

SONIA RIVEST

**Investigation des modes d'intégration physique entre un serveur
de base de données multidimensionnelle et un SIG**

Essai
présenté
à la Faculté des études supérieures
de l'Université Laval
pour l'obtention
du grade de maître ès sciences (M.Sc.)

Département des sciences géomatiques
FACULTÉ DE FORESTERIE ET DE GÉOMATIQUE
UNIVERSITÉ LAVAL

Décembre 2000

TABLE DES MATIÈRES

| | |
|--|------------|
| TABLE DES MATIÈRES | I |
| LISTE DES FIGURES..... | III |
| LISTE DES TABLEAUX | IV |
| 1 INTRODUCTION..... | 1 |
| 1.1 Mise en contexte..... | 1 |
| 1.2 Problématique..... | 1 |
| 1.3 Objectifs..... | 2 |
| 1.4 Méthodologie | 3 |
| 2 LE PROCESSUS D'ANALYSE INTERACTIF (OLAP) – REVUE DES CONCEPTS | 5 |
| 2.1 Définition et composantes d'OLAP | 5 |
| 2.2 Architectures OLAP | 6 |
| 2.2.1 Architecture ROLAP | 6 |
| 2.2.2 Architecture MOLAP..... | 7 |
| 2.2.3 Architecture HOLAP | 8 |
| 2.2.4 Comparaison des architectures ROLAP, MOLAP et HOLAP | 8 |
| 2.3 Structures de données multidimensionnelles pour bases de données relationnelles .. | 9 |
| 2.3.1 Schéma en étoile..... | 10 |
| 2.3.2 Schéma en flocon..... | 13 |
| 2.3.3 Autres schémas..... | 14 |
| 2.4 Opérations d'analyse OLAP..... | 16 |
| 3 LE PROJET OLAP DU MINISTÈRE DES TRANSPORTS DU QUÉBEC..... | 18 |
| 3.1 Description du projet..... | 18 |
| 3.2 Données utilisées | 18 |
| 3.2.1 Modélisation pour l'implantation relationnelle des données sémantiques..... | 18 |
| 3.2.2 Modélisation multidimensionnelle des données sémantiques | 19 |

| | | |
|----------|--|-----------|
| 3.2.3 | Données spatiales géométriques | 20 |
| 3.3 | Exemples d'analyses à produire | 21 |
| 4 | VERS UN PROCESSUS D'ANALYSE SPATIO-TEMPORELLE INTERACTIF (SOLAP). | 22 |
| 4.1 | Concepts de SOLAP | 22 |
| 4.2 | Caractéristiques d'un outil SOLAP idéal | 24 |
| 4.3 | Solutions technologiques potentielles | 30 |
| 4.3.1 | Solutions avec serveur OLAP et SIG | 30 |
| 4.3.2 | Solutions avec serveur OLAP et autre logiciel de visualisation | 34 |
| 4.3.3 | Solutions sans serveur OLAP et avec SIG | 34 |
| 4.3.4 | Solutions sans serveur OLAP et avec autre logiciel de visualisation | 35 |
| 5 | INTÉGRATION OLAP - SIG | 36 |
| 5.1 | Différentes possibilités d'intégration OLAP - SIG | 36 |
| 5.2 | Particularités des données du projet OLAP - MTQ | 39 |
| 5.3 | Considérations sur l'affichage cartographique | 39 |
| 5.4 | Développement d'une application d'essai | 41 |
| 5.5 | Solutions d'intégration OLAP - SIG implantées | 43 |
| 5.6 | Évaluation des solutions d'intégration implantées | 49 |
| 5.6.1 | Critères d'évaluation | 49 |
| 5.6.2 | Comparaison des solutions d'intégration | 50 |
| 5.7 | Application du mode d'intégration optimal aux données du projet OLAP - MTQ | 51 |
| 5.7.1 | Exemple d'analyse réalisée | 53 |
| 6 | CONCLUSION | 57 |
| 7 | REMERCIEMENTS | 58 |
| 8 | BIBLIOGRAPHIE | 59 |

LISTE DES FIGURES

| | |
|---|----|
| Figure 1 : Phases de l'essai | 4 |
| Figure 2 : Architecture ROLAP. Adaptée de Bourgon et <i>al.</i> [1999]..... | 6 |
| Figure 3 : Architecture MOLAP [Bourgon et <i>al.</i> 1999]..... | 7 |
| Figure 4 : Architecture HOLAP. Adaptée de Bourgon et <i>al.</i> [1999]..... | 8 |
| Figure 5 : Schéma en étoile du cube « accidents » du projet OLAP - MTQ | 12 |
| Figure 6 : Schéma en flocon du cube « accidents » du projet OLAP - MTQ | 13 |
| Figure 7 : Schéma en constellation des cubes « accidents » et « pentes » du projet OLAP - MTQ | 15 |
| Figure 8 : Schéma mixte du cube « accidents » du projet OLAP - MTQ | 15 |
| Figure 9 : Schéma d'une table contenant toutes les données du cube « accidents » du projet OLAP - MTQ | 16 |
| Figure 10 : Exemple de dimension « temps » en notation MDS..... | 20 |
| Figure 11 : Exemple d'interface SOLAP offert par ProClarity avec le plugiciel KMapX..... | 32 |
| Figure 12 : Exemple d'interface disponible à l'intérieur de Visualizer 1.5..... | 33 |
| Figure 13 : Interface de l'application d'essai développée. | 42 |
| Figure 14 : Exemple d'affichage cartographique réalisé avec la solution d'implantation 1 – une seule dimension spatiale géométrique. | 45 |
| Figure 15 : Exemple d'affichage cartographique réalisé avec la solution d'implantation 3 – plusieurs dimensions spatiales géométriques. | 48 |
| Figure 16 : Affichage des seuils d'intervention par route pour l'année 2000. | 53 |
| Figure 17 : Affichage des seuils d'intervention par sous-route pour l'année 2000. | 54 |
| Figure 18 : Affichage des seuils d'intervention par segment pour l'année 2000..... | 55 |
| Figure 19 : Affichage des seuils d'intervention par segment pour l'année 2002..... | 56 |

LISTE DES TABLEAUX

| | |
|--|----|
| Tableau 1 : Résumé des caractéristiques des architectures ROLAP, MOLAP et HOLAP | 9 |
| Tableau 2 : Comparaison des solutions d'intégration - une seule dimension spatiale géométrique | 50 |
| Tableau 3 : Comparaison des solutions d'intégration - plusieurs dimensions spatiales géométriques | 51 |

1 INTRODUCTION

1.1 Mise en contexte

Les organisations recueillent des volumes de plus en plus importants de données. Emmagasinées dans un entrepôt, celles-ci deviennent la base des processus d'analyse et orientent les décisions stratégiques. Toutefois, ces données ne sont pas toujours exploitées à leur plein potentiel et une partie de leur richesse est tout simplement laissée de côté. Cette richesse, parfois insoupçonnée, est apportée par la composante spatiale des données. « Dissimulée dans la plupart des données se trouve une composante géographique qui peut être liée à un lieu : une adresse, un code postal, (...), une ville, une région ou un pays » [Traduction libre, ESRI 2000]. En fait, il a été constaté qu'environ 80 pour cent des données comprennent une composante spatiale [Franklin 1992]. Toute implantation d'entrepôt de données devrait donc comprendre une dimension spatiale, en plus d'une dimension temporelle [Gonzales 1999]. Mais, pour tirer parti de ces deux dimensions, des outils dédiés à l'analyse spatio-temporelle devraient alors être incorporés aux processus d'analyse existants. Parmi les outils potentiels les plus connus figurent les systèmes d'information géographiques (SIG). Malgré les possibilités intéressantes offertes, les SIG seuls ne présentent pas l'efficacité requise par les applications analytiques et des solutions alternatives doivent être envisagées [Bédard et al. 2000], telles que l'OLAP.

1.2 Problématique

Il a été démontré que l'OLAP (« On-Line Analytical Processing »), ou processus d'analyse interactif¹, possède un potentiel réel pour supporter l'analyse spatio-temporelle [Caron 1998]. En effet, de par son architecture multidimensionnelle et son fonctionnement, un outil d'analyse OLAP permet d'effectuer des analyses spatio-temporelles avec facilité et rapidité, ces dernières étant des caractéristiques clés du processus. Cependant, pour exploiter tout le potentiel d'analyse spatio-temporelle de cette technologie, l'outil OLAP conventionnel, destiné jusqu'à récemment à l'analyse de données sémantiques exclusivement, doit aussi comporter un « volet cartographique » [Bédard 1997] permettant, au minimum, la visualisation cartographique des éléments spatiaux géométriques sur lesquels porte l'analyse. Des systèmes commerciaux, couplant des fonctionnalités OLAP et des fonctionnalités de systèmes d'information géographiques (SIG), ont été mis sur le marché récemment [BusinessObjects 1998, Cognos 2000, Knosys 2000]. Ces systèmes, certains OLAP dominants, certains SIG dominants [LGS

¹ L'acronyme OLAP sera utilisé tout au long du document afin d'alléger le texte.

Group Inc. 2000], offrent un sous-ensemble plus ou moins élaboré des fonctionnalités souhaitables d'une nouvelle famille d'outils OLAP, les outils OLAP spatiaux, ou SOLAP. Les systèmes commerciaux existants présentent cependant plusieurs limitations, soit au niveau de leurs modes de fonctionnement et d'utilisation, soit au niveau des capacités offertes. Par exemple, ces systèmes ne permettent la visualisation que d'une dimension spatiale géométrique à la fois, ce qui peut être suffisant dans certains cas, mais qui peut représenter un obstacle majeur à une pleine analyse spatio-temporelle et à la découverte de corrélations spatiales significatives. Un outil SOLAP idéal devrait donc présenter une flexibilité, non offerte jusqu'à présent, qui lui permettrait de s'adapter à tous les types de besoins d'analyse spatio-temporelle multidimensionnelle.

Un pas vers le développement d'un tel outil SOLAP est d'étudier les possibilités d'établir un lien dynamique entre un serveur OLAP et un SIG, ce dernier servant entre autres d'outil de visualisation et de navigation cartographique. L'intégration dynamique ferait en sorte qu'une opération effectuée dans le volet OLAP conventionnel serait automatiquement répercutée dans le volet cartographique et inversement. Les considérations théoriques et pratiques d'une telle intégration font l'objet du présent essai.

1.3 Objectifs

L'objectif principal de l'essai est d'analyser et de tester différents types d'architectures OLAP et de structures de données multidimensionnelles, afin de déterminer, parmi les solutions analysées, la meilleure façon d'établir un lien dynamique avec les bases de données spatiales d'un SIG.

Cet objectif comporte trois sous objectifs :

- analyser les possibilités de relier les architectures MOLAP (OLAP multidimensionnelle), ROLAP (OLAP relationnelle) et HOLAP (OLAP hybride) à la base de données d'un SIG;
- tester les modes d'intégration OLAP - SIG, dans le contexte du projet OLAP du Ministère des transports du Québec (MTQ), avec des structures de données en étoile, en flocon et multidimensionnelle propriétaire et à l'aide du logiciel SQL Server et d'un SIG à déterminer;
- parmi les solutions testées, sélectionner la meilleure selon les critères définis, l'implanter et la tester concrètement dans un système fonctionnel.

1.4 Méthodologie

La réalisation de l'essai a été divisée en cinq principales phases.

Phase 1. Une étude comparative des structures de données multidimensionnelles (cube multidimensionnel, schéma en étoile et schéma en flocon), ainsi que des différentes architectures OLAP, a été effectuée. Les résultats de cette étude ont été utilisés lors de l'évaluation des solutions développées à la phase 4 d'élaboration et d'essai de solutions d'intégration OLAP - SIG.

Phase 2. En parallèle avec la première phase, une recherche de systèmes existants intégrant des fonctionnalités OLAP et SIG a été réalisée. Cette recherche avait pour but de présenter les caractéristiques, avantages et limitations de solutions actuellement disponibles sur le marché, par rapport aux caractéristiques souhaitables d'un outil SOLAP idéal.

Phase 3. Une application d'essai, destinée à tester les différentes solutions d'intégration élaborées, a été développée en langage Visual Basic à l'aide des bibliothèques d'objets du logiciel ProClarity 2.0 de Knosys et de GeoMedia 4 d'Intergraph. Cette phase comprenait aussi une investigation sommaire des modes d'affichage des données, selon le nombre de dimensions sémantiques et spatiales géométriques visualisées.

Phase 4. Différentes solutions d'intégration OLAP - SIG furent élaborées. Ces solutions ont été testées, à l'aide d'un sous-ensemble des données du cube d'accidents du projet OLAP du MTQ emmagasinées dans SQL Server, dans le contexte des différentes architectures OLAP (MOLAP, ROLAP et HOLAP). Dans le cas des architectures ROLAP et HOLAP, les structures de données multidimensionnelles en étoile et en flocon furent utilisées. Les essais ont été réalisés sur un Pentium II, 300 MHz, 288 MO RAM.

Phase 5. Les solutions développées lors de la phase 4 ont été évaluées en fonction de critères définis par l'auteur dans le cadre des expérimentations, selon le processus de mise en œuvre des solutions d'intégration et à partir des caractéristiques d'un outil SOLAP idéal. La meilleure solution développée dans le cadre de l'essai a été appliquée aux données du projet OLAP - MTQ afin de produire un système fonctionnel.

La figure 1 présente schématiquement les différentes phases de l'essai ainsi que le lien entre elles.

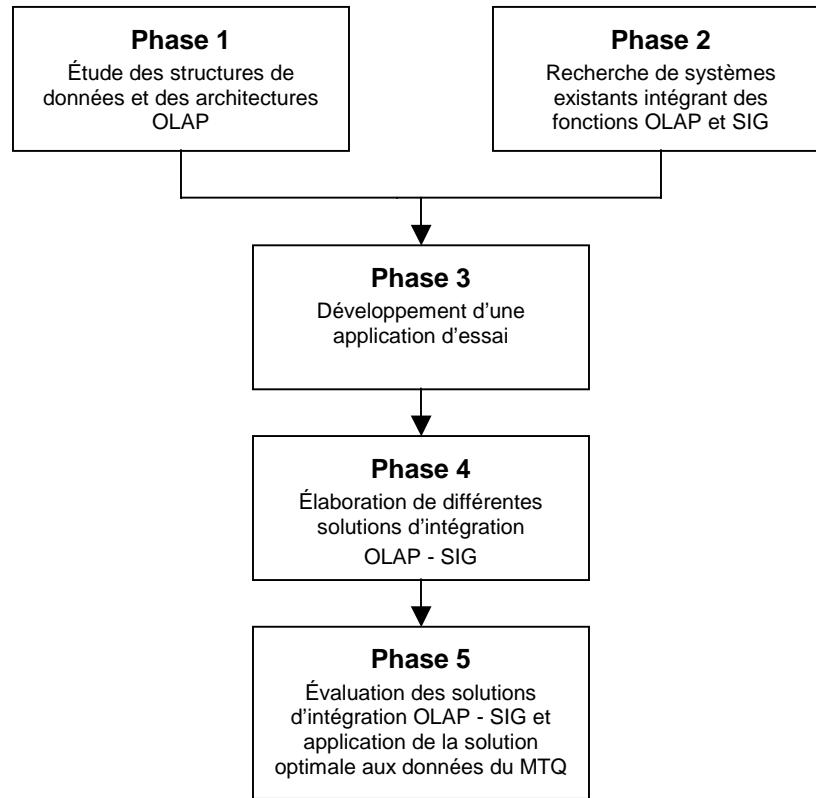


Figure 1 : Phases de l'essai

La première partie du document a introduit le contexte et les objectifs de l'essai. Les sections suivantes en présentent la théorie, les détails de la réalisation et les résultats. La section 2 contient une brève revue des concepts d'OLAP ainsi qu'une comparaison de ses différentes architectures et structures multidimensionnelles. La section 3 décrit le projet OLAP du MTQ et les données utilisées dans le cadre de l'essai. Par la suite, la définition et les caractéristiques souhaitables d'un outil SOLAP sont exposées. La section 5 décrit les solutions d'intégration OLAP - SIG analysées, développées et testées. De plus, cette section contient les résultats de l'évaluation des solutions et la description de la solution optimale. La discussion et la conclusion suivent, à la section 6.

2 LE PROCESSUS D'ANALYSE INTERACTIF (OLAP) – REVUE DES CONCEPTS

2.1 Définition et composantes d'OLAP

Cette section consiste en une brève révision des concepts d'OLAP. Pour plus de détails, le lecteur est invité à consulter des références de base sur le sujet [Codd et *al.* 1993, Olap Council 1995, Thomsen 1997].

OLAP est défini comme « une catégorie de logiciels axés sur l'exploration et l'analyse rapides des données selon une approche multidimensionnelle à plusieurs niveaux d'agrégation » [Caron 1998]. L'architecture d'un système OLAP est constituée de trois services : la base de données, le serveur et le module client [Bédard et *al.* 1997].

La base de données d'une architecture OLAP doit posséder une structure multidimensionnelle. Celle-ci peut être implantée à l'intérieur d'un système de gestion de base de données (SGBD) relationnel (et objet-relationnel) ou multidimensionnel. Lorsqu'un SGBD relationnel est utilisé, les données doivent être structurées selon une approche multidimensionnelle en utilisant des schémas particuliers. Il s'agit principalement des schémas en étoile et en flocon, qui seront décrits aux sections 2.3.1 et 2.3.2 respectivement. Lorsqu'un SGBD multidimensionnel est utilisé, les données doivent être chargées dans la structure multidimensionnelle propriétaire à partir d'une base de données relationnelle ou d'un fichier texte. La base de données d'une architecture OLAP doit pouvoir supporter des données agrégées.

Le serveur OLAP est le logiciel qui gère la base de données OLAP et l'accès des utilisateurs à celle-ci. Il comprend un engin de calcul et permet habituellement à plusieurs utilisateurs concurrents d'effectuer des requêtes de type agrégatif sur la base de données. Le serveur OLAP est conçu spécifiquement pour manipuler des données structurées de façon multidimensionnelle.

Le module client est un logiciel d'accès, de manipulation et d'exploration des données de nature multidimensionnelle. Il permet d'effectuer des requêtes, d'obtenir des données du serveur OLAP et de les visualiser en deux ou trois dimensions. Le module client propose habituellement plusieurs modes de visualisation, allant du simple tableau aux différents types de diagrammes tels que les diagrammes à barres ou les diagrammes circulaires, pour ne nommer que les plus courants.

2.2 Architectures OLAP

Le type de base de données gérée par le serveur OLAP définit le type de serveur et ainsi le type d'architecture OLAP. Il existe trois principales architectures : ROLAP, MOLAP et HOLAP. La présente section contient un survol des caractéristiques importantes de ces architectures.

2.2.1 Architecture ROLAP

L'architecture ROLAP (« Relational OLAP » ou OLAP relationnelle) utilise une base de données relationnelle, ou objet-relationnelle, pour le stockage des données de base et des agrégations destinées à l'analyse. La base de données relationnelle doit être structurée suivant un schéma plus ou moins dénormalisé afin de faciliter les processus d'analyse. Pour ce faire, les schémas en étoile et en flocon sont les plus utilisés. Ces schémas seront décrits en détail aux sections 2.3.1 et 2.3.2. Le serveur ROLAP extrait les données de la base de données à l'aide de requêtes SQL [Bédard et *al.* 1997], les interprète selon une vue multidimensionnelle, et les transfère au module client. La figure 2 présente le fonctionnement de l'architecture ROLAP.

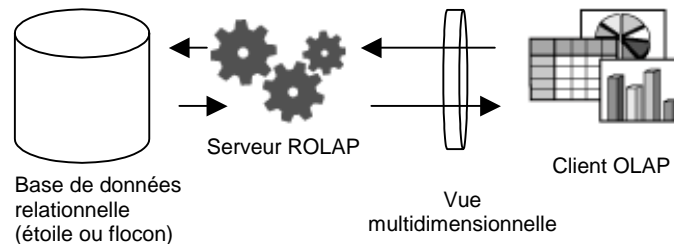


Figure 2 : Architecture ROLAP. Adaptée de Bourgon et *al.* [1999]

Cette architecture, du fait que toutes les données soient emmagasinées dans une base de données relationnelle mieux adaptée aux processus transactionnels [Bédard et Proulx 1998], présente généralement les temps de réponse les plus longs [Thomsen et *al.* 1999]. Les partisans de cette architecture affirment cependant qu'elle permet d'éviter d'emmagasiner de grandes quantités de données de base dans une structure parallèle et de manière redondante, comme le fait l'architecture MOLAP [Hettler 1997]. En pratique, certains suggèrent de n'utiliser l'architecture ROLAP que dans le cas de très grands volumes de données rarement exploités et pour lesquels la performance d'interrogation n'est pas un facteur critique [Microsoft Corporation 2000].

2.2.2 Architecture MOLAP

L'architecture MOLAP (« Multidimensional OLAP » ou OLAP multidimensionnelle) emmagasine les données de base ainsi que les données agrégées dans une base de données multidimensionnelle possédant une structure propriétaire optimisée pour l'analyse, souvent nommée cube ou hypercube. Selon la requête effectuée, le serveur MOLAP extrait les données directement de l'hypercube et les présente à l'utilisateur via le module client. La figure 3 illustre le fonctionnement de l'architecture MOLAP.

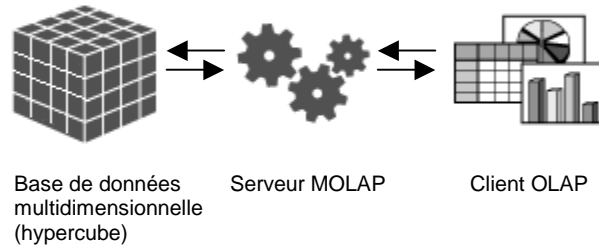


Figure 3 : Architecture MOLAP [Bourgon et al. 1999]

Cette approche est généralement la plus performante puisque les données, à l'intérieur de la base de données multidimensionnelle, sont organisées d'une manière similaire à la manière dont elles seront utilisées lors de l'analyse [Thomsen et al. 1999], c'est-à-dire sous forme de matrices. Certains pourraient penser que le fait d'emmagasiner toutes les données de base dans une nouvelle structure augmente de beaucoup l'espace de stockage nécessaire. En fait, des technologies telles que le « *sparse-matrix technology* » permettent d'éliminer les cellules vides des matrices et de compresser ces dernières afin de réduire l'espace requis au minimum [Gill and Rao 1996]. Des tests ont effectivement démontré qu'en utilisant ce genre de technologie, l'architecture MOLAP nécessite un espace de stockage additionnel moins élevé que l'architecture ROLAP pour laquelle les agrégations sont emmagasinées dans la base de données relationnelle [Thomsen et al. 1999]. L'architecture MOLAP demande toutefois une maintenance plus élevée que les autres architectures puisque des modifications aux données de base nécessitent leur rechargement dans la base de données multidimensionnelle et le recalcul des agrégations, tandis que seulement ce recalcul est requis dans le cas des autres architectures.

2.2.3 Architecture HOLAP

L'architecture HOLAP (« Hybrid OLAP » ou OLAP hybride) constitue un croisement des deux architectures ROLAP et MOLAP présentées précédemment aux sections 2.2.1 et 2.2.2 respectivement. Au sein de cette architecture, un serveur HOLAP accède deux bases de données différentes. La première, multidimensionnelle propriétaire, contient les données agrégées et agit comme point de départ pour l'analyse. La seconde, relationnelle, contient les données détaillées de base. Lorsqu'accédées, les données détaillées sont transformées selon une vue multidimensionnelle avant d'être transférées au module client. La figure 4 présente le fonctionnement de l'architecture HOLAP.

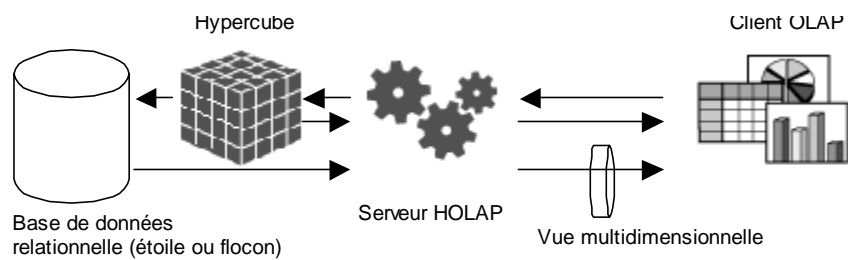


Figure 4 : Architecture HOLAP. Adaptée de Bourgon et al. [1999]

L'architecture HOLAP hérite des caractéristiques des architectures ROLAP et MOLAP selon le type de données accédées par les requêtes, ROLAP dans le cas des données de base et MOLAP dans le cas des données agrégées. Cette architecture est particulièrement utile lorsque la majeure partie des requêtes touchent des données agrégées et que les données détaillées représentent un volume important [Microsoft Corporation 2000, Tiedrich 2000]. Comme seulement les données agrégées sont emmagasinées dans la base de données multidimensionnelle, les données de base demeurant dans la base de données relationnelle, cette architecture est celle qui demande l'espace de stockage supplémentaire le plus faible, à condition d'utiliser les technologies de compression citées en 2.2.2.

2.2.4 Comparaison des architectures ROLAP, MOLAP et HOLAP

Le tableau 1 résume les caractéristiques de chacune des architectures OLAP. Ces caractéristiques seront prises en compte lors de l'évaluation des solutions d'intégration OLAP - SIG à la section 5.6.

Tableau 1 : Résumé des caractéristiques des architectures ROLAP, MOLAP et HOLAP

| Critère de comparaison | ROLAP | MOLAP | HOLAP |
|---|------------------|---|---|
| Stockage des données de base (détaillées) | BD relationnelle | BD multidimensionnelle | BD relationnelle |
| Stockage des agrégations | BD relationnelle | BD multidimensionnelle | BD multidimensionnelle |
| Performance des requêtes (en ordre croissant de performance, 3 correspondant à la meilleure performance) | 1 | 3 | 2 |
| Espace de stockage nécessaire (en ordre croissant d'espace supplémentaire nécessaire, 3 correspondant au plus grand volume) | 3 | 2 (si des technologies de compression sont utilisées) | 1 (si des technologies de compression sont utilisées) |
| Maintenance de la structure (en ordre croissant du niveau de maintenance nécessaire) | 1 | 2 | 1 |

2.3 Structures de données multidimensionnelles pour bases de données relationnelles

Cette section décrit et compare les schémas utilisés pour fournir une structure multidimensionnelle aux données emmagasinées dans une base de données relationnelle.

Les bases de données relationnelles destinées aux opérations transactionnelles sont habituellement normalisées de telle sorte que la redondance des données est à son minimum. Cette approche est particulièrement utile pour maintenir l'intégrité de la base de données et simplifier sa mise à jour. Cependant, une très forte normalisation des données complexifie l'analyse de ces dernières car la base de données résultante présente un nombre élevé de tables. Un nombre élevé de joints entre les tables est donc nécessaire, lors de l'exécution d'une requête, afin d'obtenir la totalité de l'information désirée, ce qui augmente la complexité de la requête et en diminue la performance [Sterling 1998]. Pour cette raison, lorsque des besoins transactionnels et analytiques doivent être comblés à partir des mêmes données de base, il est de mise d'utiliser deux bases de données différentes, la première dédiée aux opérations transactionnelles et la

seconde dédiée aux activités d'analyse [Inmon 1996]. À l'intérieur d'une base de données relationnelle destinée à l'analyse, la dénormalisation des données est utilisée afin de limiter le nombre de tables et le nombre de joints nécessaires entre les différentes tables, facilitant ainsi l'élaboration des requêtes et améliorant la rapidité d'exécution. De plus, afin de diminuer les calculs à effectuer lors de l'exécution de requêtes, et donc d'améliorer la performance, les bases de données analytiques contiennent une certaine quantité de données dérivées précalculées.

Au sein des architectures ROLAP ou HOLAP, la base de données relationnelle doit donc être modélisée d'une manière dénormalisée afin de d'optimiser les processus d'analyse. Les modèles de données utilisés à cette fin, c'est-à-dire les schémas en étoile et en flocon, sont habituellement basés sur les concepts suivants : les dimensions, les mesures et les faits, qui définissent la structure logique d'une base de données OLAP.

Une dimension peut être définie comme un axe, ou un thème, selon lequel les données seront analysées. Le temps et le découpage administratif du territoire peuvent être des exemples de dimensions. Une dimension contient des membres organisés en hiérarchie [Thomsen et al. 1999], chacun des membres appartenant à un niveau hiérarchique, c'est-à-dire présentant un niveau de granularité particulier. Pour la dimension « temps », les années, les mois et les jours sont des exemples de niveaux hiérarchiques et « 1998 » est un exemple de membre appartenant au niveau « année ».

Une mesure est un élément de donnée numérique sur lequel portent les analyses, en fonction des différentes dimensions [Thomsen et al. 1999]. Par exemple, dans un contexte d'analyse de données de transport, le coût des travaux et le nombre d'accidents sont des mesures possibles.

Le fait représente la valeur d'une mesure, mesurée ou calculée, selon un membre de chacune des dimensions. Par exemple, « le coût des travaux en 1995 pour la région 02 est 250 000 \$ » est un fait qui exprime la valeur de la mesure « coût des travaux » pour le membre « 1995 » du niveau « année » de la dimension « temps » et le membre « 02 » du niveau « région » de la dimension « découpage administratif ».

2.3.1 Schéma en étoile

Le schéma en étoile tire son nom de sa configuration. En effet, un objet central, nommé « table des faits », est connecté à un certain nombre d'objets de manière radiale, les « tables de dimension » [Gill and Rao, 1996]. Le schéma en étoile présente une approche intuitive de

représentation des données, similaire à la vue qu'un analyste a de ces données et des requêtes à effectuer sur ces données.

La table des faits :

- contient des données quantitatives de base à analyser, les faits (*cf.* section 2.3), et pouvant être analysés selon les différentes perspectives que sont les dimensions;
- contient autant de clés étrangères qu'il y a de dimensions, c'est-à-dire une clé étrangère pour chacune des tables de dimension;
- peut contenir des données de différentes granularités [Archer Decision Sciences 1995], cette dernière étant définie comme « le niveau de détail (...) des données stockées dans une base de données » [Bédard et al. 1997];
- est dénormalisée.

Les tables de dimension :

- contiennent les attributs définissant chacun des membres des niveaux hiérarchiques d'une dimension;
- contiennent une clé primaire;
- sont dénormalisées.

La figure 5 montre le schéma en étoile du cube « accidents » du projet OLAP - MTQ. Autour de la table des faits se retrouvent huit tables de dimensions. Les dimensions et les mesures de ce cube seront décrites aux annexes 2 et 3. Le schéma présenté ici ne montre que le schéma théorique du cube, et non la manière optimale de l'implanter à l'intérieur de la base de données. Une table de dimension ne contenant aucun attribut autre que la clé primaire, telle que par exemple la table de la dimension « temps » de la figure, ne nécessite effectivement pas de matérialisation dans la base de données. La discussion des solutions d'implantation optimales des données du projet MTQ n'est pas pertinente ici et elle sera présentée dans un rapport de projet à venir.

Le schéma de la figure 5 présente donc la table des faits contenant 5 mesures : le nombre d'accidents, le nombre de véhicules impliqués, le nombre de blessés léger, le nombre de blessés grave ainsi que le nombre de morts. Chaque enregistrement de la table des faits correspond à un enregistrement de chacune des tables de dimension. Par exemple, pour l'année 1995, la pente 1234, la courbe 2345, la sous-route 22, l'accident 123, la structure b234, l'identifiant de circulation 345 et le numéro de centre de service 3476 (ces dernières données correspondant à un enregistrement dans chaque table de dimension), on retrouve les valeurs de mesures suivantes : un accident, deux véhicules impliqués, aucun blessé léger, un blessé grave et aucun mort.

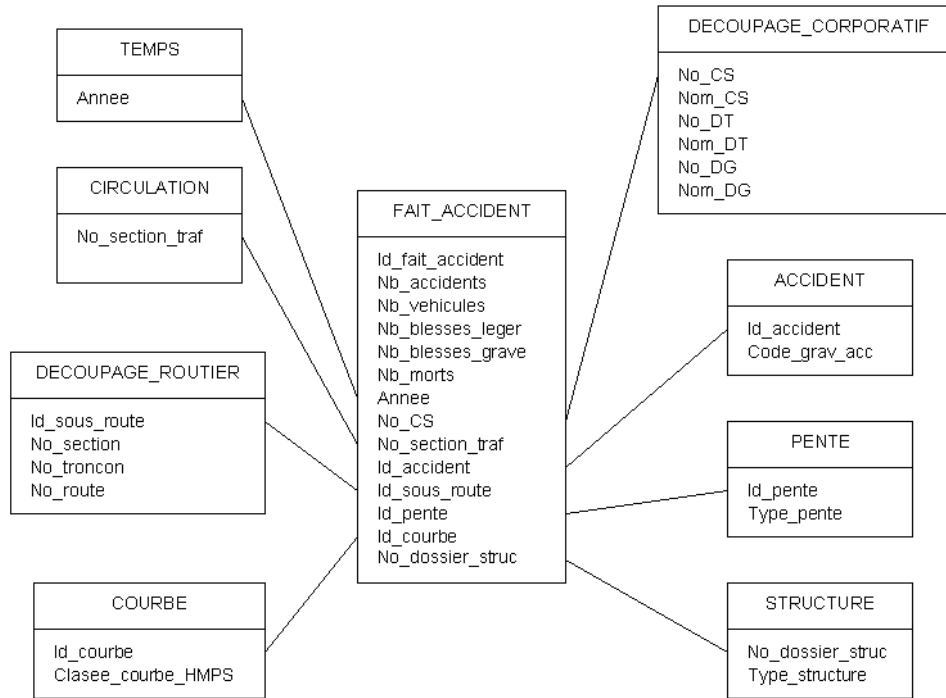


Figure 5 : Schéma en étoile du cube « accidents » du projet OLAP - MTQ

La table des faits est habituellement triée en ordre chronologique, puis à l'intérieur de cet ordre, selon un ordre dépendant des analyses les plus courantes. Cette table est habituellement volumineuse.

Avantages du schéma en étoile :

- navigation facile entre les tables par rapport à la structure normalisée du modèle conceptuel, placé à l'annexe 1 pour permettre la comparaison, et par rapport au schéma en flocon;
- réduction du nombre de tables et donc du nombre de joints entre les tables nécessaires pour effectuer des requêtes. L'élaboration de requêtes est donc relativement simple;
- modèle de données simple car comportant un nombre réduit de tables;
- volume réduit de métadonnées et maintenance relativement facile car le nombre de tables et le nombre de liens entre elles sont restreints.

Désavantage du schéma en étoile :

- même si peu de joints sont nécessaires pour effectuer les requêtes, les tables à joindre peuvent être volumineuses, ce qui peut affecter la performance des joints.

2.3.2 Schéma en flocon

Le schéma en flocon est dérivé du schéma en étoile où les tables de dimension sont normalisées. Avec ce schéma, chacune des dimensions est décomposée selon sa (ou ses) hiérarchie(s). L'expression « hiérarchies alternatives » est utilisée pour désigner les hiérarchies différentes et parallèles d'une même dimension. Par exemple, la dimension « structure » présentée à l'annexe 2 contient 4 hiérarchies alternatives, ayant en commun le niveau des structures. Pour le projet OLAP - MTQ, le schéma en flocon du cube « accidents » est illustré à la figure 6. Les dimensions et les mesures de ce cube seront décrites aux annexes 2 et 3. Ici encore, la figure présente le schéma théorique et non le schéma d'implantation optimale.

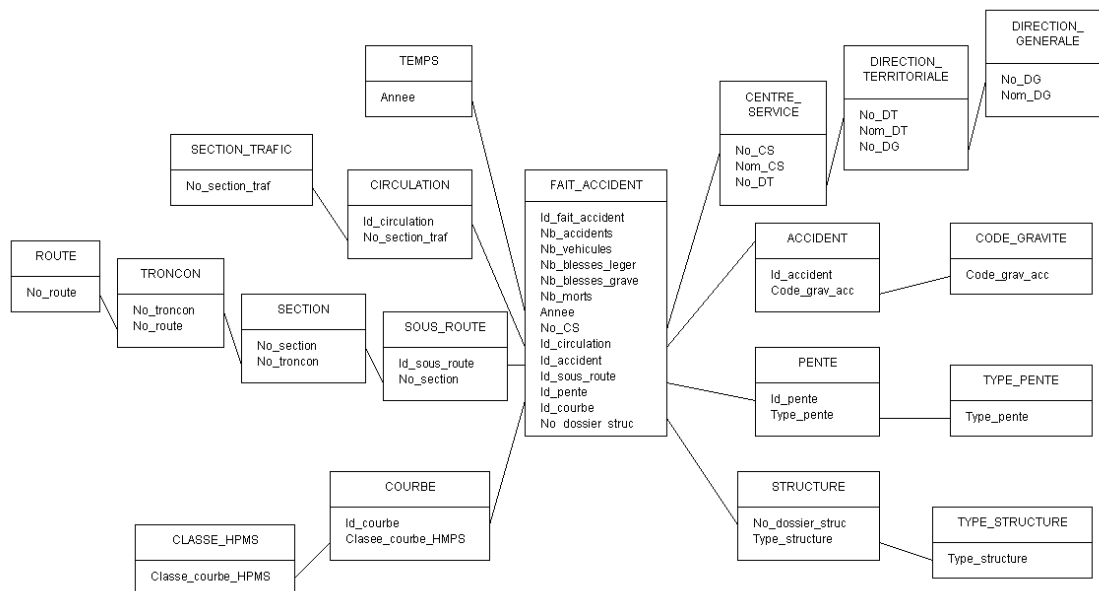


Figure 6 : Schéma en flocon du cube « accidents » du projet OLAP - MTQ

Avantage du schéma en flocon :

- minimise le volume des tables à joindre lors de l'exécution des requêtes par rapport au schéma en étoile.

Désavantages du schéma en flocon :

- modèle de données plus complexe que le schéma en étoile car comportant un nombre plus élevé de tables;
- volume plus élevé de métadonnées et maintenance plus complexe que le schéma en étoile car le nombre de tables et le nombre de relations entre elles sont plus élevés;

- même si le volume des tables à joindre lors de l'exécution des requêtes est bas, le nombre de joints nécessaires est plus élevé que pour le schéma en étoile. L'élaboration de requêtes est donc plus complexe.

Certains affirment que le schéma en flocon ne devrait être utilisé que lorsque les tables de dimension contiennent un très grand volume de données [Rinn 1998], la performance des joints étant alors améliorée en décomposant chaque table de dimension en plusieurs tables de volume moindre, chacune contenant les données relatives à un niveau hiérarchique de la dimension. Le choix du schéma doit donc être fonction d'un équilibre entre le nombre de tables et le volumes des tables qui seront jointes lors des analyses.

Des variations des deux schémas présentés plus haut existent, permettant d'optimiser la performance d'applications spécifiques. Ces variations sont présentées à la section 2.3.3.

2.3.3 Autres schémas

Schéma en constellation

Le schéma en constellation est dérivé du schéma en étoile et contient plusieurs tables des faits partageant ou non des tables de dimension [Bédard et *al.* 2000]. La figure 7 présente un schéma en constellation des cubes « accidents » et « pentes » du projet OLAP - MTQ.

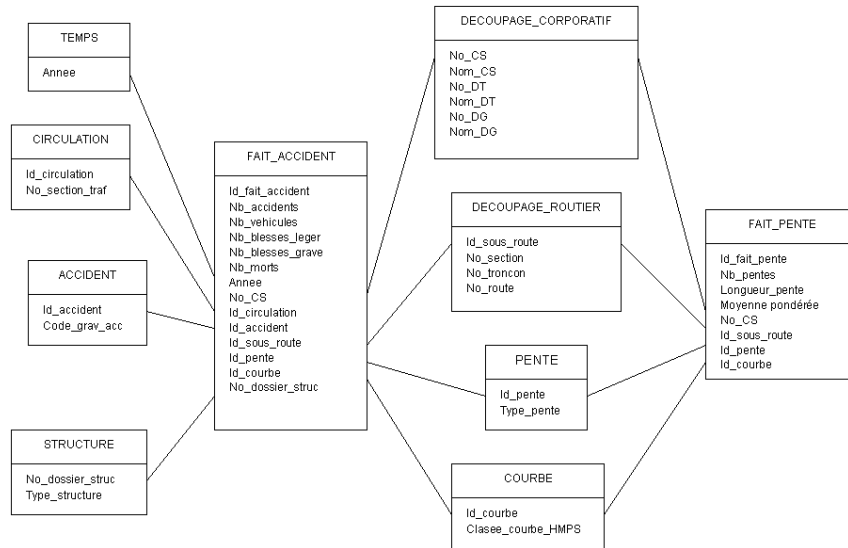


Figure 7 : Schéma en constellation des cubes « accidents » et « pentes » du projet OLAP - MTQ

Schéma mixte

Le schéma mixte, comme son nom l'indique, est un croisement entre le schéma en étoile et le schéma en flocon où seulement les tables de dimension contenant le plus grand volume de données, et le plus de redondance, sont normalisées [Gill and Rao 1996]. La figure 8 présente un exemple de schéma mixte pour le cube « accidents ».

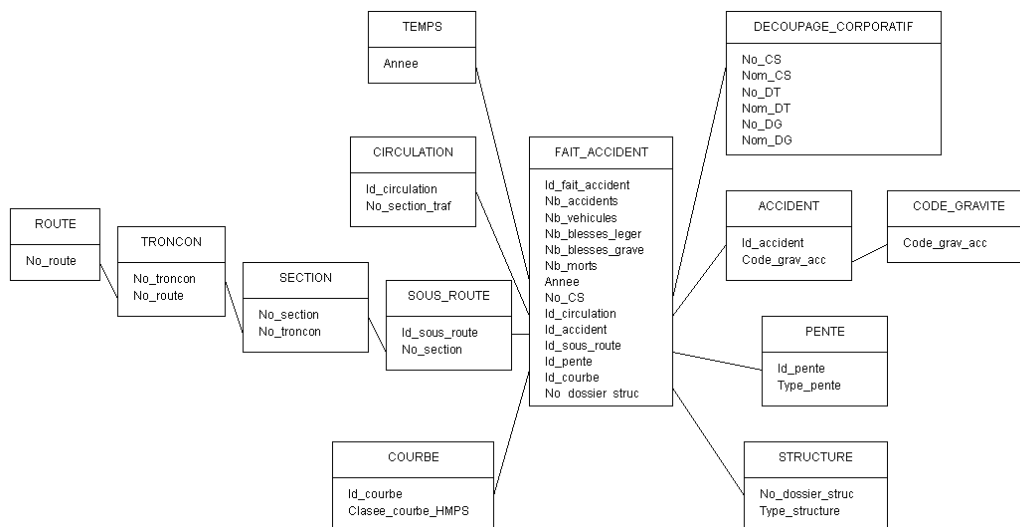


Figure 8 : Schéma mixte du cube « accidents » du projet OLAP - MTQ

Une seule table

Lorsque le volume de données est petit, il est possible d'emmagasiner les données de la table des faits, ainsi que les données des tables de dimension, à l'intérieur d'une même table, ce qui présente le maximum de redondance, mais qui ne nécessite aucun joint entre tables. La figure 9 présente un exemple de schéma de table contenant les données de la table des faits et des tables de dimension.

| FAIT_ACCIDENT |
|-----------------------|
| Id_fait_accident |
| Nb_accidents |
| Nb_vehicules |
| Nb_blesses_leger |
| Nb_blesses_grave |
| Nb_morts |
| Annee |
| No_CS |
| Nom_CS |
| No_DT |
| Nom_DT |
| No_DG |
| Nom_DG |
| Id_circulation |
| No_section_traf |
| Id_accident |
| Code_gravité_accident |
| Id_sous_route |
| No_section |
| No_troncon |
| No_route |
| Id_pente |
| Type_pente |
| Id_courbe |
| Classe_HPMS |
| No_dossier_struc |

Figure 9 : Schéma d'une table contenant toutes les données du cube « accidents » du projet OLAP - MTQ

2.4 Opérations d'analyse OLAP

À l'intérieur d'un module client OLAP, l'analyse s'effectue principalement à partir des opérations « pivoter », « forer », « forer latéralement » et « remonter ». Celles-ci seront brièvement décrites. Pour des exemples concrets d'utilisation de ces opérations, le lecteur est invité à consulter le mémoire de Caron [1998].

L'opération « pivoter » consiste à interchanger deux dimensions de manière à permettre, ou à restreindre, l'utilisation des autres opérations sur ces dimensions. Comme toutes les dimensions

ne peuvent être visualisées à la fois, cette opération permet de modifier les axes de l'analyse en cours.

L'opération « forer » consiste à passer d'un niveau de granularité donné à l'intérieur d'une dimension au niveau de granularité immédiatement inférieur dans la même dimension.

L'opération « forer latéralement » consiste, en fait, en trois opérations différentes :

- passer d'une mesure à une autre;
- passer d'un niveau hiérarchique alternatif à un autre à l'intérieur d'une dimension;
- passer d'une strate à une autre, c'est-à-dire, changer de membre analysé dans une dimension qui n'est pas visualisée.

L'opération « remonter » consiste à passer d'un niveau de granularité donné à l'intérieur d'une dimension au niveau de granularité immédiatement supérieur dans la même dimension.

Certains modules client offrent aussi des outils de requête ad-hoc, permettant d'interroger directement la base de données OLAP à l'aide d'un langage d'interrogation.

3 LE PROJET OLAP DU MINISTÈRE DES TRANSPORTS DU QUÉBEC

3.1 Description du projet

Le Service des orientations stratégiques (SOS) du Ministère des transports du Québec (MTQ) utilise actuellement plusieurs bases de données à partir desquelles il effectue une gamme de traitements et d'analyses. Comme ces bases de données ont été conçues de manière indépendante et pour des applications spécifiques au départ, des problèmes d'intégration et de performance sont rencontrés lorsque des analyses utilisant l'ensemble des données contenues dans les différentes bases de données sont effectuées.

Le SOS a mandaté le Centre de recherche en géomatique de l'Université Laval afin de réaliser une étude comportant deux volets. Le premier volet portait sur l'analyse de la structure actuelle des différentes bases de données routières utilisées. Le second volet portait sur la proposition d'une nouvelle structure globale pour répondre aux différents besoins d'analyse et de gestion et ce, dans une optique multidimensionnelle.

La réalisation des deux volets a été complétée et un modèle de classes d'objets, ainsi qu'un dictionnaire complet de données ont été produits [Lemieux 2000].

Dans le cadre d'un second mandat accordé par le SOS, des données couvrant une zone pilote ont été intégrées dans la nouvelle structure.

Les étapes subséquentes du projet visaient à démontrer, à l'aide de prototypes conçus avec des outils OLAP et des outils SIG, les possibilités d'analyses, d'agrégations de résultats et d'intégration entre outils OLAP et SIG.

3.2 Données utilisées

3.2.1 Modélisation pour l'implantation relationnelle des données sémantiques

Le modèle initial de classes d'objet du projet du MTQ [Lemieux 2000] a été réalisé à l'aide de Perceptory, un outil de modélisation visuelle de bases de données développé à partir du formalisme UML (« Unified Modeling Language ») et enrichi pour la référence spatiale en tenant compte de la norme ISO/TC 211 [Bédard 2000b]. Il a été reproduit à l'annexe 1. Ce modèle a subi quelques modifications mineures depuis sa première implantation.

Le modèle de classes d'objets réalisé est totalement compatible avec le système de référence linéaire RTSC : Route – Tronçon – Section – Chaînage dont le MTQ s'est doté afin de positionner des objets ou des événements en relation avec la géométrie du réseau routier, ce dernier étant évidemment à la base des différentes analyses. Il permet la segmentation des sections du réseau routier, chacune de ces sections étant uniforme par rapport aux entités de découpage territorial de base (municipalité, centre de service, circonscription électorale et division de recensement), en fonction des caractéristiques suivantes [Lemieux 2000] :

- milieu (rural, semi-urbain, urbain);
- nombre de voies;
- largeur et type de chaussée;
- largeur et type d'accotement gauche et droit;
- largeur d'emprise;
- type de terrain;
- ponts et autres structures;
- sections de trafic (indicateurs de circulation);
- classes de pente et courbe;
- état de chaussée (indice de rugosité et ornières);
- travaux programmés et réalisés.

Le processus de segmentation utilisant tous les critères ci-dessus produit les « segments atomiques ». Un segment atomique représente la plus petite unité uniforme de réseau routier. En fonction du type d'analyse à réaliser, une agrégation de segments atomiques peut s'effectuer selon certains critères particuliers pour produire les « segments d'analyse ».

Les données utilisées dans le cadre du présent essai consistent en un sous-ensemble des données maintenant implantées dans la structure décrite par le modèle de classes et couvrent deux directions territoriales, unités territoriales administratives du MTQ, celle de Québec (39) et celle de Chaudières-Appalaches (34). Ces données sont emmagasinées dans une base de données SQL Server.

3.2.2 Modélisation multidimensionnelle des données sémantiques

Suite à des discussions avec les analystes en transport, les dimensions et les mesures nécessaires aux analyses de nature multidimensionnelle ont été définies. Celles-ci ont été regroupées en différents cubes, un cube étant ici vu comme un ensemble de mesures organisées par un ensemble de dimensions [Thomsen et *al.* 1999] et qui sera physiquement implanté dans une base de données ayant une structure multidimensionnelle en vue d'une utilisation à partir

d'un outil OLAP. Les différents cubes ont été définis en fonction des axes d'analyse qui seront brièvement présentés à la section 3.3.

Ces dimensions, mesures et cubes ont été définis en dehors du cadre du présent essai. Cependant, leur description [Proulx et Bernier 2000] est reproduite aux annexes 2 et 3 car plusieurs références y seront faites au cours de sections subséquentes. L'annexe 2 présente l'ensemble des dimensions définies pour le projet. Ce document utilise la notation MDS « Multidimensional Domain Structure » [Thomsen 1998], qui représente une dimension comme un axe avec ses différents niveaux hiérarchiques. Le niveau détaillé est placé à la base et les niveaux hiérarchiques supérieurs, ou niveaux d'agrégation, lui sont superposés. La figure 10 montre un exemple de dimension « temps » représentée à l'aide de la notation MDS.

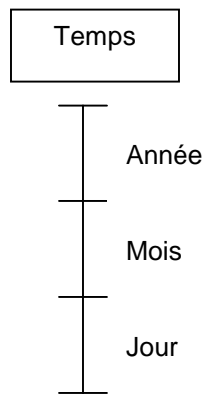


Figure 10 : Exemple de dimension « temps » en notation MDS

L'annexe 3 contient un tableau décrivant les mesures et les dimensions formant chacun des cubes.

3.2.3 Données spatiales géométriques

Les données spatiales géométriques utilisées dans le cadre de l'essai comprennent principalement les données linéaires représentant le réseau routier géré par le MTQ, pour les directions territoriales de Québec et de Chaudières-Appalaches, compatibles avec le système de référence RTSC.

Une segmentation atomique a été réalisée, au départ, sur les éléments linéaires afin de les préparer aux essais des différentes solutions d'intégration OLAP - SIG. L'utilisation particulière des segments de réseau routier sera décrite pour chacune des solutions d'intégration implantées, à la section 5.5.

Des données surfaciques, représentant les différents découpages territoriaux utilisés par le MTQ, ont été employées comme données de contexte dans certains cas.

Toutes les données géométriques utilisées sont en format TAB de MapInfo.

3.3 Exemples d'analyses à produire

Les éléments routiers généralement analysés se regroupent selon deux axes d'intervention : l'axe de conservation des infrastructures et l'axe de fonctionnalité. L'axe de conservation comprend l'état de la chaussée et celui des différentes structures. L'axe de fonctionnalité s'intéresse à la fluidité, la sécurité et la fonctionnalité des structures.

Pour l'axe conservation, par exemple, l'analyste en transport doit être capable, sur un horizon de vingt ans, de prévoir la dégradation des routes, d'en évaluer leurs déficiences ainsi que les budgets nécessaires pour les corriger [Lemieux 2000]. Un cadre simplifié pour ce genre d'analyse pourrait comprendre les étapes générales suivantes :

- prévoir la variation du débit journalier moyen annuel;
- vérifier l'effet de cette variation sur l'indice de rugosité de la surface routière, en fonction de la classe de route;
- vérifier si l'indice de rugosité atteint les seuils d'intervention, en fonction de la classe de route;
- déterminer les coûts d'intervention en fonction des seuils d'intervention, de la classe de route et du milieu;
- en fonction des budgets, répartir les investissements de manière équitable entre les différentes directions territoriales et de manière à ne pas dépasser les niveaux de déficiences fixés.

Les analyses doivent pouvoir être agrégées selon différentes entités territoriales, telles que la direction territoriale, le centre de service, la circonscription électorale provinciale et la municipalité régionale de comté, ou selon d'autres types d'entités d'agrégation telles que le réseau stratégique, le milieu, la route et le type de courbe ou de pente.

4 VERS UN PROCESSUS D'ANALYSE SPATIO-TEMPORELLE INTERACTIF (SOLAP)

4.1 Concepts de SOLAP

La présente section décrit les caractéristiques essentielles qui permettent de définir un outil SOLAP. La section 4.2 exposera des caractéristiques supplémentaires de ce qui pourrait être considéré comme un outil SOLAP idéal.

Il a été mentionné à la section 1 que dans le but d'exploiter la composante spatiale des données, lors d'analyses OLAP, il est nécessaire d'ajouter, à l'outil existant, des fonctionnalités dédiées à l'affichage et à la manipulation de données spatiales géométriques. Ce besoin est à la base du concept de SOLAP (« Spatial On-Line Analytical Processing »), ou processus d'analyse spatio-temporelle interactif² [Caron 1998]. Ce concept est récent et peu de définitions formelles existent. La définition d'OLAP proposée par Caron [1998] pourrait être modifiée afin de préciser son application à l'analyse spatio-temporelle. La définition d'un outil SOLAP serait donc « Une catégorie de logiciels axés sur l'exploration et l'analyse spatio-temporelle rapides des données selon une approche multidimensionnelle à plusieurs niveaux d'agrégation ». Une autre définition a été suggérée par Marchand [1998]. Selon ce dernier, un outil SOLAP est « Une plate-forme visuelle d'analyse spatio-temporelle supportant les besoins intuitifs de navigation multidimensionnelle d'un utilisateur ». La première partie de cette dernière définition, c'est-à-dire « plate-forme visuelle d'analyse » explicite la « catégorie de logiciels » qu'est un outil OLAP. Nous proposons donc de combiner les deux définitions citées ci-dessus. Un outil SOLAP pourrait alors être défini comme « Une plate-forme visuelle supportant l'exploration et l'analyse spatio-temporelle rapides des données selon une approche multidimensionnelle à plusieurs niveaux d'agrégation ».

De par sa définition, un outil SOLAP permet donc de manipuler une dimension particulière : l'espace, en plus des dimensions temporelles et strictement descriptives, ou sémantiques. Trois types de dimensions spatiales ont été identifiés : les dimensions spatiales descriptives, les dimensions spatiales géométriques vers descriptives et les dimensions spatiales géométriques [Han et al. 1998, Bédard et al. 2000].

Les dimensions spatiales descriptives ne contiennent aucune composante géométrique et la référence spatiale est effectuée à l'aide de données nominales seulement. Ce type de dimension

² L'acronyme SOLAP sera utilisé tout au long du document afin d'alléger le texte.

spatiale peut être implanté comme une dimension descriptive, mais aucune visualisation cartographique n'est alors possible [Bédard et al. 2000]. Un exemple de dimension spatiale descriptive serait une dimension de découpage routier où seulement les numéros de routes, de tronçons et de sections seraient disponibles pour l'analyse. Ce type de dimension spatiale peut évidemment être exploité par un outil SOLAP, en présentant toutefois les mêmes caractéristiques et limitations qu'une exploitation à partir d'un outil OLAP conventionnel.

Les dimensions spatiales géométriques vers descriptives comprennent une composante géométrique associée aux données des niveaux hiérarchiques inférieurs. Cependant, à partir d'un certain niveau, les agrégations deviennent descriptives seulement et la référence spatiale s'effectue à l'aide de données nominales exclusivement. Par exemple, une telle dimension pourrait contenir, au niveau hiérarchique inférieur, les provinces du Canada, chacune étant associée à une composante géométrique. Ces provinces pourraient être agrégées en valeurs considérées comme nominales telles que le Canada pacifique et le Canada atlantique [Bédard et al. 2000]. Ce type de dimension expose toute sa richesse spatiale seulement lorsqu'exploité à partir d'un outil SOLAP.

Les dimensions spatiales géométriques possèdent une composante géométrique associée à chacun des niveaux hiérarchiques. Le niveau hiérarchique le plus bas contient les primitives spatiales géométriques, tandis que les niveaux supérieurs contiennent des agrégations géométriques de ces primitives. Un exemple de ce type de dimension contiendrait des entités géométriques polygonales pour les municipalités, celles-ci représentant le niveau hiérarchique le plus bas de la dimension, ainsi que des agrégations de ces entités polygonales, formant de nouvelles entités polygonales potentiellement disjointes, pour les niveaux d'agrégation des municipalités régionales de comté et des régions administratives. Tout comme le type de dimension spatiale géométrique vers descriptive, ce type de dimension est pleinement exploité à l'aide d'un outil SOLAP.

Dans une base de données SOLAP, non seulement les dimensions, mais aussi les mesures peuvent posséder une composante spatiale géométrique [Han et al. 1998]. Il existe donc deux types de mesures : la mesure descriptive numérique, utilisée dans les systèmes conventionnels, et la mesure spatiale, comportant une composante géométrique. Un exemple de mesure spatiale serait l'ensemble des polygones agrégés résultant de la combinaison des membres du niveau « peuplements forestiers » d'une dimension de découpage forestier avec les membres du niveau « sous-bassins » d'une dimension de bassins versants. L'outil SOLAP doit être en mesure de manipuler ces deux types de mesures.

Un outil SOLAP est composé de deux volets, l'un descriptif, similaire à ce que les outils client OLAP existants offrent, et l'autre cartographique exploitant la composante spatiale géométrique des dimensions et mesures. Les deux volets sont intégrés et accessibles à partir d'une même interface qui doit être conviviale afin de faciliter son utilisation. L'utilisateur peut donc s'engager dans l'analyse spatio-temporelle de ses données par l'un ou l'autre des volets. Chacun des volets doit comprendre un espace de visualisation où s'afficheront les tableaux et diagrammes dans le cas du volet descriptif ainsi que les cartes dans le cas du volet cartographique. L'outil SOLAP devrait permettre de visualiser les résultats de l'analyse via les volets descriptif et cartographique simultanément, les deux volets étant liés, c'est-à-dire qu'une opération dans un des volets serait répercutée immédiatement dans l'autre. Les opérations OLAP de forage et de remontage, qui sont possibles directement via les types de visualisations descriptives (tableaux et diagrammes), doivent aussi être possibles via la carte directement. Il est alors question d'opérations OLAP descriptives dans le premier cas et d'opérations OLAP spatiales dans le second. L'opération de forage spatial sur un membre d'une dimension spatiale géométrique consiste à afficher les données géométriques du niveau de granularité immédiatement inférieur à l'intérieur de la dimension. L'opération de remontage spatial consiste, quant à elle, à afficher les données géométriques du niveau de granularité immédiatement supérieur à l'intérieur de la dimension.

Un système SOLAP doit évidemment hériter des avantages d'utilisation des outils OLAP par rapport aux autres technologies d'exploitation de données, c'est-à-dire, la facilité et la rapidité d'interrogation et ce, afin de permettre à l'utilisateur de se concentrer sur ses analyses et non sur le processus d'analyse lui-même.

Différentes possibilités d'implantation d'un outil SOLAP peuvent être envisagées. Les principales seront décrites à la section 4.3.

4.2 Caractéristiques d'un outil SOLAP idéal

Au fil des explorations et expérimentations réalisées au cours de l'essai, des caractéristiques ou fonctions souhaitables d'un outil SOLAP, en plus des caractéristiques essentielles exposées à la section 4.1, ont pu être déterminées, en vue d'assurer une utilité et une flexibilité maximales. La présente section contient la description de ces caractéristiques. La liste n'est pas exhaustive et d'autres caractéristiques souhaitables pourront être ajoutées au fur et à mesure que de tels outils seront développés et que de nouveaux besoins apparaîtront.

a. Opérations d'analyse OLAP

Il a été mentionné à la section 4.1 que les opérations de forage et de remontage doivent être possibles via la visualisation (descriptive ou cartographique) directement, en pointant sur un ou

plusieurs éléments représentés. Elles pourraient aussi, avec les autres opérations que sont le forage latéral et le pivot, être disponibles à partir d'éléments d'interface (boutons, menus) s'appliquant aux deux volets à la fois. De plus, il serait éventuellement possible de permettre à l'utilisateur d'effectuer des opérations de pivot et de forage latéral en lui donnant la possibilité de sélectionner des éléments affichés en contexte dans la carte, ces éléments de contexte étant des membres de dimensions ne faisant pas partie de l'analyse en cours, et de glisser les éléments sélectionnés dans la légende afin de les ajouter à l'analyse. De la même façon, il serait possible de sélectionner une entrée de la légende et de la glisser à l'extérieur de celle-ci afin de retirer les éléments qu'elle représente de l'analyse en cours.

b. Analyse via le volet descriptif et le volet cartographique de façon indépendante

La section 4.1 spécifie que l'outil SOLAP doit permettre d'effectuer l'analyse via les volets descriptif et cartographique simultanément, les deux volets étant liés. L'outil devrait aussi permettre de désactiver, et même de fermer, un des volets pour mener l'analyse dans l'autre volet seulement. Par exemple, une analyse des coûts des travaux routiers pourrait débuter dans le volet cartographique seulement afin de localiser les données d'intérêt sur le réseau routier. Ayant atteint le niveau de granularité le plus bas dans ce volet, l'analyse pourrait alors se poursuivre dans le volet descriptif exclusivement afin d'obtenir le détail de la répartition des coûts en fonction de dimensions descriptives.

c. Une ou plusieurs dimensions spatiales géométriques

L'outil SOLAP devrait avoir la capacité de manipuler plusieurs dimensions spatiales géométriques provenant d'un même cube. De même, l'outil devrait permettre de manipuler des dimensions spatiales géométriques possédant des hiérarchies alternatives. À ces fins, l'outil devrait donc fournir des fonctions pour la création d'agrégations géométriques en fonction des différentes dimensions spatiales géométriques de la base de données. De cette façon, il serait possible d'étudier les corrélations spatiales entre des phénomènes appartenant à des dimensions différentes, tels que les peuplements forestiers et les bassins versants, par exemple.

d. Traitement de tous les types de primitives géométriques

L'outil SOLAP devrait permettre de traiter les types de primitives géométriques de base, en deux dimensions, tels que définis par le comité ISO/TC 211 : point, courbe (ligne), surface (polygone) [ISO/TC 211 1999], de même que les géométries composées ou complexes, formées de plusieurs primitives géométriques connectées ou disjointes respectivement, telles que, par exemple, un ensemble de polygones disjoints formant une même occurrence d'archipel.

e. Nombre de mesures visualisables à la fois

Il devrait être possible, à l'intérieur d'un outil SOLAP, de visualiser plus d'une mesure à la fois dans les deux volets. La visualisation descriptive de plusieurs mesures pourrait être réalisée en affichant un diagramme différent pour chacune des mesures. De la même façon, la visualisation cartographique de plusieurs mesures pourrait être réalisée en affichant simultanément une carte différente par mesure, ou en ajoutant des diagrammes superposés aux éléments de la carte, un diagramme par mesure pour chacun des éléments, à condition dans ce dernier cas que le nombre de mesures soit relativement restreint.

f. Flexibilité de visualisation des données

Un outil SOLAP devrait permettre la présentation d'une dimension spatiale géométrique sous forme cartographique avec la présentation d'une dimension thématique sous forme de graphiques descriptifs, tels que des diagrammes circulaires, superposés aux éléments géométriques de la carte. Par exemple, pour visualiser la répartition du nombre d'accidents de différents types sur les segments du réseau routier, l'outil SOLAP pourrait d'abord afficher les segments sur une carte, puis afficher un diagramme circulaire montrant la répartition des types d'accidents sur chacun des segments. De plus, il devrait être possible de visualiser plusieurs cartes à la fois, par exemple, une carte pour chacune des mesures que l'on désire visualiser, comme mentionné au paragraphe précédent.

g. Modification de la symbologie

L'utilisateur d'un outil SOLAP devrait être en mesure de modifier la symbologie utilisée dans le volet cartographique, en fonction de ses besoins spécifiques, de manière à faire ressortir l'information d'intérêt. Par exemple, un utilisateur du milieu forestier qui utilise toujours les mêmes couleurs pour la représentation des groupements d'essences forestières devrait pouvoir utiliser la même symbologie dans son outil SOLAP afin de lui permettre de repérer ses informations du premier coup d'oeil. Cette fonction permettrait aussi de créer des cartes thématiques à des fins de diffusion.

h. Glissière temporelle

Afin d'exploiter la dimension temporelle, un élément d'interface spécifique, la glissière temporelle, pourrait être inclus dans l'outil SOLAP. Cette glissière faciliterait la navigation temporelle en permettant à l'utilisateur de se positionner à l'époque ou à la période d'intérêt simplement en glissant un curseur sur une ligne du temps couvrant l'étendue temporelle des données. Un exemple de glissière temporelle est décrit dans Caron [1998].

i. Affichage en contexte des données des secteurs voisins au secteur analysé

Un outil SOLAP devrait permettre d'afficher, en contexte, des données spatiales géométriques pour les secteurs adjacents au secteur analysé afin de permettre une meilleure compréhension des phénomènes présents en bordure de ce secteur. Par exemple, une concentration élevée d'accidents à l'extrémité d'un élément du réseau routier pourrait s'expliquer par la présence d'une intersection, localisée sur le même élément routier, mais située dans un autre secteur d'analyse.

j. Niveaux de granularité différents pour chacun des volets

Dans le but de manipuler des dimensions de type géométrique vers descriptif, c'est-à-dire des dimensions présentant des niveaux de granularité différents pour leurs composantes descriptives et géométriques, l'outil SOLAP devrait permettre de visualiser des niveaux de granularité différents dans les deux interfaces. Par exemple, lorsque le niveau de granularité limite, supérieur ou inférieur, serait atteint dans l'un des volets mais pas dans l'autre, le premier volet demeurerait fixe jusqu'à ce qu'une opération, dans ce volet ou dans l'autre, permette d'y afficher de nouvelles informations.

k. Choix des agrégations à afficher

L'outil SOLAP devrait offrir la possibilité de visualiser seulement les agrégations significatives, de manière à ne pas confondre l'utilisateur avec de grandes quantités de données ne présentant aucun intérêt. Cette fonction permettrait donc de filtrer les cellules vides du cube consulté qui correspondent chacune à une combinaison impossible de membres des différentes dimensions. Pour illustrer ce propos, supposons qu'un cube contienne, entre autres, une dimension de découpage administratif et une dimension de découpage électoral. Plusieurs combinaisons de membres de ces dimensions ne sont pas possibles, par exemple la combinaison de la circonscription électorale de Trois-Rivières et de la région administrative de Québec, et les mesures correspondant à cette combinaison sont nulles ou vides. Ce genre de combinaison ne devrait donc pas être affiché dans une visualisation.

l. Légende interactive

La légende d'un outil SOLAP devrait, par défaut, s'appliquer aux deux volets à la fois, c'est-à-dire que la même symbologie devrait être assignée aux différentes visualisations, qu'elles soient descriptives ou cartographiques, dans le but de faciliter la compréhension et de mettre en évidence le lien entre les données descriptives et les données spatiales géométriques. De plus, la légende devrait être modifiable d'une manière interactive, les changements à la légende s'appliquant automatiquement aux données visualisées. L'utilisateur pourrait d'abord modifier la symbologie des classes d'éléments ou des éléments individuels selon le type de symbolisation utilisé. Ensuite, lorsque des classes sont utilisées pour assigner une symbologie à des groupes

d'éléments géométriques, l'utilisateur pourrait modifier le nombre de classes, regrouper certaines classes ou modifier le type de classification en assignant par exemple un nombre égal d'éléments par classes ou en utilisant des intervalles de classes égaux. De plus, il serait possible, à l'aide de glissières, de modifier l'étendue d'une ou plusieurs classes afin de mettre en évidence certains groupes d'éléments par rapport aux autres. Par exemple, si les résultats d'une requête sont regroupés en 5 classes de valeurs ayant chacune une symbologie, l'utilisateur pourrait, à l'aide d'un simple glissement, étirer la classe inférieure pour ne visualiser que deux classes, c'est-à-dire la première classe originale ainsi que le regroupement des 4 autres classes dans une nouvelle classe. Enfin l'utilisateur devrait avoir la possibilité d'afficher ou non la légende selon ses besoins de visualisation.

m. Navigation via la légende

Un outil SOLAP pourrait offrir la possibilité d'effectuer des opérations de forage et remontage à partir de la légende directement, les opérations pouvant alors être qualifiées de thématiques.

n. Ajout de mesures calculées

Il devrait être possible, à l'intérieur d'un outil SOLAP, de définir de nouvelles mesures calculées à partir de mesures existantes emmagasinées dans le cube. Cet ajout pourrait se faire via la légende où la symbologie associée à la nouvelle mesure serait définie, en utilisant différentes variables visuelles telles que la trame, permettant ainsi de raffiner l'information affichée. Par exemple, si un cube possède une mesure du nombre d'accidents, une nouvelle mesure calculée pourrait consister en un pourcentage d'augmentation de ce nombre d'accidents par rapport à la période précédente. L'affichage cartographique de la première mesure pourrait se faire à l'aide de la couleur, tandis que celui de la mesure calculée pourrait être superposé au premier affichage à l'aide d'une trame.

o. Filtrage sur les membres

Un outil SOLAP pourrait comprendre des fonctions de filtrage sur les membres des dimensions visualisées, afin de restreindre l'affichage aux données d'intérêt pour une analyse particulière. Par exemple, dans une dimension accidents, il pourrait être souhaitable de ne visualiser que certains membres du niveau hiérarchique « type d'accidents », tels que les accidents graves et les accidents mortels. Un filtre pourrait donc être appliqué sur les types d'accidents affichés.

p. Opérations d'analyse spatiale

À l'intérieur d'un outil SOLAP, il pourrait être utile d'intégrer des opérations d'analyse spatiale, métriques et topologiques, telles que la mesure de distances et de superficies, la création de corridors ou l'analyse de réseaux afin de raffiner l'analyse OLAP spatio-temporelle en fonction de besoins particuliers. Ces opérations d'analyse spatiale devraient pouvoir s'effectuer sur différents niveaux d'agrégation. Par exemple, dans une application forestière qui permet d'analyser le volume de bois produit en fonction de différentes dimensions dont celle des groupements d'essences, l'utilisation de corridors pourrait permettre d'analyser le volume de bois pour des régions particulières, non définies par l'une des dimensions, telles que les bandes vertes de largeur variable autour des lacs et des rivières.

Afin d'optimiser la réalisation des analyses topologiques en particulier, les relations spatiales topologiques entre les éléments pourraient être emmagasinées directement à l'intérieur de la base de données multidimensionnelle. Des travaux en cours au Centre de recherche en géomatique de l'Université Laval portent sur ce volet d'analyse spatiale SOLAP [Marchand 2000].

q. Généralisation automatique et représentation multiple

Un outil SOLAP devrait permettre d'offrir différentes vues alternatives sur les données cartographiques, chacune pouvant, ou non, présenter un niveau de généralisation particulier, en fonction de l'échelle d'affichage de ces données. Une diminution de l'échelle d'affichage pourrait, ou non, amener une augmentation du niveau de généralisation et inversement. Des recherches ont démontré que la généralisation automatique d'éléments cartographiques demande de lourds traitements et que seules des machines très puissantes permettent d'obtenir des temps de réponse satisfaisants dans un contexte SOLAP [Martel 1999]. En réponse à cette limitation, l'utilisation de la représentation multiple est souhaitable [Bédard, Devillers et al. 2000]. En effet, l'outil SOLAP pourrait permettre le stockage, possiblement de manière hiérarchique, de plusieurs représentations, ou versions géométriques, pour un même membre d'une dimension spatiale géométrique ou une même mesure spatiale. L'outil SOLAP gérerait alors lui-même les vues permettant d'accéder à ces différentes représentations. Par exemple, un bâtiment pourrait être représenté à l'aide d'un point à petite échelle et à l'aide d'un polygone à plus grande échelle. Le concept de « Vuel » développé par Bédard [Bédard, Devillers et al. 2000] vise à gérer ces vues multiples.

r. Facilité à emmagasiner des données géométriques historiques

L'outil SOLAP devrait comprendre des facilités pour le stockage des données géométriques historiques afin de permettre une correspondance facile et rapide avec leur contrepartie descriptive.

s. Connexions possibles

Un outil SOLAP devrait pouvoir exploiter des données provenant d'un serveur OLAP, mais aussi d'un SGBD simulant le fonctionnement d'un serveur OLAP à l'aide de requêtes sur une base de données relationnelle structurée à l'aide d'un des schémas présentés à la section 2.3. Cette dernière option offre certains avantages qui seront décrits à la section 4.3.3.

4.3 Solutions technologiques potentielles

Cette section contient la description de différentes familles de solutions technologiques potentielles pour le développement et l'implantation d'un outil SOLAP.

Le regroupement des solutions d'implantation en différentes familles est basé sur le fait que plusieurs technologies différentes peuvent être utilisées afin de remplir les fonctions des volets descriptif et cartographique d'un outil SOLAP. Les fonctions du volet descriptif peuvent d'abord évidemment être supportées par un outil OLAP conventionnel ne gérant que la composante descriptive des données. Ces fonctions peuvent aussi être offertes par un SGBD relationnel ou objet-relationnel. Pour ce volet, les catégories « avec serveur OLAP » et « sans serveur OLAP » seront conservées pour la classification des solutions. Les fonctions du volet cartographique peuvent, quant à elles, être supportées par un logiciel de visualisation, un logiciel de cartographie assistée par ordinateur (CAD) ou un SIG, par exemple. Pour ce volet, les catégories « SIG » et « autre logiciel de visualisation » seront utilisées pour le regroupement des solutions.

4.3.1 Solutions avec serveur OLAP et SIG

À l'intérieur de cette catégorie, il est possible de distinguer quatre sous-catégories : les solutions SIG dominant, les solutions OLAP dominant, l'intégration totale OLAP - SIG et les applications hybrides [LGS Group Inc. 2000]. Les deux dernières sous-catégories étant similaires, elles sont ici regroupées en une seule. Au sein de cette classification, l'outil dominant est celui dans lequel des fonctions de l'autre outil ont été intégrées. L'outil dominant fournit l'interface graphique principale de l'application SOLAP.

Solutions SIG dominant

Les solutions SIG dominant offrent toutes les fonctionnalités de l'outil SIG, mais un sous-ensemble seulement des fonctionnalités de l'outil OLAP. Les fonctions OLAP qui ne sont pas présentes sont, par exemple, les possibilités d'effectuer un pivot en glissant des dimensions dans l'interface. Ces fonctions peuvent alors être disponibles via un menu. Ce type d'intégration est utile lorsque l'application nécessite l'utilisation de fonctions spécifiques au SIG.

Un exemple de logiciel commercial qui pourrait faire partie de cette catégorie est BusinessQuery for ArcView GIS, une solution issue d'un partenariat entre les compagnies Business Objects et ESRI [Business Objects 1998]. La documentation au sujet de ce logiciel étant quasi inexistante, ses particularités ne seront pas décrites ici.

Des implantations personnalisées de solutions de cette catégorie sont possibles, en intégrant à l'intérieur d'un SIG existant, des fonctionnalités OLAP par le biais de programmes développés à l'aide de bibliothèques de fonctions ou d'objets. En effet, plusieurs logiciels client OLAP, tels que ProClarity de Knosys, Brio.Enterprise de Brio Technology, Essbase d'Hyperion et Decision de Comshare rendent disponibles leurs bibliothèques de fonctions et d'objets pour la réalisation d'applications spécifiques à l'aide de langages de programmation courants comme Visual Basic ou C++. Il est alors possible de développer une extension OLAP à intégrer au logiciel SIG. Bien entendu, il est aussi possible de développer une toute nouvelle application en utilisant, non seulement les bibliothèques de fonctions des logiciels client OLAP, mais aussi celles des logiciels SIG. MapInfo, ArcView et GeoMedia sont parmi les SIG à permettre l'utilisation de leurs bibliothèques de fonctions avec les produits MapX, MapObjects et GeoMedia respectivement. De plus, le produit OpenMap qui consiste en un ensemble de composants java dédiées à la manipulation des données spatiales géométriques [OpenMap 2000], peut aussi être utilisé pour le développement du volet SIG.

Solutions OLAP dominant

Ce type de solution procure toutes les fonctionnalités d'un outil OLAP, ainsi qu'un sous-ensemble des fonctions disponibles à l'intérieur d'un SIG, telles que les fonctions d'affichage et de sélection d'élément spatiaux géométriques. Les fonctions d'analyse spatiale ne sont pas disponibles pour ce type de solution, qui peut être qualifiée d'application géospatiale périphérique où la référence spatiale n'est utilisée que comme support à la visualisation d'analyses non spatiales [Bédard *et al.* 1997].

Un premier exemple de logiciel appartenant à cette catégorie est le logiciel ProClarity, de la compagnie Knosys [Knosys 2000b]. ProClarity est un logiciel client OLAP qui permet de manipuler des bases de données OLAP (cubes) créées à l'aide d'Analysis Services de Microsoft SQL Server. ProClarity permet de visualiser les données descriptives d'un cube sous différentes formes graphiques telles que des tableaux et autres diagrammes. Le plugiciel KMapX permet la visualisation, sous forme cartographique, des données géométriques associées à une dimension spatiale géométrique du cube. Cet utilitaire est basé sur la technologie MapX de MapInfo. Un fichier de configuration permet de définir les données spatiales géométriques à coupler à une des dimensions d'un cube. La figure 11 présente un exemple d'interface, incluant les volets descriptif et cartographique, offert par le logiciel ProClarity 2.0.

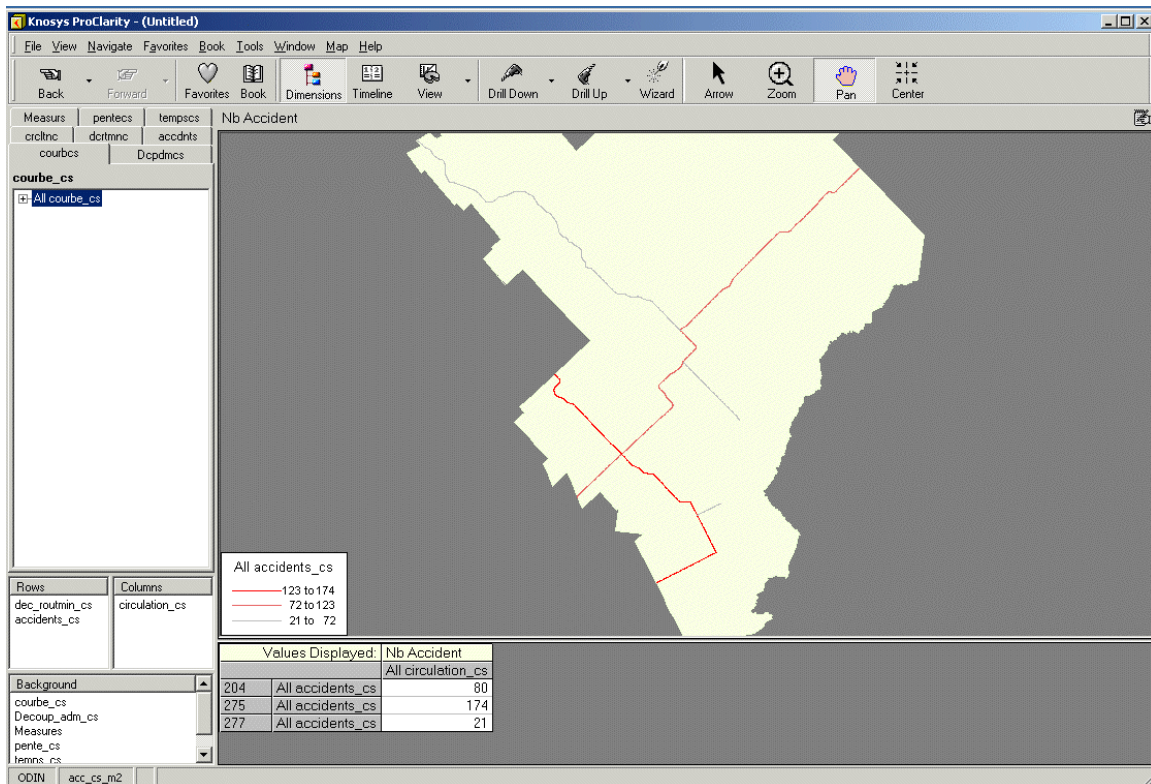


Figure 11 : Exemple d'interface SOLAP offert par ProClarity avec le plugiciel KMapX.

Un autre produit disponible sur le marché et faisant partie de cette catégorie est le logiciel Visualizer de Cognos [Cognos 2000]. Visualizer est un logiciel de visualisation de données pouvant provenir de sources diverses telles qu'un serveur OLAP ou une base de données relationnelle. Visualizer permet d'afficher les données descriptives d'une base de données multidimensionnelle sous forme de différents types de diagrammes. Le logiciel permet aussi l'affichage cartographique des données spatiales géométriques d'une dimension. Visualizer utilise la technologie MapX de MapInfo. La figure 12 présente un exemple d'interface offert par Visualizer et comprenant les deux volets d'un outil SOLAP.

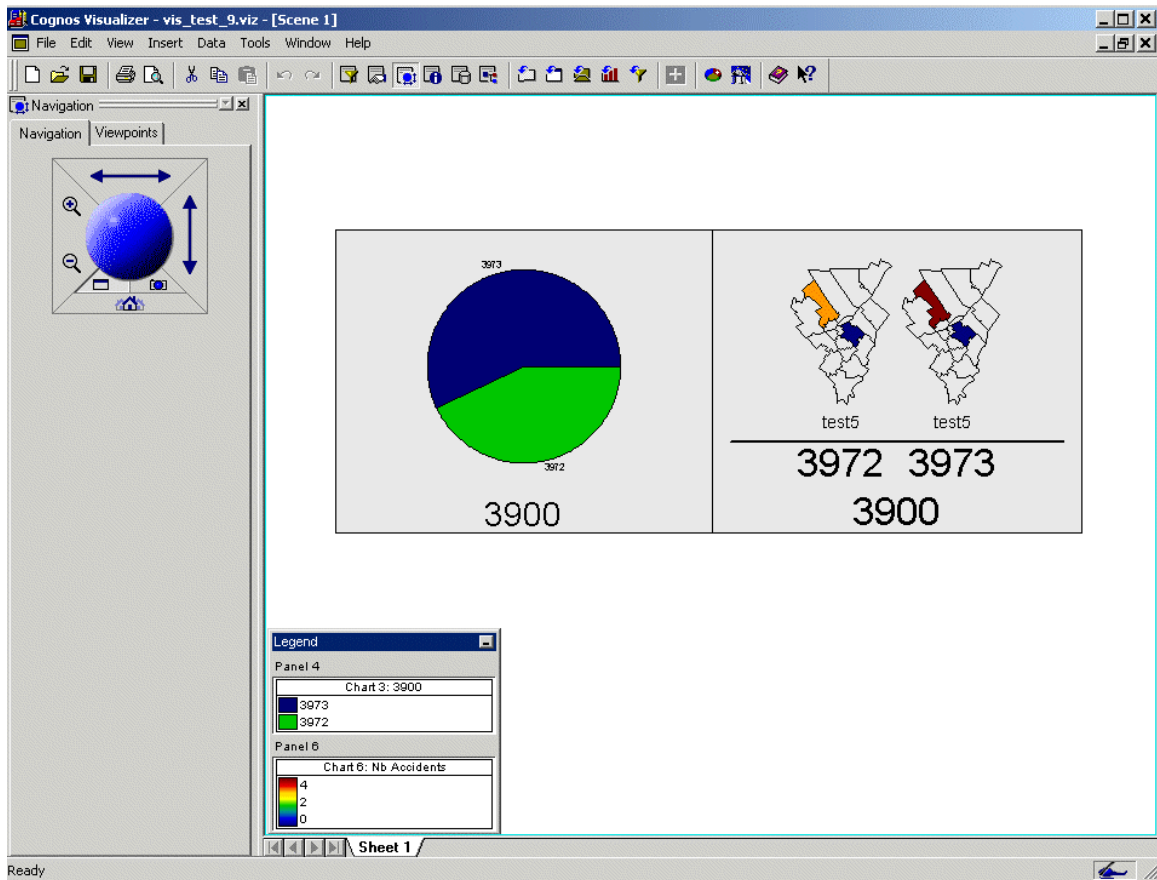


Figure 12 : Exemple d'interface disponible à l'intérieur de Visualizer 1.5.

Les deux logiciels cités plus haut ont été testés et comparés par rapport à leurs fonctionnalités ainsi qu'aux caractéristiques d'un outil SOLAP idéal présentées à la section 4.2. L'annexe 4 contient une description détaillée du fonctionnement de ces logiciels ainsi qu'un tableau comparatif résumant les résultats des essais. Les principales faiblesses de ces produits se situent au niveau du nombre de dimensions spatiales géométriques supportées. En effet, les deux logiciels ne permettent de visualiser qu'une seule dimension de ce type à la fois, ce qui élimine la possibilité d'étudier des corrélations spatiales. De plus, Visualizer n'est pas très flexible au niveau de la construction du volet de visualisation cartographique car chacune des opérations OLAP spatiales doit être prédéfinie et associée à une nouvelle carte.

Solutions d'intégration totale

Ce type de solution, intégrant les fonctionnalités d'un outil OLAP et les fonctionnalités d'un SIG, pourrait être qualifié d'application centrée-géospatiale où la référence spatiale des objets est utilisée constamment dans l'exploration et l'analyse des données [Bédard et al. 1997]. Ce type de solution n'est actuellement pas disponible sur le marché, mais pourra éventuellement provenir

d'une compagnie offrant à la fois des produits SIG et OLAP [LGS Group Inc. 2000]. Cependant, à l'aide des bibliothèques de fonctions de logiciels client OLAP et de logiciels SIG, le développement d'une telle solution est possible, à l'intérieur d'un cadre spécifique. Cette dernière option a été nommée l'approche hybride. Seule cette option permet pour l'instant d'offrir toutes les caractéristiques souhaitables d'un outil SOLAP décrites à la section 4.2, à condition de les programmer. C'est cette option qui a été utilisée dans le cadre du présent essai dans le but de tester les modes d'intégration OLAP et SIG.

4.3.2 Solutions avec serveur OLAP et autre logiciel de visualisation

La composante SIG des solutions avec serveur OLAP et SIG décrites à la section 4.3.1 peut être remplacée par un logiciel de visualisation cartographique, que ce soit un CAD ou même un logiciel de visualisation scientifique tel qu'AVS (Advanced Visual Systems). Ce dernier est un puissant logiciel de visualisation offrant un environnement de développement, AVS Express Developer Edition [Advanced Visual Systems 1998], ainsi que des fonctionnalités d'affichage de données spatiales géométriques stockées dans le SGBD Oracle 8 (avec le module spatial).

Des solutions de ce type, visant la visualisation cartographique et non seulement la visualisation descriptive, ne semblent pas être disponibles commercialement. Elles doivent donc être développées à l'aide des bibliothèques d'objets de logiciels existants.

4.3.3 Solutions sans serveur OLAP et avec SIG

Le serveur OLAP des solutions d'intégration décrites en 4.3.1 peut être simulé à l'intérieur d'une base de données relationnelle. Lorsque le volume de données à consulter est peu élevé, cette solution peut s'avérer très avantageuse, puisque les calculs d'agrégation peuvent s'effectuer de manière sélective et contrôlée à l'aide de requêtes SQL sur la base de données. Ces requêtes peuvent alors être adaptées en fonction des besoins d'un projet particulier, en évitant par exemple de calculer les agrégations non-significatives ou en permettant de joindre les tables impliquées dans les requêtes de manière plus flexible que ne le permettent habituellement les serveurs OLAP qui utilisent des fonctions d'« *inner join* ». Cette solution doit cependant inclure, dans la base de données, des éléments permettant de gérer la hiérarchie des membres des dimensions utilisées, en vue de la réalisation d'opérations OLAP telles que le forage et le remontage.

4.3.4 Solutions sans serveur OLAP et avec autre logiciel de visualisation

La dernière famille de solutions couple une base de données relationnelle simulant un serveur OLAP à un logiciel autre qu'un SIG pour la manipulation des données spatiales géométriques. Les particularités des solutions décrites en 4.3.2 et 4.3.3 s'appliquent.

Un exemple de ce type d'implantation est le prototype SOLAP du projet ICEM-SE (« Interface cartographique pour l'exploration multidimensionnelle des indicateurs de santé environnementale sur le World Wide Web »), développé au Centre de recherche en géomatique de l'Université Laval à partir du SGBD Access et du logiciel de visualisation cartographique SoftMap, en langage Visual Basic.

Rappelons que toutes les familles de solution présentées, certaines n'étant présentement intégrées dans aucun logiciel commercial, peuvent être développées à partir de bibliothèques d'objets de logiciels actuellement disponibles sur le marché.

5 INTÉGRATION OLAP - SIG

5.1 Différentes possibilités d'intégration OLAP - SIG

Plusieurs familles de solutions d'implantation d'un outil SOLAP ont été présentées à la section 4.3. Dans la présente section, les détails d'intégration OLAP - SIG des solutions décrites à la section 4.3.1 seront discutés.

Pour débiter, il est possible de faire un parallèle entre la gestion des données spatiales géométriques d'un SOLAP et la gestion des données descriptives à l'intérieur d'un tel système ou d'un système OLAP conventionnel. De la même façon que des agrégations de données descriptives sont calculées et emmagasinées dans une structure multidimensionnelle, des agrégations de la composante spatiale géométrique des données doivent être préparées et emmagasinées. Certains serveurs OLAP, pour une question d'espace mémoire, ne calculent et n'emmagasinent pas toutes les agrégations de données descriptives possibles en prétraitement. Seulement certaines agrégations sont calculées en fonction de critères qu'il est possible de déterminer, par exemple l'espace disque maximum utilisé, ou le gain de performance lors de l'exécution des requêtes [Microsoft Corporation, 2000]. Le serveur OLAP détermine alors les agrégations optimales à calculer et à emmagasiner afin de respecter, approximativement, les critères définis par l'utilisateur. Il est possible d'appliquer le même procédé à la composante spatiale géométrique des données. Les agrégations spatiales géométriques correspondant aux agrégations descriptives pourraient alors être pré-calculées, possiblement durant le même processus que le calcul des agrégations descriptives, et emmagasinées dans la structure dédiée au stockage des données géométriques qui sera discutée plus loin. À l'exploration de la base de données multidimensionnelle par le module client, la requête dont le résultat n'aurait pas été préalablement calculé et emmagasiné serait effectuée à la volée, le serveur OLAP se servant alors des agrégations pré-calculées, les agrégeant pour former le résultat de la requête en cours. Cela semble suffisamment efficace et performant en ce qui concerne les données descriptives. Pour ce qui est de la composante géométrique du résultat de la requête, l'application aurait à calculer une nouvelle agrégation en unissant, à l'intérieur d'une même dimension, ou en intersectant, entre deux dimensions, des agrégations existantes.

Ce principe est discuté par Stefanovic [1997]. Ce dernier affirme que le temps de réponse des requêtes est satisfaisant seulement si un pré-calcul approprié des agrégations est effectué et que, d'un autre côté, la création de toutes les agrégations possibles demande un espace de stockage volumineux. Il présente donc le concept de « matérialisation sélective », c'est-à-dire le pré-calcul de seulement certaines agrégations de façon à obtenir un équilibre entre les avantages de la

génération d'une agrégation particulière lors de l'exécution des requêtes, et l'espace de stockage nécessaire pour emmagasiner les agrégations pré-calculées. La matérialisation sélective peut être basée sur trois critères : la fréquence d'accès aux agrégations, déterminée à l'aide de statistiques d'accès, la taille des agrégations pré-calculées, ainsi que le bénéfice de la construction des agrégations par rapport au gain de performance général. Des algorithmes ont été développés pour sélectionner les agrégations à pré-calculer selon les trois critères [Stefanovic 1997]. Les agrégations n'ayant pas été pré-calculées doivent être effectuées à la volée, à l'aide de ce qui peut être considéré comme un opérateur d'agrégation géométrique. Des recherches ont présenté des solutions afin d'optimiser la performance de calculs effectués à la volée, entre autres en ce qui concerne la fusion de polygones [Zhou et al. 1999]. Cependant, les algorithmes présentés ne sont pas encore implantés dans les logiciels commerciaux. Les technologies actuellement disponibles sur le marché sont incapables d'effectuer de tels calculs d'agrégation de manière assez performante pour respecter une des caractéristiques fondamentales d'un outil SOLAP, la rapidité d'exécution.

Puisque les opérations d'agrégation, de fragmentation, d'union et d'intersection actuellement disponibles dans les logiciels commerciaux demandent des temps de traitement inacceptables dans le contexte d'un outil SOLAP, la seule solution viable avec ces outils commerciaux existants est le pré-calcul de toutes les agrégations. Les logiciels SIG peuvent être utilisés pour créer ces agrégations. De même, des logiciels de conversion tels que FME (« Feature Manipulation Engine ») offrent aussi des fonctions pour la création d'agrégations.

Il est possible de distinguer l'agrégation physique, où plusieurs éléments agrégés deviennent un seul élément distinct qui est stocké dans la structure de données géométriques, par rapport à l'agrégation qu'on peut qualifier de virtuelle, où plusieurs éléments agrégés forment un ensemble d'éléments distincts mais possédant une symbolisation identique, l'utilisateur ayant alors l'impression de voir un seul élément. Ce second cas peut être envisagé lorsque le logiciel utilisé pour la manipulation des données spatiales géométriques ne permet pas la gestion des entités spatiales complexes telles que les nuages de points et les ensembles de polygones ou d'arcs disjoints. L'application doit alors contenir des mécanismes, tels qu'une table de correspondance contenant l'identifiant de l'élément complexe simulé, et l'identifiant des éléments primitifs le composant, pour gérer l'ensemble d'éléments comme un seul élément.

Les considérations sur les agrégations descriptives et géométriques ayant été discutées, les considérations sur leur stockage suivent. Deux solutions sont possibles dans ce cas : utiliser deux structures parallèles, la structure multidimensionnelle pour les agrégations descriptives et une structure SIG destinée spécifiquement au stockage des agrégations géométriques, ou bien

n'utiliser qu'une structure, la structure multidimensionnelle. Ces deux solutions seront présentées dans les lignes qui suivent.

Utilisation de deux structures parallèles

Tel que mentionné plus haut, cette solution implique que les données descriptives demeurent emmagasinées dans la structure multidimensionnelle, tandis que les données géométriques soient emmagasinées dans une structure spatiale géométrique provenant d'un SIG. Le lien entre les deux structures peut être fait en emmagasinant un identifiant dans la base de données de structure multidimensionnelle, pour chacun des éléments descriptifs ayant une composante géométrique. Cet identifiant peut alors correspondre, dans la structure SIG, à un élément géométrique possiblement complexe dans le cas d'une agrégation, à une liste de pointeurs vers un ensemble d'éléments géométriques ou à une carte pré-définie [Shekhar et al. 2000].

L'utilisation d'une structure SIG pour emmagasiner les données géométriques peut permettre d'effectuer facilement des requêtes spatiales, non SOLAP, à partir du volet cartographique. La structure SIG est spécifiquement développée pour le stockage et la manipulation de données géométriques et fournit des fonctions à cet effet.

Ce type de solution pourrait aussi permettre de définir une dimension spatiale géométrique directement à l'intérieur de la structure SIG dans le cas des architectures ROLAP ou HOLAP, la structure emmagasinant la table de dimension (ou les tables de dimension dans le cas d'un schéma en flocon). Le cube résultant serait alors distribué sur deux bases de données : celle de la structure SIG et celle contenant les données descriptives.

Utilisation d'une seule structure multidimensionnelle

Cette solution implique que les données descriptives, ainsi que les données géométriques associées soient emmagasinées dans la même structure multidimensionnelle. Si le cube ne contient qu'une seule dimension spatiale géométrique, le stockage de la géométrie dans la structure de données multidimensionnelle pourrait être réalisé lors de sa création, en définissant, par exemple, une propriété pour chacun des membres de la dimension spatiale géométrique. Lorsque le cube contient plusieurs dimensions spatiales géométriques, les agrégations de membres de ces dimensions peuvent alors être considérées comme des mesures spatiales et leur stockage nécessite un outil capable d'écrire dans la base de données.

Cette solution implique que le logiciel utilisé pour la visualisation cartographique lise directement les données géométriques dans la base de données de structure multidimensionnelle. Cela n'est présentement possible que par programmation.

De plus, le fait d'emmagasiner la géométrie à l'intérieur du cube permettrait d'effectuer des corrélations de cette dernière avec des données d'autres dimensions, si de tels besoins se présentent.

SOLAP web

Des serveurs OLAP Web existent sur le marché. Ce type de serveur pourrait être couplé à un serveur SIG Web afin de produire un outil d'analyse SOLAP Web. Le présent essai s'est limité à l'étude d'applications locales. Le volet Web fera l'objet d'études à venir.

5.2 Particularités des données du projet OLAP - MTQ

Les données fournies pour le projet OLAP - MTQ peuvent être vues de deux manières différentes. La première, et la plus intuitive, consiste à ne considérer qu'une seule dimension spatiale géométrique pour chacun des différents cubes de données décrits à l'annexe 3, celle du découpage routier qui est à la base des analyses. Cette dimension sert alors de localisation à l'information contenue dans les autres dimensions descriptives des cubes, cette dernière information venant se superposer, de différentes façons qui seront vues à la section 5.3, aux données géométriques affichées.

La seconde façon d'utiliser les données du MTQ est de considérer toutes les dimensions comme étant des dimensions spatiales géométriques. Cette manière de procéder découle directement du fait que les données contenues dans chacune des dimensions du projet servent à segmenter le réseau routier en éléments de caractéristiques uniformes, tel que mentionné à la section 3.2.1, certaines de ces caractéristiques variant en fonction de la dimension temps. Seule la dimension des accidents ne contient aucune donnée segmentant le réseau routier. Dans ce cas particulier, il est quand même possible de considérer cette dimension comme étant de type spatial géométrique, en permettant d'agréger les segments de base selon le code de gravité d'accident dominant de chacun des segments, par exemple. La table des faits d'un cube construit selon ce principe peut être considérée comme la table servant à une segmentation statique des données géométriques linéaires du réseau routier.

5.3 Considérations sur l'affichage cartographique

Une étude sommaire des possibilités d'affichage cartographique particulières à un outil SOLAP a été effectuée en vue du développement de l'application d'essai décrite à la section 5.4. L'affichage cartographique différera en fonction du nombre de dimensions spatiales géométriques présentes à l'intérieur du cube visualisé. Ces considérations s'ajoutent aux règles existantes

gérant l'utilisation des différentes variables visuelles telles que la couleur, l'orientation, la forme ou la texture, en fonction des types de phénomènes à représenter.

Lorsqu'une seule dimension spatiale géométrique est présente, celle-ci peut être considérée comme la base de l'affichage cartographique. Un sous-ensemble de ses données est donc affiché en permanence. Selon le nombre d'autres dimensions visualisées ainsi que le nombre de membres actifs pour chacune, différentes possibilités se présentent. Premièrement, lorsqu'un seul membre de chacune des autres dimensions est actif et qu'une seule mesure est visualisée, les éléments affichés de la dimension spatiale géométrique sont symbolisés selon la valeur de la mesure associée à chacun de ceux-ci. Lorsque le nombre de valeurs différentes de la mesure est inférieur ou égal à 7, chacune des valeurs différentes de la mesure peut se voir attribuer une symbologie propre qui est alors associée aux éléments géométriques correspondant à cette valeur. Lorsque le nombre de valeurs différentes de la mesure est supérieur à 7, il est préférable de regrouper celles-ci en 7 classes ou moins [Bédard 2000a]. Chacune des différentes classes créées reçoit une symbologie différente et cette dernière est associée aux éléments géométriques faisant partie de la classe. En second lieu, lorsqu'une seule des autres dimensions contient plusieurs membres actifs, et qu'une seule mesure est visualisée, plusieurs cartes peuvent alors être créées, une pour chacun des membres de cette autre dimension, ce procédé facilitant la lecture des cartes [Bédard 2000a]. Une alternative est d'utiliser des diagrammes statistiques montrant les informations relatives aux membres de l'autre dimension, et de les associer aux éléments de la carte. Enfin, les autres cas nécessitent la création de plusieurs cartes, en plus de l'ajout de diagrammes statistiques selon l'approche décrite précédemment. Ces possibilités sont similaires à celles utilisées par le logiciel ProClarity [Knosys 2000a] et qui sont décrites à l'annexe 4.

La visualisation de plusieurs mesures à la fois peut se traduire par la création de plusieurs cartes, une carte par mesure. Une alternative, applicable seulement si le nombre de mesures visualisées est petit, serait d'ajouter un diagramme statistique par mesure. Chacun des éléments géométriques de la carte aurait donc plusieurs diagrammes d'associés, un diagramme par mesure. De plus, la visualisation d'un nombre restreint de mesures pourrait être réalisée par l'utilisation de plusieurs variables visuelles telles que la couleur, la trame ou l'orientation. Un exemple de ce dernier cas est donné à la section 4.2 (n).

Lorsque toutes les dimensions sont considérées comme des dimensions spatiales géométriques, les agrégations géométriques affichées peuvent recevoir une symbologie propre, si le nombre d'agrégations est relativement faible. Dans le cas contraire, les valeurs de mesures pour chacune des agrégations, individuelles ou regroupées en classes, reçoivent une symbologie propre, cette symbologie étant par la suite assignée aux différentes agrégations géométriques.

Pour ce cas où toutes les dimensions sont des dimensions spatiales géométriques, la sélection de plusieurs membres d'autres dimensions ne fait que modifier les agrégations à afficher. Il n'est donc pas nécessaire d'utiliser des diagrammes statistiques à superposer aux agrégations géométriques affichées, à moins de visualiser plusieurs mesures à la fois.

Lorsque le cube contient plusieurs dimensions spatiales géométriques, mais qu'il contient aussi des dimensions descriptives, l'affichage est alors composé d'une combinaison des deux possibilités décrites plus haut.

Il est à noter que les dimensions spatiales géométriques vers descriptives sont gérées comme des dimensions spatiales géométriques lorsque les données affichées proviennent de niveaux de granularité comportant une composante géométrique. Ces dimensions sont gérées comme des dimensions descriptives lorsque les données affichées appartiennent à des niveaux strictement descriptifs.

La symbologie utilisée pour représenter les objets peut être stockée à l'intérieur de l'application, ou même de la base de données multidimensionnelle. Elle peut aussi être calculée en fonction de différents paramètres tels que la valeur de la mesure associée. Le choix dépend des besoins particuliers des utilisateurs. Par exemple, une application en santé environnementale pourrait nécessiter que les valeurs de mesures soient toujours regroupées en fonction des mêmes classes, symbolisées de la même façon, afin de pouvoir effectuer des comparaisons entre différents résultats d'analyse.

Les opérations OLAP peuvent s'accompagner d'une modification de l'échelle d'affichage, le forage s'accompagnant d'un agrandissement d'échelle tandis que le remontage s'accompagnant d'une réduction d'échelle. Il est alors possible d'utiliser le rectangle englobant minimal afin d'effectuer un cadrage sur les éléments affichés.

La carte affichée devrait contenir des données de contexte lorsque ces dernières facilitent la localisation des phénomènes analysés.

5.4 Développement d'une application d'essai

Une application d'essai a été développée afin d'implanter et de tester les solutions d'intégration qui seront décrites à la section 5.5. Cette application a été programmée en utilisant le langage Visual Basic. Le volet cartographique de l'application a été réalisé à l'aide des bibliothèques d'objets de GeoMedia 4 d'Intergraph. Le volet descriptif a été développé avec les bibliothèques d'objets du

logiciel ProClarity 2.0 qui a été introduit à la section 4.3.1 et dont le fonctionnement est décrit en détail à l'annexe 4. La figure 13 présente l'interface intégrant les deux volets.

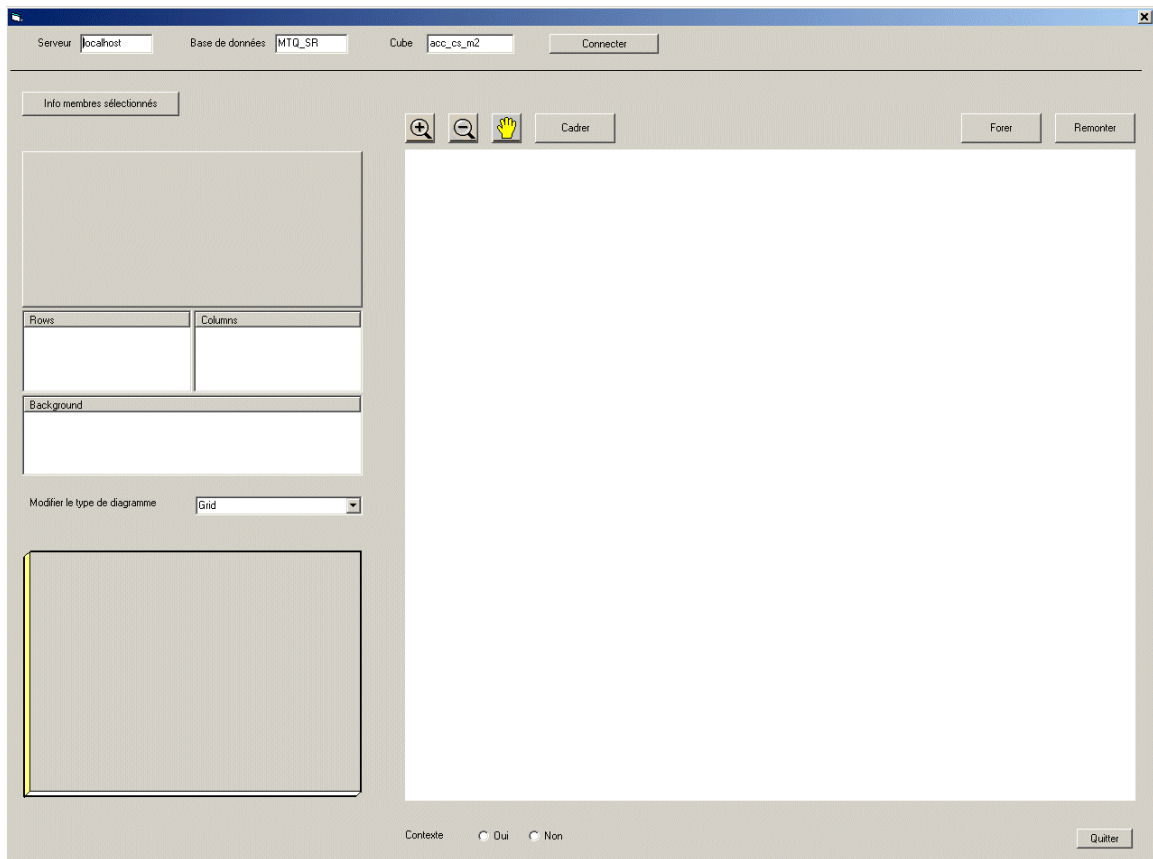


Figure 13 : Interface de l'application d'essai développée.

Le volet cartographique, à droite, en plus de contenir un espace de visualisation destiné à afficher les cartes, contient des fonctions minimales de navigation cartographique telles que l'agrandissement ou la réduction d'échelle (« zoom ») et le déplacement (« pan »). De plus, des boutons permettent d'effectuer des opérations de forage ou de remontage spatial. Le volet descriptif, à gauche, contient deux éléments : un contrôle permettant de manipuler les dimensions et les membres de celles-ci, c'est-à-dire permettant d'effectuer des opérations de pivot et de forage latéral, ainsi qu'un espace de visualisation affichant la composante descriptive des données sous forme de tableaux, de diagrammes circulaires, de diagrammes à barres ou autres. Dans cet espace de visualisation, le forage descriptif est l'opération par défaut. L'opération de remontage est aussi disponible. Des facilités de connexion à une base de données multidimensionnelle sont disponibles au haut de l'interface. D'autres fonctions ont été ajoutées à des fins de démonstration seulement et ne seront pas décrites ici.

Le volet cartographique met en application, de manière simplifiée, les principes d'affichage cartographique décrits à la section 5.3. Des exemples d'affichage seront présentés aux sections 5.5 et 5.7.

Les données utilisées pour les essais comprennent un sous-ensemble des données du cube d'accidents, pour le centre de service de Lac-Etchemin. Les dimensions utilisées, ainsi que leurs niveaux de granularité, sont :

- dimension temps : année
- dimension découpage administratif : municipalité
- dimension circulation : section de trafic
- dimension accidents : code de gravité
- dimension découpage routier : route, tronçon, section, sous-route, segment
- dimension pentes : classe HPMS
- dimension courbes : classe HPMS

Ces données sont emmagasinées dans SQL Server. Les schémas en étoile et en flocon ont été implantés pour réaliser les essais sur les architectures ROLAP et HOLAP. Le chargement des données dans la structure multidimensionnelle de l'architecture MOLAP a été réalisé à partir du schéma en étoile. Le module Microsoft Analysis Services a été utilisé comme serveur OLAP. Ce module offre des assistants très conviviaux pour la construction rapide de dimensions et de cubes de données.

5.5 Solutions d'intégration OLAP - SIG implantées

Quatre solutions d'intégration ont été implantées et testées à l'aide de l'application décrite en 5.4.

Les solutions présentées ici sont basées sur le fait qu'un serveur OLAP est utilisé, ce serveur contrôlant directement les opérations effectuées sur le cube. C'est donc par lui que se fait toute la gestion multidimensionnelle du cube concernant entre autres la navigation hiérarchique dans les données. C'est aussi par lui que passent toutes les opérations OLAP, qu'elles soient spatiales ou descriptives. À chaque nouvelle requête sur le cube, si un résultat est retourné, les volets descriptif et cartographique prennent en charge chacun leur composante des données résultantes et mettent à jour leur affichage.

La description des solutions d'intégration présentées ici mettra l'accent sur le fonctionnement du volet cartographique.

Solution 1 : lien implicite entre la structure multidimensionnelle et la structure spatiale géométrique

La première solution développée s'applique à un cube contenant une seule dimension spatiale géométrique. Elle utilise deux structures : la structure multidimensionnelle pour le stockage des données descriptives et la structure SIG, pour le stockage des données spatiales géométriques.

Cette solution utilise chacun des membres de la dimension spatiale géométrique comme identifiant implicite permettant de faire le lien entre les données descriptives de la base de données OLAP et les données géométriques. Dans le cas présent, il s'agit du numéro de l'élément routier. Afin que le lien soit possible, il est nécessaire que chacun des membres soit unique à l'intérieur d'un même niveau hiérarchique. Chacun des niveaux de la dimension peut alors correspondre à une couche de données géométriques. Dans le cas où chacun des membres serait unique à l'intérieur de la dimension, toutes les données géométriques pourraient être emmagasinées dans une même couche. Cette solution correspond à ce qui est implanté dans le logiciel ProClarity avec le plugiciel KMapX.

À l'intérieur de cette implantation, une opération OLAP spatiale récupère l'identifiant de l'élément géométrique pointé, cet identifiant correspondant directement à un membre de la dimension spatiale géométrique. À l'aide des fonctions des bibliothèques OLAP, les données correspondant au membre sélectionné (parent ou enfants selon que l'opération est un remontage ou un forage respectivement) sont obtenues. Ces données sont ensuite insérées dans une requête au cube et un nouveau résultat est renvoyé, prêt à être affiché dans les deux volets. Le volet cartographique vérifie alors quels sont les membres actifs de la dimension spatiale géométrique et les affiche dans son espace de visualisation.

Cette solution s'implante de la même façon, quelle que soit l'architecture OLAP utilisée. Les avantages et inconvénients des différentes architectures OLAP dans ce contexte sont donc ceux discutés à la section 2.2.4. Les différentes architectures ont été testées à l'aide d'essais chronométrés et l'architecture MOLAP, tel qu'attendu, est celle qui a présenté les meilleures performances.

Cette solution peut nécessiter des manipulations mineures des données, dans le cas où les données géométriques ne sont pas consistantes par rapport aux données descriptives emmagasinées dans le cube. Elle ne requiert aucun prétraitement automatique au démarrage de l'application.

Un exemple d'affichage cartographique réalisé à l'aide de cette solution est présenté à la figure 14. Cet affichage contient trois routes, associées chacune à un diagramme circulaire présentant la proportion du nombre d'accidents de différents types ayant eu lieu sur celles-ci.

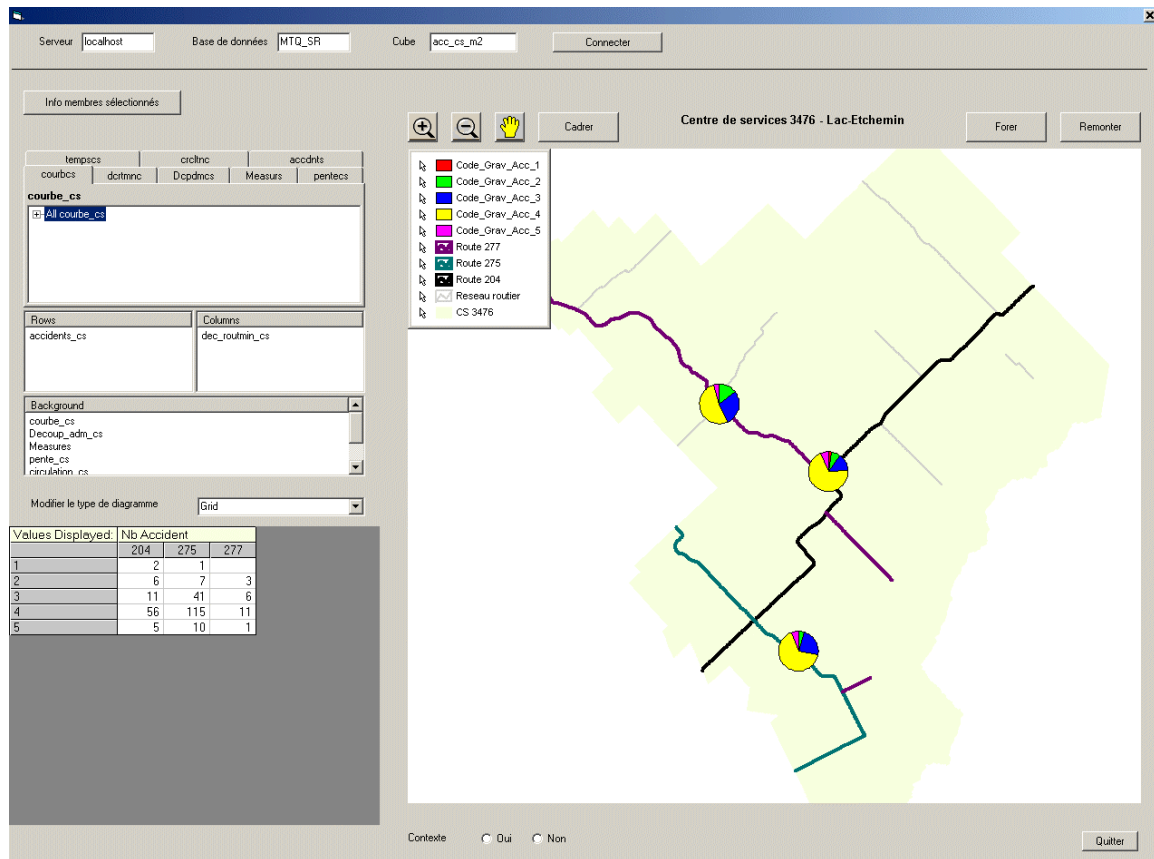


Figure 14 : Exemple d'affichage cartographique réalisé avec la solution d'implantation 1 – une seule dimension spatiale géométrique.

Solution 2 : simulation d'une seule structure multidimensionnelle

Tout comme la solution 1, cette solution s'applique aussi à un cube contenant une dimension spatiale géométrique. Elle simule cependant l'utilisation d'une seule structure de données, la structure multidimensionnelle.

Cette solution permet le stockage de la géométrie associée à chacun des membres de la dimension spatiale géométrique à l'intérieur d'une structure propriétaire gérée par le serveur OLAP, comme une propriété du membre. Ce stockage n'est cependant pas permanent, d'où l'emploi du terme « simulation », et doit être effectué en prétraitement automatique à chaque nouvelle connexion au cube. Ce prétraitement consiste à obtenir, pour chacun des membres de

la dimension spatiale géométrique, sa géométrie, emmagasinée dans GeoMedia comme une donnée de type Binary Large Object (« blob »), à l'aide d'une requête. Cette géométrie est ensuite emmagasinée dans la structure propriétaire à l'aide de fonctions des bibliothèques OLAP. Comme ce stockage n'est pas permanent, il n'affecte pas la structure multidimensionnelle originale des données. L'implantation de cette solution est donc, ici encore, indépendante de l'architecture OLAP utilisée. Des essais ont montré que l'architecture MOLAP est celle qui présente les meilleures performances.

Cette solution peut nécessiter des manipulations mineures des données afin d'assurer la consistance entre les données spatiales géométriques et les données descriptives.

À l'intérieur de cette implantation, une opération OLAP spatiale récupère la géométrie de l'élément géométrique pointé, recherche le membre correspondant à cette géométrie et de la même façon que la solution 1, détermine ses données correspondantes (parent ou enfants) puis les insère dans une requête au cube. Les résultats de l'affichage cartographique pour cette solution, sont les mêmes que pour la solution 1.

Cette solution aurait le potentiel d'offrir une meilleure performance d'affichage que la solution 1 s'il était possible d'afficher directement la géométrie stockée sous forme de « blob ». Cela n'est pas le cas avec GeoMedia et de nombreuses manipulations sont nécessaires afin de créer une couche de données et d'y stocker la géométrie préalablement transformée.

Il est possible d'emmagasiner des propriétés de membres dans un cube de manière permanente, lors de sa construction. Le stockage de la géométrie a été tenté, mais sans succès avec Microsoft Analysis Services. Cependant, cette option est peut-être possible en utilisant d'autres outils.

Solution 3 : lien explicite emmagasiné, entre la structure multidimensionnelle et la structure spatiale géométrique

Cette solution peut s'appliquer à plusieurs dimensions spatiales géométriques. Elle utilise deux structures : la structure multidimensionnelle pour le stockage des données descriptives et la structure spatiale géométrique, fournie par le SIG, pour le stockage des données géométriques.

Puisqu'elle permet de manipuler plusieurs dimensions spatiales géométriques, cette solution doit fournir des fonctions pour créer les agrégations correspondant aux combinaisons de membres des différentes dimensions. Les agrégations doivent recevoir un nouvel identifiant permettant de retracer chacun des membres la composant. Cela est réalisé en emmagasinant, dans les tables relationnelles de base qui vont servir à la constitution des dimensions spatiales géométriques, et

pour chacun des membres allant composer ces dimensions, un identifiant numérique, déterminé à partir de l'ordre du membre à l'intérieur de chacun des niveaux de chacune des dimensions spatiales géométriques. Cet ordre est l'ordre alphanumérique dans le cas de Microsoft Analysis Services. L'identifiant est emmagasiné dans le cube comme une propriété de chacun des membres des dimensions. À la première connexion au cube, un prétraitement recherche, dans le cube, les agrégations géométriques à créer. Une agrégation est créée seulement lorsqu'au moins une valeur de mesure lui étant associée est non nulle ou non vide. Chaque agrégation reçoit comme identifiant, la concaténation, selon l'ordre des dimensions à l'intérieur du cube, de l'identifiant des membres agrégés. Ce prétraitement doit être effectué seulement lors de la première consultation du cube, puisqu'il emmagasine les agrégations de façon permanente dans une couche de données géométriques qui sera réutilisée lors des consultations subséquentes. Ce processus de création des agrégations nécessite que toutes les primitives spatiales géométriques à agréger soient disponibles dans une même couche de données, celle-ci contenant tous les attributs des tables de dimension permettant la définition des membres. La préparation de cette couche peut nécessiter des manipulations mineures. De plus, l'identifiant de chacun des membres doit être au préalable emmagasiné dans la table de dimension correspondante, ce qui peut entraîner la modification de plusieurs tables, incluant leurs métadonnées, dans le cas d'un schéma en flocon.

Lorsqu'une opération OLAP spatiale est demandée, la dimension dans laquelle elle sera réalisée doit être spécifiée par l'utilisateur. L'identifiant de l'élément pointé est ensuite récupéré. Cet identifiant doit être décomposé pour déterminer l'identifiant du membre faisant partie de la dimension choisie pour subir l'opération OLAP. À l'aide des fonctions des bibliothèques OLAP, les données (parent ou enfants) correspondant au membre affecté sont déterminées. Ces données sont ensuite insérées dans une requête au cube et un nouveau résultat est renvoyé, prêt à être affiché dans les deux volets. Le volet cartographique vérifie quels sont les membres actifs de chacune des dimensions, crée l'identifiant géométrique à partir de l'identifiant de ces membres actifs, recherche les éléments géométriques correspondants dans la base de données spatiale et les affiche dans l'espace de visualisation cartographique.

Cette solution d'implantation nécessite la manipulation des tables de dimension en vue de l'ajout des identifiants. Cette manipulation est donc affectée par le type de schéma utilisé, selon les avantages et inconvénients des schémas présentés à la section 2.3. Dans ce contexte, le schéma en étoile nécessite moins de manipulations car il contient moins de tables à modifier que le schéma en flocon.

La figure 15 présente un exemple d'affichage cartographique fourni par ce type de solution. Cet affichage permet de voir la distribution spatiale des phénomènes analysés, la répartition des segments affectés par les différents types d'accidents dans le cas présent.

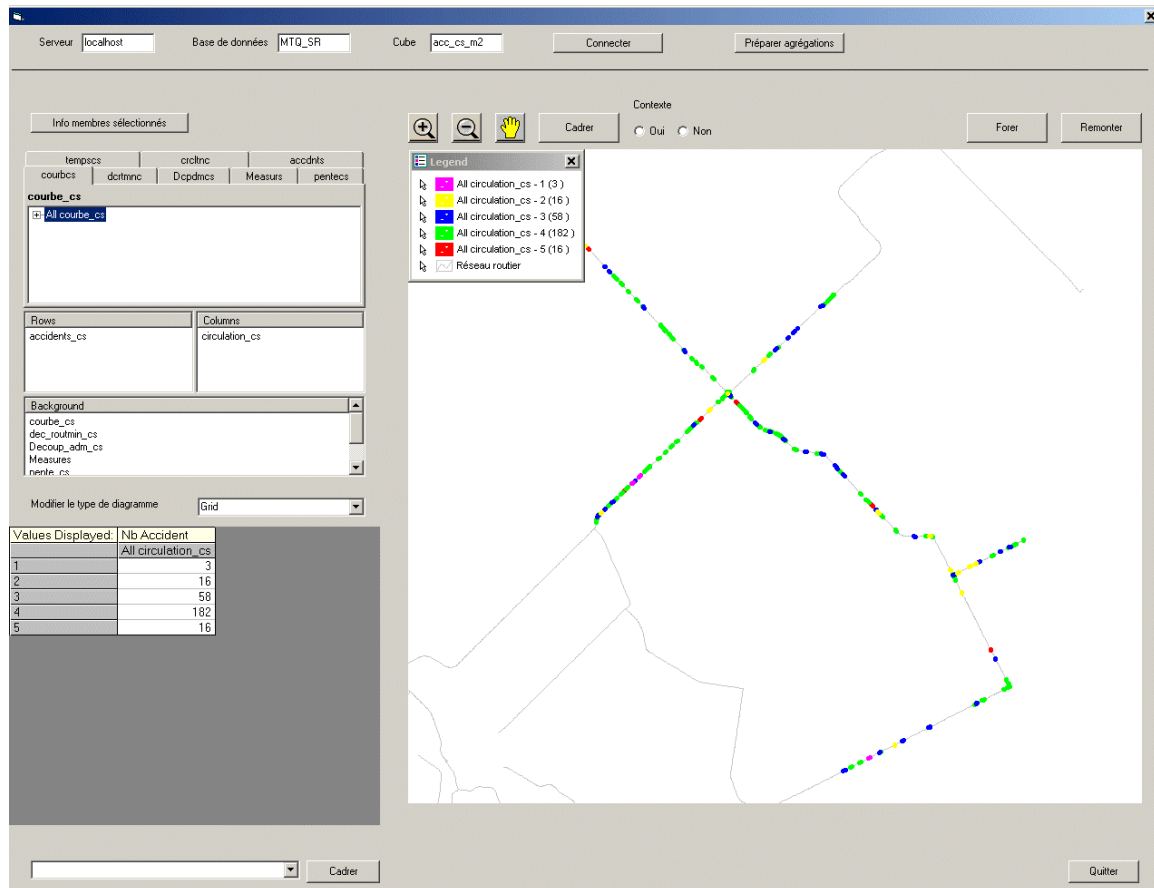


Figure 15 : Exemple d'affichage cartographique réalisé avec la solution d'implantation 3 – plusieurs dimensions spatiales géométriques.

Solution 4 : lien explicite généré, entre la structure multidimensionnelle et la structure spatiale géométrique

Le fonctionnement de cette solution est similaire à celui de la solution 3, sauf en ce qui concerne le stockage et la gestion des identifiants. La solution 3 nécessitait que l'identifiant de chacun des membres des dimensions spatiales géométriques soit présent dans la table de dimension correspondante, afin qu'il puisse être emmagasiné dans le cube comme une propriété de ce membre. La présente solution ne nécessite aucun stockage d'identifiant, ce qui permet de diminuer les manipulations nécessaires sur les tables de dimensions, et de potentiellement réduire leur volume.

Cette solution nécessite un prétraitement, générant automatiquement l'identifiant de chacun des membres des dimensions spatiales géométriques. Cet identifiant est dans le cas présent emmagasiné dans un vecteur servant de table de correspondance. Cette solution s'implante de la même façon quelle que soit l'architecture OLAP utilisée. Des essais ont été effectués et ont démontré qu'ici encore, et tel que l'on peut s'y attendre, l'architecture MOLAP est la plus performante.

5.6 Évaluation des solutions d'intégration implantées

5.6.1 Critères d'évaluation

Durant l'implantation des solutions d'intégration OLAP - SIG, des critères d'évaluation ont été définis. La présente section contient une description de ces différents critères.

a. Manipulation des données originales

Ce critère permet d'évaluer la quantité, ainsi que la nature, des manipulations à effectuer sur les données originales afin de pouvoir les utiliser à l'intérieur de la solution implantée. Une solution idéale nécessiterait peu ou pas de modifications aux données.

b. Considérations d'utilisation

Ce critère évalue les conditions d'utilisation de la solution, c'est-à-dire les traitements automatisés préalables à l'utilisation, au démarrage de l'application. Idéalement, si la solution comporte de tels traitements, ils devraient s'exécuter rapidement afin que le temps d'attente de l'utilisateur soit réduit au minimum.

c. Facilité d'intégration des nouvelles données à analyser

Ce critère évalue la quantité de travail requis afin d'intégrer de nouvelles données dans les structures implantées. Les deux composantes, descriptive et géométrique, des données sont prises en compte. Idéalement, les mises à jour aux données devraient pouvoir s'effectuer de manière rapide, ou même transparente, afin que les données les plus récentes soient disponibles en tout temps.

d. Performance

Ce critère s'intéresse strictement à la vitesse d'exécution d'opérations OLAP réalisées sur les données. Il a été testé à l'aide d'un chronomètre de haute performance, fourni par une librairie d'objets de Windows 2000. Plusieurs opérations de forage et de remontage sur la dimension du découpage routier ont été effectuées, à plusieurs reprises, en vue d'obtenir une moyenne du temps d'exécution des opérations.

e. Flexibilité à s'adapter à un nombre variable de dimensions spatiales géométriques

Ce critère évalue si la solution développée est capable de manipuler un nombre variable de dimensions spatiales géométriques.

f. Intégration des caractéristiques souhaitables d'un SOLAP

Ce critère évalue si les caractéristiques souhaitables d'un outil SOLAP telles que décrites à la section 4.2 sont, ou peuvent aisément être, intégrées à la solution. Cependant, il est nécessaire de mentionner que le développement des solutions a été axé sur l'intégration des volets descriptif et cartographique, et non sur la réalisation d'un outil SOLAP complet.

5.6.2 Comparaison des solutions d'intégration

Le tableau 2 contient les résultats comparatifs de l'évaluation des solutions d'intégration 1 et 2 implantées dans le cadre de l'essai. Le tableau 3 contient les résultats comparatifs des solutions d'intégration 3 et 4.

Comme toutes les solutions d'intégration présentent une performance optimale lorsque l'architecture MOLAP est utilisée, la comparaison est donc basée sur cette architecture seulement.

Tableau 2 : Comparaison des solutions d'intégration - une seule dimension spatiale géométrique

| Critères d'évaluation | Solution 1 | Solution 2 |
|---|---|---|
| Manipulation des données originales | Peu ou pas de modifications nécessaires | Peu ou pas de modifications nécessaires |
| Considérations d'utilisation | Aucun prétraitement nécessaire | Un prétraitement est nécessaire dans le but de stocker la géométrie dans une structure propriétaire, à chaque nouvelle connexion |
| Facilité d'intégration de nouvelles données | Aucune manipulation autre que l'ajout de données comme tel | Aucune manipulation autre que l'ajout de données comme tel |
| Performance | Plus performante que la solution 2 | Performance affectée par les manipulations nécessaires pour l'affichage |
| Flexibilité – nombre de dimensions spatiales géométriques | Pas flexible, une seule dimension spatiale géométrique possible | Pas flexible, une seule dimension spatiale géométrique possible |
| Intégration des caractéristiques souhaitables d'un SOLAP | Intègre les caractéristiques a, d, f, i, n, o, totalement ou en partie. Les autres caractéristiques, sauf c, peuvent être ajoutées, totalement ou en partie, par programmation. | Intègre les caractéristiques a, d, f, i, n, o, totalement ou en partie. Les autres caractéristiques, sauf c, peuvent être ajoutées, totalement ou en partie, par programmation. |

Tableau 3 : Comparaison des solutions d'intégration - plusieurs dimensions spatiales géométriques

| Critères d'évaluation | Solution 3 | Solution 4 |
|---|--|---|
| Manipulation des données originales | Des manipulations sont nécessaires afin de générer un identifiant pour chacun des membres et de l'emmagasiner dans les tables de dimension | Peu ou pas de modifications nécessaires |
| Considérations d'utilisation | Le calcul des agrégations est nécessaire à la première connexion à un cube | Un prétraitement est nécessaire à chaque nouvelle connexion afin de générer la table de correspondance pour les identifiants. Le calcul des agrégations est nécessaire à la première connexion à un cube |
| Facilité d'intégration de nouvelles données | Les manipulations d'identifiants doivent être refaites à chaque fois que de nouvelles données sont ajoutées. De plus, les agrégations doivent être recalculées | Les agrégations doivent être recalculées |
| Performance | Performance équivalente à la solution 4 | Performance équivalente à la solution 3 |
| Flexibilité – nombre de dimensions spatiales géométriques | Flexible | Flexible |
| Intégration des caractéristiques souhaitables d'un SOLAP | Intègre les caractéristiques a, c, d, f, i, j, n, o, totalement ou en partie. Les autres caractéristiques peuvent être ajoutées, totalement ou en partie, par programmation. | Intègre les caractéristiques a, c, d, f, i, j, n, o, totalement ou en partie. Les autres caractéristiques peuvent être ajoutées, totalement ou en partie, par programmation. |

Deux solutions optimales peuvent être déterminées, une dans le cas où le cube ne contient qu'une dimension spatiale géométrique et une autre dans le cas où le cube en contient plusieurs. Dans le premier cas, la solution 1 représente la meilleure option. En plus d'être la plus performante, elle est simple à mettre en oeuvre et à maintenir et ne requiert aucun prétraitement. Dans le cas où plusieurs dimensions spatiales géométriques peuvent être contenues dans le cube, la solution 4 est la plus simple à implanter car nécessitant le moins de manipulations de la part de l'utilisateur.

5.7 Application du mode d'intégration optimal aux données du projet OLAP - MTQ

Afin de réaliser une analyse concrète telle que décrite à la section 3.3, la solution optimale pour une dimension spatiale géométrique, la solution 1 décrite en 5.5, a été utilisée pour le cube d'état de chaussée. Ce cube représente le point de départ des analyses réalisées sur le réseau routier. Un cube général a été construit, couvrant l'ensemble des données disponibles, c'est-à-dire les données des directions territoriales de Québec et de Chaudière-Appalaches. Par la suite, un cube spécifique a été construit pour le centre de service de Lac-Etchemin. La division des données à analyser en différents cubes permet de réduire le volume de données à l'intérieur de chacun. Le cube général sert alors aux analyses de plus haut niveau, tandis que le cube

spécifique permet d'explorer en profondeur les pistes intéressantes découvertes dans le cube général.

La dimension spatiale géométrique utilisée est évidemment celle du découpage routier. Cette dimension, dans le cadre du cube général, ne contient que le niveau des routes. Dans le cadre du cube spécifique, cette dimension contient tous les niveaux définis à l'annexe 2. Lors de sa définition, au début du projet OLAP - MTQ, la dimension du découpage routier comprenait des hiérarchies alternatives pour les classes fonctionnelles de routes ainsi que pour le milieu. Ces hiérarchies alternatives ont été redéfinies comme des dimensions indépendantes dans le cadre de l'essai en vue de pouvoir les utiliser conjointement.

Les mesures qui ont été implantées dans le cube sont les suivantes :

- valeur d'état minimum;
- valeur d'état maximum;
- valeur d'état moyen;
- djma.

La dernière mesure (djma) avait été définie pour le cube « circulation ». Elle a été ajoutée au cube d'état de chaussée car elle est nécessaire aux calculs de dégradation de l'état de la chaussée. Les valeurs de ces mesures sont emmagasinées dans la table des faits. Des mesures calculées ont été ajoutées pour simuler la détérioration du réseau routier en fonction du temps :

- djma estimé;
- valeur d'état moyen estimé;
- seuil d'intervention;
- coût d'intervention.

La valeur de ces mesures calculées devrait varier en fonction du temps, de la classe fonctionnelle de route, ainsi que du milieu. Dans le cadre de la présente application, seulement la variation en fonction du temps a été implantée. La valeur de ces mesures n'est pas emmagasinée. Elle est recalculée dynamiquement à chaque utilisation.

L'architecture MOLAP a été utilisée pour la construction des cubes. La section 5.7.1 présente un exemple d'analyse réalisée à l'aide de la présente implantation.

5.7.1 Exemple d'analyse réalisée

Cette section présente une mise en situation ainsi qu'un exemple d'analyse simple réalisée à l'aide de l'application décrite à la section 5.7.

Les responsables du centre de service de Lac-Etchemin, préparant les demandes de budget de travaux routiers pour les années 2001 et 2002, effectuent une analyse sur les données d'état de chaussée du réseau sous leur juridiction. Après des opérations de pivot et de forage latéral appropriées, les données sur l'état moyen des routes, pour l'année en cours, sont affichées à l'écran. Chacune des routes est colorée selon la valeur du seuil d'intervention qui détermine si la route est en bon état (vert), nécessite une nouvelle couche de revêtement (jaune) ou une reconstruction complète (rouge). La figure 16 illustre cette situation.

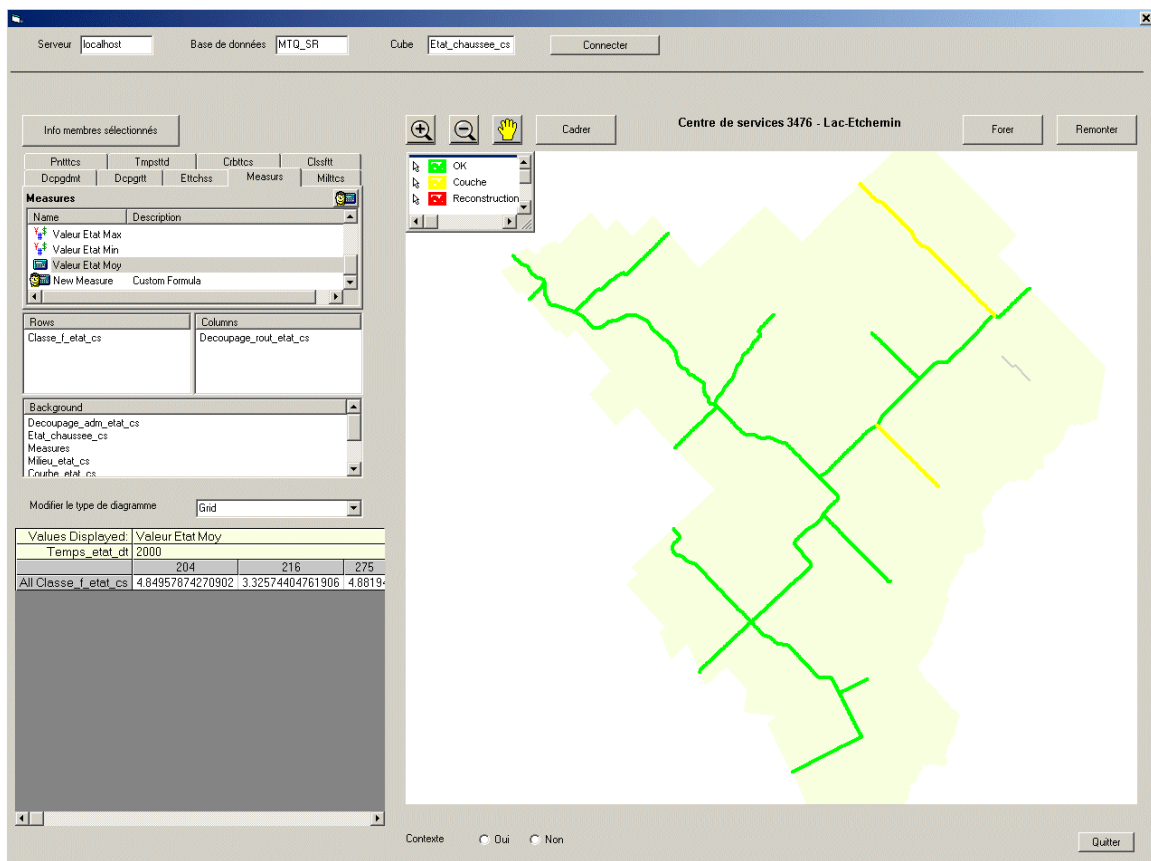


Figure 16 : Affichage des seuils d'intervention par route pour l'année 2000.

L'affichage permet de distinguer deux routes demandant potentiellement une nouvelle couche de revêtement. Un forage spatial sur la route située la plus au nord, jusqu'au niveau des sous-

routes, permet de déterminer quelle partie plus spécifique du réseau est touchée. La figure 17 illustre cette nouvelle situation.

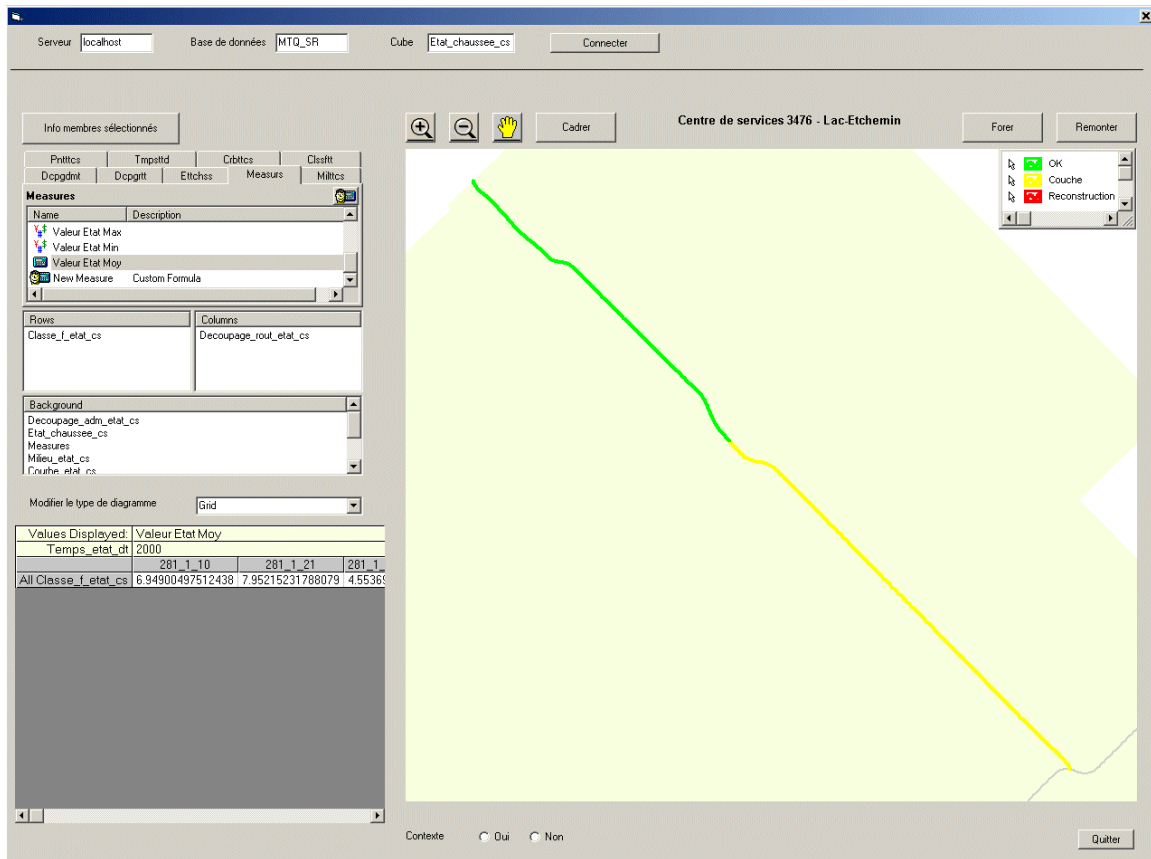


Figure 17 : Affichage des seuils d'intervention par sous-route pour l'année 2000.

Ce nouvel affichage permet donc de déterminer quelle sous-route présente un état de chaussée nécessitant une intervention. Un nouveau forage spatial, jusqu'au niveau des segments, permet de déterminer, de façon détaillée, les portions du réseau qui sont en mauvais état. La figure 18 montre les segments avec leur seuil d'intervention.

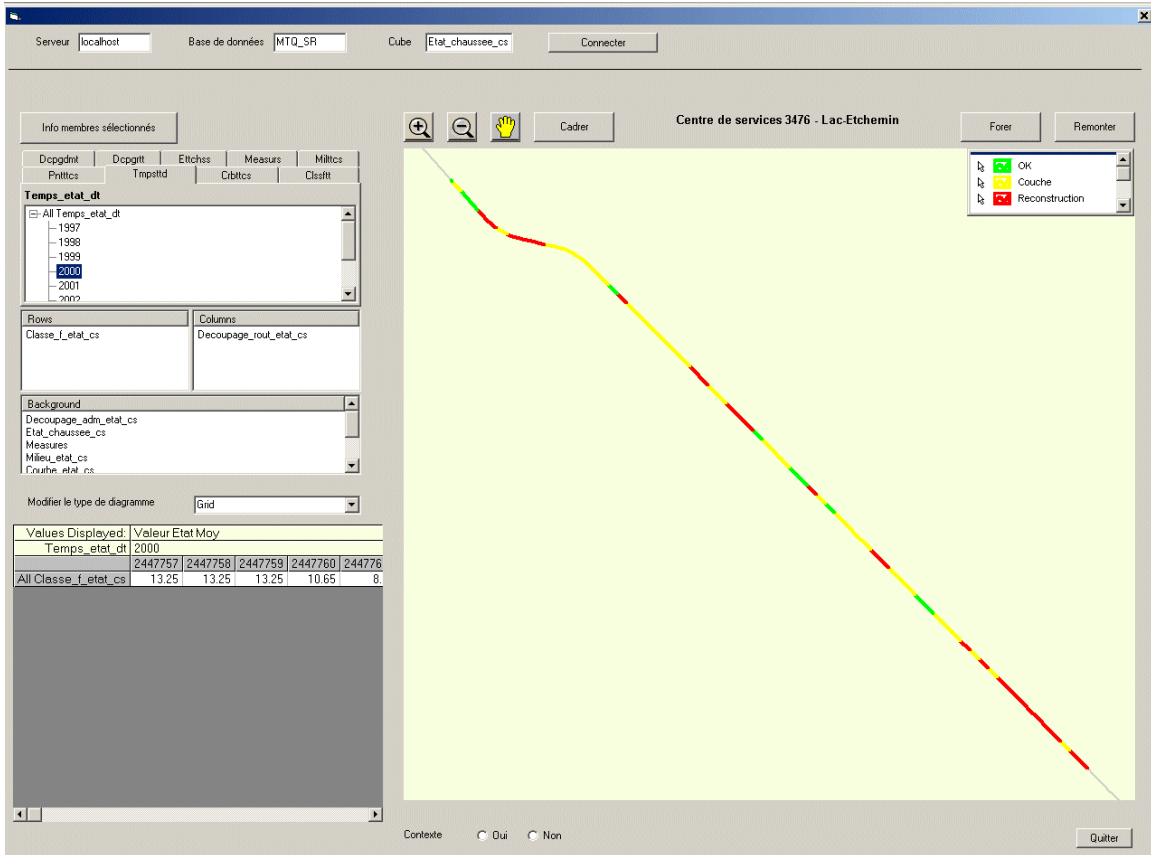


Figure 18 : Affichage des seuils d'intervention par segment pour l'année 2000.

Une opération de forage latéral, sur l'année 2002, permet ensuite de voir quel sera l'état estimé du réseau routier si aucun investissement n'est réalisé sur cette route pour l'année 2001. La figure 19 montre cet état estimé.

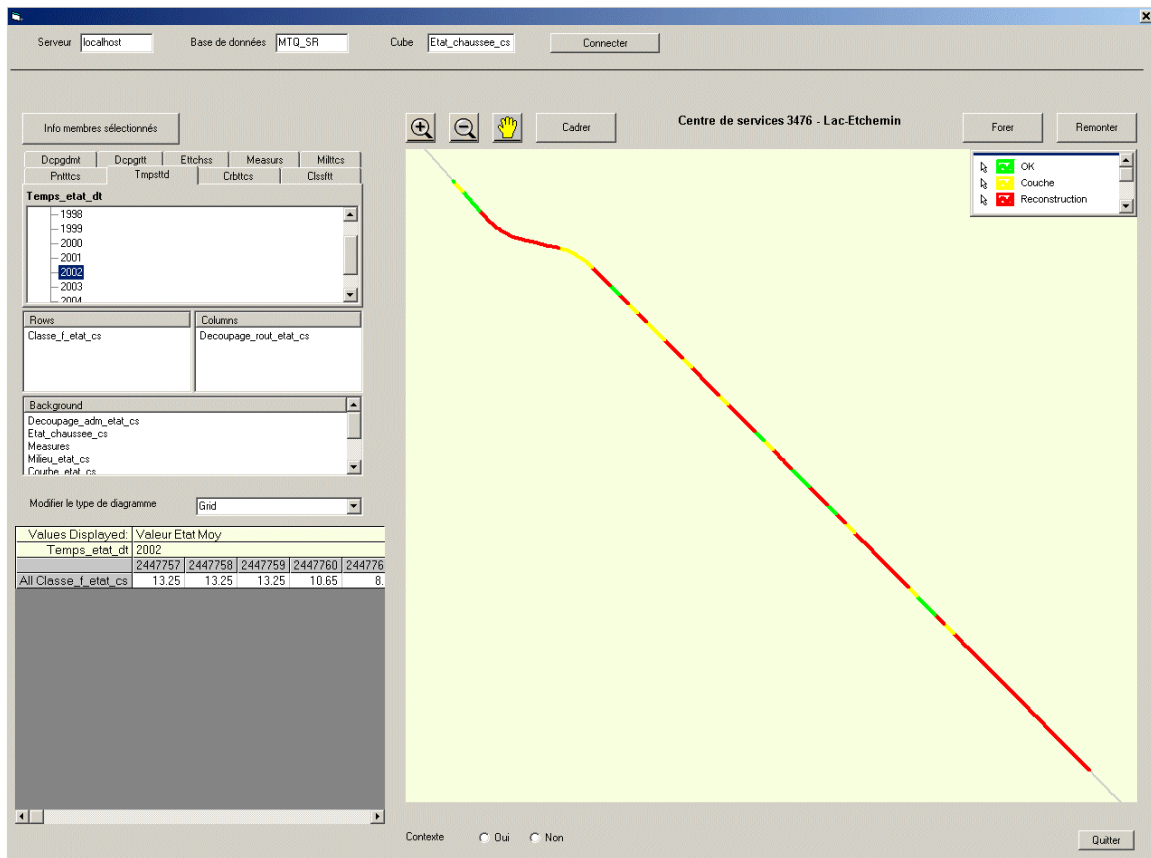


Figure 19 : Affichage des seuils d'intervention par segment pour l'année 2002.

L'analyse pourrait ensuite se poursuivre afin de déterminer le coût des interventions nécessaires afin de minimiser la dégradation de la chaussée. Cet exemple ne montre qu'une des analyses possibles réalisables à l'aide d'un outil SOLAP et a pour but de démontrer le potentiel qu'offre l'intégration du volet cartographique pour ce type d'analyse.

6 CONCLUSION

Ce document a d'abord présenté les concepts d'OLAP et le vocabulaire propre à cette technologie. Les différentes architectures, ROLAP, MOLAP et HOLAP ont été comparées, de même que les schémas utilisés à l'intérieur des architectures ROLAP et HOLAP afin de fournir une structure multidimensionnelle aux données contenues dans une base de données relationnelle. Par la suite, le projet OLAP du MTQ, projet qui a servi de cadre aux expérimentations d'intégration OLAP - SIG, a été décrit brièvement afin d'en exposer les particularités intéressantes. La section suivante a été consacrée au concept de SOLAP. Les caractéristiques essentielles d'un tel outil, ainsi que des caractéristiques supplémentaires souhaitables ont été présentées. Différentes solutions technologiques potentielles pour la réalisation d'un outil SOLAP ont aussi été décrites. Finalement, la section 5 a présenté les détails de modes d'intégration OLAP – SIG possibles et de ceux implantés et testés. Une évaluation de ces derniers a permis de déterminer deux solutions plus avantageuses à différents points de vue. Ces solutions diffèrent selon le nombre de dimensions spatiales géométriques présentes au sein du cube à analyser. La solution la plus appropriée lorsqu'une seule dimension spatiale géométrique est utilisée a été implantée sur un ensemble de données du MTQ afin de présenter un exemple d'analyse pertinent.

Toutes les solutions d'intégration discutées n'ont pas pu être implantées et testées dans le cadre de cet essai, en raison de limitations technologiques ou de contraintes de temps. De plus, d'autres possibilités d'intégration apparaîtront fort probablement à court terme au fur et à mesure que les technologies de l'information évolueront.

Les prochaines étapes à franchir en vue du développement d'un outil SOLAP idéal concernent le raffinement et l'intégration des concepts de généralisation automatique et de représentation multiple. Des recherches à ce sujet sont en cours au Centre de recherche en géomatique de l'Université Laval. De plus, des études sont également en cours concernant la gestion de la topologie spatiale à l'intérieur d'une base de données multidimensionnelle.

7 REMERCIEMENTS

Cet essai s'insère dans le cadre de deux projets de recherche et développement : premièrement le projet « Analyse multidimensionnelle et système d'information géographique appliqués à l'analyse de données de gestion routière » du Ministère des transports du Québec, et deuxièmement le projet « Conceptualisation des fondations technologiques pour la prise de décision à l'aide du World Wide Web », projet numéro deux du réseau Geoide.

Je désire premièrement remercier sincèrement le Dr Yvan Bédard pour sa disponibilité et son support tout au long du projet. Ses conseils et son aide précieuse ont été grandement utiles et appréciés. En second lieu, je désire offrir mes remerciements à M. Simon Plante, du Ministère des transports du Québec, qui a fourni les données ayant servi aux différents essais des solutions d'intégration OLAP – SIG. Enfin, j'aimerais remercier Marie-Josée Proulx, Suzie Larrivée, François Lemieux et Alexandre Brisebois pour leur implication dans les projets mentionnés plus haut.

8 BIBLIOGRAPHIE

- Advanced Visual Systems, 1998, *AVS Express Product Information*, <http://www.avs.com/products/expovr.htm>.
- Archer Decision Sciences, 1995, *Star Schema 101: A pictorial Introduction to Dimensional Modeling*, <http://www.archer-decision.com/knowpre.htm>.
- Bédard, Y., 1997, *Spatial OLAP*, Vidéoconférence, 2ème Forum annuel sur la R-D, Géomatique VI: Un monde accessible, 13-14 novembre 1997, Montréal.
- Bédard, Y., 2000a, *Fondements des SIG*, Notes du cours GMT-20843, Faculté de foresterie et géomatique, Université Laval.
- Bédard, Y., 2000b, *Qu'est-ce que Perceptory?*, <http://sirs.scg.ulaval.ca/perceptory/introduc.htm>.
- Bédard, Y. & M.-J. Proulx, 1998, *Les outils informatiques d'exploitation de données*, Notes de cours pour la formation continue, Faculté de foresterie et géomatique, Université Laval.
- Bédard, Y., R. Devillers, E. Bernier & B. Moulin, 2000, Automatic Generalisation or Multiple Representations: Towards a combination of both approaches. GEOIDE DEC#9 Workshop on feature extraction and cartographic generalization, Université Laval, Québec, November 17th.
- Bédard, Y., S. Larrivée, M.-J. Proulx, P.-Y. Caron & F. Létourneau, 1997, *Gespatial Data Warehousing : Positionnement technologique et stratégique*, Rapport préparé pour le Centre de recherche de la défense de Valcartier, Université Laval, 79 pp.
- Bédard, Y., T. Merrett & J. Han, 2000, *Fundamentals of Spatial Data Warehousing for Geographic Knowledge Discovery*, à paraître dans : *Geographic Data Mining and Knowledge Discovery*, H. Miller and J. Han (éditeurs), Research Monographs in GIS series, Taylor & Francis.
- Bourgon, J.-F., J. Brodeur, F. Lemieux, P. Marchand, & S. Rivest, 1999, *Fondements théoriques et pratiques des serveurs universels pour l'exploitation multidimensionnelle des très grandes bases de données*, Rapport préparé dans le cadre du cours SCG-62058, Faculté de foresterie et géomatique, Université Laval, 67 pp.
- Business Objects, 1998, *Business Objects and ESRI Announce Strategic Partnership*, Business Objects Press Releases, <http://www.businessobjects.com/news/>.
- Caron, P.Y., 1998, *Étude du potentiel de OLAP pour supporter l'analyse spatio-temporelle*, Mémoire de M. Sc., Faculté de foresterie et géomatique, Université Laval, 132 pp.
- Codd, E.F., S.B. Codd & C.T. Salley, 1993, *Providing OLAP (On-Line Analytical Processing) to User-Analysts: An IT Mandate*, Hyperion white papers, <http://www.hyperion.com>, 20 pp.
- Cognos, 2000, *Cognos Visualizer Product Information*, <http://www.cognos.com/products/visualizer>.
- ESRI, 2000, *ArcView GIS Brochure*. http://www.esri.com/library/whitepapers/av_lit.html, 20 pp.
- Franklin, C., 1992, *An Introduction to Geographic Information Systems: Linking Maps to Databases*, Database, April 1992, pp. 13-21.

- Gill, H.S. & P.C. Rao, 1996, *The official Guide to Data Warehousing*, QUE Corporation, 382 pp.
- Gonzales, M.L., 1999, *Spatial OLAP – Conquering Geography*, DB2 Magazine, Spring 1999, http://www.db2mag.com/db_area/archives/1999/q1/99sp_gonz.shtml.
- Han, J., N. Stefanovic, & K. Koperski, 1998, *Selective materialization: An Efficient Method for Spatial Data Cube Construction*, Proceedings of the 1998 Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD'98), Melbourne, Australia, pp. 144-158.
- Hettler, M., 1997, *OLAP Serves Up Your Data*. Byte, March 1997, p.130.
- Inmon, W.H., 1996, *Building the Data Warehouse*, Second Edition, John Wiley & Sons, 410 pp.
- ISO/TC 211, 1999, Draft International Standard, Geographic Information – Spatial Schema, 143 pp.
- Knosys, 2000a, *Geo Spatial Mapping: MapInfo MapX Plug-In for Knosys ProClarity 3.0*, <http://www.knosysinc.com/productextent1.html>.
- Knosys, 2000b, *ProClarity OLAP Client Fact Sheet*, http://www.knosysinc.com/fact_sheet/olap_client.pdf.
- Lemieux, F., 2000, *Conception d'une base de données pour l'analyse et la planification stratégique du réseau routier québécois*, Essai de M. Sc., Faculté de foresterie et géomatique, Université Laval, 35 pp.
- LGS Group Inc., 2000, *Analysis of Health Surveillance Business Intelligence Tools and Applications*, Final Draft, 111 pp.
- Marchand, P., 1998, *Identification des capacités et limitations relatives au développement d'un SOLAP avec les technologies Geomedia et Powerplay*, Rapport de stage, Faculté de foresterie et géomatique, Université Laval.
- Marchand, P., 2000, *Concepts fondamentaux pour un nouveau type d'interface exploitant la multidimensionnalité des phénomènes à référence spatiale (avec application à la santé environnementale)*, Proposé d'études de doctorat, Faculté de foresterie et géomatique, Université Laval.
- Martel, C., 1999, *Développement d'un cadre théorique pour la gestion des représentations multiples dans les bases de données spatiales*, Mémoire de M. Sc., Faculté de foresterie et de géomatique, Université Laval, 128 pp.
- Microsoft Corporation, 2000, *Microsoft SQL Server 2000 documentation*.
- OLAP Council, 1995, *OLAP and OLAP Server Definitions*, <http://www.olapcouncil.org/research/glossaryfy.htm>.
- OpenMap, 2000, *What is OpenMap?*, <http://openmap.bbn.com/whatis.html>.
- Proulx, M.-J. & E. Bernier, 2000, *Document interne du projet OLAP-MTQ*, Centre de recherche en géomatique, Université Laval.
- Rinn, N., 1998, *Making the case for Star Schema*, Data Warehouse Report, July 28, 1998, <http://datawarehouse.dci.com/articles/1998/07/28star.htm>.
- Shekhar, S., C. Lu, X. Tan, S. Chawla, & R.R. Vatsavai, 2000, *Map Cube: A Visualization Tool for Spatial Data Warehouses*, à paraître dans : *Geographic Data Mining and Knowledge*

Discovery, H. Miller and J. Han (éditeurs), Research Monographs in GIS series, Taylor & Francis.

Stefanovic, N., 1997, *Design and Implementation of On-Line Analytical Processing (OLAP) of Spatial Data*. Mémoire de M. Sc., Simon Fraser University, 118 pp.

Sterling, S., 1998, *Essential Modeling Option*, Teradata Review, Fall 1998.
<http://www.teradatareview.com/Fall98/sterling.html>.

Thomsen, E., 1997, *OLAP solutions: building multidimensional information systems*. John Wiley & Sons, 608 pp.

Thomsen, E., G. Spofford, & D. Chase, 1999, *Microsoft OLAP Solutions*. John Wiley & Sons, 509 pp.

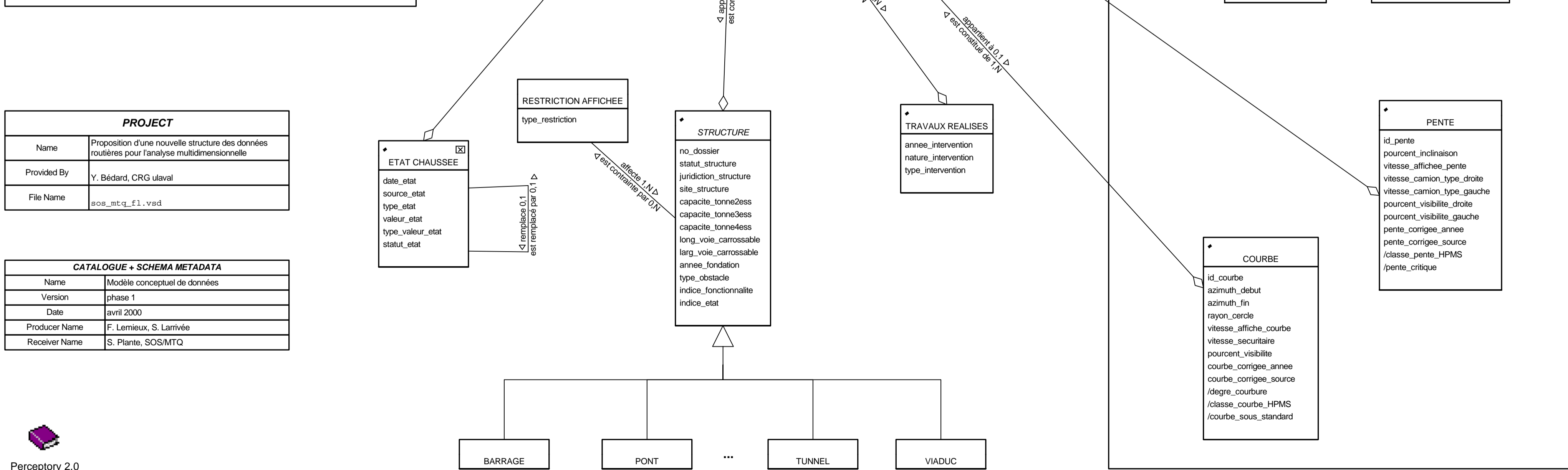
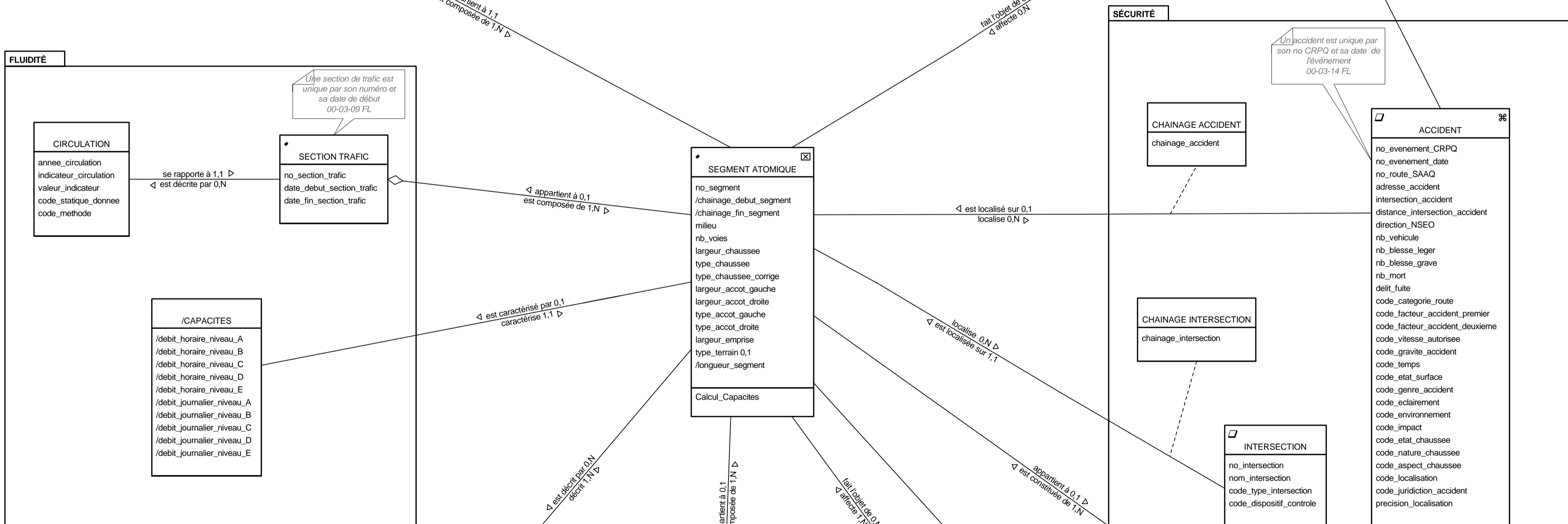
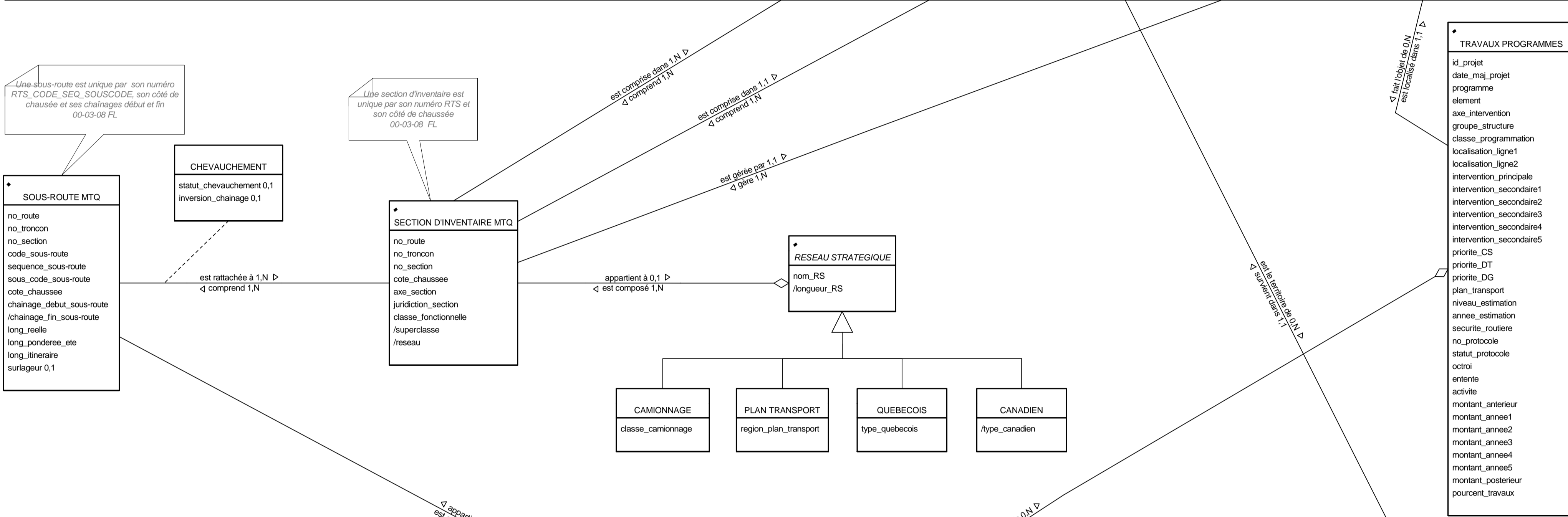
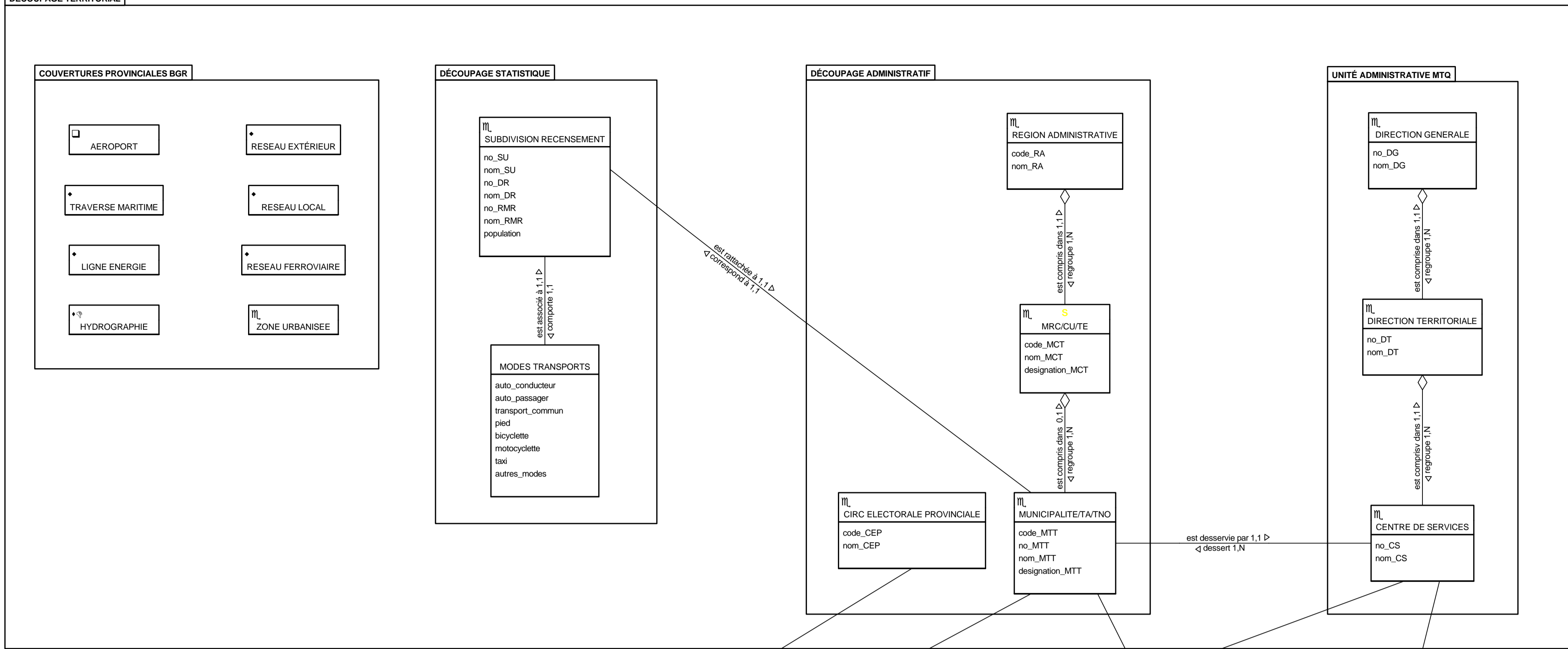
Tiedrich, A.H., 2000, *Microsoft Corp. SQL Server 2000 Analysis Services white paper*,
<http://www.microsoft.com/sql/productinfo/SQL2000.pdf>.

Zhou, X., D. Truffet, & J. Han, 1999, *Efficient Polygon Amalgamation Methods for Spatial OLAP and Spatial Data Mining*, Proceedings of the 6th International Symposium on Large Spatial Databases (SSD'99), Hong Kong, pp. 167-187.

ANNEXE 1

Modèle conceptuel de classes d'objets pour le projet du MTQ [Lemieux 2000], reproduit ici afin de permettre la comparaison avec les structures multidimensionnelles de la section 2.3.

DECOUPAGE TERRITORIAL



ANNEXE 2

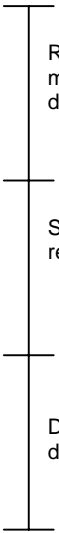
Ce document contient la description des dimensions d'analyse du projet OLAP du Ministère des transports du Québec. Adapté de Proulx et Bernier [2000].

Temps



Année

Découpage de recensement



Région
métropolitaine
de recensement

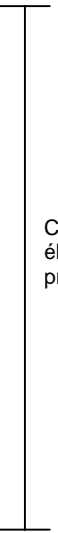


Subdivision de
recensement



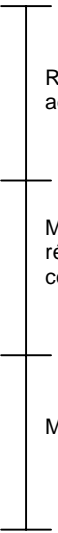
Division
de recensement

Découpage électoral (CEP)



Circonscription
électorale
provinciale

Découpage administratif



Région
administrative

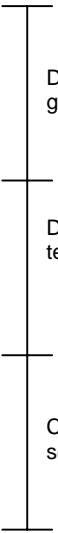


Municipalité
régionale de
comté (MRC)



Municipalité

Découpage corporatif



Direction
générale

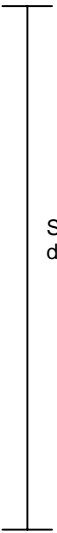


Direction
territoriale



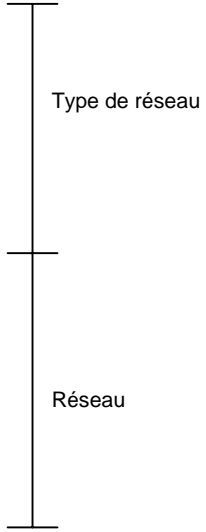
Centre de
service

Circulation

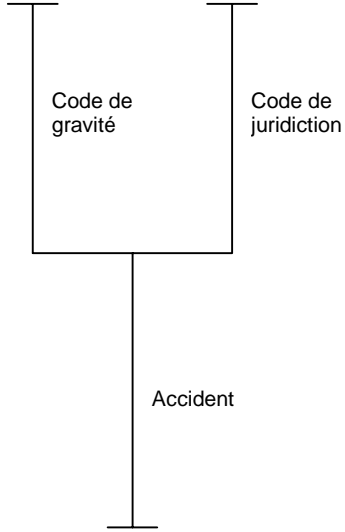


Section
de trafic

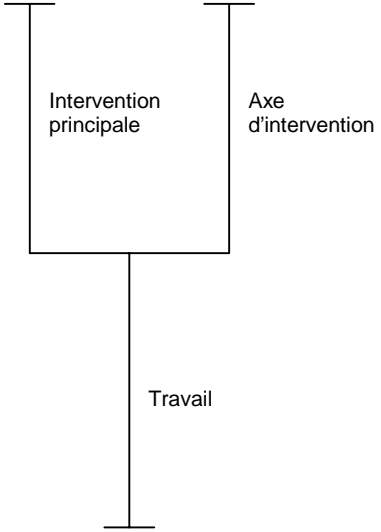
**Réseau
stratégique**



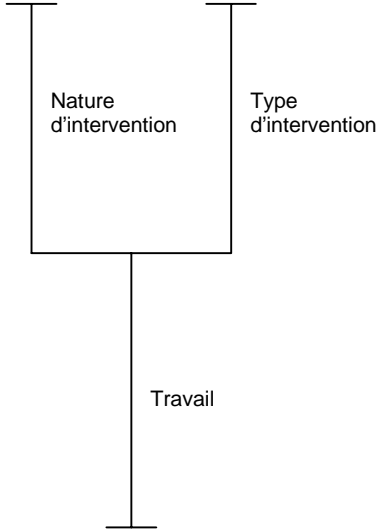
Accidents



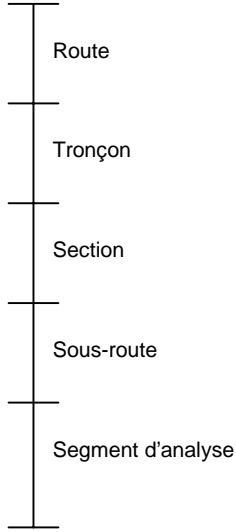
**Travaux
programmés**



**Travaux
réalisés**



**Découpage
routier**

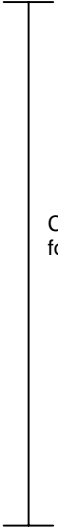


Juridiction



Juridiction

Super-classification



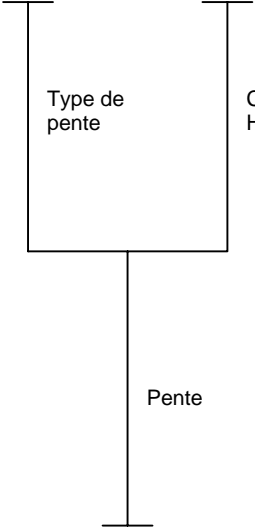
Classe
fonctionnelle

Milieu



Milieu

Pentes

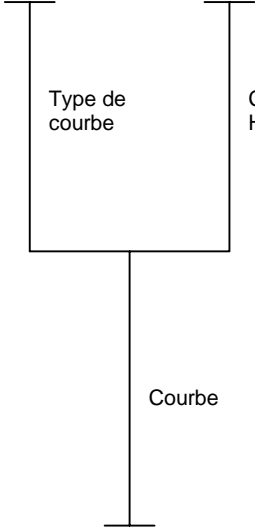


Type de
pente

Classe
HPMS

Pente

Courbes

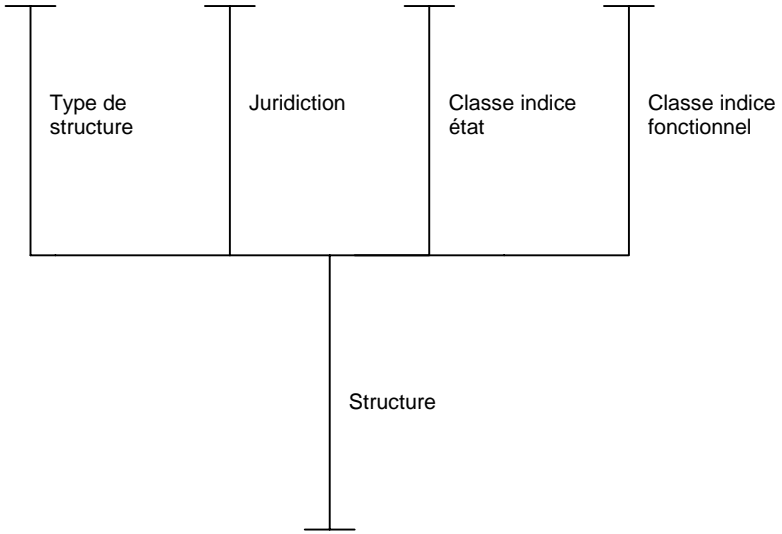


Type de
courbe

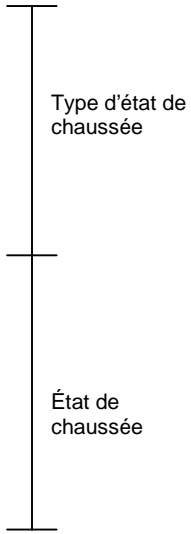
Classe
HPMS

Courbe

Structure



État chaussée



ANNEXE 3

Ce document contient la description des mesures et des cubes d'analyse du projet OLAP du Ministère des transports du Québec. Adapté de Proulx et Bernier [2000].

| Cubes | | Dimensions | | | | | | | | | | | | | | |
|--|-------------------------|------------|--------------------------|---------------------|-------------------------|----------------------|-------------|--------------------|-----------|--------------------|------------------|-------------------|--------|---------|------------|---------------|
| Opérateurs d'agrégation Σ : sommation M : moyenne Mi : minimum Ma : maximum Re : recalcul | Opérateurs d'agrégation | Temps | Découpage de recensement | Découpage électoral | Découpage administratif | Découpage corporatif | Circulation | Réseau stratégique | Accidents | Travaux programmés | Travaux réalisés | Découpage routier | Pentes | Courbes | Structures | État chaussée |
| Travaux réalisés | | | | | | | | | | | | | | | | |
| Longueur (km) | Σ | ↑ | | | | ↑ | | ↑ | | | ↑ | ↑ | | | ↑ | |
| Travaux programmés | | | | | | | | | | | | | | | | |
| Longueur (km) | Σ | ↑ | | | | ↑ | ↑ | ↑ | | ↑ | | ↑ | | | ↑ | |
| Coût (\$/km) | Σ | | | | | | | | | | | | | | | |
| Circulation | | | | | | | | | | | | | | | | |
| Débit JMA MOY | M | ↑ | ↑ | ↑ | | ↑ | ↑ | ↑ | | | | ↑ | | | | |
| Débit JMH MOY | M | | | | | | | | | | | | | | | |
| Débit JME MOY | M | | | | | | | | | | | | | | | |
| 30 ^e heure MOY | M | | | | | | | | | | | | | | | |
| Heure de pointe MOY | M | | | | | | | | | | | | | | | |
| % directionnel MOY | M | | | | | | | | | | | | | | | |
| % camion | M | | | | | | | | | | | | | | | |
| Taux de variation annuel MOY | M | | | | | | | | | | | | | | | |
| Facteur de l'HP MOY | M | | | | | | | | | | | | | | | |
| Distance parcourue : (DJMA * long.) | Σ | | | | | | | | | | | | | | | |
| Distance parcourue pond. : (DP * nb. Voie) /2 | Σ | | | | | | | | | | | | | | | |
| Accident | | | | | | | | | | | | | | | | |
| Nombre d'accidents | Σ | ↑ | | | | ↑ | ↑ | ↑ | ↑ | | | ↑ | ↑ | ↑ | ↑ | |
| Nombre de véhicules | Σ | | | | | | | | | | | | | | | |
| Nombre de blessés léger | Σ | | | | | | | | | | | | | | | |
| Nombre de blessés grave | Σ | | | | | | | | | | | | | | | |
| Nombre de morts | Σ | | | | | | | | | | | | | | | |
| Structure | | | | | | | | | | | | | | | | |
| Nombre de structures | Σ | | | | ↑ | ↑ | | ↑ | | | | ↑ | | | ↑ | |
| Longueur de structure | | | | | | | | | | | | | | | | |
| Superficie de structure | | | | | | | | | | | | | | | | |
| Indice d'état MOY, MIN, MAX | M Mi Ma | | | | ↑ | ↑ | | ↑ | | | | ↑ | | | ↑ | |
| Indice de fonctionnalité MOY, MIN, MAX | M Mi Ma | | | | ↑ | ↑ | | ↑ | | | | ↑ | | | ↑ | |

| Cubes | | Dimensions | | | | | | | | | | | | | | |
|--|-------------------------|------------|--------------------------|---------------------|-------------------------|----------------------|-------------|--------------------|-----------|--------------------|------------------|-------------------|--------|---------|------------|---------------|
| Opérateurs d'agrégation Σ : sommation M : moyenne Mi : minimum Ma : maximum Re : recalcul | Opérateurs d'agrégation | Temps | Découpage de recensement | Découpage électoral | Découpage administratif | Découpage corporatif | Circulation | Réseau stratégique | Accidents | Travaux programmés | Travaux réalisés | Découpage routier | Pentes | Courbes | Structures | État chaussée |
| Réseau routier | | | | | | | | | | | | | | | | |
| Longueur réelle Longueur pondérée été Longueur d'itinéraire | Σ | | | ↑ | ↑ | ↑ | | ↑ | | | | ↑ | ↑ | ↑ | | |
| Largeur de chaussée MOY, MIN, MAX Largeur d'emprise MOY, MIN, MAX Largeur d'acc. D MOY, MIN, MAX Largeur d'acc. G MOY, MIN, MAX | M Mi Ma | | | | | | | | | | | ↑ | ↑ | ↑ | | |
| Nombre d'intersections | Σ | | | | | ↑ | | ↑ | | | | ↑ | | | | |
| Courbe | | | | | | | | | | | | | | | | |
| Nombre de courbes Longueur de courbe | Σ | | | | | ↑ | | ↑ | | | | ↑ | ↑ | ↑ | | |
| Moyenne pondérée degré courbure/longueur | Re | | | | | ↑ | | ↑ | | | | ↑ | ↑ | ↑ | | |
| Pente | | | | | | | | | | | | | | | | |
| Nombre de pentes Longueur de pente | Σ | | | | | ↑ | | ↑ | | | | ↑ | ↑ | ↑ | | |
| Moyenne pondérée % inclin./long. | Re | | | | | ↑ | | ↑ | | | | ↑ | ↑ | ↑ | | |
| État de chaussée | | | | | | | | | | | | | | | | |
| Valeur d'état MOY, MIN, MAX | M Mi Ma | ↑ | | ↑ | ↑ | ↑ | | ↑ | | | | ↑ | ↑ | ↑ | | ↑ |
| Prof. IRI été MOY, MIN, MAX Prof. IRI hiver MOY, MIN, MAX Longueur d'ornièrè MOY, MIN, MAX | M Mi Ma | ↑ | | ↑ | ↑ | ↑ | | ↑ | | | | ↑ | ↑ | ↑ | | ↑ |
| % Couverture d'ornièrè % Couverture d'IRI | Re | ↑ | | ↑ | ↑ | ↑ | | ↑ | | | | ↑ | ↑ | ↑ | | ↑ |

ANNEXE 4

Ce document contient un tableau comparatif des fonctionnalités de deux logiciels appartenant à la famille de solutions d'intégration avec serveur OLAP et SIG (OLAP dominant) : ProClarity 2.0 avec le plugiciel KMapX de Knosys et Visualizer 1.5 de Cognos.

Description et comparaison des fonctionnalités des logiciels ProClarity 2.0, avec le plugiciel KMapX (Knosys), et Visualizer 1.5 (Cognos)

ProClarity 2.0

ProClarity, de la compagnie Knosys, est un logiciel client OLAP qui permet de manipuler des cubes de données créés à l'aide d'Analysis Services de Microsoft SQL Server. ProClarity permet de visualiser les données descriptives d'un cube sous différentes formes graphiques telles que des tableaux ou d'autres diagrammes. Le plugiciel KMapX, développé à l'aide de MapX de la compagnie MapInfo, permet la visualisation, sous forme cartographique, des données géométriques associées à une dimension spatiale géométrique d'un cube.

Les données géométriques nécessaires doivent être emmagasinées sous forme de couches dans un « *geoset* ». Un utilitaire, le *Geoset Manager*, est livré avec KMapX. Cet utilitaire est basé sur la technologie MapX de MapInfo et permet d'intégrer des fichiers de format TAB (MapInfo) dans un « *geoset* » utilisable par KMapX.

Le plugiciel KMapX offre un utilitaire qui permet de définir le fichier de configuration nécessaire pour la visualisation des données spatiales géométriques. Cet utilitaire se nomme *ProClarity MapX Configuration Tool*. Le fichier de configuration contient les paramètres suivants :

- le cube accédé;
- la dimension spatiale du cube à visualiser sous forme de carte;
- le « *geoset* » correspondant à la dimension spatiale géométrique à visualiser cartographiquement.

Le fichier de configuration permet aussi d'associer chacun des niveaux hiérarchiques de la dimension spatiale géométrique spécifiée en paramètre à une couche du « *geoset* ». Par exemple, si l'on désire visualiser des données à l'aide d'une carte des États-Unis et pouvoir forer sur chacun des états afin d'obtenir des informations sur les comtés appartenant à ces états, il faut tout d'abord associer une couche des états au niveau « État » de la dimension spatiale. Par la suite, il est nécessaire d'associer une couche des comtés, une seule couche pour tous les comtés des États-Unis, au niveau « Comté » de la dimension spatiale.

À l'ouverture de ProClarity, l'utilisateur doit spécifier le cube auquel il désire avoir accès, ainsi que le type de graphique à visualiser. Si la visualisation est de type cartographique, le nom du fichier de configuration doit aussi être spécifié. L'utilisateur peut alors effectuer ses analyses.

Visualizer 1.5

Visualizer, de la compagnie Cognos, est un logiciel de visualisation de données pouvant provenir de sources diverses telles qu'un serveur OLAP ou une base de données relationnelle.

Visualizer permet de visualiser les données descriptives d'un cube OLAP sous différentes formes graphiques telles que les tableaux et autres diagrammes. Visualizer permet l'affichage cartographique des données spatiales géométriques associées à une dimension, à l'aide de fonctionnalités de MapX de la compagnie MapInfo.

Les données géométriques à afficher doivent être emmagasinées sous forme de couches dans un « *geoset* ». L'utilitaire *Geoset Manager* est livré avec Visualizer.

Dans Visualizer, chacune des visualisations, ainsi que chacune des opérations de forage possibles doivent être définies à l'avance. Une couche du « *geoset* » est associée à un niveau hiérarchique d'une dimension spatiale, par exemple le niveau supérieur. Ensuite, il est possible d'ajouter une opération de forage sur un ou plusieurs éléments de la carte, un « *map drill* ». Cette opération de forage doit être reliée à une autre couche du « *geoset* ». C'est cette couche complète qui s'affichera lorsque l'utilisateur exécutera l'opération de forage prédéfinie. Par exemple, si l'on désire visualiser des données à l'aide d'une carte des États-Unis et permettre à l'utilisateur de forer sur chacun des états pour obtenir des informations sur les comtés, il faut tout d'abord associer une couche des états au niveau « État » de la dimension spatiale. Par la suite, il faut ajouter un « *map drill* » pour chacun des états, un à un, et associer à cette opération une couche des comtés contenus dans l'état traité, c'est-à-dire une nouvelle couche de comtés pour chacun des états.

La visualisation définie est sauvegardée dans un fichier qui peut être ouvert à l'intérieur de Visualizer.

| Fonctionnalité | ProClarity 2.0 avec KMapX | Visualizer 1.5 |
|---|---|--|
| <i>Dimensions et mesures</i> | | |
| Nombre de dimensions | Limite non documentée. | Limite non documentée. |
| Nombre de mesures | Limite non documentée. | Limite non documentée. |
| Nombre de dimensions spatiales géométriques | Une, mais elle peut posséder des hiérarchies alternatives. | Plusieurs, mais la visualisation ne s'applique qu'à un niveau d'une seule dimension à la fois. |
| Traitement des hiérarchies alternatives d'une même dimension spatiale géométrique | KMapX va permettre d'associer une couche géométrique d'un « <i>geoset</i> » à chacun des niveaux des hiérarchies alternatives d'une dimension spatiale géométrique. ProClarity utilise la première hiérarchie d'une dimension par défaut. Le logiciel permet à l'utilisateur de sélectionner la hiérarchie dans laquelle il veut effectuer des opérations. | Visualizer utilise la première hiérarchie par défaut. Il permet cependant à l'utilisateur de sélectionner la hiérarchie dans laquelle il veut effectuer des opérations. |
| Nombre de dimensions descriptives visualisables à la fois | Limite non documentée (le nombre d'axes dans un diagramme, ainsi que le nombre de diagrammes sont ajustés selon le nombre de dimensions visualisées). | 4 dimensions au maximum, selon le type de visualisation choisi (différents types de diagrammes 2D ou 3D). |
| Nombre de dimensions spatiales géométriques visualisables à la fois | Une (et une seule hiérarchie à la fois). | Une. Visualizer permet de définir les données spatiales géométriques affichées au départ et correspondant à un niveau hiérarchique de la dimension spatiale géométrique. Par la suite, des « <i>map drill</i> » peuvent être ajoutés. Chacun de ces « <i>map drill</i> » permet d'afficher une nouvelle carte qui peut représenter un membre d'une autre dimension spatiale géométrique. |
| Nombre de mesures visualisables à la fois | Limite non documentée (cependant, avec la visualisation cartographique offerte par KMapX, lorsque le nombre de mesures augmente, la clarté des informations affichées diminue). | 2 au maximum (si on utilise une visualisation 3D). |
| <i>Visualisation</i> | | |
| Nombre de visualisations à la fois | 2 au maximum. | 80 cartes par panneau, 80 panneaux par scène, 80 scènes par feuille, 80 feuilles par visualisation. |
| Type de visualisation cartographique (statique ou dynamique) | Dynamique. Il est possible d'effectuer des opérations OLAP directement dans la carte. | Dynamique. |
| Lien entre le volet cartographique et le volet descriptif | Une opération dans un des volets met à jour la visualisation dans l'autre volet et inversement. | Le lien est possible seulement si des axes sont inclus dans la visualisation cartographique (c'est-à-dire qu'il y aura un certain nombre de cartes différentes selon les membres des dimensions placées sur les axes). |

| Fonctionnalité | ProClarity 2.0 avec KMapX | Visualizer 1.5 |
|--|---|---|
| <i>Visualisation (suite)</i> | | |
| Possibilité de visualiser plusieurs cartes à la fois (par exemple une carte pour chacune des mesures que l'on désire visualiser) | Oui, pour un maximum de 4 cartes à la fois, selon le choix des membres de dimensions et des mesures à afficher (le logiciel affiche automatiquement les cartes) (Voir note 1). | Oui, mais chacune des cartes doit être construite à l'avance. |
| <i>Entités géométriques</i> | | |
| Types d'entités géométriques vectorielles supportés | Points, lignes, polygones. | Points, polygones. |
| Entités géométriques complexes supportées | Oui, mais ne supporte pas plusieurs éléments ayant la même valeur pour l'attribut utilisé pour faire le lien avec les données de la base de données multidimensionnelle (par exemple si deux polygones représentent un même comté, ces deux polygones doivent être regroupés en un seul polygone complexe). | Oui, mais ne supporte pas plusieurs éléments ayant la même valeur pour l'attribut utilisé pour faire le lien avec les données de la base de données (par exemple si deux polygones représentent un même comté, ces deux polygones doivent être regroupés en un seul polygone complexe). |
| Possibilité de combiner les types d'entités géométriques vectorielles à l'intérieur d'une même dimension spatiale géométrique | Oui. Un type par niveau. hiérarchique. | Oui. Un type par niveau. hiérarchique. |
| Format matriciel supporté | Non (car le logiciel utilise des « <i>geosets</i> » MapInfo. Aucun format matriciel actuellement supporté par MapInfo ne permet de relier des attributs aux pixels). | Non (car le logiciel utilise des « <i>geosets</i> » MapInfo. Aucun format matriciel actuellement supporté par MapInfo ne permet de relier des attributs aux pixels). |
| <i>Traitement des données</i> | | |
| Possibilité de choisir les agrégations significatives à afficher | Non. | Non. |
| Possibilité de filtrage sur des attributs | Une forme de filtrage, à la visualisation, permet d'isoler ou d'éliminer des membres de dimensions des visualisations en cours. | Oui, les filtres doivent être définis à l'avance. |
| <i>Traitement des données spatiales géométriques</i> | | |
| Possibilité d'utiliser ses propres cartes | Oui, en incluant les couches nécessaires dans un « <i>geoset</i> » MapInfo. Les propriétés « visible » et « selectable » doivent être désactivées pour chacune des couches. | Oui, en incluant les couches nécessaires dans un « <i>geoset</i> » MapInfo. L'ordre des attributs de chacune des couches est important. Il est conseillé de placer le champ qui fait le lien avec les données OLAP en première position. |
| Possibilité de modifier la symbologie des cartes visualisées (à la visualisation) | Seulement la symbologie des éléments liés à des informations OLAP visualisées peut être modifiée. | Seulement la symbologie des éléments liés à des données OLAP visualisées peut être modifiée, ainsi que la couleur des éléments ne correspondant à aucune donnée de la base de données. |
| Organisation des données spatiales géométriques en couches | Chaque niveau hiérarchique est associé à une couche différente. | Chaque niveau hiérarchique est associé à une ou plusieurs couches. |

| Fonctionnalité | ProClarity 2.0 avec KMapX | Visualizer 1.5 |
|--|---|--|
| <i>Traitement des données spatiales géométriques (suite)</i> | | |
| Organisation des couches de données spatiales géométriques | Les couches sont regroupées dans un seul « <i>geoset</i> » et doivent être placées en ordre de granularité croissante (du plus détaillé au plus général) si on veut voir tous les éléments. | Les couches sont regroupées en un ou plusieurs « <i>geosets</i> ». À l'intérieur d'un « <i>geoset</i> », les couches doivent être placées en ordre de granularité croissante. |
| Possibilité d'afficher des cartes en contexte | Oui, en les incluant dans le « <i>geoset</i> ». | Oui, en les incluant dans un ou des « <i>geosets</i> ». |
| Affichage des données spatiales géométriques lors de la visualisation | La carte présentant les données du niveau supérieur de toutes les dimensions est affichée par défaut. Suite à une opération, la carte est mise à jour et présente les éléments correspondant au résultat de l'opération. | Une carte de départ est affichée. Cette carte est celle qui était affichée lors de la fermeture précédente de la visualisation. Chaque opération de forage ou de remontage est associée à une nouvelle couche. Celle-ci est affichée au complet lorsque l'opération est effectuée. |
| Passage d'un niveau hiérarchique à un autre dans la dimension spatiale géométrique | Passage d'une couche du « <i>geoset</i> » à une autre. Seulement les éléments correspondant au résultat de l'opération OLAP sont affichés. | Lors d'une opération de forage ou de remontage, une nouvelle couche complète est affichée (et non seulement les enfants du membre qui a subi le forage, par exemple). |
| Concepts de généralisation automatique et de représentation multiple supportés | Non. | Non. |
| Possibilité d'emmagasiner des données spatiales géométriques historiques | Aucune facilité. | Aucune facilité. |
| Gestion des relations spatiales topologiques | Non. | Non. |
| <i>Opérations</i> | | |
| Opérations OLAP (forage, remontage, forage latéral et pivot). | <p>Les opérations de pivot et de forage latéral s'effectuent dans le contrôle « dimension » qui est commun aux deux volets.</p> <p>Les opérations de forage et remontage peuvent s'effectuer directement dans la visualisation (en pointant sur un élément du diagramme ou de la carte). Ces opérations sont aussi disponibles à partir d'un menu (qui permet de choisir le membre qui subira l'opération).</p> <p>D'autres opérations sont aussi disponibles à partir du menu seulement :</p> <ul style="list-style-type: none"> expansion (« <i>expand</i> »), contraction (« <i>collapse</i> »), élimination (« <i>eliminate</i> »), isolement (« <i>isolate</i> »). <p>(Voir note 2).</p> <p>Toutes ces opérations se répercutent dans les volets descriptif et cartographique simultanément.</p> | <p>Dans le volet cartographique, les opérations de forage et de remontage sont possibles directement (en utilisant le bouton de droite de la souris et en choisissant le membre cible), mais chacune des opérations doit être prévue à l'avance en définissant un « <i>map drill</i> ».</p> <p>Les opérations de pivot ou de forage latéral ne sont pas possibles directement dans une visualisation cartographique. Une nouvelle visualisation cartographique doit être construite pour afficher des données différentes.</p> |
| Opérations OLAP sur plusieurs éléments à la fois. | Possible. | Possible mais les opérations doivent être prédéfinies. |

| Fonctionnalité | ProClarity 2.0 avec KMapX | Visualizer 1.5 |
|--|---|---|
| <i>Opérations (suite)</i> | | |
| Analyse sur les deux volets de manière parallèle ou indépendante (on peut décider de naviguer seulement dans l'interface cartographique ou dans l'interface descriptive) | Même si une des visualisations est fermée, elle reste synchronisée avec celle qui est ouverte. Pour une navigation complètement indépendante, deux sessions différentes doivent être ouvertes. | Les deux sont possibles. Les visualisations d'un même ensemble de données ne sont pas nécessairement liées. |
| Niveaux de granularité descriptifs et cartographiques différents | Possible. Cependant, lorsque la visualisation cartographique demeure affichée et qu'une opération mène à un niveau qui n'est pas associé à des données spatiales géométriques, un changement d'échelle s'effectue (la carte s'affiche à très petite échelle). | Possible. |
| <i>Requêtes</i> | | |
| Possibilité d'effectuer des requêtes directement sur la base de données OLAP | Oui, à l'aide d'un assistant ou d'un éditeur MDX. Nécessite une dimension « temps ». | Oui, à l'aide de Cognos Impromptu. |
| <i>Légende</i> | | |
| Possibilité de modifier la légende (légende interactive) | Choix du nombre de classes ainsi que leur symbologie. Possibilité d'une classification automatique : - quantiles - intervalles égaux - coupures naturelles - écart-type. | Modification interactive des classes et de la palette de couleurs. Possibilité d'ajouter un seuil (« water line »). |
| Opérations OLAP via la légende | Forage et Remontage. | Non. |
| <i>Connexions</i> | | |
| Types de connexions possibles (SGBD, serveur OLAP) | Cube Analysis Services de Microsoft SQL Server exclusivement. | Plusieurs types de connexions possibles (serveurs OLAP, BD relationnelles, fichiers Microsoft Excel et fichiers texte). |
| Possibilité de connecter à un cube virtuel | Oui. | Oui. |
| <i>Environnement de développement</i> | | |
| Environnement de développement intégré (si oui, quels outils de développement sont supportés) | Oui, toutes les bibliothèques de ProClarity sont utilisables lors du développement d'applications personnalisées avec Visual Basic, Visual Interdev, C++ et Delphi. | Non. |
| <i>Maintenance</i> | | |
| Passage du transactionnel au analytique (facilité de mise à jour). | Des modifications aux données de la dimension spatiale géométrique peuvent nécessiter la modification des couches MapInfo, la modification du « geoset », ainsi que la modification du fichier de configuration. | Des modifications aux données de la dimension spatiale géométrique peuvent nécessiter la création d'une nouvelle visualisation. |
| <i>Déploiement sur le Web</i> | | |
| Possibilité de déploiement sur le Web. | Non (mais possible avec la nouvelle version 3.0) | Oui, nécessite cependant Visualizer Web edition. |

Note 1.

Dans ProClarity, les dimensions peuvent être organisées selon 6 axes, chacun pouvant contenir plusieurs dimensions. Les trois premiers sont l'axe des rangées, l'axe des colonnes et l'axe des

filtres. L'axe des rangées et l'axe des colonnes correspondent aux axes des différents diagrammes visualisables (tableau, diagramme à barres, ...). Ils peuvent être représentés de façon orthogonale. L'axe des filtres contient les dimensions en arrière-plan. Pour les dimensions qui sont placées sur les axes des rangées et des colonnes, plusieurs membres peuvent être actifs (visualisés) à la fois. Pour chacune des dimensions de l'axe des filtres, un seul membre doit être actif à la fois.

ProClarity classe les visualisations cartographiques en quatre types : « *single ranged* », « *multiple ranged* », « *single chart* » et « *multiple chart* » qui dépendent du nombre de mesures à visualiser et du nombre de dimensions visualisables (placées sur les axes).

Lorsque seulement un membre de chacune des dimensions autres que la dimension spatiale géométrique, et que seulement une mesure, sont actifs, le type « *single ranged* » est utilisé. ProClarity n'affiche qu'une seule carte montrant les différentes valeurs de mesure à l'aide d'un symbolisme particulier pour chacun des éléments géométriques.

Lorsque plusieurs membres d'une dimension située sur le même axe que la dimension spatiale géométrique sont actifs, le type « *multiple ranged* » est utilisé. ProClarity affiche plusieurs cartes, une carte pour chacun des membres actifs de la dimension située sur le même axe que la dimension spatiale géométrique. Les autres dimensions doivent comprendre seulement un membre actif.

Lorsque plusieurs membres d'une dimension située sur l'axe opposé à l'axe de la dimension spatiale géométrique sont actifs, le type utilisé est le « *single chart* ». ProClarity affiche une seule carte contenant les éléments de la dimension spatiale géométrique, chacun de ces éléments étant associé à un diagramme (circulaire ou à barres) montrant la valeur de la mesure pour chacun des membres actifs de la dimension située sur l'axe opposé à l'axe contenant la dimension spatiale géométrique.

Dans les autres cas, le type utilisé est le « *multiple chart* ». Plusieurs cartes du type « *single chart* », pour un maximum de quatre, sont affichées à la fois à l'écran.

Note 2.

L'opération d'expansion permet d'effectuer un forage sur un des membres d'une dimension visualisée. Tous les autres membres de la dimension demeurent aussi affichés (contrairement au forage où seulement les enfants du membre qui a subi le forage sont affichés).

L'opération de contraction est l'opération inverse de l'expansion.

L'opération d'élimination permet d'éliminer un membre d'une dimension de la visualisation.

L'opération d'isolement permet de visualiser seulement un membre particulier d'une dimension.