

Proceedings

17th International Conference on Data Engineering

2-6 April 2001

Heidelberg, Germany

Sponsored by

IEEE Computer Society Technical Committee on Data Engineering (TCDE)

ABB



invent

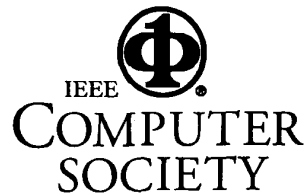


Microsoft

SOFTWARE AG
THE XML COMPANY



software | design & management



Los Alamitos, California

Washington □ Brussels □ Tokyo

Copyright © 2001 by The Institute of Electrical and Electronics Engineers, Inc.
All rights reserved

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries may photocopy beyond the limits of US copyright law, for private use of patrons, those articles in this volume that carry a code at the bottom of the first page, provided that the per-copy fee indicated in the code is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

Other copying, reprint, or republication requests should be addressed to: IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, P.O. Box 133, Piscataway, NJ 08855-1331.

The papers in this book comprise the proceedings of the meeting mentioned on the cover and title page. They reflect the authors' opinions and, in the interests of timely dissemination, are published as presented and without change. Their inclusion in this publication does not necessarily constitute endorsement by the editors, the IEEE Computer Society, or the Institute of Electrical and Electronics Engineers, Inc.

IEEE Computer Society Order Number PR01001
ISBN 0-7695-1001-9
ISBN 0-7695-1002-7 (case)
ISBN 0-7695-1003-5 (microfiche)
ISSN 1063-6382

Additional copies may be ordered from:

IEEE Computer Society
Customer Service Center
10662 Los Vaqueros Circle
P.O. Box 3014
Los Alamitos, CA 90720-1314
Tel: + 1 714 821 8380
Fax: + 1 714 821 4641
<http://computer.org/>
csbooks@computer.org

IEEE Service Center
445 Hoes Lane
P.O. Box 1331
Piscataway, NJ 08855-1331
Tel: + 1 732 981 0060
Fax: + 1 732 981 9667
[http://shop.ieee.org/store/
customer-service@iecc.org](http://shop.ieee.org/store/customer-service@iecc.org)

IEEE Computer Society
Asia/Pacific Office
Watanabe Bldg., 1-4-2
Minami-Aoyama
Minato-ku, Tokyo 107-0062
JAPAN
Tel: + 81 3 3408 3118
Fax: + 81 3 3408 3553
tokyo.ofc@computer.org

Editorial production by Danielle C. Young

Cover art production by Joe Daigle/Studio Productions

Printed in the United States of America by The Printing House

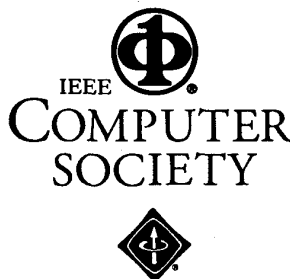




Table of Contents

17th International Conference on Data Engineering (ICDE 2001)

Message from the Program Co-Chairs	xii
Conference Officers	xiii
Program Committee	xiv
Tutorials Program	xvii
Demos Program	xx
External Referees	xxi

Session 1: Middleware

Exactly-Once Semantics in a Replicated Messaging System	3
<i>Y. Huang and H. Garcia-Molina</i>	
CORBA Notification Service: Design Challenges and Scalable Solutions	13
<i>R. Gruber, B. Krishnamurthy, and E. Panagos</i>	
Measuring and Optimizing a System for Persistent Database Sessions	21
<i>R. Barga and D. Lomet</i>	

Session 2: Temporal Databases

A Temporal Algebra for an ER-Based Temporal Data Model	33
<i>J. Lee and R. Elmasri</i>	
A Split Operator for Now-Relative Bitemporal Databases	41
<i>M. Agesen, M. Böhlen, L. Poulsen, and K. Torp</i>	
Incremental Computation and Maintenance of Temporal Aggregates	51
<i>J. Yang and J. Widom</i>	

Industry Session 1: E-Commerce

The Importance of Extensible Database Systems for e-Commerce	63
<i>S. DeFazio, R. Krishnan, J. Srinivasan, and S. Zeldin</i>	
E-Business Applications for Supply Chain Management: Challenges and Solutions	71
<i>F. Casati, U. Dayal, and M.-C. Shan</i>	
Delivering E-Business Faster with Database Technology	
<i>P. Shum</i>	

Session 3: Integration and Query Processing in Distributed Environments

Model-Based Mediation with Domain Maps.....	81
<i>B. Ludäscher, A. Gupta, and M. Martone</i>	
Processing Queries with Expensive Functions and Large Objects in Distributed Mediator Systems	91
<i>L. Bouganim, F. Fabret, F. Porto, and P. Valduriez</i>	
Tuning an SQL-Based PDM System in a Worldwide Client/Server Environment	99
<i>E. Müller, P. Dadam, J. Enderle, and M. Feltes</i>	

Session 4: Issues in Information Processing

Bundles in Captivity: An Application of Superimposed Information	111
<i>L. Delcambre, D. Maier, S. Bowers, M. Weaver, L. Deng, P. Gorman, J. Ash, M. Lavelle, and J. Lyman</i>	
High-Level Parallelisation in a Database Cluster: A Feasibility Study Using Document Services.....	121
<i>T. Grabs, K. Böhm, and H.-J. Schek</i>	
Efficient Sequenced Temporal Integrity Checking.....	131
<i>W. Li, R. Snodgrass, S. Deng, V. Gattu, and A. Kasthurirangan</i>	

Industry Session 2: XML

XML Data and Object Databases: The Perfect Couple?	143
<i>A. Renner</i>	
Tamino — A DBMS Designed for XML	149
<i>H. Schöning</i>	
The Nimble XML Data Integration System	155
<i>D. Draper, A. HaLevy, and D. Weld</i>	

Session 5: Database Engines

High-Performance, Space-Efficient, Automated Object Locking	163
<i>L. Daynès and G. Czajkowski</i>	
Differential Logging: A Commutative and Associative Logging Scheme for Highly Parallel Main Memory Database	173
<i>J. Lee, K. Kim, and S. Cha</i>	
Efficient Bulk Deletes in Relational Databases	183
<i>A. Gärtner, A. Kemper, D. Kossmann, and B. Zeller</i>	

Session 6: Data Mining from Patterns

On Dual Mining: From Patterns to Circumstances, and Back	195
<i>G. Grahne, L. Lakshmanan, X. Wang, and M. Xie</i>	
Mining Partially Periodic Event Patterns with Unknown Periods.....	205
<i>S. Ma and J. Hellerstein</i>	
PrefixSpan: Mining Sequential Patterns Efficiently by Prefix-Projected Pattern Growth	215
<i>J. Pei, J. Han, B. Mortazavi-Asl, H. Pinto, Q. Chen, U. Dayal, and M.-C. Hsu</i>	

Industry Session 3: Mobility and Pervasive Computing

Mobile Data Management: Challenges of Wireless and Offline Data Access	227
<i>E. Giguère</i>	
Microsoft Server Technology for Mobile and Wireless Applications	229
<i>P. Seshadri</i>	
IBM DB2 Everyplace: A Small Footprint Relational Database System	230
<i>J. Karlsson, A. Lal, C. Leung, and T. Pham</i>	

Session 7: Workflow Management

Dependable Computing in Virtual Laboratories	235
<i>G. Alonso, W. Bausch, C. Pautasso, A. Kahn, and M. Hallett</i>	
Workflow and Process Synchronization with Interaction Expressions and Graphs	243
<i>C. Heinlein</i>	
Inter-Enterprise Collaborative Business Process Management	253
<i>Q. Chen and M. Hsu</i>	

Session 8: Time Series

Duality-Based Subsequence Matching in Time-Series Databases	263
<i>Y.-S. Moon, K.-Y. Whang, and W.-K. Loh</i>	
Variable Length Queries for Time Series Data	273
<i>T. Kahveci and A. Singh</i>	
TAR: Temporal Association Rules on Evolving Numerical Attributes.....	283
<i>W. Wang, J. Yang, and R. Muntz</i>	

Industry Session 4: E-Commerce/Internet Applications

Database Performance for Next Generation Telecommunications	295
<i>M. Cochinwala</i>	
SpinCircuit: A Collaborative Portal Powered by E-Speak	661
<i>R. Pathak</i>	
Data Management Support of Web Applications.....	299
<i>D. Fishman</i>	

Session 9: XML

An Automated Change-Detection Algorithm for HTML Documents Based on Semantic Hierarchies.....	303
<i>S.-J. Lim and Y.-K. Ng</i>	
An XML Indexing Structure with Relative Region Coordinate.....	313
<i>D. Kha, M. Yoshikawa, and S. Uemura</i>	
Querying XML Documents Made Easy: Nearest Concept Queries	321
<i>A. Schmidt, M. Kersten, and M. Windhouwer</i>	
A Graph-Based Approach for Extracting Terminological Properties of Elements of XML Documents.....	330
<i>L. Palopoli, G. Terracina, and D. Ursino</i>	

Session 10: Indexing and Clustering

B ⁺ -Tree Indexes with Hybrid Row Identifiers in Oracle8i	341
<i>E. Chong, S. Das, C. Freiwald, J. Srinivasan, A. Yalamanchi, M. Jagannath, A.-T. Tran, and R. Krishnan</i>	
B-Tree Indexes and CPU Caches	349
<i>G. Graefe and P.-Å. Larson</i>	
Spatial Clustering in the Presence of Obstacles	359
<i>A. Tung, J. Hou, and J. Han</i>	
Selectivity Estimation for Spatial Joins	368
<i>N. An, Z.-Y. Yang, and A. Sivasubramaniam</i>	

Industry Session 5: Data Warehousing and Data Mining

Integrating Data Mining with SQL Databases: OLE DB for Data Mining	379
<i>A. Netz, S. Chaudhuri, U. Fayyad, and J. Bernhardt</i>	

SAP Business Information Warehouse — From Data Warehousing to an E-Business Platform	388
<i>T. Zurek and K. Kreplin</i>	
fAST Refresh Using Mass Query Optimization	391
<i>W. Lehner, B. Cochrane, H. Pirahesh, and M. Zaharioudakis</i>	
Session 11: Similarity Processing and Special Operators	
Quality-Aware and Load-Sensitive Planning of Image Similarity Queries	401
<i>K. Böhm, M. Mlivonic, and R. Weber</i>	
A Cost Model and Index Architecture for the Similarity Join	411
<i>C. Böhm and H.-P. Kriegel</i>	
The Skyline Operator	421
<i>S. Börzsönyi, D. Kossmann, and K. Stocker</i>	
Session 12: Data Mining/Frequent Item Sets	
Mining Frequent Itemsets with Convertible Constraints.....	433
<i>J. Pei, J. Han, and L. Lakshmanan</i>	
MAFIA: A Maximal Frequent Itemset Algorithm for Transactional Databases	443
<i>D. Burdick, M. Calimlim, and J. Gehrke</i>	
An Efficient Approximation Scheme for Data Mining Tasks.....	453
<i>G. Kollios, D. Gunopoulos, N. Koudas, and S. Berchtold</i>	
Industry Session 6: Applications	
Bringing the Internet to Your Database: Using SQL Server 2000 and XML to Build Loosely-Coupled Systems.....	465
<i>M. Rys</i>	
Infrastructure for Web-Based Application Integration	473
<i>D. Gawlick</i>	
Developing Web Services	477
<i>A. Bosworth</i>	
Session 13: Nearest Neighbor Queries	
An Index Structure for Efficient Reverse Nearest Neighbor Queries.....	485
<i>C. Yang and K.-I. Lin</i>	

Distinctiveness-Sensitive Nearest-Neighbor Search for Efficient Similarity Retrieval of Multimedia Information	493
<i>N. Katayama and S. Satoh</i>	
Approximate Nearest Neighbor Searching in Multimedia Databases	503
<i>H. Ferhatosmanoglu, E. Tuncel, D. Agrawal, and A. El Abbadi</i>	
Session 14: OLAP	
Rewriting OLAP Queries Using Materialized Views and Dimension Hierarchies in Data Warehouses	515
<i>C.-S. Park, M. Kim, and Y.-J. Lee</i>	
The MD-Join: An Operator for Complex OLAP	524
<i>D. Chatziantoniou, M. Akinde, T. Johnson, and S. Kim</i>	
Overcoming Limitations of Sampling for Aggregation Queries	534
<i>S. Chaudhuri, G. Das, M. Datar, R. Motwani, and V. Narasayya</i>	
Industry Session 7: DBMS Industrial Issues	
Pseudo Column Level Locking	545
<i>N. Ponnekanti</i>	
Discovery and Application of Check Constraints in DB2	551
<i>J. Gryz, B. Schiefer, J. Zheng, and C. Zuzarte</i>	
Database Managed External File Update	557
<i>N. Mittal and H.-I. Hsiao</i>	
Session 15: Query Processing	
Block Oriented Processing of Relational Database Operations in Modern Computer Architectures	567
<i>S. Padmanabhan, T. Malkemus, R. Agarwal, and A. Jhingran</i>	
Integrating Semi-Join-Reducers into State-of-the-Art Query Processors	575
<i>K. Stocker, D. Kossmann, R. Braumandl, and A. Kemper</i>	
Using EELs: A Practical Approach to Outerjoin and Antijoin Reordering	585
<i>J. Rao, B. Lindsay, G. Lohman, H. Pirahesh, and D. Simmen</i>	
Counting Twig Matches in a Tree	595
<i>Z. Chen, H. Jagadish, F. Korn, N. Koudas, S. Muthukrishnan, R. Ng, and D. Srivastava</i>	

Session 16: Similarity Searches

An Index-Based Approach for Similarity Search Supporting Time Warping in Large Sequence Databases.....	607
<i>S.-W. Kim, S. Park, and W. Chu</i>	
High Dimensional Similarity Search with Space Filling Curves	615
<i>S. Liao, M. Lopez, and S. Leutenegger</i>	
Similarity Search without Tears: The OMNI-Family of All-Purpose Access Methods	623
<i>R. Filho, A. Traina, C. Traina, Jr., and C. Faloutsos</i>	

Session 17: Caching

Cache-On-Demand: Recycling with Certainty.....	633
<i>K.-L. Tan, S.-T. Goh, and B. Ooi</i>	
Cache-Aware Query Routing in a Cluster of Databases	641
<i>U. Röhm, K. Böhm, and H.-J. Schek</i>	
Prefetching Based on the Type-Level Access Pattern in Object-Relational DBMSs.....	651
<i>W.-S. Han, K.-Y. Whang, Y.-S. Moon, and I.-Y. Song</i>	
Author Index	665