The VLDB Journal

The International Journal on Very Large Data Bases

Volume 2(3) (1993)

The *VLDB* Journal

The International Journal on Very Large Data Bases

Volume 2(3) (1993)

THE BOXWOOD PRESS
PACIFIC GROVE, CA, USA

Submission of a paper to *The VLDB Journal* is understood to imply that it is not being considered for publication elsewhere and that the author's permission to publish his/her article(s) in this journal implies the exclusive authorization of the publisher to deal with all issues concerning the copyright therein.

Submission of multi-authored manuscripts to this journal implies the consent of *each* of the authors. The publisher will assume that the senior or corresponding author has specifically obtained the approval of all other co-authors to submit the manuscript to this journal.

No responsibility is assumed by the Publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein.

Special regulations for authors. Upon acceptance of an article by the journal, the author(s) will be asked to transfer copyright of the article to the publisher. This transfer will ensure the widest possible dissemination of information.

Published quarterly by:

THE BOXWOOD PRESS
183 Ocean View Blvd.
Pacific Grove, CA 93950, USA
Telephone: (408) 375-9110/Fax: (408) 375-0430

PREFACE BY THE EDITORS-IN-CHIEF

It gives us great pleasure to introduce the first special issue of the *VLDB Journal* to be devoted to Parallel and Distributed Information Systems. While this special issue was handled "regularly" by the invited editors, namely Mike Carey and Patrick Valduriez, it is special for two reasons. First, in view of the explosion of interest in this field during the last two years and following the recommendation of the invited editors, the call for articles was not announced separately as usual. Because of the obvious overlap in subject area, the call for articles was linked to the call for papers for the Second International Conference on Parallel and Distributed Information Systems (PDIS). Second, this issue does not contain a survey by the invited editors because a number of such surveys have been published recently. A short introduction to the four contributions of this issue provides references to the most important surveys.

These four articles present the leading edge research results of four different and important aspects of PDIS. We hope that the growing readership of the *VLDB Journal* will find this an interesting issue and we would like to thank the invited editors for their great effort and patience during its production.

Hans -J. Schek
Fred Maryanski
May, 1993

PREFACE TO SPECIAL ISSUE ON PDIS

Commercial database technology has moved from the earlier hierarchical and network models to the relational model. The main advantages of relational database systems over their predecessors are data independence and high-level query languages (e.g., SQL). These advantages increase programmer productivity and favor automatic optimization. Today, after a decade of optimization and tuning, they can provide a performance level reaching that of earlier systems. Therefore, they are now being used extensively in commercial data processing for decision-support or on-line transaction processing applications.

Distributed and parallel database technologies are among the most important computing developments of the past decade. The set-oriented nature of the relational model, which allows simple relation partitioning, initially encouraged such developments. A distributed database system permits the management of a collection of multiple, logically interrelated databases distributed over a computer network with advantages like client-server connections and distribution transparency. A parallel database system strives to exploit multiprocessor computer architectures in order to build high-performance and highly available database servers at a much lower price than equivalent mainframe computers.

Both distributed and parallel database systems have to deal with data partitioning, either across sites (or a computer network) or across processing nodes (or a multiprocessor), and they therefore share common problems such as data placement, query processing, and transaction management. However, the solutions have been quite different traditionally because of different assumptions regarding their respective computing environments (e.g., network speed) and the need to scale up to much larger numbers of nodes in parallel database systems. The recent advances in high-speed network technology (e.g., FDDI, ATM) may weaken such differences in the future. However, distributed database systems also face specific problems such as geographical distribution, local autonomy, and heterogeneity.

The initial promises of distributed and parallel database systems—transparent management of distributed and replicated data, improved data availability, improved performance through parallelism, and easier system expansion—are met to various degrees by today's commercial relational products. However, a number of important open problems touch on issues ranging from those of parallel data processing to distributed multidatabase management. Furthermore, it is still an open issue to decide which of the possible architectures (e.g., shared-disk and shared-nothing) is best for parallel database management under various conditions. These open issues are discussed in recent survey articles on distributed databases (Oszu and Valduriez, 1991) and parallel database systems (DeWitt and Gray, 1992; Valduriez, 1993). Finally, there are also new issues raised by the introduction of higher functionality, such as knowledge-based or object-oriented capabilities, to produce next-generation database systems (Stonebraker et al., 1990).

This special issue includes four articles, each of which represents a significant extension of research presented at the Second International Conference on Parallel and Distributed Information Systems (PDIS), held January 20–22, 1993 in San Diego, California. The PDIS conference is an international forum for the dissemination and discussion of ideas related to distributed or parallel systems for managing large volumes of information; its

scope includes the design, development, and utilization of parallel or distributed systems for the management of databases, knowledge bases, files, and textual information.

This special issue begins with three articles focusing on aspects of query processing in parallel database systems. The first, "Query Processing and Inverted Indices in Shared-Nothing Text Document Information Retrieval Systems" by A. Tomasic and H. Garcia-Molina, compares the performance impact on query processing of various physical organizations for inverted lists through a new probabilistic model and through simulation experiments. It shows that the choice of an index organization depends mainly on the access time of the storage devices and on the network bandwidth.

The next article, "Parallel Query Processing with Zig-Zag Trees" by M. Ziane, M. Zait, and P. Borla-Salamet, presents an approach to compile-time optimization and parallelization of queries for execution in either shared-memory or shared-nothing parallel database systems. The authors propose the use of zig-zag trees and show that they provide better response time than right-deep trees in cases of limited memory. The reported performance measurements in the article are from the DBS3 parallel database system prototype.

The third article, "Considering Data Skew Factor in Multi-Way Join Query Optimization for Parallel Execution," is by K.A. Hua, Y.-L. Lo, and H.C. Young. In this article a dynamic load balancing strategy is proposed for multi-way join processing in shared-nothing parallel database systems. The proposed strategy is designed to address load imbalances that can arise due to data skew in such systems. Performance results from a simulation and analysis of the proposed join ordering strategy show very good improvements over conventional join processing techniques.

The fourth and final article in this special issue deals with transaction management in distributed multidatabase systems. Entitled "A Theory of Global Concurrency Control in Multidatabase Systems," the article is by A. Zhang and A.K. Elmagarmid, and presents a theoretical basis for global transaction serializability in multidatabase systems. The new types of serializability are proposed, and a new criterion, termed hybrid serializability, is derived by combining the two and is proven to be optimal.

In closing, we wish to express our appreciation to the authors, who had to revise their articles under tight time constraints, and to the reviewers, who provided invaluable help. We are also grateful to Prof. H.-J. Schek for providing us with the opportunity to produce this special issue and for his help throughout the review process.

Michael J. Carey
Patrick Valduriez

DeWitt, D.J. and Gray, J. Parallel database systems: The future of high performance database systems. *Communications of the ACM,* 35:6, 1992.

Ozsu, T. and Valduriez, P. Distributed databases: Where are we now? *IEEE Computer,* 24:8, 1991.

Stonebraker, M., Rowe, L.A., Lindsay, R., Gray, J., Carey, M., Brodie, M., Bernstein, P., and Beech, D. Third-generation database system manifesto. *ACM SIGMOD Record,* 19:3, 1990.

Valduriez, P. Parallel database systems: Open problems and new issues. *Distributed and Parallel Databases* (an international journal), 1:2, 1993.