Lecture Notes in Computer Science 2169
Edited by G. Goos, J. Hartmanis, and J. van Leeuwen

Springer
Berlin
Heidelberg
New York
Barcelona
Hong Kong
London
Milan
Paris
Tokyo

Michael Jaedicke

New Concepts for Parallel Object-Relational Query Processing

Springer

Series Editors

Gerhard Goos, Karlsruhe University, Germany
Juris Hartmanis, Cornell University, NY, USA
Jan van Leeuwen, Utrecht University, The Netherlands

Author

Michael Jaedicke
Donnersbergerstr. 34
80634 München
Germany
E-mail: mjaedicke@yahoo.com

Cataloging-in-Publication Data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Jaedicke, Michael: New concepts for parallel object relational query processing / Michael Jaedicke. - Berlin ; Heidelberg ; New York ; Barcelona ; Hong Kong ; London ; Milan ; Paris ; Tokyo : Springer, 2001
(Lecture notes in computer science ; Vol. 2169)
Zugl.: Stuttgart, Univ., Diss., 1999
ISBN 3-540-42781-3

CR Subject Classification (1998): E.2, H.2

ISSN 0302-9743
ISBN 3-540-42781-3 Springer-Verlag Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable for prosecution under the German Copyright Law.

Springer-Verlag Berlin Heidelberg New York
a member of BertelsmannSpringer Science+Business Media GmbH

http://www.springer.de

© Springer-Verlag Berlin Heidelberg 2001
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Christian Grosche, Hamburg
Printed on acid-free paper SPIN: 10840389 06/3142 5 4 3 2 1 0

Preface

When I started this research, no commercial database system supported both object-relational features and parallelism. In the meantime this situation has changed dramatically. All major vendors now offer a parallel, object-relational DBMS as their high-end product. However, there is still a lot to do. Firstly, object-relational (or extensible) DBMS have yet to mature fully. Secondly, the integration of parallelism and extensibility has not yet been completed. This work is my attempt to make a contribution to both issues.

Some of the concepts and techniques developed have been implemented in a prototypical parallel database system called MIDAS. This system is the result of a team effort to which many people have contributed. My own contributions to the system are the user-defined functions and user-defined table operators, the extension of the system tables for those user-defined objects and for detailed statistics, and the new query optimizer, for which I worked with Clara Nippl. She contributed especially to the cost model, the physical operators, and implementation rules. Furthermore, I provided support for the development of concepts for the query parallelizer, the query execution control system, and the send/receive operators of the execution system.

Acknowledgments

I am very thankful to Professor Dr.-Ing. habil. Bernhard Mitschang, who gave me the opportunity to carry out this research. His continuous support, encouragement, and useful comments during many fruitful discussions and during the preparation of joint publications are especially acknowledged.

I am very grateful to Prof. Dr.-Ing. Theo Härder who introduced me thoroughly to database systems and taught me a lot about engineering. I am also grateful to him for the analysis of my results and his valuable remarks.

I also acknowledge the help of my colleagues Giannis Bozas, Clara Nippl, Angelika Reiser, and Stephan Zimmermann with whom I worked together on the MIDAS project. Stephan provided software engineering support for the whole group and was always helpful. His work centered on PVM communication, the process architecture, buffer and transaction management, the parallelization of the execution system, the scheduler and benchmarking, and performance. Clara worked on the parallelization of the execution system, the scheduler, the parallelizer, and the optimizer. Giannis focussed on the lock manager and the buffer management. I also thank Prof. Rudolf Bayer, who led the project together with Prof. Mitschang, for his support.

I also enjoyed giving practical courses on RDBMS together with Angelika. I thank all my colleagues in Munich and Stuttgart, especially Aiko Frank, Volker Markl, Jürgen Sellentin, and Martin Zirkel for their help and comradeship during the last years. I am grateful to my colleague Ralf Rantzau and to Prof. Härder for improving my English.

Special thanks also go to the students that I have supervised: Pascal Frantz, Stefan Haas, Karl Hahn, Sebastian Heupel, Bernd Holzhey, Kay Krueger-Barvels, Sabine Perathoner, Ralf Schumacher, Robert Seehafer, and Susanne Stamp. It was a pleasure to work with them. Many of them made significant contributions to this work. I also thank the many other students who worked on the MIDAS project.

I also gratefully acknowledge the valuable comments of the anonymous referees of diverse papers, which have also improved this work. The feedback from C. Mohan, G. Lohman, and M. Carey on the issue of interfaces for query execution plans and the feedback from Harald Schöning on the parallel execution of user-defined functions is much appreciated.

Finally, I would like to thank the Deutsche Forschungsgemeinschaft (DFG) for funding the MIDAS project (SFB 342, B2). I also acknowledge the support of the Studienstiftung des Deutschen Volkes during my university course.

Special thanks go to Springer-Verlag, especially to Alfred Hofmann, and the series editors, for publishing my work in LNCS.

Last but not least I am grateful to my parents for their affection and continuous care.

July 2001 Michael Jaedicke

Abstract

During the last few years parallel object-relational database management systems have emerged as the leading data management technology on the market place. These systems are extensible by user-defined data types and user-defined functionality for the data. This work focuses on the efficient parallel execution of user-defined functionality. The main contributions describe techniques to support data parallelism for user-defined scalar and aggregate functions, to support intra-function parallelism for the execution of a scalar function on a large object, and a new technology to provide extensibility with regard to new set-oriented database operations that can efficiently implement user-defined functionality in parallel object-relational database management systems. Some of these techniques have been implemented in the MIDAS prototype or on top of a commercial object-relational database management system.

Table of Contents

CHAPTER 5 *The Multi-operator Method*

CHAPTER 6 *User-Defined Table Operators*

CHAPTER 7 *Implementation of UDTO*