Multimedia Federated Databases on Intranets: Web-Enabling IRO-DB

Georges Gardarin

Laboratoire PRiSM URA CNRS 1525
Université de Versailles-St Quentin
45 avenue des Etats-Unis
78035 Versailles Cedex - FRANCE
email : (1) <firstname>.<lastname>@prism.uvsq.fr

Abstract. Integrating semantically heterogeneous databases requires rich data models to homogenize disparate distributed entities with relationships and to access them through consistent views using high level query languages. In this paper, we first survey the IRO-DB system, which federates object and relational databases around the ODMG data model. Then, we point out some technical issues to extend IRO-DB to support multimedia databases on the Web. We propose to make it evolve towards a three-tiered architecture including local data sources with adapters to export objects, a mediator to integrate the various data sources, and an interactive user interface supported by a Web browser. We show through an example that new heuristics and strategies for distributed query processing and optimization have to be incorporated.

Key words. Interoperable database, Federated database, Object-oriented database, Remote data access, Schema integration, Multimedia, Web, Query processing.

1. Introduction

Object-oriented multidatabase systems (also referred to as federated databases or heterogeneous databases) represent the confluence of various trends in computer science and technology [1], among them object-orientation [2], distributed databases, and interoperability. Recently, the Internet has become the major vehicle in networking industry for information access and dissemination. The Web as a service on top of the Internet or Intranets focuses on transparent navigation and hypermedia document oriented information access. Thus, today there is a need to integrate the object-oriented multidatabase technology within multimedia Web-based systems, both for Intranet applications and Internet services. This paper discusses the integration of multimedia and Web techniques within the IRO-DB federated database system.

While much of the early work in federated databases concentrated on relational technology [3], some projects have been developed at the beginning of the 90's based on object models. Pegasus [4] was one of the first developed at HP Lab. It is centered around a global object model to which local objects (e.g., tables from a relational DB) are mapped. The global model is close to the object-relational one and the query language is SQL3+, an adapted version of SQL3. Started in 1993, the IRO-DB Esprit project [5] has developed a similar approach, but based on the ODMG "standard" [6].
IRO-DB supports an interactive schema integrator workbench to design integrated views of the federated database, and to automatically generate mappings to exported schemas. IRO-DB also focuses on relationship traversal and complex objects handling through collection support. This paper first gives an overview of the IRO-DB system architecture and describes some of the main components of the system.

Recently, a new generation of heterogeneous database systems has appeared in research. This generation typically focuses on a better support of multimedia objects and an integration with web-based technology. Some of the most well known projects are Garlic of IBM Almaden [7], Tsimmis of Standford University [8], Information Manifold of AT&T [9], and Disco from INRIA [10]. While some of these projects model data as labeled graphs with more or less typed nodes (Tsimmis), most still use object-relational (IM) or pure object models (Disco, Garlic). As with relational systems which have been extended to support object technology (e.g., Oracle 8 or Informix), we believe that operational federated object-oriented database systems can be extended to support multimedia objects and Web technology.

This paper first describes the IRO-DB object-oriented federated database system. Developed by a consortium of European partners, IRO-DB is now operational and an application was recently demonstrated. Thus, the difficulty is now to extend it to support multimedia objects, such as geographical data and images, and to make it Web-enabled. In the next section, we survey the IRO-DB architecture and main components. In the third section, we try to isolate the main issues to address for extending it to support Web technology with multimedia database servers. We propose a three-tiered architecture and discuss some query processing issues. In conclusion, we summarize the contributions of this paper and our future plans.

2. The IRO-DB Project

IRO-DB is an object-oriented federated database system. A version is currently operational. It interconnects a relational system INGRES, and three object-oriented DBMS : O2, Matisse and Ontos. In this section, we briefly describe the main system features and some key components.

2.1 Project Overview

IRO-DB (Interoperable Relational and Object-Oriented DataBases) is an ESPRIT project developed in Europe from 1993 to 1996. The novelty of the IRO-DB architecture is to use the ODMG'93 standard [6] as a common object model supporting the ODL definition language and the OQL query language to federate various object-oriented and relational data sources. The IRO-DB architecture is clearly divided into three layers, thus facilitating the cooperative development of the project in several research centers. The local layer adapts local data sources to the ODMG standard ; the communication layer efficiently transfers OQL requests and the resulting collections of objects; the interoperable layer provides schema integration tools, security management, transaction management, object management, as well as a global query processor.

Accordingly, IRO-DB follows an architecture with three layers of schemas, i.e., with local schemas, import/export schemas, and interoperable schemas (also referred to as