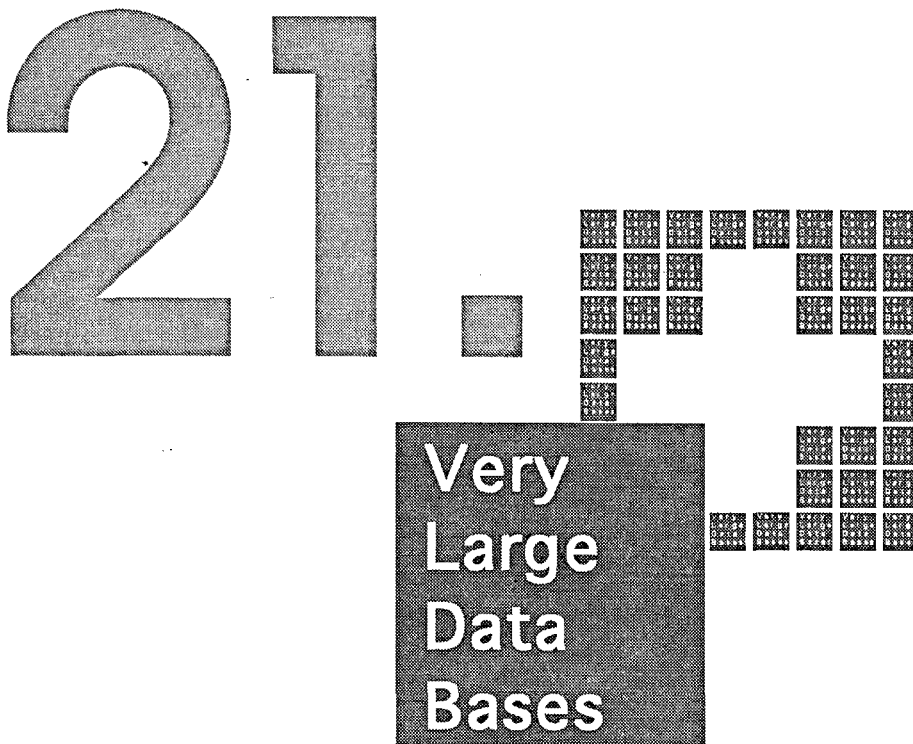


VLDB '95

Proceedings of the
21st International
Conference on
Very Large Data Bases

Zurich, Switzerland
Sept. 11-15, 1995



Editors:
Umeshwar Dayal
Peter M. D. Gray
Shojiro Nishio

Contents

Invited Talks

- Database Systems: Road Kill on the Information Superhighway?
David J. DeWitt (University of Wisconsin)
- The Double Life of the Transaction Abstraction: Fundamental Principle and Evolving System Concept.....2
Henry F. Korth (Panasonic Technologies, Inc.)
- Document Management as a Database Problem7
Rudolf Bayer (Technical University of Munich)
- From VLDB to VMLDB (Very MANY Large Data Bases): Dealing with Large-Scale Semantic Heterogeneity 11
Stuart E. Madnick (Massachusetts Institute of Technology)

Session 1: Object Stores

- A Performance Evaluation of OID Mapping Techniques 18
A. Eickler, C.A. Gerlhof, D. Kossmann (Germany, USA)
- OODB Bulk Loading Revisited: The Partitioned-List Approach 30
J.L. Wiener, J.F. Naughton (USA)
- Efficient Incremental Garbage Collection for Client-Server Object Database Systems 42
L. Amsaleg, M. Franklin, O. Gruber (France, USA)

Session 2: Distributed Information Retrieval

- W3QS: A Query System for the World-Wide Web 54
D. Konopnicki, O. Shmueli (Israel)
- Duplicate Removal in Information Dissemination 66
T.W. Yan, H. Garcia-Molina (USA)
- Generalizing GLOSS to Vector-Space Databases and Broker Hierarchies 78
L. Gravano, H. Garcia-Molina (USA)

Session 3: Disk Management

- Hot Block Clustering for Disk Arrays with Dynamic Striping 90
K. Mogi, M. Kitsuregawa (Japan)
- A Performance Study of Workfile Disk Management for Concurrent
Mergesorts in a Multiprocessor Database System 100
K.-L. Wu, P.S. Yu, J.-Y. Chung, J.Z. Teng (USA)
- Declustering Databases on Heterogeneous Disk Systems 110
L.T. Chen, D. Rotem, S. Seshadri (USA)

Session 4: Concurrency Control

Value-cognizant Speculative Concurrency Control	122
<i>A. Bestavros, S. Braoudakis (USA)</i>	
High-Concurrency Locking in R-Trees	134
<i>M. Kornacker, D. Banks (Germany, USA)</i>	
Index Concurrency Control in Firm Real-Time DBMS	146
<i>B. Goyal, J. Haritsa, S. Seshadri, V. Srinivasan (India, USA)</i>	

Session 5: Object Technology

A Data Transformation System for Biological Data Sources	158
<i>P. Buneman, S.B. Davidson, K. Hart, C. Overton, L. Wong (USA, Singapore)</i>	
Schema and Database Evolution in the O ₂ Object Database System	170
<i>F. Ferrandina, T. Meyer, R. Zicari, G. Ferran, J. Madec (Germany, France)</i>	
Procedures in Object-Oriented Query Languages	182
<i>K. Subieta, Y. Kambayashi, J. Leszczykowski (Japan, Poland)</i>	

Session 6: Transaction Models

Towards a Cooperative Transaction Model - The Cooperative Activity Model	194
<i>M. Rusinkiewicz, W. Klas, T. Tesch, J. Waesch, P. Muth (USA, Germany)</i>	
A Practical and Modular Method to Implement Extended Transaction Models	206
<i>R. Barga, C. Pu (USA)</i>	
Using Formal Methods to Reason about Semantics-Based Decompositions of Transactions	218
<i>P. Ammann, S. Jajodia, I. Ray (USA)</i>	

Session 7: Query Optimization

Bypassing Joins in Disjunctive Queries	228
<i>M. Steinbrunn, K. Peithner, G. Moerkotte, A. Kemper (Germany)</i>	
Coloring Away Communication in Parallel Query Optimization	239
<i>W. Hasan, R. Motwani (USA)</i>	
The Fittest Survives: An Adaptive Approach to Query Optimization	251
<i>H. Lu, K.-L. Tan, S. Dao (Singapore, USA)</i>	

Session 8: Multimedia

Type Classification of Semi-Structured Documents	263
<i>M. Tresch, N. Palmer, A. Luniewski (USA)</i>	

L/MRP: A Buffer Management Strategy for Interactive Continuous Data Flows in a Multimedia DBMS ...	275
<i>F. Moser, A. Kraiß, W. Klas (Germany)</i>	
Avoiding Retrieval Contention for Composite Multimedia Objects	287
<i>S. Chaudhuri, S. Ghandeharizadeh, C. Shahabi (USA)</i>	

Session 9: Cost Models

Estimating the Selectivity of Spatial Queries Using the ‘Correlation’ Fractal Dimension	299
<i>A. Belussi, C. Faloutsos (Italy, USA)</i>	
Sampling-Based Estimation of the Number of Distinct Values of an Attribute	311
<i>P.J. Haas, J.F. Naughton, S. Seshadri, L. Stokes (USA, India)</i>	
A Cost Model for Clustered Object-Oriented Databases	323
<i>G. Gardarin, J.-R. Gruser, Z.-H. Tang (France)</i>	

Session 10: Query Processing

Processing Object-Oriented Queries with Invertible Late Bound Functions	335
<i>S. Flodin, T. Risch (Sweden)</i>	
Eager Aggregation and Lazy Aggregation	345
<i>W.P. Yan, P. Larson (Canada)</i>	
Aggregate-Query Processing in Data Warehousing Environments	358
<i>A. Gupta, V. Harinarayan, D. Quass (USA)</i>	

Session 11: Parallel Databases

A Non-Uniform Data Fragmentation Strategy for Parallel Main-Memory Database Systems	370
<i>N. Bassiliades, I. Vlahavas (Greece)</i>	
Managing Intra-operator Parallelism in Parallel Database Systems	382
<i>M. Mehta, D.J. DeWitt (USA)</i>	
Dynamic Multi-Resource Load Balancing in Parallel Database Systems	395
<i>E. Rahm, R. Marek (Germany)</i>	

Session 12: Mining Association Rules

Mining Generalized Association Rules	407
<i>R. Srikant, R. Agrawal (USA)</i>	
Discovery of Multiple-Level Association Rules from Large Databases	420
<i>J. Han, Y. Fu (Canada)</i>	
An Efficient Algorithm for Mining Association Rules in Large Databases	432
<i>A. Savasere, E. Omiecinski, S. Navathe (USA)</i>	

Session 13: Recovery and Availability

Improving Performance in Replicated Databases through Relaxed Coherency	445
<i>R. Gellersdoerfer, M. Nicola (Germany)</i>	
Redo Recovery after System Crashes.....	457
<i>D. Lomet, M.R. Tuttle (USA)</i>	
The ClustRa Telecom Database: High Availability, High Throughput, and Real-Time Response.....	469
<i>S.-O. Hvasshovd, Ø. Torbjørnsen, S.E. Bratsberg, P. Holager (Norway)</i>	

Session 14: Data Mining

NeuroRule: A Connectionist Approach to Data Mining.....	478
<i>H. Lu, R. Setiono, H. Liu (Singapore)</i>	
Fast Similarity Search in the Presence of Noise, Scaling, and Translation in Time-Series Databases.....	490
<i>R. Agrawal, K.-I. Lin, H.S. Sawhney, K. Shim (USA)</i>	
Querying Shapes of Histories	502
<i>R. Agrawal, G. Psaila, E.L. Wimmers, M. Zait (USA, Italy)</i>	

Session 15: Schema Management

A Scalable Architecture for Autonomous Heterogeneous Database Interactions.....	515
<i>S. Milliner, A. Bouguettaya, M. Papazoglou (Australia)</i>	
OPOSSUM: Desk-Top Schema Management through Customizable Visualization	527
<i>E.M. Haber, Y.E. Ioannidis, M. Livny (USA)</i>	
Flexible Relations - Operational Support of Variant Relational Structures	539
<i>C. Kalus, P. Dadam (Germany)</i>	

Session 16: Access Methods

The hB ^{II} -tree: A Modified hB-tree Supporting Concurrency, Recovery and Node Consolidation	551
<i>G. Evangelidis, D. Lomet, B. Salzberg (USA)</i>	
Generalized Search Trees for Database Systems	562
<i>J.M. Hellerstein, J.F. Naughton, A. Pfeffer (USA)</i>	
Near Neighbor Search in Large Metric Spaces.....	574
<i>S. Brin (USA)</i>	

Session 17: Mass Storage

Query Processing in Tertiary Memory Databases.....	585
<i>S. Sarawagi (USA)</i>	

Applying Database Technology in the ADSM Mass Storage System.....	597
<i>L.-F. Cabrera, R. Rees, W. Hineman (USA)</i>	

Session 18: Spatial Data

Benchmarking Spatial Join Operations with Spatial Output	606
<i>E.G. Hoel, H. Samet ((USA)</i>	
Similarity based Retrieval of Pictures Using Indices on Spatial Relationships	619
<i>A.P. Sistla, C. Yu, C. Liu, K. Liu (USA)</i>	

Panels

Scientific Journals: Extinction or Explosion?	631
<i>Chair: R. Ramakrishnan (Wisconsin University)</i>	
Databases and Workflow Management: What is it All About?	632
<i>Chair: A. Reuter (University of Stuttgart)</i>	
Metrics for Accessing Heterogeneous Data: Is There Any Hope?	633
<i>Chair: L. Seligman (MITRE Corporation)</i>	

Application Session 1

Providing Database Migration Tools - A Practitioner's Approach	635
<i>A. Meier (Switzerland)</i>	
Promises and Realities of Active Database Systems	642
<i>E. Simon, A. Kotz-Dittrich (France, Switzerland)</i>	

Application Session 2

Database De-Centralization - A Practical Approach	654
<i>T. Didriksen, C. Galindo-Legaria, E. Dahle (Norway)</i>	
A Product Specification Database for Visual Prototyping	666
<i>K. Sumiya, K. Yasutake, H. Tanaka, N. Sanada, Y. Imai (Japan)</i>	

Vendor Session 1

Very Large Databases: How Large, How Different?	677
<i>D. Vaskevitch (Microsoft Corporation)</i>	

Vendor Session 2

DB2 Query Parallelism: Staging and Implementation.....	686
<i>Y. Wang (IBM Corporation)</i>	
Managing a DB2 Parallel Edition Database.....	692
<i>G. Fecteau (IBM corporation)</i>	

Vendor Session 3

DB2 Common Server:Technology Progress & Directions.....	694
<i>B. Lindsay (IBM Corporation)</i>	
OS Support for VLDBs: Unix Enhancements for the Teradata Database.....	696
<i>J. Catozzi, S. Rabinovici (ATT, Teradata)</i>	

Vendor Session 4

Accessing a Relational Database through an Object-Oriented Database Interface	702
<i>J. Orenstein (Object Design Incorporated)</i>	

Vendor Session 5

Informix-Online XPS: A Dynamically Scalable RDBMS for Open Parallel Platforms.....	706
<i>H. Spintzik (Informix Incorporated)</i>	
The Oracle Warehouse	707
<i>G. Hallmark (Oracle Corporation)</i>	

Vendor Session 6

Efficient Search of Multidimensional B-Tree.....	710
<i>H. Leslie, R. Jain, D. Birdsall, H. Yaghmai (Tandem Computers)</i>	
BigSur: A System for the Management of Earth Science Data.....	720
<i>M. Stonebraker, P. Brown (Illustra)</i>	