



UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Information integration among Heterogeneous and Autonomous Applications

Benabdelkader, A.

Publication date
2002

[Link to publication](#)

Citation for published version (APA):

Benabdelkader, A. (2002). *Information integration among Heterogeneous and Autonomous Applications*. Febo Druk.

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Résumé

L'intégration de l'Information entre des Applications Heterogènes et Autonomes

Une grande variété d'applications réparties émergent de nos jours dans des domaines divers. Ces applications déploient divers systèmes de base de données pour la gestion de leur information, dont la diversité provient des contraintes spécifiques de la gestion de l'information et des objectifs à atteindre.

Certaines applications diffèrent dans leurs caractéristiques principales et des fonctionnalités exigées. Entre autre, ces applications diffèrent dans leur architecture (distribuée/centralisée), leur taille, leur complexité, et le type de données qu'elles manipulent. Leurs besoins sont définis par les fonctionnalités globales qu'elles doivent fournir et du niveau exigé d'interopération avec d'autres sites géographiquement distribués. Les Systèmes de Gestion de Base de Données utilisés (SGBDs) sont choisis pour répondre aux caractéristiques et aux contraintes spécifiques de chaque application. Malheureusement, les SGBDs disponible actuellement n'offrent pas la possibilité de les adapter efficacement pour tout type d'applications. Certains SGBDs s'adaptent mieux aux petites applications, alors que d'autres ont été spécialement conçu pour les applications complexes ; et par conséquent sont plus adaptés à la gestion de l'information multimedia et aux données de grandes tailles. Ainsi, *toute tentative dans la direction de contraindre différentes applications à utiliser le même système de gestion de base de données pour gérer tous leurs services d'information est peu réaliste*. Même dans le cadre d'un même environnement, il arrive souvent que l'on soit amené à utiliser plusieurs systèmes de gestion de bases de données.

Il est clair d'après les études faites sur les applications, décrites dans la présente thèse, que les nouvelles applications exigent l'accès à des données stockées dans plusieurs bases de données préexistantes, détenues à plusieurs emplacements géographiquement distribués et distantes. Par conséquent, un des critères principaux exigé par la plupart des systèmes complexes, est d'avoir la possibilité à la collaboration et de fournir des mécanismes d'intégration de l'information entre les systèmes répartis, hétérogènes, et autonomes. Il s'est avéré d'après les études faites sur des applications réelles que le développement des systèmes permettant l'interopérabilité et l'intégration des données entre des systèmes répartis et autonomes est loin d'être trivial. Nous avons également appris du développement de ces applications que fournir l'intégration de l'information et l'interopérabilité entre les systèmes répartis, par l'intermédiaire du déploiement des bases de données standards et des nouvelles technologies d'Internet, est l'une des approches les plus prometteuses dans le domaine d'intégration de l'information hétérogènes entre des sites autonomes. Afin de répondre aux nouvelles contraintes de gestion de l'information dans des systèmes avancés et complexes, un système

robuste d'intégration de l'information doit être conçu et développé, servant les besoins d'intégration de l'information et d'interopération entre ces organismes.

Dans ce contexte, le travail décrit dans cette thèse est dédié à la conception et au développement d'un Système Générique et Flexible d'Intégration de l'Information. La conception de **GFI₂S** est basée sur la recherche, l'évaluation, et la validation des méthodologies et approches discutées dans le chapitre 2. En plus, le développement de **GFI₂S** est motivé par l'expertise acquise lors du développement des divers projets de R&D adressés en chapitres 3, 4, et 5 de la présente thèse. La flexibilité de **GFI₂S** réside dans sa capacité à ajouter/enlever un nouveau système à/de la fédération avec le minimum d'effort. La flexibilité dans **GFI₂S** est obtenue grâce à l'utilisation de l'architecture spécifique de deux-composants, alors que, sa Généricité est assurée par le déploiement des base de données standards et des nouvelles technologies de l'Internet.

Le travail décrit dans le chapitre 6 illustre les composants du système **GFI₂S** qui permettent l'intégration de différents types de données entre des applications hétérogènes. L'architecture de **GFI₂S** est constituée de deux composants principaux : (1) la couche locale d'adaptation qui facilite l'accès aux sources de données dans les nœuds locaux, et (2) la couche de fédération de nœuds qui fournit des liens à l'information et aux applications externes et supporte le partage d'information et d'interopération. L'architecture de **GFI₂S** permet l'intégration d'une grande variété d'applications existantes avec des moyens efficaces, tout en préservant leur hétérogénéité, distribution, et pleine autonomie.

- l'hétérogénéité vient du fait que chaque application peut appliquer son propre système de gestion de bases de données, et sa propre représentation de données hétérogènes en terme de structures et de sémantique.
- la distribution est liée au stockage et au traitement d'information sur des sources de données distribués, situés sur différents sites géographiquement distribués.
- l'autonomie est liée au fait que chaque base de données au niveau de chaque nœud de la fédération est un système indépendant. Typiquement, une base de données locale est préexistante à la création d'un réseau de coopération et possédant sa propre politique d'administration, et ses groupes d'utilisateurs.

Les fonctionnalités principales de l'approche d'intégration de l'information de **GFI₂S** résident dans: (a) la combinaison spécifique des base de données standards et des nouvelles technologies d'Internet avec les approches fondamentales de recherches, et (b) le moyen par lequel elles sont déployés et liés aux éléments spécifiques du **GFI₂S**. Ces deux fonctionnalités rendent l'approche de **GFI₂S** distincte de toutes autres approches existantes de fédération/intégration, et présentent **GFI₂S** comme une *solution générique* fournissant une *architecture flexible et évolutive* pour l'intégration de l'information et l'interopération entre des sites hétérogènes, distribués et autonomes.