorerie dans une structure et TOTAL_DBM permet de calculer le DBM total pour un mandat.

Les mesures des finances sont des éléments de données qui contribuent aux analyses fournissant de l’information pour aider les décideurs dans les organisations à prendre

de meilleures décisions. Dans le cadre de ce travail, l'étude des données obtenues à partir de nos sources a permis faire ressortir un certain nombre de mesure :

- NOMINAL_DEPENSE et NOMINAL_DEPENSE pour le cube trésorerie.
- TOTAL_DBM pour le cube budget

3.3 Spécification détaillée des restitutions

La restitution est la dernière phase technique du projet, elle va permettre de mettre à la disposition des utilisateurs les données du cube et du datawarehouse sous forme de rapports. Ainsi dans cette section nous détaillerons d'abord les méthodes d'accès aux différents rapports puis revenir sur les méthodes d'exploitation des rapports à savoir les critères de sélections

3.3.1 Paramètres

Les paramètres permettent de filtrer des données, de faire varier la présentation et d'établir des liens avec des rapports associés. Cependant, sur chaque rapport nous avons mis des paramètres pour permettre à l'utilisateur d'exploiter au maximum le rapport.

La figure suivante montre l'utilisation des paramètres dans les rapports.

EXERCICE	ENTITE	BANQUE TRANSACTION	NOMINAL DEPENSE	NOMINAL RECET
2002	ACU	BCEAO		
			10000000	
			8583816	
			10500000	
			20000000	
			0	
			0	
			0	
			0	

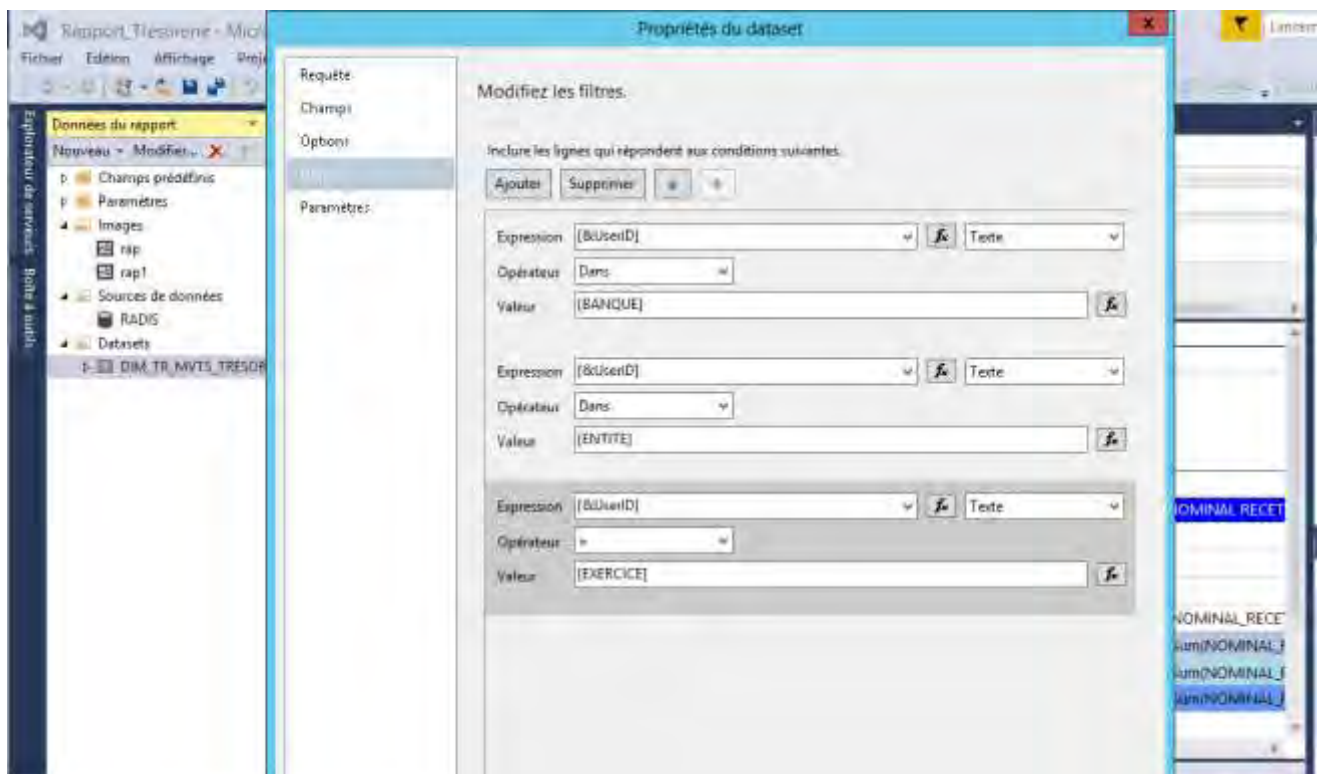
FIGURES 4.18 – Utilisation des paramètres

L'image ci-dessus présente le rapport de la situation d'exécution de trésorerie. Cependant un paramètre a été utilisés pour permettre à l'utilisateur de personnaliser son rendu. Par exemple pour voir la situation d'exécution de trésorerie, l'utilisateur

peut choisir la banque qui lui convient Ainsi les résultats sont en fonction de ce paramètre choisi.

3.3.2 Filtres

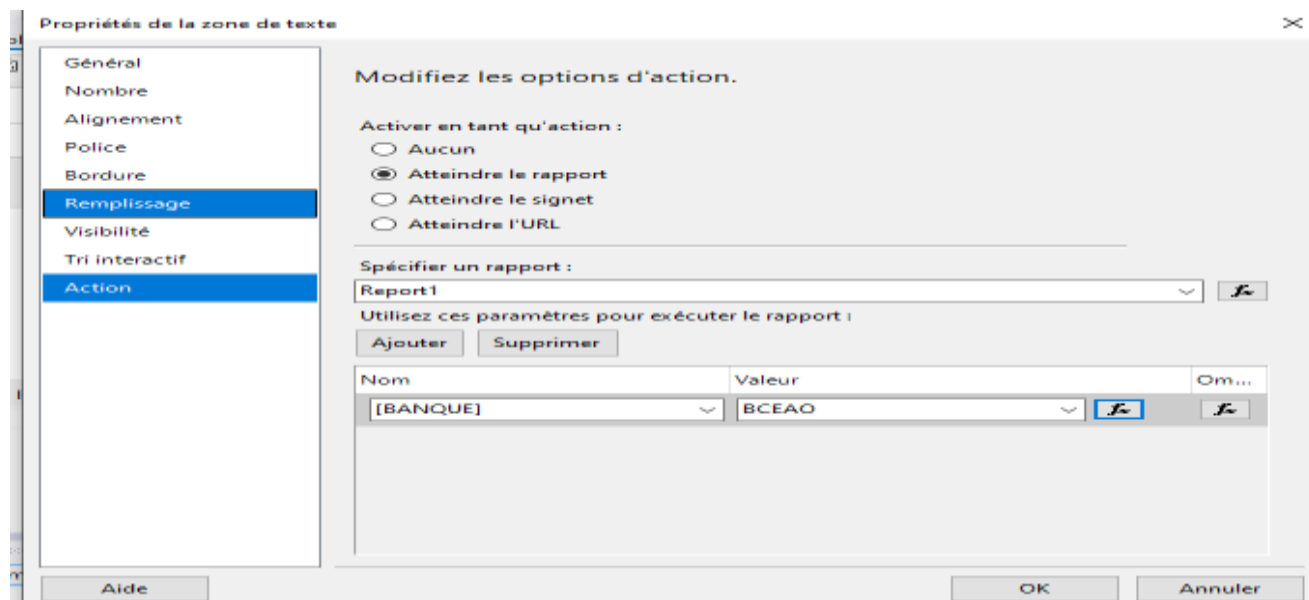
Comme son nom l'indique, un filtre permet de filtrer le résultat d'une requête selon un paramètre donné. Cependant nous avons utilisé le filtre pour mettre en place la gestion des confidentialités sur les rapports c'est-à-dire, nous assurons que l'information est lue par utilisateur ayant les droits requis. Ainsi, nous filtrons chaque résultat en fonction des droits de l'utilisateur connecté. La figure suivante montre comment est mis en place ce filtre.



FIGURES 4.19 – Utilisation des filtres

3.3.3 Navigation inter rapport

Afin de faciliter l'exploitation des rapports, un système de navigation inter rapport est mis en place. Ainsi, à partir d'un rapport principal, l'utilisateur est capable d'atteindre n'importe quel autre rapport. Le schéma qui suit illustre les faits.



FIGURES 4.20 – Navigation inter rapport

Chapitre 5 : Présentation et exploitation des résultats

À présent, nous présentons les résultats de notre solution afin d'évaluer sa conformité avec les besoins recensés au début. Cependant, l'exploitation des données s'est faite via plusieurs technologies. Nous avons utilisé SSRS (SQL Server Reporting Services) pour nous connecter à la base multidimensionnelle et mettre en place des rapports que les autorités ont le plus besoins. Étant donné que

Dans la mesure où tous les rapports ne peuvent pas se faire au préalable, nous utilisons Microsoft Excel avec les tableaux croisés dynamiques pour permettre d'avoir d'autres rapports qui répondraient aux futurs besoins.

1. Rapports

Plusieurs rapports ont été mis en place en réponse aux besoins en termes de prise de décision ou pour le suivi de l'évolution des activités de l'UCAD. Parmi ces rapports nous avons :

- Situation d'exécution de trésorerie des finances à l'ucad

Ce premier rapport donne la situation d'exécution de trésorerie par année « **EXERCICE** », par « **STRUCTURE** » par banque « **BANQUE DE TRANSACTION** » en fonction des dépenses et des recettes.

Université Cheikh Anta Diop



Situation d'exécution de trésorerie par banque				
EXERCICE	ENTITE	BANQUE TRANSACTION	NOMINAL DEPENSE	NOMINAL RECETTE
2002	ACU	BCEAO		
			10000000	0
			8583816	0
			10500000	0
			20000000	0
			0	6000000
			0	100000000
			0	310000
			0	7165000
			0	644946701
			0	1446200
			0	1500000
			3490	0
			150000000	0
		Total	199087306	761367901
	ESP	BCEAO		

FIGURES 4.21 – Situation d'exécution de trésorerie par banque

L'image ci-dessus présente le rapport de la situation d'exécution de trésorerie par année, par structure et par banque d'une manière plus détaillée. Cependant un paramètre a été

utilisé pour permettre à l'utilisateur de personnaliser son rendu. Par exemple faire une comparaison des situations d'exécution par rapport à plusieurs années universitaires, plusieurs établissements et plusieurs banques etc. il suffit pour cela de jouer sur les paramètres.

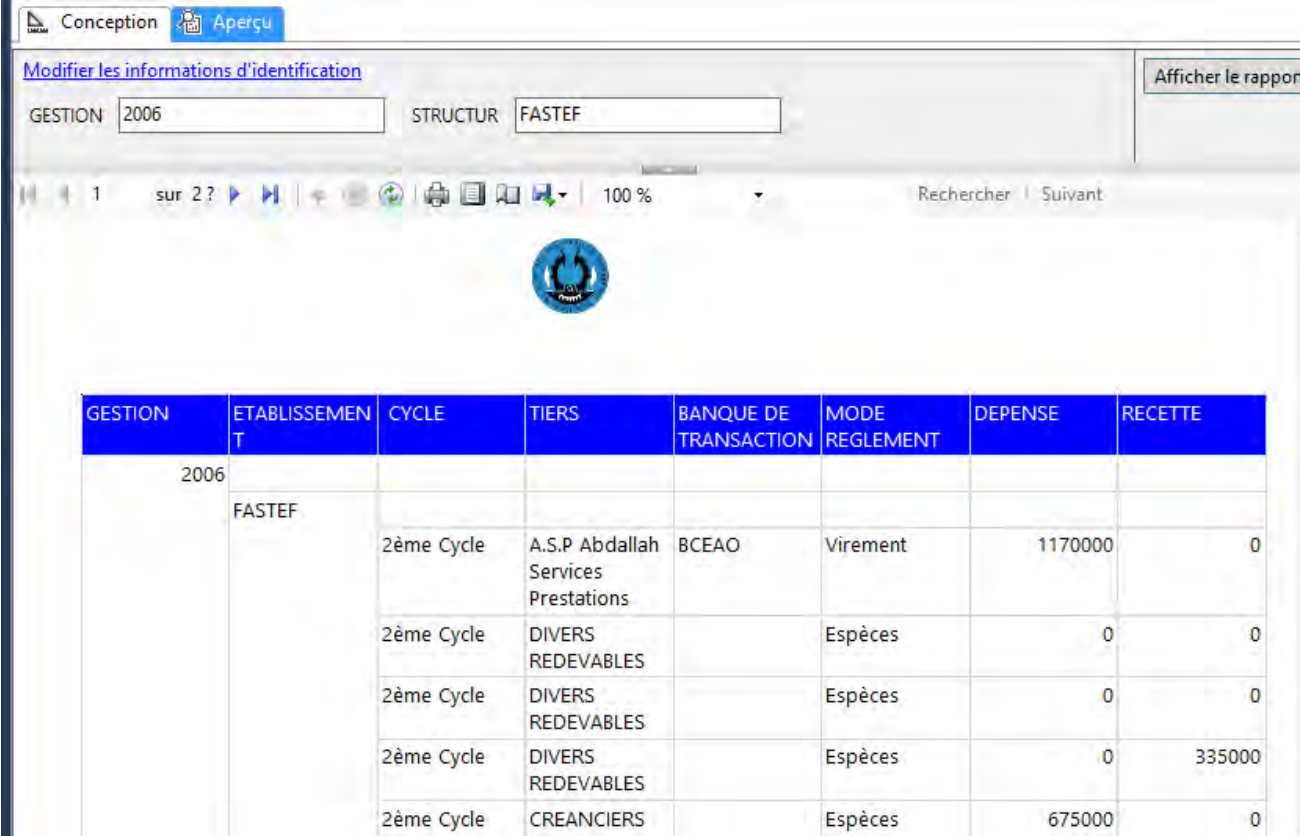
- **Situation d'exécution du budget par rubrique**

NOM RUBRIQUE	EXERCICE	Ets	numeroMandat	objetMandat	numeroEngagem ent	libelleEngagemen t	mod
Achat de cartes de t�l�phone	2019	FLSH	18-2019/000129	achat de cartes de credit	18-2019/000137	achat cartes de credit	Virer
	2019	FLSH	18-2019/000155	achat de carte de credit /communication gidel	18-2019/000161	achat de cartede credit/coomunicat ion master Gidel	Virer
	2019	FLSH	18-2019/000360	ACHAT DE CREDIT ORANGE	18-2019/000364	ACHAT DE CARTE CREDIT ORANGE	Virer

FIGURES 4.22 – Situation d'ex cution du budget par rubrique

L'image ci-dessus pr sente le rapport de la situation d'ex cution du budget par rubrique par rubrique pour l'ann e 2019. Toutefois, nous pouvons avoir la situation d'ex cution du budget selon n'importe quel  tablissement ou ann e, il suffit de faire varier les param tres.

- Répartition du budget par année, par établissement et par cycle



GESTION	ETABLISSEMENT	CYCLE	TIERS	BANQUE DE TRANSACTION	MODE REGLEMENT	DEPENSE	RECETTE
2006	FASTEF						
		2ème Cycle	A.S.P Abdallah Services Prestations	BCEAO	Virement	1170000	0
		2ème Cycle	DIVERS REDEVABLES		Espèces	0	0
		2ème Cycle	DIVERS REDEVABLES		Espèces	0	0
		2ème Cycle	DIVERS REDEVABLES		Espèces	0	335000
		2ème Cycle	CREANCIERS		Espèces	675000	0

FIGURES 4.23 – Répartition du budget par établissement et par année

Sur ce rapport nous avons fait un croisement des données des finances et celles des scolarités pour voir le budget en dépenses et en recettes par année et par établissement donné en fonction du cycle, du tiers qui peut être un fournisseur, de la banque de transaction, du mode de paiement. Nous avons utilisé deux paramètres pour permettre au recteur, le doyen, ou même un responsable de formation d'avoir une vision globale de ce qui se passe sur la répartition du budget au niveau de ce qu'il dirige.

2. Rapports sur Excel

EXERCICE	Ent	numeroMandat	objetDuMand	numeroEngagement	libelleEngagement	modeDePaiement	numeroBordere	BANQUE DE	montantMand	precompteMand	paiementEm Tiers
2019	ESP	21-2019/000001	Règlement	21-2019/000001	Règlement diverses factures	Chèque	21-2019/000001		2 445 000	0	2 445 000 K&S Travel +
2019	ESP	21-2019/000002	Rembourser	21-2019/000002	Remboursement achat billets	Virement	21-2019/000002		454 437	0	454 437 MARI LABAT
2019	ESP	21-2019/000003	Primes de re	21-2019/000004	Primes de rend Ens	Virement	21-2019/000003		1 200 000	0	1 200 000 DIVERS CREA
2019	ESP	21-2019/000004	Vac Ens Vac	21-2019/000003	Vac Ens Vac Vir Décembre 2018	Virement	21-2019/000004		6 380 000	0	6 380 000 DIVERS CREA
2019	ESP	21-2019/000005	Entretien &	21-2019/000005	Entretien & sélection 2017/20	Virement	21-2019/000005		8 700 000	0	8 700 000 DIVERS CREA
2019	ESP	21-2019/000006	Participation	21-2019/000006	Participation aux frais d'obsèques	Chèque	21-2019/000006		150 000	0	150 000 Marie Pierre
2019	ESP	21-2019/000007	Participation	21-2019/000007	Participation aux frais d'obsèques	Chèque	21-2019/000007		130 000	0	130 000 Penda Mouh
2019	ESP	21-2019/000008	Vac Ens Perm	21-2019/000008	Vac Ens Perm Vir Décembre 2	Virement	21-2019/000008		7 020 000	0	7 020 000 DIVERS CREA
2019	ESP	21-2019/000009	Gestion de d	21-2019/000009	Gestion de dossiers 2017/2018	Virement	21-2019/000009		1 500 000	0	1 500 000 DIVERS CREA
2019	ESP	21-2019/000010	Vac Ens Bil	21-2019/000010	Vac Ens Bil Décembre 2017	Virement	21-2019/000010		70 000	0	0 DIVERS CREA
2019	ESP	21-2019/000011	Séminaire	21-2019/000011	Séminaire organisé du 18 au 1	Virement	21-2019/000011		1 200 000	0	1 200 000 DIVERS CREA
2019	ESP	21-2019/000012	Vac Ens Vac	21-2019/000012	Vac Ens Vac Bil Décembre 201	Virement	21-2019/000012		550 000	0	0 DIVERS CREA
2019	ESP	21-2019/000013	Prestations	21-2019/000013	Prestations effectuées (partie	Exploies	21-2019/000013		228 000	0	342 000 DIVERS CREA
2019	ESP	21-2019/000014	Vac Ens Vir	21-2019/000014	Vac Ens Vir Décembre 2017	Virement	21-2019/000014		9 530 000	0	9 530 000 DIVERS CREA
2019	ESP	21-2019/000015	Vac Ens Vac	21-2019/000015	Vac Ens Vac Vir Décembre 201	Virement	21-2019/000015		6 180 000	0	6 180 000 DIVERS CREA
2019	ESP	21-2019/000016	Gestion des	21-2019/000016	Gestion des dossiers 2017/201	Exploies	21-2019/000016		98 000	0	0 Ibrahim KA
2019	ESP	21-2019/000017	Tests d'entr	21-2019/000017	Tests d'entrée DST correction	Exploies	21-2019/000017		274 500	0	0 DIVERS CREA
2019	ESP	21-2019/000018	Remboursem	21-2019/000018	Remboursement frais de scolar	Chèque	21-2019/000018		920 000	0	920 000 Adama SENE
2019	ESP	21-2019/000019	Tests d'entr	21-2019/000019	Tests d'entrée DST concors & s	Virement	21-2019/000019		915 000	0	980 000 DIVERS CREA
2019	ESP	21-2019/000020	Remboursem	21-2019/000020	Remboursement frais de scolar	Chèque	21-2019/000020		115 000	0	115 000 Aric Mouss
2019	ESP	21-2019/000021	Remboursem	21-2019/000021	Remboursement frais de scolar	Chèque	21-2019/000021		463 000	0	928 100 Tahar Sabati
2019	ESP	21-2019/000022	Remboursem	21-2019/000022	Remboursement frais de scolar	Chèque	21-2019/000022		180 000	0	180 000 Adam Soule
2019	ESP	21-2019/000023	Remboursem	21-2019/000023	Remboursement frais de scolar	Chèque	21-2019/000023		230 000	0	230 000 Maimouna K
2019	ESP	21-2019/000024	Primes de re	21-2019/000024	Primes de rend 4ème trimestre	Virement	21-2019/000024		7 000 000	0	7 000 000 DIVERS CREA
2019	ESP	21-2019/000025	Indemnités	21-2019/000025	Indemnités allouées membri	Virement	21-2019/000025		1 610 000	0	1 610 000 DIVERS CREA

FIGURES 4.24 – Exemple de rapport sur Excel

Ce rapport traduit l'état d'exécution du budget pour l'année 2019. A partir de ce rapport, les utilisateurs pourront déterminer des reportings en rapport avec leurs besoins. Par exemple, il est possible à partir de là de créer des tableaux dynamiques croisés. L'image suivante illustre un exemple de tableau croisé obtenu à partir du rapport ci-dessus.

Étiquettes de lignes	Somme de total	Virement	Somme de total	Liquidation	Somme de budget	Primitif
ENSETP	457209842		4249246190		3927007622	
ESP	17938300000		4,71198E+11		4,83929E+11	
FASEG	519330000		1,44556E+11		2,42495E+11	
FASTEF	932446000		8777399847		8557790000	
FLSH	240200000		293900000		15760425194	
FMPOS	-542503172		54138838469		83143140293	
FSJP	1672610000		21607643542		20379777352	
FST	169000000		17108313619		17963652560	
IFAN/CAD	695000000		1052817192		1105634202	
INSEPS	0		571884367		617500000	
RECTORAT	2,35138E+11		2,61814E+12		2,49024E+12	
Total général	2,56408E+11		3,35142E+12		3,36811E+12	

FIGURES 4.25 – Total mandat par établissement



FIGURES 4.25 – Total mandat par établissement

Chapitre 6 : Conclusion et perspectives

1. Conclusion

Au terme de ce mémoire, nous rappelons qu'il s'agit d'un projet de fin d'études que nous avons effectué au sein de la Direction de l'Informatique et des Systèmes d'Information de l'UCAD, pendant une durée de 6 mois.

Le travail que nous avons mené jusqu'ici se situe dans le cadre des systèmes d'aide à la décision. Dans un premier temps, nous avons identifié trois problématiques. La première consiste en une préparation des données décisionnelles. La deuxième se centre sur l'exploitation des données multidimensionnelles au travers d'analyses exploratoires. La troisième vise à la présentation des données aux décideurs.

Notre mission consistait donc à étudier, concevoir et mettre en place un Système d'Information Décisionnel pour la gouvernance à l'UCAD. Dans ce cas, notre projet est découpé en trois sections : la première consiste à intégrer les données dans l'entrepôt de données à l'aide de l'outil Talend Open Studio for Data Integration après leur mise en cohérence, la deuxième est la création des cubes de données à partir de l'entrepôt de données grâce à l'outil SQL Server Analysis Services et la troisième est la mise en œuvre des rapports et des tableaux de bords nécessaires pour une prise de décision à partir de l'entrepôt de données ou des cubes de données à l'aide de l'outil de reporting SQL Server Reporting Service. L'intégration de sources de données hétérogènes dans l'entrepôt de données est un problème d'actualité. La disposition de rapports et de tableaux de bord est vitale pour une structure comme l'UCAD.

Ce stage nous a alors permis d'une part, de mettre en exergue des connaissances diverses acquises en classe dans les champs de l'informatique décisionnelle. Il nous a donc été bénéfique sur le plan technique vu que nous avons pu maîtriser la chaîne de mise en œuvre d'un projet BI en utilisant divers outils. D'autre part, également, il nous a donné l'occasion de pratiquer des méthodes souples pour la gestion du projet et d'acquérir une expérience du travail collaboratif.

2. Perspectives

Tout système étant appelé à évoluer dans le temps, il est alors possible de faire des améliorations qui peuvent l'élargir. Dans le but de garantir une continuité agile au projet, il est alors important de proposer certaines perspectives.

En guise de perspectives, notre solution a l'avantage d'être ouverte et extensible. En effet, les résultats de ce travail constituent les bases d'un travail à poursuivre et à améliorer dans d'autres périmètres fonctionnels. Ainsi, les perspectives futures sont, d'abord, l'intégration des données de l'application de gestion de la paie, des données de l'application des voyages d'études et des données de la recherche dans l'entrepôt de données afin de mettre l'exhaustivité des données dans le Data warehouse. Faire une refonte des applications RADIS et IRIS qui vont se baser sur l'entrepôt. Et avec la

Direction des études et statistiques de l'UCAD revisiter l'ensemble des indicateurs des établissements, du rectorat, du ministère de l'enseignement supérieur et de la recherche, de l'UNESCO, de la Banque mondiale et de l'ensemble des partenaires de l'UCAD pour les mettre dans la future application.

BIBLIOGRAPHIES ET WEBOGRAPHIES

- Breslin, M. (2004). Data Warehousing Battle of the Giants: Comparing the Basics of the Kimball and Inmon Models. *BUSINESS INTELLIGENCE JOURNAL*, 15. Consulté le 10 01, 2019, sur <https://www.semanticscholar.org/paper/DW-MODELS-Data-Warehousing-Battle-of-the-Giants-%3A-Breslin/c80f8aaea5bf58846b0125b460401fed8230c2d2>
- Inmon, W. H. (2002). *Building the Data Warehouse Third Edition* (Vol. 3). (J. W. Sons, Éd.) New York, Etats Unis d'Amérique: Wiley Computer Publishing. Consulté le 09 10, 2019, sur <https://leseprobe.buch.de/images-adb/e2/e7/e2e7f926-314c-4e83-86a1-e481fafaf580.pdf>
- John Wiley & Sons, I. (Éd.). (s.d.). New York, Etats Unis d'Amérique: Wiley Computer Publishing. Consulté le 10 11, 2019, sur <https://leseprobe.buch.de/images-adb/e2/e7/e2e7f926-314c-4e83-86a1-e481fafaf580.pdf>
- Ralph Kimbal, M. R. (2002). *The Data Warehouse Toolkit Second Edition* (Vol. 2). (I. John Wiley & Sons, Éd.) New York, Etats Unis d'Amérique: Wiley Computer Publishing. Consulté le 09 27, 2019, sur <http://users.itk.ppke.hu/~szoer/DW/Kimball%20&%20Ross%20-%20The%20Data%20Warehouse%20Toolkit%202nd%20Ed%20%5BWiley%202002%5D.pdf>
- Ralph Kimball, J. C. (2004). *The Data Warehouse ETL Toolkit*. (J. W. Sons, Éd.) New York, Etats Unis d'Amérique: Wiley Publishing, Inc. Consulté le 10 22, 2019, sur <http://users.itk.ppke.hu/~szoer/DW/Kimball%20&%20Caserta%20-The%20Data%20Warehouse%20ETL%20Toolkit%20%5BWiley%202004%5D.pdf>
- [Chouder, 2007] : Lamri Chouder ; « Entrepôt Distribué de Données » ; Thèse de Magistère Option : SI; Institut National de Formation en Informatique (I.N.I) 2007.
- [Chuck, 1998] : Chuck Ballard, Dirk Herreman, Don Schau, Rhonda Bell, Eunsang Kim, Ann Valencic; Data Modeling Techniques for Data Warehousing; International Technical Support Organization; <http://www.redbooks.ibm.com>; février 1998.
- [Codd, 93] : E. F. Codd ; « Providing OLAP (On-Line Analytical Processing) to User- Analysts : an IT mandate. » ; Technical report ; E.F. Codd & Associates; 1993.