

Cas d'utilisation réelle de Nautilus : calculs d'indicateurs chez un opérateur télécom

Adrien Schmidt, Serge Fantino

Squid Solutions
83, rue de Paris
92100 Boulogne-Billancourt
adrien@squidsolutions.com
serge@squidsolutions.com

Résumé. Nautilus est un logiciel d'analyse de bases de données. Le but de cette application est de généraliser l'utilisation de données clients au sein des entreprises. Elle facilite l'accès aux données en permettant de visualiser et manipuler les données du SGBD sous forme de concepts métiers. Elle inclut un générateur de requêtes SQL et un outil de gestion de tâches désignées pour l'agrégation de grands volumes de données. Le principe de fonctionnement est basé sur l'enchaînement de phases permettant la création des données d'analyse : importation des métadonnées du SGBD ; construction d'un dictionnaire de des concepts métiers ; spécification des champs à calculer. Les différents traitements tels que les jointures et l'alimentation des tables sont optimisés afin de rendre l'application utilisable sur des SGBD d'entreprise.

1 Introduction

Nautilus est un outil d'analyse de bases de données qui facilite la préparation de données agrégées et formatées mises à disposition des outils de datamining et de restitution.

Nautilus est bâti autour d'un système d'abstraction des données couplé à un moteur de requêtes. L'application modélise l'environnement de données en construisant une pile de métadonnées, permettant de visualiser, documenter, et de manipuler des données du SGBD sous la forme de concepts métiers (produits, trafic, revenus, segment, etc.). Ce système de paramétrage est relié à un générateur de requêtes SQL et un gestionnaire des tâches conçus pour l'agrégation de volumes importants de données.

Nautilus permet ainsi de produire de manière rapide et fiable de grands volumes de données d'analyse « à l'intérieur » des SGBD, sans nécessiter d'extractions de données. De ce fait, Nautilus remplace avantageusement les scripts et les datamarts métiers créés ad hoc tant pour les performances que la sécurité.

Nautilus, disponible depuis mai 2007, est développé sous la forme de plug-in Eclipse.

La construction de jeux de données avec Nautilus se déroule en deux étapes :

- Modélisation des données du SGBD ;
- Construction des indicateurs

2 Modélisation des données

La première étape consiste à modéliser les concepts métiers sur les données du SGBD. Nautilus se connecte à la base de données via JDBC et effectue les opérations de rétro-ingénierie du SGBD en important ses métadonnées.

Une fois que le modèle de données est intégré, l'utilisateur crée le dictionnaire des objets métier dont il se servira pour spécifier les indicateurs : par ex. la notion "d'appel entrant".

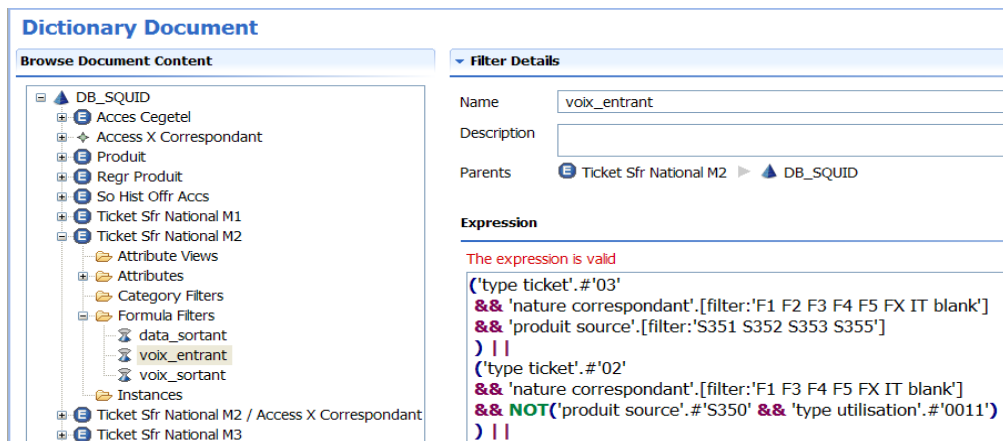


FIG. 1 – Edition d'une règle métier dans le dictionnaire Nautilus.

3 Construction d'indicateurs

La définition des éléments métier du dictionnaire permet de spécifier les champs à calculer. Ces calculs se font avec les fonctions natives du SGBD : agrégations, comptages, statistiques ainsi que les fonctions de datamining de plus en plus intégrées aux SGBD.

Nautilus génère automatiquement le SQL pour créer les indicateurs au sein du SGBD. Afin de répondre aux contraintes des départements informatiques de grandes entreprises, l'application décompose les calculs en une série de tâches qui rendent plus lisibles les calculs et permettent d'optimiser les traitements.

L'application maintient les paramètres des calculs effectués pour faciliter leur partage et leur réutilisation dans l'entreprise.

Summary

Nautilus is a database analysis software. The purpose of the application is to generalize the usage of customer data within the enterprise. It helps business users access data by mapping business concepts onto the database. Nautilus includes a metadata management module and an optimized SQL generator to compute large volumes of data. Nautilus is developed as Eclipse plug-ins.