The Garlic Project

M. Tork Roth, M. Arya, L. Haas, M. Carey, W. Cody, R. Fagin, P. Schwarz, J. Thomas, and E.Wimmers
IBM Almaden Research Center
{torkroth, manish, laura, carey, cody, fagin, schwarz, wimmers}@almaden.ibm.com, jthomas@cruzio.com

Overview

The goal of the Garlic [1] project is to build a multimedia information system capable of integrating data that resides in different database systems as well as in a variety of non-database data servers. This integration must be enabled while maintaining the independence of the data servers, and without creating copies of their data. "Multimedia" should be interpreted broadly to mean not only images, video, and audio, but also text and application specific data types (e.g., CAD drawings, medical objects, ...). Since much of this data is naturally modeled by objects, Garlic provides an object-oriented schema to applications, interprets object queries, creates execution plans for sending pieces of queries to the appropriate data servers, and assembles query results for delivery back to the applications. A significant focus of the project is support for "intelligent" data servers, i.e., servers that provide media-specific indexing and query capabilities [2]. Database optimization technology is being extended to deal with heterogeneous collections of data servers so that efficient data access plans can be employed for multi-repository queries.

A prototype of the Garlic system has been operational since January 1995. Queries are expressed in an SQL-like query language that has been extended to include object-oriented features such as reference-valued attributes and nested sets. In addition to a C++ API, Garlic supports a novel query/browser interface called PESTO [3]. This component of Garlic provides end users of the system with a friendly, graphical interface that supports interactive browsing, navigation, and querying of the contents of Garlic databases. Unlike existing interfaces to databases, PESTO allows users to move back and forth seamlessly between querying and browsing activities, using queries to identify interesting subsets of the database, browsing the subset, querying the content of a set-valued attribute of a particularly interesting object in the subset, and so on.

The Demo

In our demo, Garlic integrates data from three repositories: a DB2 Client/Server database, an ObjectStore object-oriented database from Object Design, Inc., and a QBIC repository [4] that provides query by image content. Queries may involve one or more repositories. When the query spans multiple repositories, Garlic manages the processing of the query across the different repositories and fills in for those repositories that have limited query processing power. For example, a multimedia query may involve images as well as relational parametric data that describe the images. Garlic will dispatch the relational predicates to the DB2 repository, the image predicates to the QBIC repository, and integrate their responses to form the end result of the query.

PESTO is used to browse and query the Garlic database. Objects (including collections) are represented by interconnected data windows, with the attributes for an object contained in a window. An important feature of the query/browser is its synchronous browse capability — as a user scrolls through a top level window, the child windows are automatically updated. At any point in the browse session, the user may ask a query at any level of nesting. PESTO's query model is known as *query-in-place* — query predicates are entered in the same fields as those that were brought up for the browse session, and upon execution of the query, the user is returned to the browse state to view the results of the query.

In summary, our demo exploits the unique features of the PESTO query/browser to explore and query the unified object-oriented view of heterogeneous data presented by Garlic. We show that the Garlic system is capable of coordinating a broad range of object-oriented, relational and multimedia queries across several repositories with very different data models and levels of sophistication in query processing. We also show PESTO's seamless integration of browsing and querying over object-oriented data, which makes it a very natural interface for the Garlic system.

References

[1] M. Carey, L. Haas, et al., "Towards Heterogeneous Multimedia Information Systems: The Garlic Approach", RIDE-DOM '95, Taiwan 3/95. (5th Int'l Workshop on Research Issues in Data Engineering: Distributed Object Management.)

[2] W. Cody, L. Haas, et al., "Querying Multimedia Data from Multiple Repositories by Content: the Garlic Project", VDS '95 (IFIP 2.6 Working Conference on Visual Database Systems -3), March 1995.

[3] M. Carey, L. Haas, et. al, "PESTO: An Integrated Query/Browser for Object Databases", submitted for publication.

[4] W. Niblack, R. Barber, et. al, "The QBIC Project: Querying Images by Content Using Color, Texture and Shape", Proc. SPIE, San Jose, CA, February 1993.

Permission to make digital/hard copy of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage, the copyright notice, the title of the publication and its date appear, and notice is given that copying is by permission of ACM, Inc. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee.

SIGMOD '96 6/96 Montreal, Canada
© 1996 ACM 0-89791-794-4/96/0006...$3.50