Chapter VIII

Database Management Issues in the Web Environment

J.F. Aldana Montes
Universidad de Málaga, Spain

A.C. Gómez Lora
Universidad de Málaga, Spain

N. Moreno Vergara
Universidad de Málaga, Spain

M.M. Roldán García
Universidad de Málaga, Spain

ABSTRACT

The focus of this chapter is on the progressive adaptation of database techniques to Web usage in a way quite similar to the evolution from integrated file management systems to database management systems. We review related and open issues, such as the semi-structured data and XML, integrity problem, query optimization problem, and integration issues in both the Web and Semantic Web environments. The representation of meta-information along with data opens up a new way to automatically process Web information due to the use of explicit semantic information. We hope that researchers will take into account traditional database techniques and how they can assist new Web technologies. In this sense, the amalgamation of Web and database technology appears to be very promising.

INTRODUCTION

The focus of this chapter is on introducing a new paradigm: the Web as the database, i.e., the progressive adaptation of database techniques to Web usage. We consider that this will be done in a way quite similar to the evolution from integrated file management systems to database management systems. In any case, this is a much more difficult goal and quite a lot of work is still to be done.

The database community has been seriously disturbed with the Web technologies expansion. Particularly, two reports have produced a special commotion in database field. The first one, the Asilomar report (Bernstein et al., 1998), postulates the new directives in databases tendencies, previewing the Web impact in this field. The second one, Breaking out of the Box (Silberschatz & Zdonik, 1996), proposes how the database community must transfer its technology to be introduced into Web technology. In this sense, we have broken out of the database box into its autonomous functional components, and we are using these to reach a solution for the problem of heterogeneous data sources integration.

It is within this context that we are going to study the different issues of data management in the Web. This chapter reviews related and open issues, such as the semi-structured data and XML, integrity problem, query optimization problem, and integration issues in both the Web and Semantic Web environments. Finally, we will briefly discuss how traditional database technologies can be used to solve these open points. The special features of the Web environment make techniques for querying or maintaining the Web different from those for traditional databases.

Thinking about the Web as a huge, highly distributed database, we may consider different dimensions to conceptually describe it. Tamer Özsu and Valduriez (1999) define a classification of database systems with respect to: (1) their distribution, (2) the autonomy of local systems and (3) the heterogeneity of database systems. The autonomy concept is considered as the distribution of control, not of data. This indicates the degree to which individual DBMSs can operate independently. Whereas autonomy refers to the distribution of control, the distribution dimension deals with the physical distribution of data over multiple sites. With respect to heterogeneity, this can range from hardware heterogeneity, differences in networking protocols, variations in DBMSs, etc., to the data model or the policy for managing integrity on the database.

Figure 1. Extending the Cube
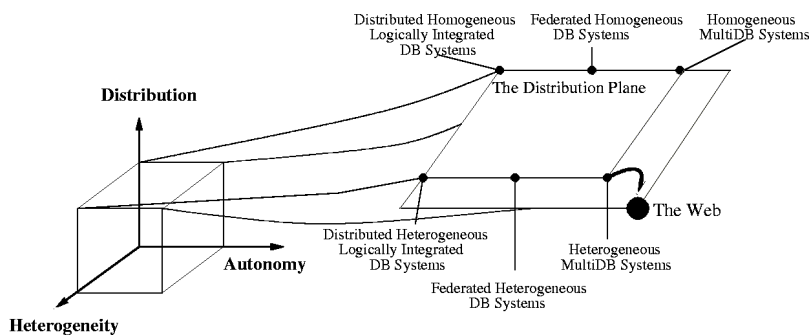
Related Content

A Scalable Algorithm for One-to-One, Onto, and Partial Schema Matching with Uninterpreted Column Names and Column Values
Boris Rabinovichand Mark Last (2014). *Journal of Database Management (pp. 1-16).*
www.irma-international.org/article/a-scalable-algorithm-for-one-to-one-onto-and-partial-schema-matching-with-uninterpreted-column-names-and-column-values/138623

A Case Study of the Use of the Viable System Model in the Organization of Software Development
Peter Kawalekand David Wastell (2002). *Advanced Topics in Database Research, Volume 1 (pp. 120-134).*
www.irma-international.org/chapter/case-study-use-viable-system/4325

Refinement Equivalence in Model-Based Reuse: Overcoming Differences in Abstraction Level
Pnina Soffer (2005). *Journal of Database Management (pp. 21-39).*
www.irma-international.org/article/refinement-equivalence-model-based-reuse/3335

OLAP with a Database Cluster
Uwe Röhm (2009). *Database Technologies: Concepts, Methodologies, Tools, and Applications (pp. 829-846).*
www.irma-international.org/chapter/olap-database-cluster/7944

Concurrency Control for Replicated Data in Distributed Real-Time Systems
Sang H. Son, Fengjie Zhangand Buhyun Hwang (1996). *Journal of Database Management (pp. 12-23).*
www.irma-international.org/article/concurrency-control-replicated-data-distributed/51162