Levels of Distribution Transparency in Distributed Database Systems.

Nguyen Manh Tho
Nguyen Dang Phu

March 1998
Abstract

Due to the large number of existing database systems in organizations, the advances in network technology and the increment of requirements in sharing data in distributed environments, it is necessary to evolve into a distributed database system in which each database system is generally created and administered independently, but the whole system can satisfy every user’s query requests in the same manner as in the local database system.

The main objective of this thesis is to research on distributed database system theoretically. We therefore consider so many related problems in distributed database system such as fundamental concepts in distributed database, levels of distribution transparency, distributed query optimization, management of distributed transactions, concurrency control, reliability, backup and recovery, distributed database administration. Although so many problems were considered theoretically, only one interesting problem will be solved in this thesis: levels of distribution transparency.

The thesis presents an approach for transforming from the global query into queries at other levels of transparency. There are four levels in distributed database system architecture: global schema, fragmentation schema, allocation schema, and local mapping schema. There will be three equivalent levels of distribution transparency: fragmentation transparency, location transparency, and local mapping transparency. The query validation is implemented at the simple case, but the distributed query optimization is out of the scope of this thesis.

We provide a 3-tier architecture in which includes client layer, application layer and database layer. The architecture supports dynamic evolution of the system in terms of adding or removing databases. It also support network error detection. The database layer can be any database system that supports ODBC. The client layer provides the graphic user interface that support user interaction. The application layer communicates to other layer through socket protocol. The socket’s structure and the communication procedure are also included in our solution.
A distributed algorithm to perform the query transformation process is also provided. In fact, there are 3 algorithms are used. The first one is used to convert the global query to queries at fragmentation schema, the second one is used to convert the fragmentation schema’s queries to allocation schema’s queries, and the last one converts the allocation schema’s queries to the query at local mapping schema. The approach supports complete execution of a query. A graphical interface is used to assist in the formulation of queries, sending a request query, receiving and displaying the results.

A prototype has been developed to demonstrate and evaluate the architecture, the concurrent execution of the distributed queries, and the dynamic evolution of the system. This prototype is built on Intranet with the sample data extracted from the real data of our University Academic Management System. We use a case study to illustrate the feasibility and applicability of our approach. The case study is based on an Academic Management System in which autonomous database systems come from different departments share and exchange data.