

Jeffrey F. Naughton
Computer Sciences Department
University of Wisconsin, Madison
Madison, WI 53706
(608) 262-8737

Education:

Stanford University. PhD (Computer Science) 1987.

University of Wisconsin at Madison. BS (Mathematics) 1982.

Experience:

University of Wisconsin-Madison. Professor. 1997 - present.

University of Wisconsin-Madison. Professor and Chair. 2010 - 2014.

University of Wisconsin-Madison. Associate Professor. 1992 - 1997.

University of Wisconsin-Madison. Assistant Professor. 1989 – 1992.

Princeton University, Princeton, New Jersey. Assistant Professor. 1987 – 1989.

Service to the Profession:

Associate Editor: Foundations and Trends in Databases, 2009 - 2013.

Associate Editor: VLDB Journal, 2003 – 2009.

Associate Editor: ACM Transactions on Database Systems, 1997 – 2002.

Program Committee Chair:

1996 ACM Symp. on Parallel and Distributed Information Systems (co-Chair with Gerhard Weikum)

2000 ACM SIGMOD Conference.

Program Committee Member:

International VLDB conference (1997, 2004, 2005, 2006, 2007, 2009, 2015), ACM SIGMOD Conference (1990, 1991, 1993, 1996, 1999, 2003, 2005, 2006, 2007, 2009), ACM PODS Conference (1992, 1995, 2002, 2004, 2010), ACM OOPSLA Conference (1994, 1995), ACM PDIS Conference (1994) International Conference on Database Theory (1992, 2003, 2004, 2005), IEEE Data Engineering (1999, 2002, 2004, 2005, 2006, 2010) CIDR Conference (2005, 2009) SOCC Conference (2012)

Selected Advisory Boards, Consultancies:

Greenplum, Johnson Controls, Microsoft, NEC, Rocket Fuel Media, Teradata.

Awards:

John P. Mogridge Chair of Computer Sciences, 2011-2016.

2008 ACM Software Systems Award, for work on the Gamma Parallel Database System Project. Given “to an institution or individuals recognized for developing software systems that have had a lasting influence, reflected in contributions to concepts and/or commercial acceptance.”

2004 SIGMOD Ten-Year “Test of Time” Award for “Shoring up Persistent Applications.” (With Michael Carey, David DeWitt, Marvin Solomon, and others.) Originally in *Proceedings of the ACM SIGMOD Symposium on Management of Data*, Minneapolis, MN, May 1994.

ACM Fellow, 2003

University of Wisconsin-Madison Vilas Associate Award for Excellence in Research, 2000.

NSF Presidential Young Investigator, 1991

Best Paper Award, SIGMOD 1990, Practical selectivity estimation. (With Richard J. Lipton and Donovan A. Schneider.)

UW-Madison Student ACM Chapter, “Students Choice Professor of the Year,” awarded for classroom teaching, 1998, 2002, 2003, 2007, 2014.

Patents:

1. US Patent 7,155,434, “Locking mechanism employing a name lock for materialized views” (joint with Gang Luo, Curt Ellmann, and Michael Watzke.)
2. US Patent 7,149,737, “Locking mechanism using a predefined lock for materialized views in a database system” (joint with Gang Luo, Michael Watzke, and Curt Ellmann.)
3. US Patent 7,092,951, “Auxiliary relation for materialized view” (joint with Gang Luo and Curt Ellmann.)
4. US Patent 7,085,769, “Method and apparatus for performing hash join” (joint with Gang Luo and Curt Ellmann.)
5. US Patent 6,990,503, “Rescheduling transactions in a database system” (Joint with Gang Luo, Curt Ellmann, and Michael Watzke.)
6. US Patent 6,804,678, “Non-blocking parallel band join algorithm” (Joint with Gang Luo and Curt Ellmann.)
7. US Patent 6,763,357, “Method for determining the computability of data for an active multi-dimensional cache in a relational database system” (Joint with Prasad Deshpande, Karthikeyan Ramasamy, and Amit Shukla.)
8. US Patent 6,745,198, “Parallel spatial join index” (Joint with Gang Luo and Curt Ellmann.)
9. US Patent 6,732,107, “Spatial join method and apparatus” (Joint with Gang Luo and Curt Ellmann.)
10. US Patent 6,728,694, “Set containment join operation in an object/relational database management system” (Joint with Karthikeyan Ramasamy and Jignesh Patel.)
11. US Patent 6,678,672, “Efficient exception handling during access plan execution in an on-line analytic processing system” (Joint with Karthikeyan Ramasamy, Prasad Deshpande, and Amit Shukla.)
12. US Patent 6,601,062, “Active caching for multi-dimensional data sets in relational database management system” (Joint with Prasad Deshpande, Karthik Ramasamy, and Amit Shukla.)
13. US Patent 6,567,804, “Shared computation of user-defined metrics in an on-line analytic processing system” (Joint with Karthikeyan Ramasamy, Prasad Deshpande, and Amit Shukla.)
14. US Patent 6,567,803, “Simultaneous computation of multiple moving aggregates in a relational database management system” (Joint with Karthikeyan Ramasamy, Prasad Deshpande, and Amit Shukla.)
15. US Patent 6,182,060, “Method and apparatus for storing, retrieving, and processing multi-dimensional customer-oriented data sets” (Joint with Robert Hedgcock and Timothy Keane.)

Research Grants and Awards:

1. “Extreme Relational Query Processing” 2013-15, \$1,500,000 Google Focus Award, joint with AnHai Doan and Jignesh Patel.

2. "Secure Data Sharing," PI David Page, NIH, 2012-2015, my share approximately one RA and one summer month per year.
3. "Enabling Technology for Best-Effort Data Integration Systems," \$499,292, NSF, for 2010 - 2013, with AnHai Doan.
4. "Microsoft Gray Systems Lab RAs," approx. \$450,000/yr, renewable annually, awarded 2009 - 2015, PI, with Jignesh Patel.
5. "A Comparative Study of Approaches to Cluster- Based Large Scale Data Analysis," \$110,000, NSF, for 2009-2011.
6. "Goal-Oriented Privacy-Preservation," NSF, \$1,600,000, for 2005-2012. (with David DeWitt and Raghu Ramakrishnan.)
7. "Best effort information integration," DARPA, \$500,000, for 2007-2008. (with AnHai Doan.)
8. "A Test-Bed for Condor DB," NSF, \$300,000, for 2007-2009. (with David DeWitt.)
9. "Condor DB: Integrating Condor and DBMS Technology," \$1,350,747 for 2005-2010, NSF (Joint with David DeWitt.)
10. "A Petabyte in your Pocket," \$4,000,000 for 2000-2005, NSF. (Joint with David DeWitt and David Maier.)
11. "Niagara: Net Data Management," \$850,000 for 1999-2002, DARPA. (Joint with David DeWitt and David Maier.)
12. "Paradise and Satellite Data Dissemination," \$3,534,000 for 1997-2000, ARPA. (Joint with David DeWitt.)
13. "NSF PACI Supercomputing Subcontract," \$1,250,000 for 1997-2002, NSF. (Joint with David DeWitt.)
14. "MIDSHIP: Managing Image Datasets with Scalable High Performance", NSF \$1,600,000 for 1996-2000. (Project Co-Director with Mark Hill.)
15. "Paradise Clusters: Inexpensive Scalable Support for Standardized Access to EOSDIS Data Sets", NASA \$1,270,000 for 1994-97. (Joint with David DeWitt).
16. "Paradise — A Parallel Information System for EOSDIS", NASA, 1994-1997, (\$457,000). (Joint with David DeWitt.)
17. "Optimization of Complex Queries." National Science Foundation. \$266,184 for 1992 - 1995. (With Yannis Ioannidis.)
18. NSF Presidential Young Investigator Award. \$500,000 for 1991 – 1996. (With full matching.)
19. SHORE, A High-Performance, Scalable, Persistent Object System for Mach and Touchstone, (joint with M. Carey, D. DeWitt, and M. Solomon), DARPA, 1991-98, \$4,172,000.
20. "Scheduling and Complex Query Processing in Highly Parallel Database Machines," IBM Research Initiation Award, \$600,000 for 1991–1994. (with Michael Carey and David DeWitt.)
21. "Logic-Based Database Query Languages." NSF Research Initiation Award. \$68,145 for 1989–1991.

Graduated Ph.D. Students (first employment):

Yueh-Hsuan Chiang, 2014 (Facebook)
Jiexing Li, 2014 (Google)
Ian Rae, 2014 (Microsoft)
Chen Zeng, 2013 (Google)
Khai Tran, 2013 (Oracle Labs)
Yeye He, 2012 (Microsoft Research)
Sun Chong, 2012 (Walmart Labs)
Akanksha Baid, 2011 (Bloom Reach).
Christine Reilly, 2010 (University of Texas Pan American).
Lakshmikant Shrinivas, 2009 (Vertica).
Eric Chu, 2008 (Microsoft).
Jiansheng Huang, 2008 (Rocket Fuel Media).
Ameet Kini, 2007 (Oracle).
Tochukwu Iwuchukwu, 2007 (Nokia Research).
Ahmed Ayad, 2006 (Microsoft).
Jennifer Beckmann, 2006 (Microsoft).
Alan Halverson, 2006 (Microsoft).
Rajasekar Krishnamurthy, 2004 (IBM Almaden).
Gang Luo, 2004 (IBM Watson).
Jaewoo Kang, 2003 (NC State).
Raghav Kaushik, 2003 (Microsoft Research).
Stratis Viglas, 2003 (University of Edinburgh).
Ashraf Aboulmaga, 2002 (IBM Almaden).
Qiong Luo, 2002 (HKUST).
Chun Zhang, 2002 (IBM Almaden).
Jayavel Shanmugasundaram, 2001 (Cornell).
Karthikeyan Ramasamy, 2001 (Desana Systems).
Prasad Deshpande, 1999 (Desana Systems).
Amit Shukla, 1999 (Desana Systems).
Yihong Zhao, 1998 (Informix).
Ambuj Shatdal, 1996 (NCR).
Srikant Ramakrishnan, 1996 (IBM Almaden).
Shivakumar Venkataraman, 1996 (IBM Santa Theresa).
Bradley Rubin, 1996 (CTO, Imation).
Joseph M. Hellerstein, 1995 (UC Berkeley).
Janet L. Wiener, 1995 (Compaq).
S. Seshadri, 1992 (IIT-Mumbai).
Manolis Tsangaris, 1992 (ATT Bell Laboratories).

Publications:

1. Yueh-Hsuan Chiang, AnHai Doan, Jeffrey F. Naughton, "Tracking Entities in the Dynamic World: A Fast Algorithm for Matching Temporal Records." VLDB 2014.
2. Wentao Wu, Xi Wu, Hakan Hacigumus, Jeffrey F. Naughton, "Uncertainty Aware Query Execution Time Prediction." VLDB 2014.
3. Jiexing Li, Jeffrey F. Naughton, Rimma V. Nehme, "Resource Bricolage for Parallel Database Systems." VLDB 2014.
4. Ian Rae, Alan Halverson, Jeffrey F. Naughton, "In-RDBMS inverted indexes revisited." ICDE 2014.

5. Yeye He, Siddharth Barman, Jeffrey F. Naughton, "On Load Shedding in Complex Event Processing." ICDT 2014.
6. Khai Q. Tran, Jeffrey F. Naughton, Bruhathi Sundarmurthy, Dimitris Tsirogiannis, "JECB: a join-extension, code-based approach to OLTP data partitioning." SIGMOD 2014.
7. Chaitanya gokhale, Sanjib Das, AnHai Doan, Jeffrey F. Naughton, Narasimhan Rampalli, Jude W. Shavlik, Xiojin Zhu, "Corleone: hands-off crowdsourcing for entity matching." SIGMOD 2014.
8. PengCheng Xiong, Hakan Hacigumus, Jeffrey F. Naughton, "A software-defined networking based approach for performance management of analytical queries on distributed data stores." SIGMOD 2014.
9. Yueh-Hsuan Chiang, AnHai Doan, Jeffrey F. Naughton, "Modeling entity evolution for temporal record matching." SIGMOD 2014.
10. Willis Lang, Rimma Nehme, Eric Robinson, Jeffrey F. Naughton, "Partial results in database systems." SIGMOD 2014.
11. Yun Chi, Hakan Hacigumus, W.-P. Hsiung, Jeffrey F. Naughton, "Distribution-based query scheduling," VLDB 2013.
12. Wentao Wu, Yun Chi, Hakan Hacigumus, and Jeffrey F. Naughton, "Towards Predicting Query Execution Time for Concurrent and Dynamic Database Workloads." VLDB 2013.
13. Jiexing Li, Rimma V. Nehme, Jeffrey F. Naughton, "Toward Progress Indicators on Steroids for Big Data Systems," CIDR 2013.
14. Wentao Wu, Yun Chi, Junichi Tatemura, Hakan Hacigumus, Jeffrey F. Naughton, "Predicting Query Execution Time: Are Optimizer Cost Models Really Unusable?", ICDE 2013.
15. Chen Zeng, Jin-Yi Cai, Pinyan Lu, Jeffrey F. Naughton, "On optimal differentially private mechanisms for count-range queries." ICDT 2013.
16. Chen Zeng, Eric Lantz, Jeffrey F. Naughton, David Page, "On Differentially Private Inductive Logic Programming." ILP 2013.
17. Di Wang, Yeye He, Elke A. Rundensteiner, Jeffrey F. Naughton, "Utility-maximizing event stream suppression." SIGMOD 2013.
18. Chen Zeng, Jeffrey F. Naughton, Jin-Yi Cai, "Towards an Optimal Differentially Private Mechanism for Threshold Queries." VLDB 2013.
19. Jiexing Li, Rimma V. Nehme, Jeffrey F. Naughton: GSLPI: A Cost-Based Query Progress Indicator. ICDE 2012: 678-689.
20. Chong Sun, Jeffrey F. Naughton, Siddharth Barman: Approximate String Membership Checking: A Multiple Filter, Optimization-Based Approach. ICDE 2012: 882-893
21. Christine F. Reilly, Yueh-Hsuan Chiang, Jeffrey F. Naughton: Instrumenting a logic programming language to gather provenance from an information extraction application. WWW (Companion Volume) 2012: 589-590.
22. Yeye He, Siddharth Barman, Jeffrey F. Naughton: Preventing equivalence attacks in updated, anonymized data. ICDE 2011: 529-540.

23. Yeye He, Siddarth Barman, Di Wang, Jeffrey F. Naughton: On the complexity of privacy-preserving complex event processing. PODS 2011: 165-174.
24. Jaeyoung Do, Donghui Zhang, Jignesh M. Patel, David J. DeWitt, Jeffrey F. Naughton, Alan Halver-son: Turbocharging DBMS buffer pool using SSDs. SIGMOD 2011: 1113-1124.
25. Chong Sun, Jeffrey F. Naughton: The Token Distribution Filter for Approximate String Membership. WebDB 2011.
26. "Impact of disk corruption on open-source DBMS" (with Sriram Subramanian, Yupu Zhang, Rajiv Vaidyanathan, Haryadi S. Gunawi, Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau.) In ICDE 2010.
27. "Toward industrial-strength keyword search systems over relational data" (with Akanksha Baid, Ian Rae, and AnHai Doan). In ICDE 2010.
28. "Transaction reordering" (with Gang Luo, Curt J. Ellmann, Michael and Watzke). In TDKE 2010.
29. "Toward scalable keyword search over relational data" (with Akanksha Baid, Ian Rae, Jiexing Li, and AnHai Doan). In VLDB 2010.
30. "Combining Keyword Search and Forms for Ad-Hoc Querying of Databases" (with Eric Chu, Akanksha Baid, Xiaoyong Chai, and AnHai Doan.) In SIGMOD 2009.
31. "Efficiently Incorporating User Feedback into Information Extraction and Integration Programs" (with Xiaoyong Chai, Ba-Quy Vuong, and AnHai Doan.) In SIGMOD 2009.
32. "Transparently Gathering Provenance with Provenance Aware Condor" (with Christine Reilly.) In Proceedings of the 2009 Workshop on the Theory and Practice of Provenance.
33. "Anonymization of Set-Valued Data via Top-Down, Local Generalization" (with Yeye He.) In VLDB 2009.
34. "The Case for a Structured Approach to Managing Unstructured Data" (with AnHai Doan and others.) In CIDR 2009.
35. "Schema Matching Using Interattribute Dependencies" (with Jaewoo Kang.) In IEEE Transactions on Knowledge and Data Engineering, October 2008.
36. "On the Provenance of Non-Answers to Queries over Extracted Data" (with Jiansheng Huang, Ting Chen, and AnHai Doan.) In VLDB 2008.
37. "Clustera: An Integrated Computation and Data Management System" (with David DeWitt, Eric Robinson, and others.) In VLDB 2008.
38. "Issues in applying data mining to grid job failure detection and diagnosis" (with Lakshmikanth Shrinivas.) In HPDC 2008.
39. "Form-based proxy caching for database-backed web sites: keywords and functions" (with Qiong Luo and Wenwei Xue.) In the VLDB Journal 17(3), 2008.
40. "The case for a wide-table approach to manage sparse data sets" (with Eric Chu and Jennifer Beckmann.) In SIGMOD 2007.
41. "K-Relevance: A Spectrum of Relevance for Data Sources Impacting a Query." (with Jiansheng Huang.) In SIGMOD 2007.

42. "K-Anonymization as Spatial Indexing: Toward Scalable and Incremental Anonymization." (with Tochukwu Iwuchukwu and David DeWitt.) In VLDB 2007.
43. "Declarative Information Extraction using Datalog with Embedded Extraction Predicates." (with Warren Shen, AnHai Doan, and Raghu Ramakrishnan.) In VLDB 2007.
44. "A Relational Approach to Incrementally Extracting and Querying Structure in Unstructured Data" (with Eric Chu, Akanksha Baid, Ting Chen, and AnHai Doan.) In VLDB 2007.
45. "Database Support for Weighted Match Joins" (with Ameet Kini), in the 19th International Conference on Scientific and Statistical Database Management, 2007.
46. "The Case for a Wide-Table Approach to Manage Sparse Relational Data Sets" (with Eric Chu and Jennifer Beckmann), in ACM SIGMOD 2007.
47. "K-Relevance: A Spectrum of Relevance for Data Sources Impacting a Query" (with Jiansheng Huang), to appear in ACM SIGMOD 2007.
48. "TRAC: Toward Recency and Consistency Reporting in a Database with Distributed Data Sources" (with Jiansheng Huang and Miron Livny), in VLDB 2006.
49. "Multi-Query SQL Progress Indicators" (with Gang Luo and Philip S. Yu), in EDBT 2006.
50. "Exploring Provenance in a Distributed Job Execution System" (with Christine Reilly.) International Provenance and Annotation Workshop (IPAW'06), Chicago, Illinois, May 2006.
51. "Database support for matching: limitations and opportunities" (with Ameet Kini, Srinath Shankar, and David J. DeWitt), in SIGMOD 2006.
52. "Extending RDBMSs to Support Sparse Datasets Using an Interpreted Attribute Storage Format" (with Jennifer Beckmann, Alan Halverson, and Rajasekar Krishnamurthy), in ICDE 2006.
53. "End-biased Samples for Join Cardinality Estimation" (with Cristian Estan), in ICDE 2006.
54. "Approximating Streaming Window Joins Under CPU Limitations" (with Ahmed Ayad, Stephen Wright, and Utkarsh Srivastava), in ICDE 2006 (poster paper.)
55. "The Lowell database research self-assessment" (with Serge Abiteboul, Rakesh Agrawal, Philip Bernstein, and many others), in Communications of the ACM 48(5).
56. "Locking Protocols for Materialized Aggregate Join Views" (with Gang Luo, Curt Ellmann, and Mike Watzke), in IEEE Trans. Knowl. Data Eng. 17(6). Extended version of 69.
57. "Form-Based Proxy Caching for Database-Backed Web Sites: Keywords and Functions" (with Qiong Luo and Wenwei Xue). Substantially extended version of 86.
58. "Integrating databases and workflow systems" (with Srinath shankar, Ameet Kini, and David J. DeWitt), in SIGMOD Record 34(3), 2005.
59. "Increasing the Accuracy and Coverage of SQL Progress Indicators" (with Gang Luo, Curt Ellmann, and Michael Watzke), in ICDE 2005.
60. "XML Views as Integrity Constraints and their Use in Query Translation" (with Rajasekar Krishnamurthy and Raghav Kaushik), in ICDE 2005.
61. "Efficient XML-to-SQL Query Translation: Where to Add the Intelligence?" (with Rajasekar Krishnamurthy and Raghav Kaushik), in VLDB 2004.

62. "On the Integration of Structure Indexes and Inverted Lists" (with Raghav Kaushik, Rajasekar Krishnamurthy, and Raghu Ramakrishnan) in SIGMOD 2004.
63. "Static Optimization of Conjunctive Queries with Sliding Windows Over Infinite Streams" (with Ahmed Ayad) in SIGMOD 2004.
64. "Toward a Progress Indicator for Database Queries" (with Gang Luo, Curt Ellmann, and Michael Watzke) in SIGMOD 2004.
65. "Recursive XML Schemas, Recursive XML Queries, and Relational Storage: XML-to-SQL Query Translation" (with Rajasekar Krishnamurthy, Raghav Kaushik, and Venkatesan Chakaravarthy) in ICDE 2004.
66. "Unraveling the Duplicate Elimination Problem in XML-to-SQL Query Translation" (with Rajasekar Krishnamurthy and Raghav Kaushik), in WebDB 2004.
67. "Building XML statistics for the hidden web" (with Ashraf Aboulmaga), in CIKM 2003.
68. "On the Difficulty of Finding Optimal Relational Decompositions for XML Workloads: A Complexity Theoretic Perspective" (with Rajasekar Krishnamurthy and Venkatesan T. Chakaravarthy) in ICDT 2003.
69. "Locking Protocols for Materialized Aggregate Join Views" (with Gang Luo, Curt Ellmann, and Michael Watzke) in VLDB 2003.
70. "Maximizing the Output Rate of Multi-Way Join Queries over Streaming Information Sources" (with Stratis Viglas and Josef Burger), in VLDB 2003.
71. "Mixed Mode XML Query Processing" (with Alan Halverson and eight others), in VLDB 2003.
72. "On Relational Support for XML Publishing: Beyond Sorting and Tagging" (with Raghav Kaushik and Surajit Chaudhuri), in SIGMOD 2003.
73. "On Schema Matching with Opaque Column Names and Data Values" (with Jaewoo Kang), in SIGMOD 2003.
74. "A Comparison of Three Methods for Join View Maintenance in a Parallel RDBMS" (with Gang Luo), in ICDE 2003.
75. "Evaluating Window Joins over Unbounded Streams" (with Jaewoo Kang and Stratis Viglas), in ICDE 2003.
76. "On the Difficulty of Finding Optimal Relational Decompositions for XML Workloads: A Complexity Theoretic Perspective" (with Rajasekar Krishnamurthy and Venkatesan T. Chakaravarthy), in ICDT 2003.
77. "Updates for Structure Indexes" (with Raghav Kaushik), in VLDB 2002.
78. "Middle-tier Database Caching for e-Business" (with Qiong Luo, Sailesh Krishnamurthy, C. Mohan, Hamid Pirahesh, Honguk Woo, and Bruce Lindsay), in SIGMOD 2002 (industrial track.)
79. "Parallel Ripple Joins" (with Gang Luo and Peter Haas), in SIGMOD 2002.
80. "Covering Indices for XML Queries" (with Raghav Kaushik, Phil Bohannon, and Hank Korth), in SIGMOD 2002.
81. "Rate-Based Query Optimization" (with Stratis Viglas), in SIGMOD 2002.

82. "A Non-blocking Parallel Spatial Join Algorithm" (with Gang Luo and Curt Ellmann), in ICDE 2002.
83. "Design and Evaluation of Alternative Selection Placement Strategies in Optimizing Continuous Queries" (with Jianjun Chen and David DeWitt), in ICDE 2002.
84. "The Niagara Internet Query System" (with David DeWitt, David Maier, and others), IEEE Data Engineering Bulletin 24(2), 2001.
85. "Estimating the Selectivity of XML Path Expressions for Internet Scale Applications" (with Ashraf Aboulnaga and Alaa R. Almaeldeen.) In VLDB 2001.
86. "Form-Based Proxy Caching for Database-Backed Web Sites" (with Qiong Luo.) In VLDB 2001.
87. "Generating Synthetic Complex-Structured XML Data" (with Ashraf Aboulnaga and Chun Zhang.) In WebDB 2001.
88. "On the Complexity of Join Predicates" (with Jin-Yi Cai, Venkatesan Chakaravarthy, and Raghav Kaushik). In PODS 2001, May 2001.
89. "On Supporting Containment Queries in Relational Database Systems" (with Chun Zhang, David DeWitt, and Guy Lohman.) In SIGMOD 2001, May 2001.
90. "Set Containment Joins: The Good, The Bad, and The Ugly." (with Karthikeyan Ramasamy, Jignesh Patel, and Raghav Kaushik.) In VLDB 2000, pp. 351 – 362.
91. "Architecting a Network Query Engine for Producing Partial Results," (with Jayavel Shanmugasundaram, Kristin Tufte, David DeWitt, and David Maier.) To appear in Lecture Notes in Computer Science, Vol. 1997, Springer-Verlag Publishers, 2001. A short version of this paper appeared in the proceedings of WebDB 2000, pp. 17–22.
92. "Active Query Caching for Database Web Servers," (with Qiong Luo, Rajasekar Krishnamurthy, Pei Cao, and Yunrui Li.) To appear in Lecture Notes in Computer Science, Vol. 1997, Springer-Verlag Publishers, 2001. A short version of this paper appeared in WebDB 2000, pp. 29-34.
93. "Active Caching for OLAP Workloads." (with Prasad Deshpande.) In EDBT00, March 2000, pp. 167–182.
94. "Precomputation for Multicube Data Models." (with Amit Shukla and Prasad Deshpande.) In EDBT00, March 2000, pp. 269–284.
95. "Accurate Estimation of the Cost of Spatial Queries." (with Ashraf Aboulnaga.) In ICDE 2000, March 2000, pp. 123–134.
96. "Relational Databases for XML Queries: Opportunities and Limitations." (with Jayavel Shanmugasundaram, Kristin Tufte, and others.) In VLDB 1999, pp. 302-314.
97. "Caching Multidimensional Queries using Chunks. (with Prasad Deshpande, Karthikeyan Ramasamy, and Amit Shukla.) In SIGMOD '98.
98. "Simultaneous Optimization and Evaluation of Multiple Dimensional Queries". (with Yihong Zhao, Prasad Deshpande, and Amit Shukla.) In SIGMOD '98.
99. Precomputation algorithms for static caching of multidimensional queries. (with Amit Shukla and Prasad Deshpande.) In *VLDB98*.
100. Caching Multidimensional Queries using Chunks. (with Prasad Deshpande, Karthikeyan Ramasamy, and Amit Shukla.) In *SIGMOD98*.

101. Simultaneous Optimization and Evaluation of Multiple Dimensional Queries. (with Yihong Zhao, Prasad Deshpande, and Amit Shukla.) In *SIGMOD98*.
102. Remote Load-Sensitive CACHing for Multi-Server Database Systems (with Shivakumar Venkataraman and Miron Livny.) In *Proceedings of ICDE '98*, Feb. 1998.
103. Array Support for Multidimensional Queries in Object-Relational DBMS. (With Yihong Zhao, Karthikeyan Ramasamy, and Kristin Tufte.) In *Proceedings of ICDE '98*, Feb. 1998.
104. The BUCKY Object-Relational Database Benchmark. (With Michael Carey, David DeWitt, Mohammed Asgarian, Johannes Gehrke, and Dhaval Shah.) In *Proceedings of SIGMOD97*, Tucson, AZ, May 1997.
105. An Array-Based Algorithm for Simultaneous Multidimensional Aggregates. (with Yihong Zhao and Prasad Deshpande.) In *Proceedings of SIGMOD97*, Tucson, AZ, May 1997.
106. Paradise: A Scalable GeoSpatial Database System (with J. Patel, J. Yu, and others.) In *Proceedings of SIGMOD97*, Tucson, AZ, May 1997.
107. Global Memory Management for Parallel Web Servers. (With Shivakumar Venkataraman.) In *Proceedings of Data Engineering 97*, England, March 1997.
108. On the Computation of Multidimensional Aggregates. (With Sammet Agarwal, Rakesh Agrawal, Prasad Deshpande, Ashish Gupta, Raghu Ramakrishnan, and Sunita Sarawagi.) In *Proceedings of VLDB96*, Mumbai (Bombay), India, 1996.
109. Selectivity and Cost Estimation for Joins Based on Random Sampling. (With Peter Haas, S. Seshadri, and Arun Swami.) In *Journal of Computer and System Sciences*, Volume 52, Number 3, June 1996, pages 550–569. (Extended version of 134.)
110. Estimating the Size Blowup in Multidimensional Aggregations. (With Prasad Deshpande and Amit Shukla.) In *Proceedings of VLDB96*, Mumbai (Bombay), India, 1996.
111. ParSets for Parallelizing OODB Traversals: Implementation and Performance. (With David J. DeWitt, John Shafer, and Shivakumar Venkataraman.) In *The VLDB Journal*, Volume 5, Number 1, January 1996, pp. 3–19. (Extended version of conference paper 122.)
112. Space optimization in the bottom-up evaluation of logic programs. (With S. Sudarshan, Divesh Srivastava, and Raghu Ramakrishnan.) In *ACM Trans. on Database Systems*, Volume 20, Number 4, December 1995, pp. 472–516. (Extended version of conference paper 148.)
113. Query Execution Techniques for Caching Expensive Methods. (With Joseph M. Hellerstein.) Appeared in *SIGMOD '95*.
114. Clocked adversaries for hashing. (With Richard J. Lipton.) In *Algorithmica*, Volume 9, Number 3, pages 239–252.
115. Sampling-Based Estimation of the Number of Distinct Values of an Attribute. (With Peter Haas, S. Seshadri, and Lynne Stokes.) In *Proceedings of the Twenty-first International Conference on Very Large Databases*, Zurich, Switzerland, August 1995.
116. Query Size Estimation by Adaptive Sampling. (With Richard J. Lipton.) In *Journal of Computer and System Sciences*, Volume 51, Number 1, 1995, pp. 18–25. (Extended version of conference paper 153.)
117. On the expected size of recursive datalog queries. (With S. Seshadri.) In *JCSS*, Volume 51, Number 2, October 1995, pages 137–148. (Extended version of conference paper 146.)

118. Argument reduction through factoring. (With Raghu Ramakrishnan, Yehoshua Sagiv, and Jeffrey D. Ullman.) In *Theoretical Computer Science*, Volume 146, Number 1-2, Pages 269–310, July 1995. (Extended version of conference paper 155.)
119. Generalized Search Trees for Database Systems. (With Joseph Hellerstein and Avi Pfeffer.) In *Proceedings of the Twenty-first International Conference on Very Large Databases*, Zurich, Switzerland, August 1995.
120. Bulk Loading Revisited. (With Janet Wiener.) In *Proceedings of the Twenty-first International Conference on Very Large Databases*, Zurich, Switzerland, August 1995.
121. Aggregate Processing in Parallel RDBMS. (With Ambuj Shatdal.) In *Proceedings of the ACM SIGMOD Symposium on Management of Data*, San Jose, California, May 1995.
122. ParSets for Parallelizing OODB Traversals: Implementation and Performance. (With David J. DeWitt, John Shafer, and Shivakumar Venkataraman.) In *Proceedings of the ACM Symposium on Parallel and Distributed Information Systems*, 1994.
123. Low-Latency, Concurrent Checkpointing for Parallel Programs. (With Kai Li and James S. Plank.) In *IEEE Transactions on Parallel and Distributed Systems*, Volume 5, Number 8, August 1994. (Extended version of conference paper 154.)
124. The OO7 Multiuser Benchmark: Description and Preliminary Experiences. (With Michael Carey, Chander Kant, and David DeWitt.) Invited paper in *Proceedings of the ACM OOPSLA Conference*, Portland, Oregon, October 1994.
125. How to forget the past without repeating it. (With Raghu Ramakrishnan.) In *Journal of the ACM*, Volume 41, Number 6, 1994, pages 1151-1177. (Extended version of conference paper 151.)
126. Bulk Loading into an OODBMS: A Performance Study. (With Janet Wiener.) In *Proceedings of the Twentieth International Conference on Very Large Databases*, Santiago, Chile, August 1994.
127. Cache Conscious Join Algorithms. (With Ambuj Shatdal and Chander Kant.) In *Proceedings of the Eighteenth International Conference on Very Large Databases*, Santiago, Chile, August 1994.
128. Shoring up Persistent Applications. (With Michael Carey, David DeWitt, Marvin Solomon, and others.) In *Proceedings of the ACM SIGMOD Symposium on Management of Data*, Minneapolis, MN, May 1994.
129. On the Relative Cost of Sampling for Join Selectivity Estimation. (With Peter Haas and S. Seshadri.) In *Proceedings of the ACM Symposium on Principles of Database Systems*, Minneapolis, MN, May 1994.
130. Garbage collection in persistent object stores. (With Voon-Fee Yong and Jie-Bing Yu.) In *Proceedings of the IEEE Conference on Data Engineering*, Houston, Texas, February 1994.
131. Using Shared Virtual Memory for Parallel Join Processing. (With Ambuj Shatdal.) In *Proceedings of the ACM SIGMOD Symposium on Management of Data*, Washington, DC, May 1993.
132. The OO7 Object-Oriented Database Benchmark. (With Michael Carey and David DeWitt.) In *Proceedings of the ACM SIGMOD Symposium on Management of Data*, Washington, DC, May 1993.
133. Efficient Sampling Strategies for Relational Database Operations. (With Richard J. Lipton and S. Seshadri and Donovan A. Schneider.) In *Theoretical Computer Science*, Number 116, 1993, Pages 195–226. (Combination and extension of conference papers 149 and 152.)

134. Fixed-Precision Estimation of Join Selectivity. (With Peter Haas, S. Seshadri, and Arun Swami.) In *Proceedings of the ACM Symposium on Principles of Database Systems*, Washington, DC, May 1993.
135. Nested loops revisited. (With David DeWitt and Joseph Burger.) In *Parallel and Distributed Information Systems '93*, San Diego, January 1993.
136. Practical skew handling in parallel joins. (With D. DeWitt and D. Schneider.) In *Proceedings of the Eighteenth International Conference on Very Large Database Systems*, Vancouver, British Columbia, August 1992.
137. A comparison of object clustering techniques. (With M. Tsangaris.) In *Proceedings of the ACM SIGMOD Symposium on Management of Data*, San Diego, California, June 1992.
138. An Efficient Checkpointing Method for Multicomputers with Wormhole Routing. (With Kai Li and James S. Plank.) In *International Journal of Parallel Processing*, Volume 20, Number 3, June 1992. (Extension of conference paper 141.)
139. Sampling Issues in Parallel Database Systems. (With S. Seshadri.) In *Proceedings of the International Conference on Extending Database Technology*, Vienna, Austria, March 1992.
140. Parallel Sorting on a Shared-Nothing Architecture using Probabilistic Splitting. (With David J. DeWitt and Donovan A. Schneider.) In *Proceedings of the Parallel and Distributed Information Systems Conference*, Miami Beach, Florida, December 1991.
141. Checkpointing multicomputer applications. (With Kai Li and James S. Plank.) In *Proceedings of the Tenth Symposium on Reliable Distributed Systems*, Pisa, Italy, September 1991.
142. A comparison of non-equijoin algorithms. (With David J. DeWitt and Donovan A. Schneider.) In *Proceedings of the Seventeenth International Conference on Very Large Databases*, Barcelona, Spain, August 1991.
143. A simple characterization of uniform boundedness for a class of recursions. (With Yehoshua Sagiv.) In *Journal of Logic Programming*, November 1991, pp. 233 - 254.
144. A counting method for a cyclic binary query. (With Ramsey W. Haddad.) *Journal of Computer and System Sciences*, 43(1):145–169, August 1991. (Extended version of conference paper 163.)
145. One-sided recursions. *Journal of Computer and System Sciences*, 42(2):199–236, April 1991. (Extended version of conference paper 164.)
146. On the expected size of recursive datalog queries. (With S. Seshadri.) In *Proceedings of the ACM Symposium on Principles of Database Systems*, Denver, Colorado, May 1991.
147. A stochastic approach for clustering in object bases. (With Manolis M. Tsangaris.) In *Proceedings of the ACM SIGMOD Symposium on Management of Data*, Denver, Colorado, May 1991.
148. Space optimization in the bottom-up evaluation of logic programs. (With S. Sudarshan, Divesh Srivastava, and Raghu Ramakrishnan.) In *Proceedings of the ACM SIGMOD Symposium on Management of Data*, Denver, Colorado, May 1991. .
149. On estimating the size of projections. (With S. Seshadri.) In *Proceedings of the International Conference on Database Theory*, Paris, France, December 1990.
150. A unified approach to logic program evaluation. (With Raghu Ramakrishnan.) In *Computational Logic, Essays in Honor of Alan Robinson*, Jean-Louis Lassez and G. Plotkin editors, MIT Press, 1991.

151. How to forget the past without repeating it. (With Raghu Ramakrishnan.) In *Proceedings of the Sixteenth International Conference on Very Large Databases*, Pages 278–289, Brisbane, Australia, August 1990.
152. Practical selectivity estimation. (With Richard J. Lipton and Donovan A. Schneider.) In *Proceedings of the ACM SIGMOD Symposium on Management of Data*, Atlantic City, New Jersey, May 1990. Awarded best paper of conference.
153. Query size estimation through random sampling. (With Dick Lipton.) In *Proceedings of the ACM Symposium on Principles of Database Systems*, April 1990, Nashville, Tennessee.
154. A real-time, concurrent checkpoint and recovery algorithm for parallel programs. With (Kai Li and James S. Plank.) In *Proceedings of the 1990 ACM Symposium on Principles and Practice of Parallel Programming*, Seattle, Washington, March, 1990.
155. Argument reduction through factoring. (With Raghu Ramakrishnan, Yehoshua Sagiv, and Jeffrey D. Ullman.) In *Proceedings of the Fifteenth International Conference on Very Large Databases*, pages 173–182, Amsterdam, The Netherlands, August 1989.
156. Data independent recursion in deductive databases. *Journal of Computer and System Sciences*, 38(2):259–289, April 1989. (Extended version of conference paper 167.)
157. Estimating the size of generalized transitive closures. (With Richard J. Lipton.) In *Proceedings of the Fifteenth International Conference on Very Large Databases*, pages 165–172, Amsterdam, The Netherlands, August 1989.
158. Efficient evaluation of right- and left-linear predicates. (With Raghu Ramakrishnan, Yehoshua Sagiv, and Jeffrey D. Ullman.) In *Proceedings of the ACM SIGMOD International Conference on Management of Data*, pages 235–242, Portland, Oregon, May 1989.
159. Minimizing expansions of recursions. (With Yehoshua Sagiv.) In Hasan Ait-Kaci and Maurice Nivat, editors, *Resolution of Equations in Algebraic Structures*, volume 1, pages 321–349, San Diego, California, 1989. Academic Press, Inc.
160. Minimizing function-free recursive inference rules. *Journal of the ACM*, 36(1):69–91, January, 1989.
161. Multiprocessor main memory transaction processing. (With Kai Li.) In *Proceedings of the International Symposium on Databases in Parallel and Distributed Systems*, December 1988. Also Princeton TR CS-TR-159-88.
162. Compiling separable recursions. In *Proceedings of the SIGMOD International Symposium on Management of Data*, pages 312–319, Chicago, Illinois, May 1988.
163. Counting methods for cyclic relations (With Ramsey W. Haddad.) In *Proceedings of the ACM Symposium on Principles of Database Systems*, pages 333–340, Austin, Texas, March 1988.
164. One-sided recursions. In *Proceedings of the ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*, pages 340–348, San Diego, California, March, 1987.
165. A decidable class of bounded recursions. (With Yehoshua Sagiv.) In *Proceedings of the ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*, pages 227–236, San Diego, California, March 1987.
166. Redundancy in function-free recursive inference rules. In *Proceedings of the IEEE Symposium on Logic Programming*, pages 236–245, Salt Lake City, Utah, September, 1986.

167. Data independent recursion in deductive databases. *Proceedings of the ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*, pages 267–279, Cambridge, MA, March, 1986.