

Enrichissement de l'architecture ANSI/SPARC pour expliciter la sémantique des données: une approche fondée sur les ontologies.

Chimène Fankam*, Stéphane Jean*, Guy Pierra*, Ladjel Bellatreche*

*LISI / ENSMA - Université de Poitiers
86961 Futuroscope Cedex
{fankamc,jean,pierra,bellatreche}@ensma.fr,

Résumé. L'architecture de bases de données ANSI/SPARC a été principalement définie pour permettre l'accès aux données indépendamment de leur représentation physique. La conception d'une base de données selon cette architecture passe par la transformation d'un modèle conceptuel en un modèle logique. Cette transformation peut engendrer une perte de sémantique de données. Ce qui pose des problèmes d'intégration et d'échange de plusieurs bases de données, ou de génération des interfaces d'accès aux données pour un utilisateur final. En tant que modèle permettant d'exprimer la sémantique des données, les ontologies constituent une solution pertinente à ces problèmes. Dans cet article, nous montrons le besoin d'étendre l'architecture ANSI/SPARC par l'ajout d'un niveau ontologique, permettant de conserver les ontologies décrivant la sémantique des données contenues dans une base de données. Notons que cette extension n'aura aucune incidence sur les applications conçues autour de l'architecture initiale. Nous analysons cette architecture en termes de besoins d'exploitation puis discutons de l'implantation d'une telle architecture.

1 Introduction

Actuellement les bases de données représentent l'outil principal de stockage de données permettant aux utilisateurs d'accéder simultanément aux données. Construire une base de données passe souvent par plusieurs étapes : (1) la modélisation conceptuelle, (2) la modélisation logique, (3) la modélisation physique et (4) la modélisation externe. Un modèle conceptuel est une représentation de la connaissance propre à un domaine. Il est caractérisé par (a) le domaine auquel il s'intéresse, (b) le formalisme qui a permis de le définir (modèle de Chen - connu sous le nom de modèle Entité-Association), et enfin (c) le contexte ou le point de vue qu'il souhaite représenter, et qui définit les questions auxquelles il vise à répondre. Le modèle conceptuel est traduit vers un modèle logique qui est une spécification des données telle quelle sera implémentée sur le système de base de données. Cette traduction passe par l'application de certaines règles. Le modèle physique définit la façon selon laquelle sont stockées les données et les méthodes pour y accéder (tous les mécanismes d'indexation par exemple). Des modèles externes ou les vues utilisateurs permettent d'adapter les données fournies aux