

Andrew Pavlo

Brown University
Department of Computer Science
Box 1910, 115 Waterman St.
Providence, RI 02912 USA

Voice: +1-443-242-4047
E-mail: pavlo@cs.brown.edu
Web: <http://www.cs.brown.edu/~pavlo/>
Twitter: @andy_pavlo

EDUCATION

Ph.D., Computer Science

Brown University

- Thesis: On the Optimization of Distributed Database Management Systems
- Advisor: Stanley Zdonik

Spring 2013 (*expected*)
Providence, RI USA

M.Sc., Computer Science

Brown University

- Thesis: Automatic Database Partitioning in Shared-Nothing, Parallel OLTP Systems
- Advisor: Stanley Zdonik

Spring 2009
Providence, RI USA

M.Sc., Computer Science

Rochester Institute of Technology

- Thesis: Interactive, Tree-based Graph Visualization
- Advisor: Christopher M. Homan

Spring 2006
Rochester, NY USA

B.Sc. with Honors, Computer Science

Rochester Institute of Technology

- Minors in Mathematics and Political Science

Spring 2005
Rochester, NY USA

RESEARCH EXPERIENCE

Graduate Research Assistant

Brown University, Department of Computer Science

September 2007 – present
Providence, RI USA

- **H-Store** – <http://hstore.cs.brown.edu>

Worked with a team of students at Brown, MIT, and Yale to develop an experimental, distributed main memory database management system. H-Store is designed to execute on-line transaction processing (OLTP) workloads efficiently on a shared-nothing cluster with minimal concurrency control. H-Store's design is being commercialized as **VoltDB**. As part of my dissertation, I have worked on several research projects based on the system:

Horticulture: Automatic database design tool that generates the best partitioning, replication, transaction routing configuration on H-Store for applications using a large-neighborhood search algorithm. The designs produced by this tool minimize the number of distributed transactions in the system while also mitigating the effects of skew. This work was in collaboration with Yahoo! Research and was published in SIGMOD.

Houdini: Integrated optimization framework that uses Markov models to predict at run time what operations transactions will execute before they start running. These models are derived from sample workload traces of previously executed transactions. Using these predictions, H-Store locks the minimum number of partitions needed by each transaction and then determines when the transaction is done with them. This work was in collaboration with MIT and was published in VLDB.

Hermes: Lock-free concurrency control scheme where the DBMS speculatively executes single-partition transactions any time that it detects that a distributed transaction is stalled in the system. Hermes uses Houdini's predictions to determine whether queued transactions interfere with a stalled transaction and then schedules the non-conflicting ones to run before that transaction resumes. This allows the DBMS to safely interleave transactions without using locks or snapshot isolation to ensure serializability.

AuctionMark/SEATS: Created two OLTP DBMS benchmarks based on the workloads of real-world applications. These benchmarks have been ported to the OLTP-Bench framework.

- **MongoDB-D4** – <http://database.cs.brown.edu/projects/mongodb>

Collaborated with students and engineers at 10gen, Inc. to develop an automatic database design tool for dis-

tributed document-oriented DBMSs. D4 can select the optimal sharding keys, index keys, and denormalization schemes for an application that minimizes the amount of network communication, disk operations, and workload skew in the system. It uses a large-neighborhood search algorithm to iteratively construct designs and then estimates the DBMS's performance using an analytical cost model that approximates its working set memory.

- **OLTP-Bench** – <http://oltpbenchmark.com>
Lead developer for a full-featured DBMS testbed that is specifically designed for transaction processing and Web-based applications. This project was created by merging the benchmark frameworks from H-Store and the Relational Cloud project at MIT. It includes two synthetic workloads, five workloads derived from popular benchmarks, and three workloads generated from real-world data sets. This work is in collaboration with the University of Fribourg and Microsoft. A research paper about the project is currently under submission.
- **MapReduce vs. DBMSs** – <http://database.cs.brown.edu/projects/mapreduce-vs-dbms>
Worked with a team of researchers at MIT, Yale University, the University of Wisconsin, and Microsoft to evaluate different programming models for executing analytical workloads on massive-scale data sets. We developed an open-source benchmark to measure the performance of Hadoop, Vertica, and DB2 for common data processing tasks. This work was published in both SIGMOD and the CACM. As of December 2012, these papers are cited by almost 450 scholarly publications.
- **Graffiti Networks** – <http://graffiti.cs.brown.edu>
Developed a file sharing system that uses abandoned web sites to store data. A tracker maintains an index of the data's location on these sites and can instruct clients on how to retrieve the pieces to reconstruct the original file. Extended BitTorrent protocol to allow clients to simultaneously participate in both torrent and graffiti swarms.

Research Intern

Vertica, Inc.

June 2008 – August 2008

Billerica, MA USA

- Part of a team that developed prototype of the H-Store/VoltDB OLTP database system presented at VLDB.
- Built H-Store's front-end query planning infrastructure, including internal catalog representation, distributed query planner, and a heuristic-based query plan optimizer
- Developed back-end C++ execution engine components, including iterator-style query plan executors, the initial storage manager, abstract expression tree evaluators, and read/write data set managers.

Systems Programmer

University of Wisconsin–Madison, Department of Computer Sciences

September 2005 – August 2007

Madison, WI USA

- **HTCondor Batch System** – <http://research.cs.wisc.edu/htcondor>
Implemented ability for to execute jobs at specific times similar to Unix's cron command. Added support in scheduling daemons to account for time synchronization differences of resources across multiple administrative domains. Assisted in porting HTCondor to BSD-based operating systems. Worked with researchers from the CMS team at CERN to deploy HTCondor clusters on global grid networks.
- **Metronome Framework** – <https://nmi.cs.wisc.edu>
Developed job migration mechanisms to allow jobs to automatically execute outside of local administrative domains based on resource availability. Improved job scheduling and workflow execution pipelines. Implemented and deployed new PHP/MySQL-based Web interface. Helped integrate Metronome with the European Commission's ETICS project at CERN.
- **Virtual Data Toolkit** – <http://vdt.cs.wisc.edu>
Create automatic configuration tools to allow users to easily deploy scientific data processing software on grid computing resources, such as the Open Science Grid and TeraGrid.

PROFESSIONAL EXPERIENCE

Non-Testifying Technical Expert

Choate, Hall & Stewart, LLP.

June 2009 – November 2011

Boston, MA USA

- **StreamServe, Inc. v. Hewlett-Packard Company** (Patent IP)
Reviewed technical documents for HP's Exstream document processing platform. Analyzed C/C++ source code in Exstream in order to understand the database-related components alleged to be violating the patent in question. Prepared presentations to assist counsel with understanding these core technical aspects of HP's system. Helped draft expert witness reports. The case was settled out of court.

- **Content Management Associates, Inc. v. Fidelity Investments** (Trade Secret IP)
Served as initial technical expert for defendant on trade secret litigation. Analyzed documentation and database-driven VB.NET source code provided by CMS, Inc. to assist counsel with identifying components of the system that correspond to alleged trade secrets. Prepared presentations to assist counsel with understanding core technical aspects of systems in question. Participated in a marathon one-day, 12 hour deposition. The case was settled out of court.
- **AvePoint, Inc. v. Janalent, Inc.** (Trade Secret IP)
Withheld pending outcome of litigation. Motion for summary judgment filed in Fall 2011.

Software Developer
PAID, Inc.

March 1999 – December 2004
Owings Mills, MD USA

- Worked on a large-scale, PHP/MySQL-based auction processing web site. Wrote larger portions of the shipping cost calculation framework, embeddable checkout plug-ins, ORM library, and store management infrastructure.
- Developed J2EE/EJB components to integrate auction payment processing platform with the eBay API and PayPal's Instant Payment Notification APIs.

TEACHING EXPERIENCE

Co-Advised Undergraduates

- Emanuel Buzek (B.Sc. 2012, Cloudera)
- Saurya Velagapudi (B.Sc. 2010, Google)

Co-Advised Master's Students

- Charles J. Lee (B.Sc./M.Sc. 2013, *Expected*)
- Xin Jia (M.Sc. 2013, *Expected*)
- Yang Zou (M.Sc. 2012, Microsoft)
- Yang Lu (M.Sc. 2012, Amazon)
- Visawee Angkanawaraphan (M.Sc. 2011, Amazon)
- Zhe Zhang (M.Sc. 2010, TripAdvisor)
- Ning Shi (M.Sc. 2010, VoltDB)

Courses Taught

- **NewSQL – The Course** (Brown University – CSC2270), Spring 2012
Instructor: Stanley Zdonik
Planned entire syllabus and reading schedule for seminar course on modern distributed OLTP database systems. Developed project plans for students based on the H-Store system. Met with students weekly to discuss their progress on the projects and in-class presentations. Scheduled speakers from NewSQL database companies.
- **Non-standard, Web-scale Databases** (Brown University – CSC2270), Spring 2011
Instructor: Stanley Zdonik
Developed graduate seminar course on NoSQL systems. Created a Python-based framework for students to implement NoSQL-based drivers for a non-transactional variant of the TPC-C benchmark.

Graduate Teaching Assistant

- **Advanced Topics in Database Management** (Brown University – CSC2270), Spring 2009
- **Introduction to Database Management Systems** (Brown University – CSC1270), Fall 2010
- **Introduction to Database Management Systems** (Brown University – CSC1270), Fall 2008

AWARDS & HONORS

- Amazon Research Grant, Winter 2012
- ACM SIGMOD Student Travel Grant, Spring 2010
- Class Clown, Mt. Hebron High School, Spring 1999

ACADEMIC SERVICE

- External Reviewer – Springer Journal of Distributed and Parallel Databases, January 2013
- External Reviewer – IBM Journal of Research and Development, July 2012
- Faculty Search Committee – Brown University, February 2011
- Demo Judge – SIGMOD, June 2010
- Graduate Student Recruitment Committee – Brown University, February 2010
- Demo Judge – SIGMOD, June 2009

PUBLICATIONS

- Andrew Pavlo, Carlo Curino, and Stanley Zdonik. Skew-Aware Automatic Database Partitioning in Shared-Nothing, Parallel OLTP Systems. In *SIGMOD*, pages 61–72, 2012.
- Carlo A. Curino, Djellel E. Difallah, Andrew Pavlo, and Philippe Cudre-Mauroux. Benchmarking OLTP/Web Databases in the Cloud: The OLTP-bench Framework. *CloudDB '12*, pages 17–20, 2012.
- Andrew Pavlo, Evan P.C. Jones, and Stan Zdonik. On Predictive Modeling for Optimizing Transaction Execution in Parallel OLTP Systems. *Proc. VLDB Endow.*, 5:85–96, October 2011.
- Michael Stonebraker, Daniel Abadi, David J. DeWitt, Sam Madden, Erik Paulson, Andrew Pavlo, and Alexander Rasin. MapReduce and Parallel DBMSs: Friends or Foes? *Communications of the ACM*, 53(1):64–71, 2010.
- Andrew Pavlo, Erik Paulson, Alexander Rasin, Daniel J. Abadi, David J. DeWitt, Samuel Madden, and Michael Stonebraker. A Comparison of Approaches to Large-Scale Data Analysis. In *SIGMOD*, pages 165–178, 2009.
- Robert Kallman, Hideaki Kimura, Jonathan Natkins, Andrew Pavlo, Alexander Rasin, Stanley Zdonik, Evan P. C. Jones, Samuel Madden, Michael Stonebraker, Yang Zhang, John Hugg, and Daniel J. Abadi. H-Store: A High-Performance, Distributed Main Memory Transaction Processing System. *Proc. VLDB Endow.*, 1(2):1496–1499, 2008.
- Ewa Deelman, Miron Livny, Gaurang Mehta, Andrew Pavlo, Gurmeet Singh, Mei-Hui Su, Karan Vahi, and R. Kent Wenger. Pegasus and DAGMan from concept to execution: Mapping scientific workflows onto today’s cyberinfrastructure. In *High Performance Computing (HPC) and Grids in Action*, volume 16 of *Advances in Parallel Computing*, pages 56–74. 2008.
- Andrew Pavlo, Peter Couvares, Rebekah Gietzel, Anatoly Karp, Ian D. Alderman, Miron Livny, and Charles Bacon. The NMI Build & Test Laboratory: Continuous Integration Framework for Distributed Computing Software. In *LISA*, pages 263–273. USENIX, 2006.
- Christopher Homan, Andrew Pavlo, and Jonathan Schull. Smoother transitions between breadth-first-spanning-tree-based drawings. In *14th International Symposium on Graph Drawing*, volume 4372 of *Lecture Notes in Computer Science*, pages 442–445, 2006.

UNDER PREPARATION/SUBMISSION

- Djellel E. Difallah, Andrew Pavlo, Carlo Curino, and Philippe Cudre-Mauroux. OLTP-Bench: An Extensible Testbed for Benchmarking Transactional Databases in the Cloud. November 2012.
- Andrew Pavlo, Xin Jia, Charles J. Lee, and Stanley Zdonik. The Art of Speculative Execution. December 2012.
- Andrew Pavlo, Yang Zou, Stanley Zdonik, Dan Pasette, Spencer T. Brody, and Eliot Horowitz. Distributed Document Database Design. December 2012.
- Justin DeBrabant, Andrew Pavlo, Stephen Tu, Michael Stonebraker, Stanley Zdonik, and Donghui Zhang. Anti-Caching in Main Memory Database Management Systems. December 2012.

MISCELLANEOUS WRITINGS

- Xin Jia, Andrew Pavlo, and Stanley Zdonik. Tastes Great, Less Filling: Low-Impact OLAP MapReduce Queries on High-Performance OLTP Systems. *Tiny Transactions on Computer Science*, vol 1, 2012.

- Andrew Pavlo and Ning Shi. Graffiti Networks: A Subversive, Internet-Scale File Sharing Model. *CoRR*, abs/1101.0350, 2011.
- Andrew Pavlo, Christopher Homan, and Jonathan Schull. A Parent-centered Radial Layout Algorithm for Interactive Graph Visualization and Animation. *CoRR*, abs/cs/0606007, 2006.
- Andrew Pavlo. Interactive, Tree-based Graph Visualization. Master’s thesis, Rochester Institute of Technology, 2006.

INVITED TALKS

- “On Predictive Modeling for Distributed Databases,” *VLDB*, August 2012.
- “Automatic Database Partitioning in Parallel OLTP Systems,” *SIGMOD*, May 2012.
- “Making Fast Databases Faster,” *Columbia University*, April 2012.
- “Making Fast Databases Faster,” *Yale University*, April 2012.
- “Methods for the Optimization of Parallel OLTP Systems,” *VoltDB, Inc.*, April 2012.
- “Making Fast Databases Faster,” *New England Database Summit 2012* – Massachusetts Institute of Technology, February 2012.
- “Magical Parallel OLTP Databases,” *Massachusetts Institute of Technology*, November 2011.
- “Life After the Stonebraker Stores,” *International Workshop on High Performance Transaction Systems*, October 2011.
- “MapReduce and Parallel DBMSs: Together At Last,” *IPP Symposium on Cloud Computing* – Brown University, May 2010.
- “MapReduce and Parallel DBMSs: Together At Last,” *New England Database Summit 2010* – Massachusetts Institute of Technology, January 2010.
- “H-Store: A Specialized Architecture for High-throughput OLTP Applications,” *International Workshop on High Performance Transaction Systems*, October 2009.
- “Graffiti Networks: A Subversive, Internet-Scale File Sharing Model,” *DC401 - Rhode Island Defcon Group*, October 2009.
- “MapReduce and Parallel DBMSs: A Comparison of Approaches to Large-Scale Data Analysis,” *University of Maryland College Park*, September 2009.
- “Challenges in Dynamic Deployment of Condor Across Distributed Environments,” *University of Wisconsin–Madison*, May 2007.
- “NMI Build & Test Laboratory: Continuous Integration Framework for Distributed Computing Software,” *USENIX LISA*, December 2006.

REFERENCES

Stanley Zdonik

Professor of Computer Science
Brown University
Box 1910, 115 Waterman St.
Providence, RI 02912 USA
E-mail: sbz@cs.brown.edu

Ugur Cetintemel

Professor of Computer Science
Brown University
Box 1910, 115 Waterman St.
Providence, RI 02912 USA
E-mail: ugur@cs.brown.edu

Michael Stonebraker

Adjunct Professor of Computer Science
Massachusetts Institute of Technology
32 Vassar St.
Cambridge, MA 02139 USA
E-mail: stonebraker@csail.mit.edu

Samuel Madden

Professor of Computer Science
Massachusetts Institute of Technology
32 Vassar St.
Cambridge, MA 02139 USA
E-mail: madden@csail.mit.edu