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.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:27 Jun 2022

Abstract

Information Integration among Heterogeneous and Autonomous Applications

A wide variety of distributed applications are nowadays emerging in diverse domains. These applications deploy various database systems for the management of their information, in which the diversity stems from the specific information management requirements and the objectives targeted by these applications.

Existing applications differ in their main characteristics and required features. On one hand, they differ in their distributed/centralized architecture, their size, complexity, and the type of data they handle. On the other hand, their requirements depend on the global functionalities that they need to provide and on the required level of interoperation with other sites. The used database management systems (DBMSs) are at best chosen to meet the specific characteristics and requirements of every application environment. However, currently available DBMSs lack the possibility to be efficiently used for all types of applications. Some DBMSs better suit smaller applications, while others are more dedicated to complex environments and focus on the management of, for example, multimedia information and large data sets. Thus, *any attempt in the direction of forcing different applications to use the same database system for the management of all their information services is unrealistic*. Even within the same environment, in certain complex applications, the use of more than one DBMS cannot be avoided.

Furthermore, from the application cases described in the thesis, it is clear that in today's organizations, new and existing applications require access to data stored in several pre-existing databases detained at several local and remote sites. Therefore, a main criterion required by most complex organizations, is the provision of collaboration possibilities and information integration mechanisms among distributed, heterogeneous, and autonomous systems. We also learned from the development of the application cases that providing interoperability and information integration among distributed systems, via the deployment of database standards and emerging Internet technologies, is one of the most challenging approaches in the area of integrating heterogeneous information from autonomous sites.

In Order to satisfy the new information management requirements of advanced and complex organizations, a strong information integration system must be designed and developed, serving the need for information integration and interoperation among these organizations.

In this context, the work described in this thesis focuses on the design and partial development of a Generic and Flexible Information Integration System (**GFI$_2$S**). The design of **GFI$_2$S** is based on the investigation, evaluation, and validation of the methodologies and approaches discussed in chapter 2; and motivated by the expertise gained within the devel-

opment of the various R&D projects addressed in chapters 3, 4, and 5 of the dissertation. *Flexibility* of **GFI$_2$S** resides in its ability to add/remove new system to/from the federation with involvement of minimum effort. Flexibility in **GFI$_2$S** is supported through the use of the specific two-component architecture, while, its *genericness* is achieved through the deployment of database standards, emerging Internet technologies, and middleware solutions.

The work described in chapter 6 illustrates the architectural components of **GFI$_2$S** that support the integration of different types of data from heterogeneous applications. The architecture of **GFI$_2$S** is composed of two main components of: (1) Local Adaptation Layer (LAL) that facilitates the access to the underlying databases in the node, and (2) Node Federation Layer (NFL) that provides links to the information and applications outside the node and supports the information sharing and interoperation. This two-component architecture of **GFI$_2$S** supports a wide variety of existing applications with efficient means for their interconnection and interoperation, while preserving their ***heterogeneity***, ***distribution***, and full ***autonomy***.

- ***Heterogeneity*** refers to the fact that each database may apply its own distinct DBMS, and data representation is heterogeneous in terms of structures and semantics at every site.
- ***Distribution*** refers to the storage and processing of information from distributed data sources, located on different host computers.
- ***Autonomy*** refers to the fact that each database within the federation community is an independent database system. Typically, a local database is pre-existing to the creation of a cooperation network and has its own administration policies, and users community.

The distinctive features of the **GFI$_2$S** integration approach resides in: (a) the ***specific combination*** of database standards and Internet middleware with the fundamental research approaches, and (b) the ***way in which they are deployed and inter-linked*** within the specific components of the **GFI$_2$S**. These two considerations make the **GFI$_2$S** approach distinct from all other existing federated/integrated approaches, and introduce **GFI$_2$S** as a *generic solution* providing a *flexible architecture*, and an *open facility* for integration/interoperation among heterogeneous, distributed, and autonomous sites.