



Le rôle clé de l'analyse et du profiling des sources, en amont de la construction de data warehouses alimentés par PowerCenter.

ENJEU

- Déploiement d'un environnement décisionnel avec création de plusieurs data warehouses et data marts

SOLUTIONS INFORMATICA

- Informatica PowerCenter®
- Axio (désormais intégré dans Informatica Data Explorer)

ARCHITECTURE

- Étude des sources : Axio/Informatica
- Intégration de données : Informatica
- Sources : divers ERP, SGBD et systèmes des filiales et partenaires
- Cible : base de données Oracle
- Restitution : SAS pour les experts ; Business Objects pour les utilisateurs avertis (chargés d'études du siège), et WEBI (de BO) pour les utilisateurs du réseau commercial

BÉNÉFICES

- Gain de temps et de productivité grâce à l'automatisation de l'analyse préalable des données sources
- Accélération du développement grâce à la cartographie exhaustive des données disponibles et exploitables pour chaque DWH
- Connaissance des sources permettant une approche participative favorisant l'appropriation par les utilisateurs
- Respect des délais du fait de l'industrialisation de la démarche et de l'élimination des mauvaises surprises sur le contenu des sources

Le Groupe SMABTP trouve son origine dans un groupement d'entrepreneurs parisiens constitué en 1859. Il a pour vocation de répondre à l'ensemble des besoins en assurance des professionnels du bâtiment et des travaux publics. Leader sur son secteur, le Groupe représente le quart du marché national de l'assurance construction. Deux sociétés mutuelles regroupent ses activités : SMABTP, pour l'assurance des dommages aux biens et de responsabilité, et SMAvie BTP, pour l'assurance des personnes (épargne, retraite, prévoyance et santé). Le Groupe, qui emploie plus de 2 000 collaborateurs sur l'ensemble du territoire, est au service de 115 300 assurés et a réalisé un chiffre d'affaires de 1 836 millions d'euros en 2006.

“Le fait d'avoir, grâce à l'analyse préalable de la qualité des sources, une connaissance exhaustive des données disponibles et exploitables pour le futur data warehouse facilite la collaboration avec les directions utilisatrices. Nous pouvons ainsi leur dire immédiatement s'il est possible ou non de construire les indicateurs qu'elles souhaitent, ce qui accélère considérablement l'élaboration et la validation des spécifications fonctionnelles, tout en évitant des allers-retours et des frustrations.”

— Pascale Chauvin, Chef de Projet Informatique Décisionnelle, Groupe SMABTP

Le contexte initial

En 1999, SMABTP décide de se doter d'un système décisionnel visant à satisfaire les besoins d'analyse et de reporting de ses différentes activités. Outre le choix d'outils de restitution, ce projet passe par la création de data warehouses (DWH) devant être alimentés par des sources hétérogènes : plusieurs ERP et SGBD, données des filiales et des partenaires... « Le décisionnel était un monde nouveau pour nous », se souvient Pascale Chauvin, aujourd'hui Chef de Projet Informatique Décisionnelle au sein du Groupe. « Pour l'extraction et l'intégration des données, nous avons étudié les produits du marché et arrêté notre choix sur la plate-forme PowerCenter d'Informatica. Déployée dans un nombre significatif d'entreprises françaises, cette solution se démarquait par la complétude de ses fonctionnalités ETL, l'étendue de sa connectivité et sa rapidité de mise en oeuvre. En tant que société, Informatica présentait aussi toutes les garanties : références prestigieuses, solidité financière et forte activité R&D ».

Comment s'assurer de la qualité des données sources ?

Première étape, la construction du DWH Commercial en 2000 oblige l'équipe mixte chargée du projet à consacrer un temps considérable à l'analyse des sources. « La documentation des systèmes sources est généralement incomplète et rarement suffisante pour se faire une idée juste de l'existant. Or, pour éviter pertes de temps et frustrations, nous tenions à savoir exactement quelles étaient les données disponibles et exploitables avant que les utilisateurs expriment leurs besoins », explique Pascale Chauvin. « Faute d'outil adapté, pour le premier data warehouse, nous avons réalisé ce travail à la main. C'est pourquoi, pour les étapes suivantes, nous avons décidé d'utiliser Axio, un logiciel d'analyse de données que nous avons découvert dans le cadre d'un récent projet de migration d'application ».

Remarque : Axio était alors édité par Evoke Software, racheté en 2005 par Similarity Systems. Lorsqu'en janvier 2006, Informatica a racheté Similarity Systems, les fonctionnalités d'Axio ont été intégrées dans la solution de profiling Informatica Data Explorer.

Un profiling systématique avant d'élaborer les modèles cibles

Pour les DWH suivants, l'équipe de SMABTP a systématisé l'utilisation de ces fonctionnalités. Cette approche élimine les suppositions et permet la conception de modèles de données et d'applications correspondant vraiment aux attentes des utilisateurs. Elle apporte des réponses précises à nombre de questions : où l'information est-elle pertinente ? Quelle sont les relations entre les tables de ce progiciel ? Où retrouver dans le modèle du système opérant l'information saisie en

front-office ? Où sont les redondances, les valeurs exogènes ? L'automatisation de ces investigations et analyses permet d'obtenir en peu de temps une vision claire de toutes les entités, tables et données (nombre de valeurs distinctes, nombre de lignes dans tables ou fichiers, % des valeurs distinctes permettant de distinguer les champs clés des champs informationnels, les formats de colonnes, les valeurs mini et maxi...) – autrement dit, une cartographie complète des données utilisables à des fins décisionnelles.

Un atout pour satisfaire les futurs utilisateurs

« Grâce à ces pré-études, nous identifions et éliminons rapidement du périmètre les tables vides et les informations liées au fonctionnement des ERP et autres sources », précise Pascale Chauvin. « Surtout, nous réduisons l'écart entre les attentes et ce qu'il est techniquement possible de faire compte tenu de l'état des données ». La pré-étude réalisée pour chaque DWH permet à l'équipe projet de travailler de façon interactive et constructive avec les futurs utilisateurs. Lors des réunions, l'équipe peut immédiatement donner l'origine des valeurs présentes dans une base de données, signaler les exceptions à prévoir et valider sur-le-champ l'insertion ou le rejet des règles discutées. Le gain de temps et d'efficacité est considérable dans la mesure où les spécifications fonctionnelles générales (SFG), rédigées par les directions utilisatrices à l'issue de ces réunions, se fondent sur les données exploitables et tiennent compte des contraintes de gestion. Le modèle de données du DWH ainsi que les règles de transformation et de gestion associées peuvent ensuite être rapidement construits et validés. Les besoins des utilisateurs sont satisfaits au travers d'indicateurs calculés dynamiquement ou stockés de façon statique dans des data

marts. Les DWH et data marts sont alimentés par PowerCenter.

Un outil devenu indispensable

Grâce à cette démarche, SMABTP a pu finaliser en 2002 le DWH de l'infocentre, en 2003 celui de l'activité Vie. Deux projets en cours pour l'activité IARD (DWH Production et DWH Recouvrement) seront livrés en 2008, date à laquelle sera lancé le DWH Règlements pour cette même activité.

« Cette approche basée sur la connaissance des sources nous permet de livrer un lot fonctionnel tous les 3 ou 4 mois » souligne Pascale Chauvin. « Les utilisateurs étant impliqués très en amont, il n'y a que quelques ajustements à faire en phase de recette. Tous les cas ayant été prévus, il n'y a pas non plus de surprise en pré-production et en production : les incidents d'exploitation sont réduits et les anomalies fonctionnelles très rares. Sans Axio et PowerCenter, nous n'aurions sans doute jamais pu industrialiser à ce point nos processus d'analyse et d'intégration de données ! ».

A PROPOS D'INFORMATICA

Informatica Corporation est le leader des fournisseurs indépendants de logiciels et de services d'intégration et de qualité de données. Les solutions Informatica permettent aux entreprises de tirer davantage de valeur de leurs informations en intégrant toutes leurs sources de données.

Plusieurs milliers d'entreprises dans le monde s'appuient sur Informatica pour réduire les coûts et les délais de réponse à leurs besoins d'intégration de données, quelles qu'en soient l'échelle et la complexité. Pour plus d'informations, consulter le site www.informatica.com/fr

INFORMATICA
The Data Integration Company™

Informatica France : Immeuble Le Linéa, rue du Général Leclerc - 92047 Paris La Défense Cedex (France)
Tél. : + 33 1 42 04 89 00 - Fax : + 33 1 42 04 89 01 - www.informatica.com/fr

Siège international : 100 Cardinal Way, Redwood City, CA 94063 (USA)
Tél. : + 1 650 385 5000 - Fax : + 1 650 385 5500 N° Vert USA : + 1 800 970 1179 - www.informatica.com

Présence internationale :

Allemagne • Amérique du Sud • Autriche • Australie • Belgique • Canada • Espagne • Etats-Unis • France • Italie • Inde • Japon • Pays-Bas • Portugal • Royaume-Uni • Singapour • Suisse

© 2009 Informatica Corporation. Tous droits réservés. Imprimé en France. Informatica, le logo Informatica, la signature « The Data Integration Company », Informatica PowerCenter et Informatica DataExplorer sont des marques commerciales ou des marques déposées d'Informatica Corporation aux Etats-Unis et/ou dans d'autres pays. Les autres noms de sociétés ou de produits cités sont la propriété de leurs détenteurs respectifs et peuvent avoir fait l'objet d'un dépôt de marque.

(20/05/2009)