Brian F. Cooper

Office
801 Atlantic Drive
Atlanta, GA 30332
(404) 385-2836

Internet
www.cc.gatech.edu/~cooperb
cooperb@cc.gatech.edu

RESEARCH INTERESTS

Distributed information systems, middleware systems, text search, event and data stream processing

EDUCATION

Ph.D. in Computer Science, 2003
Stanford University, Stanford, California
Dissertation: Information Preservation in Networks of Autonomous Archives
Advisor: Professor Hector Garcia-Molina

M.S. in Computer Science, 2000
Stanford University, Stanford, California

B.S. in Computer Science (highest honors), 1998
B.A. in Chemistry, 1998
University of Colorado, Boulder, Colorado

RESEARCH EXPERIENCE

Georgia Institute of Technology
Division of Computing Sciences and Systems, College of Computing
Assistant Professor, August 2003-present
Directing an active research program in distributed data and information systems involving students and faculty collaborators. Projects include:

- *Overlay Dynamic Information Networks (ODIN)* – Techniques for automatically adapting the structure of information network overlays to changing workloads, environments and requirements to result in higher performance
- *IFLOW/inTransit* – Techniques for distributed data stream processing, including resource-aware deployment of processing tasks, automatic adaptation of deployment to respond to resource changes or failures, and responsiveness of the system to business concerns and workloads
- *InfoBeacons* – Techniques for routing user text searches through large networks to information sources, even when the sources are dynamic and uncooperative

Stanford University
Department of Computer Science
Research Assistant, September 1998- August 2003
Conducted research in the archival repository project. Contributions include:

- *Stanford Archival Vault (SAV)* – Designed and implemented a reliable archival database system based on remote replication and repeated content checking. Implemented an ingest tool, *InfoMonitor*, which automatically scanned a filesystem for updated content
- *Data Trading* – Designed and tested algorithms for trading content between distributed, networked data archives; trading efficiently allocates resources to replicas to ensure fault tolerance
- *SIL model* – Developed a model of search overlays to allow resilient and efficient searching of content in scattered repositories

IBM Almaden Research Center
Database Technology Institute
Research Intern, July-September 2002
Examined integration of SOAP request capability into DB2 database engine:

- Implemented features and enhanced stability for SOAP requestor user defined function
- Implemented SOAP requestor as a built in function of the DB2 engine
- Conducted performance studies of SOAP request/response and identified optimizations for latency reduction

RightOrder, Inc.
Enterprise software corporation, San Jose, California
Research Engineer, September 2000-June 2002
Developed algorithms for efficient indexing and query processing over XML data, using trie-based path indexing

IBM T.J. Watson Research Center
Programming Technologies Department
Research Intern, May-August 1998
Developed program profiling techniques for an adaptive Java compiler

- Designed and implemented run-time program profiler for an adaptive Java compiler
- Implemented program profile data structure based on calling context tree

University of Colorado
Department of Computer Science
Undergraduate Researcher, September 1997-May 1998
Developed Java runtime profiling tools for the Compiler Optimization Group:

- Developed ProfBuilder, a package for rapid construction of runtime Java program profilers
- Designed and implemented program analysis tools based on dynamic construction of calling context tree and control flow graph structures

SERVICE

- Associate Editor, Conference and Workshop Reports, ACM SIGMOD Record (2004-present)
- Editor, ACM SIGMOD DiSC Electronic Literature Collection (2003-2006)
- Program committee member of SIGMOD 2006, ICDE 2006, ICDCS 2006, and 14 other conferences and workshops
- Co-program chair of the IEEE Workshop on Workflow and Data Flow for Scientific Applications (SciFlow 2006) in conjunction with ICDE
- Local arrangements chair, ICDE 2006
- Conference session chair at SIGMOD 2006, ICDE 2006, NetDB workshop 2006, ICDCS 2005
- Reviewer for ACM Computing Surveys, ACM/IEEE Transactions on Networking, IEEE Transactions on Parallel and Distributed Systems, Software Practice and Experience, Information Systems, IEEE Computer, IEEE Internet Computing, and 7 other journals

PUBLICATIONS

Papers are available at http://www.cc.gatech.edu/~cooperb/pubs/

Journal publications

1. Brian F. Cooper, Neal Sample, Michael J. Franklin, Joshua Olshansky and Moshe Shadmon. *Middle-Tier Extensible Data Management*. World Wide Web Journal (Kluwer), Volume 4, issue 3, 2001.
2. Brian F. Cooper and Hector Garcia-Molina. *Peer to peer data trading to preserve information*. ACM Transactions on Information Systems 20(2), pp. 133-170, April 2002.
3. Brian F. Cooper and Hector Garcia-Molina. *Peer-to-peer data preservation through storage auctions*. IEEE Transactions on Parallel and Distributed Systems 16(3), pp. 246-257, March 2005.
4. Brian F. Cooper and Hector Garcia-Molina. *Ad hoc, self-supervising peer-to-peer search networks.* ACM Transactions on Information Systems, 23(2), pp. 169-200, April 2005.
5. Brian F. Cooper and Hector Garcia-Molina. *InfoMonitor: Unobtrusively archiving a World Wide Web server*. International Journal on Digital Libraries, 5(2), pp. 106-119, April 2005.
6. Brian F. Cooper and Hector Garcia-Molina. *SIL: A model for analyzing scalable peer-to-peer search networks.* Journal of Computer Networks (Elsevier), to appear.
7. Umakishore Ramachandran, Rajnish Kumar, Matthew Wolenetz, Brian F. Cooper, Bikash Agarwalla, Junsuk Shin, Phillip W. Hutto and Arnab Paul. *Dynamic Data Fusion for Future Sensor Networks.* ACM Transactions on Sensor Networks, to appear.
8. Vibhore Kumar, Brian F. Cooper, Zhongtang Cai, Greg Eisenhauer, Karsten Schwan. *Middleware for Enterprise Scale Data Stream Management using Utility-Driven Self-Adaptive Information Flows.* Cluster Computing Journal (Springer Publishing), to appear.

Book chapters

1. Karsten Schwan, Brian F. Cooper, Greg Eisenhauer, Ada Gavrilovska, Matt Wolf, Hasan Abbasi, Sandip Agarwala, Zhongtang Cai, Vibhore Kumar, Jay Lofstead, Mohamed Mansour, Balasubramanian Seshasayee, and Patrick Widener. *AutoFlow: Autonomic Information Flows for Critical Information Systems.* Autonomic Computing: Concepts, Infrastructure, and Applications, ed. Manish Parashar and Salim Hariri, CRC Press, 2006.

Conference and workshop publications

1. Brian F. Cooper, Arturo Crespo and Hector Garcia-Molina. *Implementing a Reliable Digital Object Archive*. European Conference on Digital Libraries 2000.
2. Brian F. Cooper, Neal Sample, Michael J. Franklin, Gisli R. Hjaltason and Moshe Shadmon. *A fast index for semistructured data*. VLDB 2001. (Acceptance rate: 17%)
3. Brian F. Cooper and Hector Garcia-Molina. *Creating trading networks of digital archives*. Joint ACM/IEEE Conference on Digital Libraries, 2001. (Acceptance rate: 16%)
4. Brian F. Cooper, Neal Sample, Michael J. Franklin, Joshua Olshansky and Moshe Shadmon. *Extensible data management in the middle tier.* Research Issues in Data Engineering Workshop (in conjunction with ICDE) 2002.
5. Brian F. Cooper, Neal Sample and Moshe Shadmon. *A parallel index for semistructured data*. ACM Symposium on Applied Computing 2002.
6. Brian F. Cooper and Hector Garcia-Molina. *Bidding for storage space in a peer-to-peer data preservation system*. International Conference on Distributed

Computing Systems 2002. (Acceptance rate: 18%)

7. Brian F. Cooper and Hector Garcia-Molina. *Peer-to-peer resource trading in a reliable distributed system*. 1st International Workshop on Peer-to-Peer Systems (IPTPS), 2002.
8. Brian F. Cooper and Hector Garcia-Molina. *Studying search networks with SIL*. 2nd International Workshop on Peer-to-Peer Systems (IPTPS), 2003. (Acceptance rate: 16%)
9. Brian F. Cooper and Hector Garcia-Molina. *SIL: Modeling and measuring scalable peer-to-peer search networks* in the International Workshop on Databases, Information Systems and Peer-to-Peer Computing (in conjunction with VLDB), Berlin, 2003.
10. Brian F. Cooper. *A content model for evaluating peer-to-peer searching techniques.* ACM/IFIP/USENIX 5th International Middleware Conference, Toronto, 2004. (Acceptance rate: 14%)
11. Brian F. Cooper. *Guiding queries to information sources with InfoBeacons.* ACM/IFIP/USENIX 5th International Middleware Conference, Toronto, 2004. (Acceptance rate: 14%)
12. Brian F. Cooper. *Using information retrieval techniques to route queries in an InfoBeacons network.* Second International Workshop On Databases, Information Systems and Peer-to-Peer Computing (in conjunction with VLDB), Toronto, 2004.
13. Vibhore Kumar, Brian F. Cooper and Shamkant B. Navathe. *Predictive Filtering: A Learning-Based Approach to Data Stream Filtering.* International Workshop on Data Management for Sensor Networks (DMSN), Toronto, 2004.
14. Vibhore Kumar, Brian F. Cooper and Karsten Schwan. *Distributed Stream Management using Utility-Driven Self-Adaptive Middleware.* 2nd IEEE International Conference on Autonomic Computing (ICAC), 2005. **Best student paper award** (Acceptance rate: 21%)
15. Vibhore Kumar, Brian F. Cooper, Zhongtang Cai, Greg Eisenhauer and Karsten Schwan. *Resource-Aware Distributed Stream Management using Dynamic Overlays.* 25th IEEE International Conference on Distributed Computing Systems (ICDCS), 2005. (Acceptance rate: 14%)
16. Brian F. Cooper. *An optimal overlay topology for routing peer-to-peer searches.* ACM/IFIP/USENIX 6th International Middleware Conference, Grenoble, 2005. (Acceptance rate: 17%)
17. Brian F. Cooper. *Quickly routing searches without having to move content.* 4th International Workshop on Peer-to-Peer Systems (IPTPS), 2005. (Acceptance rate: 20%)
18. David Roberts, Sooraj Bhat, Charles Isbell, Brian F. Cooper and Jeff Pierce. *A Decision-Theoretic Approach to File Consistency in Constrained Peer-to-Peer Device Networks (short paper).* Fifth International Joint Conference On Autonomous Agents and Multiagent Systems, 2006.
19. Yong Yang, Rocky Dunlap, Michael Rexroad and Brian F. Cooper. *Performance of Full Text Search in Structured and Unstructured Peer-to-Peer Systems*. IEEE INFOCOM, Barcelona, 2006. (Acceptance rate: 18%)
20. Vibhore Kumar, Zhongtang Cai, Brian F. Cooper, Greg Eisenhauer, Karsten Schwan, Mohamed Mansour, Balasubramanian Seshasayee, Patrick Widener. *Implementing Diverse Messaging Models with Self-Managing Properties using IFLOW*. 3rd IEEE International Conference on Autonomic Computing (ICAC), 2006. (Acceptance rate: 21%)
21. Sangeetha Seshadri, Vibhore Kumar and Brian F. Cooper. *Optimizing Multiple Queries in Distributed Data Stream Systems.* 2nd IEEE International Workshop on Networking Meets Database (NetDB), in conjunction with ICDE 2006.

22. Zhongtang Cai, Vibhore Kumar, Brian F. Cooper, Greg Eisenhauer, Karsten Schwan and Robert E. Strom. *Utility-Driven Proactive Management of Availability in Enterprise-Scale Information Flows.* ACM/IFIP/USENIX 7th International Middleware Conference, 2006. (Acceptance rate: 17%)
23. Brian F. Cooper. *Trading off Resources Between Overlapping Overlays.* ACM/IFIP/USENIX 7th International Middleware Conference, 2006. (Acceptance rate: 17%)

Submitted for publication

1. Sangeetha Seshadri and Brian F. Cooper. *Routing Queries Through a Peer-to-Peer InfoBeacons Network Using Information Retrieval Techniques.* Submitted to IEEE Transactions on Parallel and Distributed Systems, January 2006; returned for revisions May 2006; currently under revision.
2. Sangeetha Seshadri, Vibhore Kumar and Brian F. Cooper. *Using Hierarchies for Optimizing Multiple Distributed Stream Queries.* Submitted to the International Conference on Data Engineering (ICDE), 2007.

Invited papers

1. Brian F. Cooper, Arturo Crespo and Hector Garcia-Molina. *The Stanford Archival Repository Project: Preserving our digital past*. Invited paper for Library and Information Research News, 26(84): pp. 17-26.
2. Mayank Bawa, Brian F. Cooper, Arturo Crespo, Neil Daswani, Prasanna Ganesan, Hector Garcia-Molina, Sepandar Kamvar, Sergio Marti, Mario Schlosser, Qi Sun, Patrick Vinograd and Beverly Yang. *Peer-to-peer research at Stanford*. Invited paper for SIGMOD Record, September 2003.
3. Brian F. Cooper, Mayank Bawa, Neil Daswani, Sergio Marti and Hector Garcia-Molina. *Authenticity and availability in PIPE networks.* Invited paper for Future Generation Computer Systems, 21(3): 391-400, March 2005.