Efendi: Federated Database System of Cadlab

E. Radeke, R. Böttger, B. Burkert, Y. Engel, G. Kachel, S. Kolmschlag, D. Nolte
Cadlab (Coop. Uni.Paderborn & Siemens Nixdorf), Bahnhofstr. 32, 33102 Paderborn, Germany
E-Mail: dbf-ma@cadlab.de, http://www.cadlab.de

Many enterprises suffer from the problems caused by heterogeneous data (base) systems. Redundancies across the mostly isolated databases often lead to data inconsistencies. Moreover, applications that need data of other/multiple databases have to use the different database interfaces which complicates code and raises maintainance costs.
In response to these problems, Efendi provides database federation services which couple multiple heterogeneous database and file systems to a federated database system (FDBS) [5]. We base on the ODMG standard [2] for the FDBS interfaces and also allow to migrate data among the different DBS while preserving global object identity.

1 System Overview

Efendi couples multiple autonomous heterogeneous database and file systems by a software layer on top of them [1]. Database specifics are encapsulated in so called DBS adapters, while other modules and interfaces are generic and usable for different Efendi installations. As canonical data model, i.e. common data format and access operations, we base on the new object-oriented standard of ODMG due to its expressive power and increasing importance. We have enhanced the ODMG interfaces (ODL, OML, OQL with C++ binding) to a multidatabase language.
Most important extension are facilities for *schema integration* in the object definition language ODL. The current ODMG ODL does not support multiple schemata and therefore has no capabilities to derive a schema from another. But schema derivation mechanisms are essential for FDBS to combine the various DBS schemata to one/some FDBS schemata, e.g. according to the five level schema architecture of [5]. Therefore we developed new language constructs for the required schema transformations. They allow to derive one schema from one or multiple other schemata whereby information can be filtered, added, and renamed. These constructs are analogous to those for class definition and inheritance of ODL which simplifies learning the extensions. The ODMG OML/OQL is extended in the semantics for providing transparent data access on multiple databases as if it was a single logical database. This enables the user to specify multidatabase queries in the current OQL and data manipulations in OML. Tasks like global transaction management, decomposition of global operations to operations of the various DBS, object caching, and redundancy control are processed in the federation kernel of Efendi. In addition, we extended the OML by some operations for the purpose of *object migration*. They enable the administrator to move, replicate, and copy data from arbitrary DBS to another one. Thereby existing DBS applications can access data of other database systems, e.g. of different departments, and do not have to be recoded to the FDBS global interface. Since the global object identity of migrated data is preserved, the data locality change is hidden for all global FDBS applications and does not enforce a code modification. In [4], we specified the functionality in more detail and [3] shows that object migration supports the stepwise reduction of the number of DBS in an enterprise.

2 Efendi Demonstration

The demonstration shows Efendi in a state of a pre-product, illustrating how an object-oriented and a relational DBS as well as the UNIX file system are coupled to an FDBS and how data can be manipulated and migrated uniformly.
Three main steps demonstrate how an enterprise can introduce and use Efendi:

1. What is necessary to install Efendi for the database systems (demo focusses on schema integration).
2. How easy multidatabase applications can be written with Efendi.
3. How easy data can be accessed, manipulated, and migrated by an example multidatabase application (demo shows concurrent work of DBS and FDBS applications).

3 Outlook

In order to validate Efendi in industrial environments, we initiated some pilot projects with industrial organizations. We plan to install Efendi in their environments. The future development of Efendi will be influenced on the one hand by requirements of our pilot project partners and on the other hand by enhancements such as performance tuning or more sophisticated transaction mechanisms.

References

[1] BÖTTGER, ENGEL, KACHEL, KOLMSCHLAG, NOLTE, RADEKE Enhancing the Data Openess of Frameworks by Database Federation Services. In *Conf. on Electronic Design Automation Frameworks (EDAF)*, Gramado, 1994.

[2] CATTELL *The object database standard: ODMG-93 (release 1.1)*. Morgan Kaufmann, 1994.

[3] RADEKE, SCHOLL Federation and stepwise reduction of database systems. In *Conf. on Applications of Databases (ADB)*, Vadstena, 1994.

[4] RADEKE, SCHOLL Functionality for object migration among distributed, heterogeneous, autonomous database systems. In *WS on Research Issues on Data Engineering - Distributed Object Management (RIDE-DOM)*, Taipei, 1995.

[5] SHETH, LARSON Federated database systems for managing distributed, heterogeneous, and autonomous databases. *ACM Computing Surveys 22(3)*, 1990.