



Research Paper No. 10

Balance of Payments  
Division

**Development of the  
Balance of Payments  
nomenclature**

by Gerard Meagher



Travail de recherche N° 10

Division de la balance  
des paiements

**Le développement de la  
nomenclature de la  
balance des paiements**

par Gerard Meagher



Statistics  
Canada

Statistique  
Canada

Canada

Balance of Payments Division

Division de la balance des paiements

**Development of the Balance of Payments nomenclature****Le développement de la nomenclature de la balance des paiements**

by Gerard Meagher

par Gerard Meagher

written in 1996

rédigé en 1996

**Abstract**

The data model for balance of payments accounting is both simple and complex. At the most elementary level, the model consists of only three: domestic transactor, foreign transactor and transaction. The complexity arises when each of these primary entities is decomposed into its hierarchical structure. The resulting tree is very wide but not very deep.

In this report, the data model is first introduced. This is followed by an overview of the Canadian implementation. The pros and cons of this implementation are then summarized. A characterization of future developments is given as a conclusion.

**Sommaire**

Le modèle des données utilisé pour comptabiliser la balance des paiements est à la fois simple et complexe. Le modèle de base consiste en trois entités: transacteur domestique; transacteur étranger et transaction. La complexité survient lorsqu'on décompose chacune de ces entités dans sa structure hiérarchique. L'arbre qui en résulte est très large quoique peu profond.

Dans ce rapport, le modèle de données est premièrement introduit. Un aperçu général est ensuite fourni sur la mise sur pied de ce modèle pour ce qui est du Canada. Le sommaire des avantages et désavantages est ensuite présenté. En conclusion, un aperçu est donné des caractéristiques qui découleront des développements futurs qui sont anticipés.

**Update:**

Since the parution of this article, data on positions has been included in the FAME database; the formulas (Appendix 2) have been revised; and a system to integrate micro-data is being implemented using an ORACLE database.

**Mise à jour:**

Depuis la rédaction de cet article, les données de positions ont été intégrées à la base FAME; les formules (Appendice 2) ont été modifiées; et un système intégré de traitement de micro-données a été commencé à la base ORACLE.

Telephone/Téléphone: (613) 951-9053 or/ou (613) 951-1856

Email Address / Courrier électronique:  
MEAGHER@STATCAN.CA

Fax / Télécopieur: (613) 951-9031

## Development of the Balance of Payments nomenclature

by Gerard Meagher<sup>1</sup>

### 1. Introduction

#### 1.1 Balance of Payments Accounting

The balance of payments accounts form part of the system of national accounts.

The balance of payments is a statistical statement, for a given time period, systematically summarizing the economic transactions of an economy with the rest of the world. Transactions, for the most part between residents and nonresidents, consist of those involving goods, services and income; those involving financial claims on, and liabilities to the rest of the world; and those classified as transfers (such as gifts), which involve offsetting entries to balance - in an accounting sense - one-sided transactions<sup>2</sup>.

The accounts are constructed using a double entry system. For each transaction a pair of entries are generated - a debit and credit - having equal values but opposite signs. In principal, the sum of all of the debits and credits should equal zero, but due to such factors as reporting deficiencies, timing problems and currency conversion, this is not true in practice. A balancing item is calculated such that the sum of the debits plus the credits plus the balancing item equals zero. This balancing item, is designated as the statistical discrepancy in balance of payments accounts.

#### 1.2 The Problem in Context

The Canadian balance of payments accounts are compiled by the Balance of Payments Division (BOPD) of the System of National Accounts Branch of Statistics Canada. The division employs a full-time staff of 36, composed of 13 economists, a support staff of 19 and a systems staff of 4.

The aim is to record every transaction between a resident of Canada and a non-resident. No easy

<sup>1</sup>. *Balance of Payments Division.*

<sup>2</sup>. *Balance of Payments Manual, Fifth Edition, 1993, International Monetary Fund, pp.6*

## Le développement de la nomenclature de la balance des paiements

par Gerard Meagher<sup>1</sup>

### 1. Introduction

#### 1.1 Comptabilisation de la balance des paiements

Les comptes de la balance des paiements sont des éléments du système de comptabilité nationale.

La balance des paiements est un état statistique où sont systématiquement résumées, pour une période donnée, les transactions économiques d'une économie avec le reste du monde. Les transactions, pour la plupart entre résidents et non-résidents, sont celles qui portent sur les biens, services et revenus; celles qui font naître des créances financières sur le reste du monde ou des engagements financiers envers celui-ci; et celles qui, telles les donations, sont considérées comme des transferts, pour lesquels il y a lieu de passer des contre-écritures de manière à solder - au sens comptable du terme - les transactions à sens unique<sup>2</sup>.

Les comptes sont enregistrés suivant la comptabilité en partie double. Autrement dit, chaque transaction donne lieu à deux inscriptions - un débit et un crédit; elles sont de valeur identique, mais affectées de signes opposés. En théorie, la somme des débits et des crédits est égale à zéro, mais tel n'est pas le cas dans la pratique, en raison de facteurs tels que les écarts en matière de déclaration, les problèmes liés au choix du moment et le change. Un poste de contrepartie est donc établi de sorte que la somme des débits, des crédits et de ce poste soit égale à zéro. On désigne le poste de contrepartie, par le terme "divergence statistique" dans les comptes de la balance des paiements.

#### 1.2 Le problème en contexte

La tenue des comptes de la balance des paiements canadienne est assurée par la Division de la balance des paiements (DBP), Direction du système de comptabilité nationale, Statistique Canada. La Division regroupe 36 employés à plein temps, soit 13 économistes, 19 employés de soutien et 4 préposés aux systèmes.

La Division doit enregistrer toutes les transactions qui surviennent entre un résident du Canada et un non-résident,

<sup>1</sup>. *Division de la balance des paiements.*

<sup>2</sup>. *Manuel de la balance des paiements, cinquième édition, 1993, Fonds monétaire international, page 6.*

task. The sources of data are very diversified. A good part of the data is collected from some 30 monthly, quarterly and annual surveys which are mailed to some 9,000 enterprises. Some of the enterprises submit their data in an electronic form. Other sources of data include: other divisions within Statistics Canada (e.g., International Trade Division and Education, Culture and Tourism Division), the Bank of Canada, other federal departments and the media, among others.

The micro data is amassed in 13 different processing systems distributed between an IBM mainframe computer and DOS based micro computers linked in a Local Area Network. These systems are written in SAS, PL/1, FoxPro and Paradox. In addition, Lotus and FoxPro are used to process data not covered by these systems.

After the micro data for a given period is edited, it is aggregated and coded into one of the some 4,000 accounts. It is then loaded into a time-series data base where all of the accounts are combined to produce the balance of payments. The time-series data base is used for analysis and to produce publications and other outputs.

The users typically spend most of their time on the micro-level systems. The time-series data base is only used to record the final aggregate accounts, compile them, perform some macro analysis and produce the publications.

The first time-series data base systemised by the Balance of Payments Division was developed internally at Statistics Canada in 1981. The system, NASH (National Accounts Series Handler) was maintained by National Accounts and Environment Division (NAED) and ran on an HP-3000 mini computer. Both the hardware and the software were starting to show their age in terms of flexibility and growth potential. A replacement was required.

### 1.3 The Software Solution

It was decided to implement the time-series data base in FAME as it would provide the required functionality with minimal additional development. Having already been adopted by the Federal Reserve Board in the United States and the Bank of Canada, FAME had proven to be a powerful solution in similar circumstances.

ce qui n'est pas une mince tâche, car les sources de données sont très diversifiées. En effet, les données proviennent d'environ 30 enquêtes mensuelles, trimestrielles et annuelles qui sont envoyées par la poste à près de 9 000 entreprises, dont certaines transmettent les données sur support électronique. Parmi les autres sources de données, mentionnons d'autres divisions de Statistique Canada (p. ex., la Division du commerce international et la Division de l'éducation, de la culture et du tourisme), la Banque du Canada, d'autres ministères fédéraux et les médias.

Les microdonnées sont extraites de 13 différents systèmes de traitement répartis entre un ordinateur principal IBM et des micro-ordinateurs à base DOS liés dans un réseau local. Ces systèmes sont rédigés en SAS, PL/1, FoxPro et Paradox. De plus, les logiciels Lotus et FoxPro sont utilisés pour traiter les données non visées par ces systèmes.

Après avoir été révisées, les microdonnées d'une période sont agrégées et codées dans l'un des quelque 4 000 comptes, puis elles sont chargées à une base de données chronologiques où tous les comptes sont combinés afin de produire la balance des paiements. La base de données chronologiques sert à des fins d'analyse et de préparation de publications et d'autres produits.

Habituellement, les utilisateurs consacrent la plus grande partie de leur temps aux systèmes de microniveau. La base de données chronologiques n'est utilisée que pour enregistrer et compiler les comptes définitifs sous forme d'agrégats, effectuer des macro-analyses et produire des publications.

La première base de données chronologiques utilisée par la Division de la balance des paiements a été élaborée à l'interne par Statistique Canada en 1981. La Division des comptes nationaux et de l'environnement (DCNE) tenait le système, appelé NASH (soit l'acronyme anglais du Système de traitement des séries des comptes nationaux), qui était exploité au moyen d'un mini-ordinateur HP-3000. Tant le logiciel que le matériel commençaient à manifester des signes de vieillesse sur le plan de la souplesse et du potentiel d'expansion. Il fallait donc les remplacer.

### 1.3 La solution fondée sur le logiciel

On a décidé de mettre en oeuvre la base de données chronologiques dans la base de données FAME, ce qui assurerait la fonctionnalité requise et exigerait peu de développement additionnel. La Réserve fédérale des États-Unis et la Banque du Canada ayant déjà adopté la base FAME, cette solution s'est révélée exceptionnelle dans les circonstances.

At Statistics Canada, FAME was first chosen by NAED and BOPD decided to follow their lead. A particularly appealing factor was the move to a commercially developed and supported software with the potential for sharing ideas and software between the two divisions as well as with other users.

Again following the footsteps of NAED, FrameMaker was selected as the desktop publishing system to produce our publications.

## 1.4 The Hardware Solution

The Balance of Payments division had a substantial investment in DOS based micro computers and software which we were not willing nor able to sacrifice. On the other hand NAED has opted to implement FAME in UNIX. A compromise solution would enable us to maintain our existing systems, some of our existing hardware and still remain compatible with NAED.

A UNIX file server, 4 UNIX compute servers and 3 UNIX work stations were purchased. Some of the division's PC's were upgraded to 486's, to give a total of 21. The 486's run an X-Terminal emulation from the compute servers. Users can use their PC's to run DOS applications, to go to FAME under UNIX or to go to the mainframe from either DOS or UNIX.

## 2. Data Model

The conversion to FAME provided an excellent opportunity to assess the structure of our data and to apply contemporary data modelling principles to our application. In early 1992 a modelling specialist from Statistics Canada was engaged to begin the modelling process.

### 2.1 The Elementary Model

The elementary model of the balance of payments is beautiful in its simplicity. It can most easily be summarized by the expression "*quid pro quo*" with a non-resident. See Appendix 1.

For example, if a Canadian resident was to purchase a foreign security from a non-resident, the quid would be the money paid and the quo would be the security received. This would be reflected in the double-entry accounting as an increase in the assets of the Canadian resident (debit) and a decrease in foreign reserves (credit).

À Statistique Canada, la base FAME a d'abord été choisie par la DCNE, puis par la DBP. Un des principaux facteurs ayant pesé en sa faveur était l'adoption d'un logiciel élaboré et appuyé commercialement qui présentait des capacités de partage des idées et de logiciels entre les deux divisions ainsi qu'avec d'autres utilisateurs.

C'est de nouveau à l'instar de la DCNE que la DBP a retenu FrameMaker à titre de système d'édition pour la production de nos publications.

## 1.4 La solution fondée sur le matériel

La Division de la balance des paiements ayant investi considérablement dans des micro-ordinateurs et des logiciels basés sur DOS, elle n'était pas disposée à s'en départir, non plus qu'elle était en mesure de le faire. En revanche, la DCNE avait choisi d'instaurer la base FAME sur UNIX. Il fallait trouver un compromis qui nous permettrait de garder nos systèmes et une partie de notre matériel tout en demeurant compatible avec la DCNE.

Un serveur de fichiers UNIX, quatre serveurs de calcul UNIX et trois postes de travail UNIX ont donc été achetés. Certains des ordinateurs personnels de la Division ont été remplacés par des 486, ce qui en porte le total à 21. Les 486 exploitent un logiciel d'émulation de terminal X à partir des serveurs de calculs. Les utilisateurs peuvent se servir de leurs ordinateurs personnels pour passer des applications DOS, accéder à la base FAME sur UNIX ou accéder à l'ordinateur principal soit par DOS soit par UNIX.

## 2. Modèle de données

La conversion à la base FAME nous a fourni l'occasion rêvée d'évaluer la structure de nos données et d'appliquer des principes contemporains de modélisation des données. Nous avons donc, au début de 1992, confié à un spécialiste de la modélisation de Statistique Canada les travaux concernant ce processus.

### 2.1 Le modèle élémentaire

Le modèle élémentaire de la balance des paiements est remarquable par sa simplicité. Il se résume en une expression «en contrepartie de», qui s'applique aux rapports avec un non-résident. Voir Appendice 1.

Par exemple, si un résident canadien achète un titre étranger d'un non-résident, il reçoit un titre en contrepartie de l'argent qu'il a versé, ce qui, selon la comptabilité en double partie, correspond à une augmentation de l'actif du résident canadien (débit) et à une diminution des réserves étrangères (crédit).

For the accounts to balance there must always be these two offsetting entries. For each entry there are three components: the Canadian transactor, the non-resident transactor and the transaction itself. The next step was to define each of these entities in more detail.

## 2.2 The Extended Model

In order to completely describe each balance of payments transaction a necessary and sufficient set of attributes was formed. There are 13 attributes in all describing the three entities.

As is to be expected, more is known about the resident transactor than about the non-resident. This is simply due to the nature of statistical reporting restrictions and not to any narcissistic bearing.

A 14<sup>th</sup> attribute called "series type" is used to identify if the series contains flows, positions, indexes, etc.. This provides a nice link between the balance of payments data and the international investment position data. By simply changing the series type of a flow series the user can have a position which corresponds exactly to the flow.

## 3. Data Base Design

### 3.1 Governing Factors

Several factors were at play in the final data base design. Each of the following factors encouraged further fragmentation of the data base.

#### 3.1.1 Series Names

Whenever possible, we wanted to limit the number of series names making it easier for the users to remember names. To this end, the same series name was used in different data bases and the data base determines the distinction. For example, if the same series existed in both a seasonally adjusted form and in a non-adjusted form the two series would have the same name but would be stored in different data bases. This has the added potential for making reports and other programs more generalized, for example, one can define a report for a list of series in the adjusted data base and by simply changing the data base run the same report for non-adjusted data.

Pour que les comptes soient en équilibre, ces deux inscriptions compensatoires doivent toujours être présentes. Chaque inscription renferme trois composantes, à savoir la partie à la transaction canadienne, la partie à la transaction non résidente et la transaction elle-même. Il fallait ensuite définir plus à fond chacune de ces composantes.

## 2.2 Le modèle étendu

Afin de décrire complètement chacune des transactions de la balance des paiements, il a fallu constituer un ensemble d'attributs suffisant. En tout, 13 attributs décrivent les trois composantes.

Comme prévu, les données sont plus abondantes sur la partie à la transaction canadienne que sur le non-résident. Il en est ainsi tout simplement en raison de la nature des restrictions en matière de déclaration statistique et non pour des motifs narcissiques quelconques.

Un 14<sup>e</sup> attribut, le «type de série», sert à déterminer si la série renferme des flux, des positions, des index, etc. Il constitue un bon lien entre les données sur la balance des paiements et celles sur la position extérieure globale. L'utilisateur n'a qu'à changer le type de série d'une série à flux pour obtenir la position qui correspond exactement à ce flux.

## 3. Conception de la base de données

### 3.1 Facteurs déterminants

Plusieurs facteurs sont entrés en ligne de compte à l'étape définitive de la conception de la base de données. Chacun des facteurs suivants a permis de fragmenter davantage la base de données.

#### 3.1.1 Noms de série

Dans la mesure du possible, nous voulions limiter le nombre de noms de série pour que les utilisateurs puissent plus facilement se les rappeler. À cette fin, nous avons utilisé le même nom de série dans différentes bases de données; ce sont ces dernières qui permettent d'établir la distinction entre les séries. Par exemple, si la même série existait à la fois sous une forme désaisonnalisée et non désaisonnalisée, les deux séries portaient le même nom, mais étaient stockées dans des bases de données différentes. Cela permettait également de généraliser davantage les rapports et autres programmes. Par exemple, il devenait possible de définir un rapport à partir d'une liste de séries dans la base de données désaisonnalisées puis, en changeant tout simplement de base de données, de produire le même rapport avec des données non désaisonnalisées.

### 3.1.2 Security

Even within a small division such as ours, there was a need to restrict the number of people who have write access to the data bases. The data base was divided along the lines of the sections where everyone had read access to all of the data in the division but write access only to the data of their own section. This was done more as a concern for accidental damage than from fear of malicious behaviour.

### 3.1.3 Performance

Dividing the data base into smaller pieces allowed for more users to update their own portion of the data base without disrupting the other users. It was also thought that retrieval time would be reduced with many people accessing several data bases rather than many people accessing a single data base. However, the time factor was not really a major concern.

Many operational procedures apply only to a particular type of data. There are time savings to be accrued by performing these operations only where necessary.

FAME formulae were used extensively to calculate all of the aggregate-level data from the entered data. These formulae form a tree-like structure with the entered data at the bottom and the statistical discrepancy at the top. See Appendix 2. Although the formulae are very simple (usually just a summation of six other series) they are nested sufficiently deep that the retrieval time for some of the upper-level series is too slow. This necessitated putting the formulae in one data base and having a parallel data base with the formulae resolved in permanent series.

### 3.1.4 Revisions Policy

We maintain a strict revision policy which allows us to make historical revisions only in the first quarter of each year. This meant being able to enter revisions whenever we discover them but only release them to the public at the designated time.

### 3.1.2 Sécurité

Même dans une division aussi petite que la nôtre, il fallait restreindre le nombre de personnes ayant un droit d'accès en écriture aux bases de données. La base de données a été divisée selon les sections où tous les employés détenaient un droit d'accès en lecture à toutes les données de la division, mais un droit d'accès en écriture visant seulement les données de leur propre section. Cette mesure a été motivée davantage par la crainte des dommages accidentels que par la crainte d'un comportement malveillant.

### 3.1.3 Rendement

En divisant la base de données en éléments plus petits, nous avons permis à plus d'utilisateurs de mettre à jour leur propre partie de la base des données sans perturber le travail des autres utilisateurs. Nous estimions également que le temps d'extraction serait réduit si nous permettions à de nombreuses personnes d'avoir accès à plusieurs bases de données plutôt qu'à une seule. Cependant, le facteur temps ne comptait pas vraiment parmi nos grandes préoccupations.

De nombreuses procédures opérationnelles ne s'appliquent qu'à un type de données particulier. Des économies de temps peuvent donc être réalisées si ces opérations ne sont exécutées qu'au besoin.

Nous nous sommes servis considérablement des formules FAME pour calculer l'ensemble des données au niveau agrégé à partir des données inscrites. Ces formules représentent une structure en forme d'arbre où les données inscrites se trouvent dans la partie inférieure et l'écart statistique, dans la partie supérieure. Voir Appendice 2. Les formules, en dépit de leur très grande simplicité (elles représentent habituellement la somme des six autres séries), étaient nichées à un niveau tellement éloigné que le temps d'extraction de certaines des séries de plus haut niveau devenait trop long. Nous avons donc dû stocker les formules dans une base de données et constituer une base de données parallèle renfermant les formules résolues d'une série permanente.

### 3.1.4 Politique de révision

Nous avons appliqué une politique de révision sévère qui nous permet d'apporter des révisions historiques seulement au cours du premier trimestre de chaque année. Il fallait donc être en mesure d'entrer les révisions lorsque nous les constatons, mais ne les diffuser auprès du grand public qu'au moment prévu.

### 3.1.5 Revision Analysis

Although the data in the data bases changes daily we need to be able to “Freeze” the data at different points in time. These times correspond to the times when we produce our publications - monthly, quarterly and annually. These copies need to be maintained for the purposes of revision analysis and to produce ad hoc tables which must correspond to data already published.

### 3.2 Main Data Bases

Our resulting data base contains three main bases: “Work”, “Draft” and “Published”. There is only one each of “Work” and “Draft” but there are many versions of “Published”. Data is migrated from “Work” to “Draft” and (at publication time) to “Published”. Formulae are only used in the “Work” data base. Changes can be made to any point in time at any time in the “Work” data base but only changes for the current revision period are moved to the “Draft” data base. Every time we produce a publication we copy the entire “Draft” data base to a “Published” data base.

Each of these main data bases is physically implemented using directories on our file server.

### 3.3 Subordinate Data Bases

Within each of these main data base are 10 subordinate data bases. They are divided by: type of data, seasonality and sources of data. The type of data can be current account, capital account, positions or other. The seasonality can be seasonally adjusted or not seasonally adjusted. The source can be entered or derived. Not all classifications apply in all cases, hence 10 subordinate data bases and not 16.

Each subordinate data base is a unique FAME data base which has the same physical file name within its main data base directory but has a unique, standard channel name. In this way data can easily be compared from main data base to main data base.

See Appendix 3.

### 3.1.5 Analyse de révision

Bien que les données de la base de données changent chaque jour, nous devons pouvoir les «geler» mensuellement, trimestriellement et annuellement, c’est-à-dire selon la fréquence de production de nos publications. Les copies doivent être conservées aux fins de l’analyse de révision et de la production de tableaux spéciaux qui doivent correspondre aux données déjà publiées.

### 3.2 Bases de données principales

Notre base de données comporte donc trois bases principales, soit une base «travaux», une base «ébauche» et plusieurs bases «publication». Les données passent de la base «travaux» à la base «ébauche», puis, au temps de la publication, à la base «publication». Les formules ne sont utilisées que dans la base «travaux». Des changements peuvent être apportés en tout temps à tous les éléments de cette base, mais seuls les changements touchant la période de révision courante passent à la base de données «ébauche». Chaque fois que nous produisons une publication, nous copions la base de données «ébauche» au complet, puis nous la versons à la base de données «publication».

La mise en oeuvre de chacune de ces bases de données principales s’effectue au moyen de répertoires du serveur de fichiers.

### 3.3 Bases de données subordonnées

Chacune de ces bases de données principales comporte 10 bases de données subordonnées qui sont divisées selon le type de données, les fluctuations saisonnières et les sources de données. Le type de données englobe des données du compte courant, du compte de capital, des positions ou d’autres types. Pour leur part, les fluctuations saisonnières renferment des données désaisonnalisées et non désaisonnalisées. Enfin, la source de données peut être inscrite ou dérivée. Puisque les classifications ne s’appliquent pas dans tous les cas, on trouve 10 bases de données subordonnées plutôt que 16.

Chaque base de données subordonnée représente une base de données FAME unique portant le même nom de fichier dans le répertoire de la base de données principale, mais un nom de canal normalisé unique. Cela facilite la comparaison des données entre les bases de données principales.

Voir Appendice 3.

### 3.4 Nomenclature - The Long and the Short of it

When adopting a nomenclature for our FAME implementation we had 2 choices: we could maintain our old series names which were relatively short, which were familiar to the users and which were used in our aggregation methodologies, or we could adopt a new series name composed of the 13 attributes identified earlier. We settled for the best of both worlds. The old names were kept as FAME series names and an alias will be built from the 13 attributes glued together by dots.

The old series names identified 1 or 2 principal components of the series. Many of the other aspects of the series were either implied, not known or ignored. The naming convention was neither consistent nor necessarily logical. This limited considerably their uses and posed particular problems to new users of the data.

For example, the series D21IN6.US contains foreign direct investment into Canadian corporations from the US. The new alias names, although long and cumbersome, will contain much more information. The same series will be coded as: F.STK##.USA.A.IN.C####.ALL#..L..L which translates into: a flow, of stocks, from US, affiliated companies, into Canadian, corporations, all industries, unknown investment type, all currencies, in an unknown province, increasing liabilities, from an unknown country of control, from an unknown province of control, of long-term maturity. See Appendix 4.

The users are not expected to memorize such names, nor even the positions or their meanings. The aliases will be accessed through special screens which will be developed to hide all of the gobbledegook and present a more coherent, logical view to the user.

With all of this classification information about each series, the user will be able to query the data base to find such things as all of the transactions in a particular currency, industry, etc.. In the future more of the data will be loaded directly from micro-data systems providing more detail to identify, for example, individual countries instead of geographic areas.

### 3.4 Nomenclature - en long et en court

Lorsqu'est venu le temps d'adopter une nomenclature pour l'ajout de notre base à la base FAME, deux choix se sont présentés. Nous pouvions soit conserver les anciens noms des séries, qui étaient relativement courts, connus des utilisateurs et utilisés dans nos méthodologies d'agrégation, soit adopter de nouveaux noms de série composés des 13 attributs susmentionnés. Nous avons choisi le meilleur des deux mondes. En effet, nous avons conservé les anciens noms pour les noms de série de la base FAME et intégré un alias constitué des 13 attributs regroupés au moyen de points.

Les anciens noms de série identifiaient une ou deux composantes principales de la série. Bon nombre des autres aspects de la série étaient soit implicites, soit inconnus, soit laissés pour compte. Les conventions de nomenclature n'étaient pas cohérentes ni forcément logiques, ce qui limitait considérablement les utilisations et causait des problèmes particuliers aux nouveaux utilisateurs des données.

À titre d'exemple, la série D21IN6.US renferme les investissements étrangers directs dans des sociétés canadiennes en provenance des États-Unis. Les nouveaux alias, bien qu'ils soient longs et encombrants, renfermeront bien plus d'information. En effet, la même série portera le code F.STK##.USA.A.IN.C####.ALL#..L..L, qui se traduit par un flux d'actions en provenance de sociétés affiliées américaines vers des sociétés canadiennes, tous les secteurs, type d'investissement inconnu, toutes les devises, dans une province inconnue, avec augmentation du passif, nationalité des capitaux majoritaires inconnue, d'une province détentrice des capitaux majoritaires inconnue, échéance à long terme. Voir Appendice 4

On ne s'attend pas à ce que les utilisateurs mémorisent ces noms ni même les positions ou leur signification. Ils pourront avoir accès aux alias grâce à des écrans spéciaux qui dissimuleront la «quincaillerie» et leur présenteront une image cohérente et logique.

Grâce à cette panoplie de renseignements sur la classification concernant chaque série, l'utilisateur pourra rechercher dans la base de données des éléments tels que l'ensemble des transactions dans une devise particulière, dans un secteur donné, et ainsi de suite. À l'avenir, davantage de données seront chargées directement aux fins des systèmes de microdonnées qui fournissent plus de détails permettant d'identifier, par exemple, des pays particuliers plutôt que des régions géographiques.

Eventually the alias names could be used to produce aggregation rules which would be used to derive the entire balance of payments. Before this can happen, some intelligence would have to be built into the codes to identify totals and sub-totals from the components.

The new nomenclature aliases will also allow us to comply with the new IMF classification system for the current account and the capital account and with the OECD/EUROSTAT classification of trade-in-services. The "Account Code" was designed to map to these standards. See Appendix 4.

#### 4. Developed Tools

It was decided early on that most users should not be expected to learn FAME. The production data bases would be built for them and they would only be expected to load their data at the appropriate times. Since they would typically only be using the time-series data base for a couple of weeks out of each quarter, a simple, graphical user interface was required, something they could intuitively use after being away from it for two months.

In addition to the developed tools outlined below, the users were given short courses on FAME input files and how to produce simple reports. In this way, they would be able to produce reports not available through the generalized tools.

The developed tools are accessed through the Balance of Payments Main Window. Here the user can open and close data bases, set the date range and frequency or access the other tools.

FAME Windows was used to create all of these windows with the exception of the "Revise" window, it was created using TeleUSE.

##### 4.1 Searcher

The Searcher is a tool for exploiting the Balance of Payments nomenclature. The design is based on a query-by-example approach. The user is presented with a window with a button for each of the attributes. When the user clicks on any button a list of the possible values for that attribute is given. For example, the "Sector" button will give a list of all of the valid sectors, plus some special groupings, plus an "any" and an "all" option. The user defines a

Avec le temps, les alias serviront à produire des règles d'agrégation qui permettront de dériver la balance des paiements intégrale. Toutefois, avant que cela ne puisse se produire, il faudra intégrer de l'intelligence aux codes pour qu'ils distinguent les totaux et les totaux provisoires des composantes.

Les nouveaux alias de la nomenclature nous permettront également de nous conformer au nouveau système de classification du FMI pour le compte courant et le compte de capital ainsi qu'avec la classification des échanges de services de l'OCDE et de l'OSCE. Le «code compte» a été créé pour établir la correspondance avec ces normes. Voir Appendice 4.

#### 4. Outils élaborés

Tout au début du processus, il a été décidé que la plupart des utilisateurs ne devraient pas être tenus d'apprendre la base FAME. Les bases de données de production seraient constituées pour eux, et ils n'auraient qu'à y charger leurs données au moment opportun. Comme ils utiliseraient la base de données chronologiques pendant quelques semaines seulement chaque trimestre, il fallait concevoir une interface graphique simple qu'ils pourraient utiliser intuitivement même s'ils n'y avaient pas touché pendant deux mois.

Outre les outils décrits ci-dessous, nous avons fourni aux utilisateurs de brèves séances de formation sur les fichiers de programmation FAME et sur la manière de produire des rapports simples, ce qui leur a permis de produire des rapports non disponibles au moyen des outils généralisés.

Les utilisateurs peuvent accéder aux outils élaborés grâce à la fenêtre principale de la balance des paiements. Ils peuvent ouvrir et fermer des bases de données, établir des fourchettes de dates et la fréquence ou avoir accès à d'autres outils.

Nous nous sommes servis du logiciel FAME Windows pour créer toutes les fenêtres, sauf la fenêtre «révision» qui, elle, a été créée au moyen du logiciel TeleUSE.

##### 4.1 Récupérateur

Le récupérateur est un outil d'exploitation de la nomenclature de la balance des paiements qui a été conçu au moyen du langage d'interrogation QBE. L'utilisateur a devant lui une fenêtre comportant un bouton pour chacun des attributs. Lorsqu'il clique sur un des boutons, une liste des valeurs possibles pour cet attribut apparaît. Par exemple, en cliquant sur le bouton «secteur», il voit apparaître une liste de tous les secteurs valables, certains groupes spéciaux de même qu'une option «l'un ou l'autre»

query by selecting the values wanted for each of the attributes.

Using the data bases that are open at the time, the "searcher" will return a list of series that satisfies the query. This list can then be used to drive analysis or arithmetic operations.

The user can now produce lists of series by doing FAME wild-carding or by using a predefined list of series known as a set. These sets are implemented as FAME namelists.

## 4.2 Analysis

The analysis tool can be used to produce reports and graphs in a number of different, standard formats using the list of series generated from the searcher, the itemlist or the sets window. The main analysis window allows the user to make various settings and then, through a series of menus and windows, produce reports and graphs.

## 4.3 Utilities

Updates are made to the "Draft" data base by using the "Revise" window. The user selects a list of series, defines the date range and frequency and then revises.

Other DBA tools have been developed. Among other things, these are used to migrate the data from "Work" to "Draft" to "Published" and to import and export data in various formats.

## 5. Pros and Cons

### 5.1 Training

The 3 most important factors for a re-engineering project of this type are: training, training and training. It seems like no matter what you do you should have done more. Some other lessons:

- Do the training as close as possible to the time that the users will begin to use the system, if not, they forget.
- Train them only for what they need as they need it, if not, they get overwhelmed and confused.

et «tous». Pour définir son interrogation, l'utilisateur choisit les valeurs qu'il désire pour chacun des attributs.

Le récupérateur, après avoir cherché dans les bases de données ouvertes, dresse une liste de séries qui répondent aux critères de l'interrogation. Cette liste peut ensuite servir pour diriger les opérations d'analyse ou de calcul.

L'utilisateur peut maintenant produire des listes de séries à partir de caractères génériques de la base FAME ou d'une liste de séries définie d'avance, appelée ensemble. Ces ensembles constituent des listes de noms de la base FAME.

## 4.2 Analyse

L'outil d'analyse peut être utilisé pour produire des rapports et des graphiques dans plusieurs présentations normalisées différentes à partir de la liste de séries générée par le récupérateur, la liste des éléments ou la fenêtre des ensembles. La fenêtre d'analyse principale permet à l'utilisateur d'établir divers paramètres, puis, au moyen d'une série de menus et de fenêtres, de produire des rapports et des graphiques.

## 4.3 Logiciels utilitaires

La fenêtre «révision» permet de mettre à jour la base de données «ébauche». L'utilisateur n'a qu'à choisir une liste de séries, à définir la fourchette de dates et la fréquence, puis à effectuer la révision.

D'autres outils de gestion de la base de données ont été conçus. Ils servent entre autres à faire passer les données de la base «travaux» à la base «ébauche» puis à la base «publication» de même qu'à importer et à exporter des données sur divers supports.

## 5. Avantages et inconvénients

### 5.1 Formation

Les trois facteurs les plus importants d'un projet de restructuration de ce genre sont : la formation, la formation et la formation. Il semble que, malgré tous les efforts déployés, on puisse toujours en faire davantage. Voici certains enseignements reçus :

- Offrez la formation le plus près possible du moment où l'utilisateur commencera à se servir du système, sinon il oubliera ce qu'il a appris.
- N'enseignez que les éléments dont l'utilisateur a besoin à mesure qu'il en a besoin, sinon il sera vite dépassé et confus.

- Instead of 1 long course for everyone make several smaller modules so they can attend only the relevant sessions.
- Use examples that they can relate to, taken from data they are familiar with. This removes an unnecessary level of complexity.
- Give homework, if they don't use it they lose it.
- Au lieu d'offrir une longue séance à l'intention de tous, présentez plusieurs petits modules, ce qui permet aux utilisateurs d'assister seulement à ceux qui les intéressent.
- Servez-vous d'exemples avec lesquels les utilisateurs peuvent s'identifier et qui sont tirés de données qu'ils connaissent. Vous enlèverez ainsi un niveau de complexité inutile.
- Donnez des devoirs, car si les utilisateurs n'appliquent pas ce qu'ils ont appris, ils l'oublieront.

## 5.2 Requirements Specification

Know what you want to build before you begin to build it. The time spent getting the requirements straight is time well invested. Unfortunately, the users are not the best source of information, they may think they know what they want, but what they want is not always what they really need.

We were lucky to be able to provide the users with a prototype early in the development phase. The final product does not look anything like the original prototype, we received a lot of useful feedback which would not have been provided if the users had not been able to try the prototype.

## 5.3 Interface

Possibly the most successful aspect of our new system has less to do with FAME itself than it has to do with the fact that we were able to provide a nice graphical interface. This, coupled with the power and openness of UNIX, allows the user to manipulate and display the data in ways never expected. In particular, the facility with which data can be displayed graphically will undoubtedly have a strong, positive effect on the quality of data.

## 5.4 Documentation

Program and system documentation aimed at developers, maintainers and data base administrators is essential and possibly as important as the programs themselves. Programmers in general, and especially new programmers, must be able to find out what a program does, what it calls, where it is called from, etc. . Failure to provide this essential information is costly in both time and resources.

## 5.2 Précision des exigences

Sachez ce que vous voulez créer avant même d'amorcer vos travaux. Le temps que vous consacrez à confirmer les exigences est très bien investi. Malheureusement, les utilisateurs ne constituent pas la meilleure source de renseignements, car, s'ils pensent savoir ce qu'ils veulent, les choses qu'ils veulent ne sont pas toujours celles dont ils ont vraiment besoin.

Nous avons eu la chance de pouvoir offrir un prototype aux utilisateurs tout au début de l'étape d'élaboration. Le produit final ne ressemble en rien au premier prototype. Nous avons reçu une foule de renseignements utiles que nous n'aurions pas pu obtenir si les utilisateurs n'avaient pas pu mettre le prototype à l'essai.

## 5.3 Interface

Il se peut fort bien que le volet le plus réussi de notre système ne porte pas tant sur la base FAME en soi, mais plutôt sur le fait que nous avons pu fournir une interface graphique agréable. Cette interface, jumelée à la puissance et à l'ouverture de UNIX, permet à l'utilisateur de manipuler et d'afficher les données dans des manières auxquelles on n'aurait jamais pensé. Surtout, la facilité d'affichage graphique des données aura sans doute un effet positif marqué sur la qualité des données.

## 5.4 Documentation

La documentation sur le programme et le système à l'intention des concepteurs, des préposés à l'entretien et des administrateurs de bases de données est essentielle, voire elle pourrait être tout aussi importante que les programmes. Les programmeurs en général et surtout les nouveaux programmeurs doivent pouvoir trouver ce qu'un programme accomplit, comment il le nomme, à partir de quoi il peut l'appeler, etc. Le défaut de fournir cette information essentielle entraîne de sérieux coûts tant au chapitre du temps qu'au chapitre des ressources.

When programming with objects it is especially important to have automated tools to document the system and find unresolved references and typing errors. Software to automatically and systematically test each module of our system was not available. We resorted to the old-fashioned method of testing: give it to the users - they will find all the bugs. We did and they at least found some of the bugs.

User documentation is another story. Do you have documentation on how to use your keyboard or your mouse? If you do, you have probably never read it. Someone showed you how to use things, you remembered a few basic principles and you started using them. Systems should be as easy to use as that. Windows, menus, buttons and lists should be intuitive and use language that the users are familiar with. After brief instructions they should be able to find their way around. If they do something wrong, the system should be robust enough to warn them of impending danger and protect them from hurting themselves too much. More to the point, have you ever met a user who reads documentation?

## 6. Future Developments

### 6.1 International Investment Position

In addition to producing the Canadian balance of payments, the Balance of Payments Division also produces Canada's international investment position. The international investment position accounts are closely related to the balance of payments accounts.

(The international investment position) is a statistical statement, as of a given date such as year-end, of the value and composition of the claims or stock of financial assets of an economy on the rest of the world, and the value and composition of the stock of the economy's liabilities to the rest of the world.<sup>3</sup>

The nomenclature system developed for the balance of payments flow accounts will also be used to define the position series. For the first time we will be able to easily link the positions to their corresponding flows. These will also be linked to investment income series providing a powerful editing and analysis tool.

<sup>3</sup> *Balance of Payments Manual, Fifth Edition, 1993, International Monetary Fund, pp.6.*

Lorsqu'on programme en se servant d'objets, il importe particulièrement de se doter d'outils automatisés pour documenter le système et trouver les références non réglées et les fautes de frappe. En l'absence d'un logiciel permettant de tester automatiquement et systématiquement chacun des modules de notre système, il restait la bonne vieille méthode de mise à l'essai, à savoir mettre le système entre les mains des utilisateurs pour qu'ils trouvent toutes les erreurs. C'est ce que nous avons fait, et les utilisateurs ont au moins trouvé certaines des erreurs.

La documentation à l'intention des utilisateurs a été une autre paire de manches. Avez-vous de la documentation sur le fonctionnement de votre clavier ou de votre souris? Dans l'affirmative, vous ne l'avez probablement jamais lue. Quelqu'un vous a montré comment faire certaines choses, vous vous êtes souvenu de quelques principes fondamentaux, puis vous avez commencé à vous en servir. Les systèmes devraient être tout aussi faciles à utiliser. Les fenêtres, les menus, les boutons et les listes devraient être intuitifs et rédigés dans des termes que les utilisateurs connaissent. Après de courtes instructions, les utilisateurs devraient pouvoir s'y retrouver. S'ils commettent une erreur, le système devrait être assez fort pour les avertir du danger imminent et empêcher qu'ils ne se blessent trop. Qui plus est, avez-vous déjà rencontré un utilisateur qui avait lu la documentation?

## 6. Évolution prévue

### 6.1 Position extérieure globale

La Division de la balance des paiements produit non seulement la balance des paiements du Canada mais aussi la position extérieure globale de ce dernier. Les comptes de la position extérieure globale sont liés de très près aux comptes de la balance des paiements.

(La position extérieure globale) est un état statistique qui donne, à une date donnée, par exemple en fin d'année, i) la valeur et la composition du stock des actifs financiers d'une économie ou de ses créances sur le reste du monde, ainsi que ii) la valeur et la composition du stock de ses engagements envers le reste du monde<sup>3</sup>.

La nomenclature élaborée pour les comptes de flux de la balance des paiements servira à définir la série des positions. Pour la première fois, nous pourrons lier facilement les positions aux flux correspondants. Celles-ci pourront également être liées aux séries du revenu de placements, ce qui offrira un outil de révision et d'analyse très puissant.

<sup>3</sup> *Manuel de la balance des paiements, cinquième édition, 1993, Fonds monétaire international, page 6.*

This work will be done in the next 12 months including the conversion of the International Investment Position publication to FrameMaker.

## 6.2 Micro Data Systems

In the long-term we plan to build a new environment in which to develop our micro data systems. These systems contain data at the company level and often at the transaction level. The current systems are written in several different languages and reside on different platforms. This makes it onerous to transfer data between systems and makes reconciliation of data from various sources particularly difficult.

The new environment will consolidate these different systems into one logical data base using the data model and nomenclature developed for the time-series dimension of our data. As the existing systems require redevelopment they will be moved into this new environment.

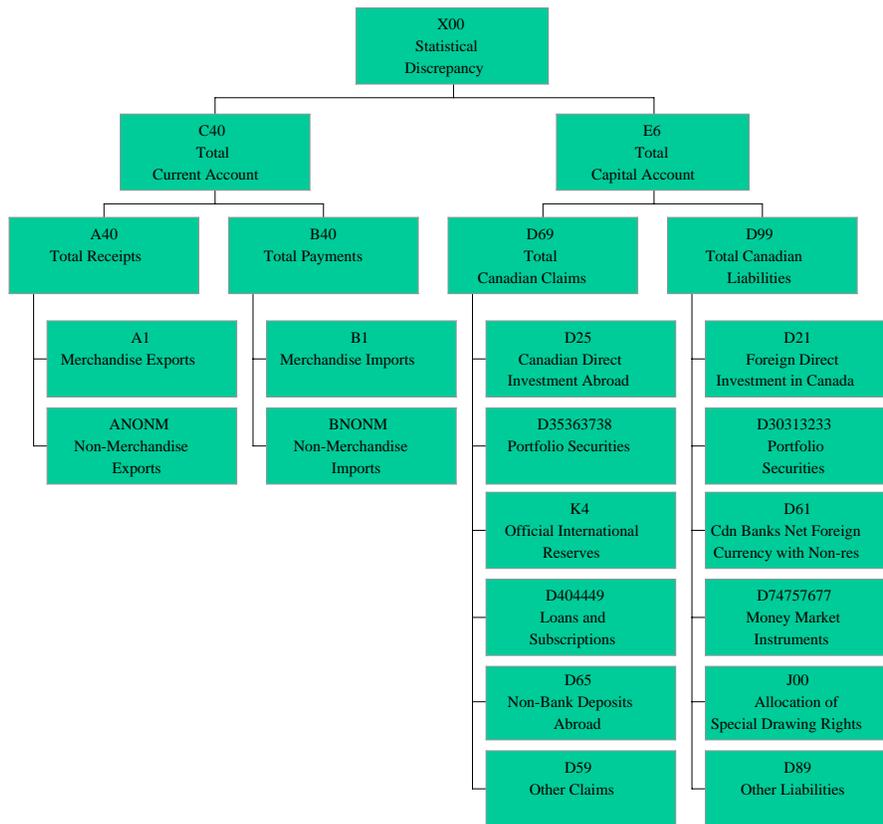
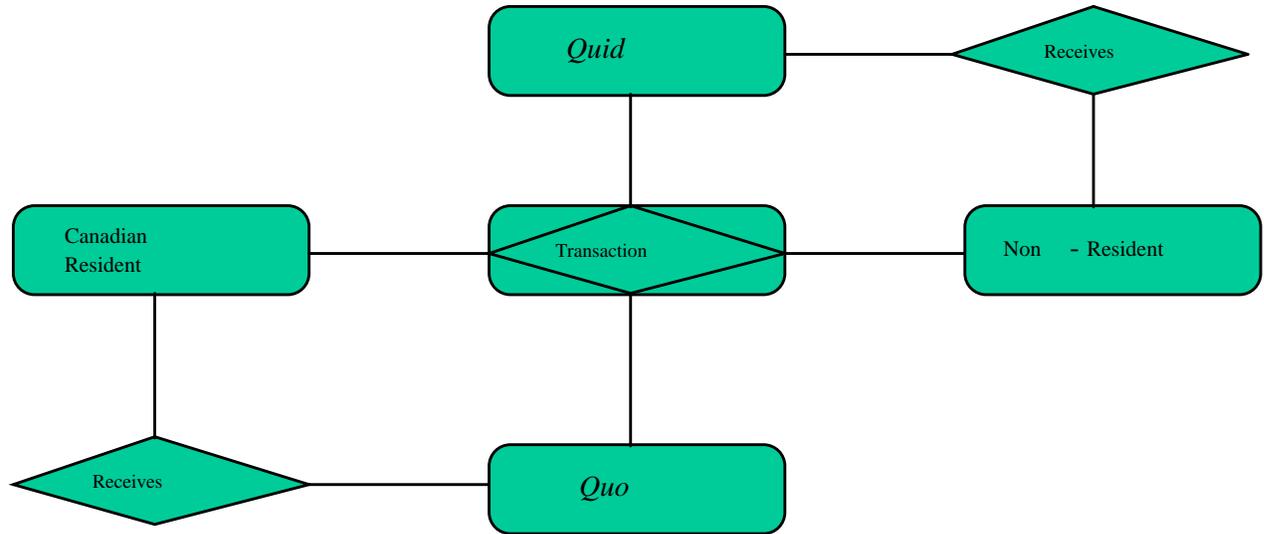
Ces travaux s'échelonnent sur les 12 prochains mois et comprendront la conversion à FrameMaker de la publication sur la position extérieure globale.

## 6.2 Systèmes de microdonnées

Nous prévoyons à long terme bâtir un nouveau milieu pour l'élaboration de nos systèmes de microdonnées. Ces systèmes renferment des données au niveau des sociétés et souvent au niveau des transactions. Les systèmes actuels ont été rédigés dans plusieurs langages différents et résident sur des plates-formes différentes. Il est donc onéreux de transférer les données entre les systèmes et particulièrement difficile de rapprocher les données provenant de sources diverses.

Le nouveau milieu groupera ces différents systèmes en une seule base de données logique utilisant le modèle de données et la nomenclature qui ont été élaborés pour la dimension des séries chronologiques de nos données. Les systèmes existants passeront au nouvel environnement à mesure qu'ils devront être restructurés.

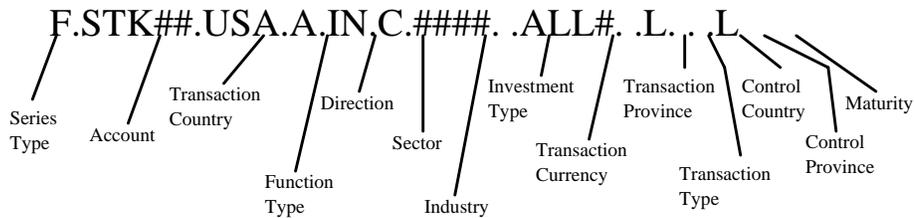
Appendix 1



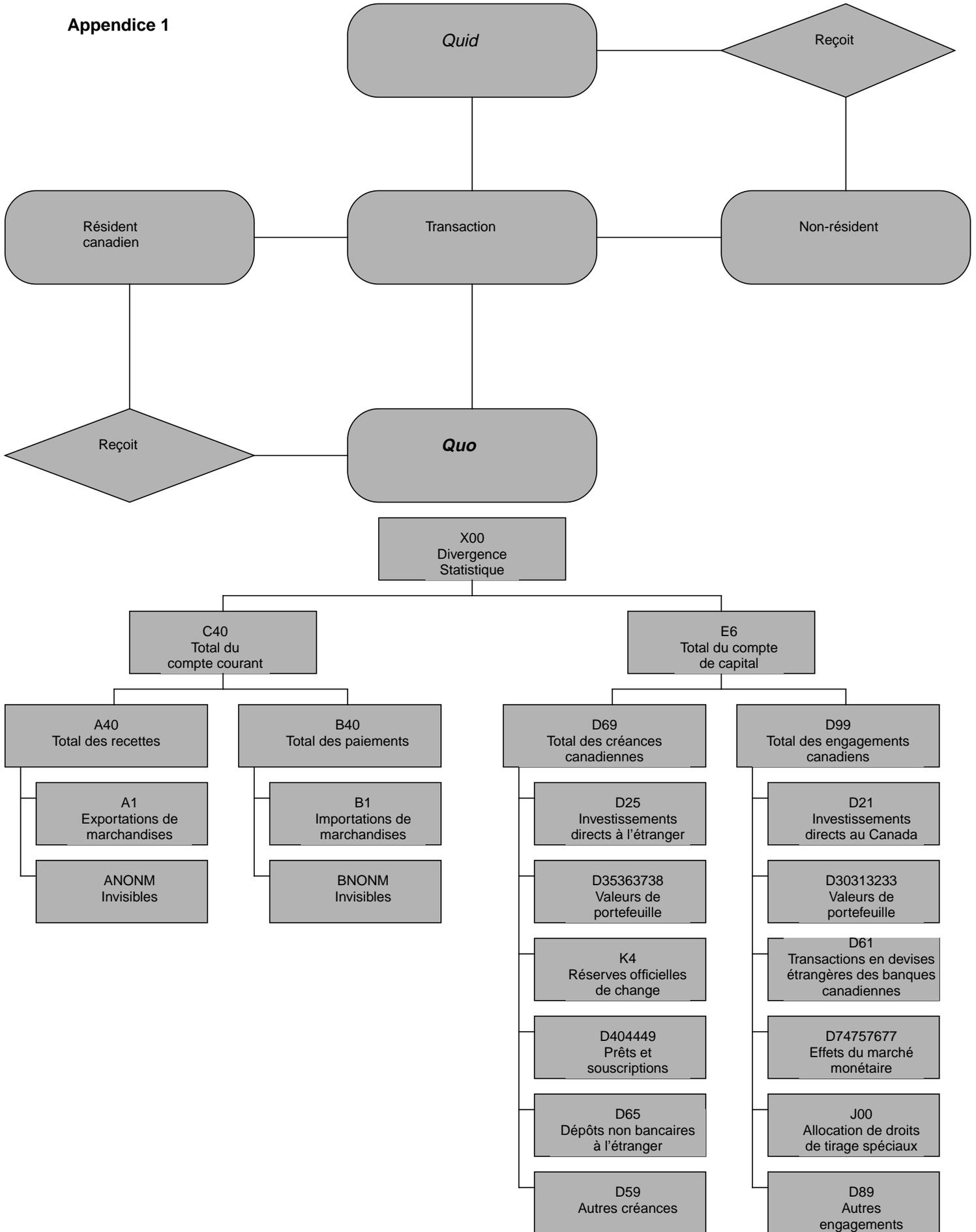
**Appendix 2**

Subordinate Data Bases			Main Data Bases		
Type of Data	Seasonality	Source of Data	Work	Draft	Published
Capital	Not Adjusted	Entered	Series	Series	Series
Capital	Not Adjusted	Derived	Formulae	Series	Series
Current	Not Adjusted	Entered	Series	Series	Series
Current	Not Adjusted	Derived	Formulae	Series	Series
Current	Adjusted	Entered	Series	Series	Series
Current	Adjusted	Derived	Formulae	Series	Series
Positions	N/A	Entered	Series	Series	Series
Positions	N/A	Derived	Formulae	Series	Series
Other	N/A	Entered	Series	Series	Series
Other	N/A	Derived	Formulae	Series	Series

**D21IN6.US**



**Appendice 1**



**Appendice 2**

Bases de données subordonnées			Bases de données principales		
Type de données	Fluctuations saisonnières	Source de données	Travaux	Ébauche	Publication
Capital	Non désaisonnalisées	Inscrite	Séries	Séries	Séries
Capital	Non désaisonnalisées	Dérivée	Formules	Séries	Séries
Courant	Non désaisonnalisées	Inscrite	Séries	Séries	Séries
Courant	Non désaisonnalisées	Dérivée	Formules	Séries	Séries
Courant	Désaisonnalisées	Inscrite	Séries	Séries	Séries
Courant	Désaisonnalisées	Dérivée	Formules	Séries	Séries
Positions	N/A	Inscrite	Séries	Séries	Séries
Positions	N/A	Dérivée	Formules	Séries	Séries
Autre	N/A	Inscrite	Séries	Séries	Séries
Autre	N/A	Dérivée	Formules	Séries	Séries

D21IN6.US

