

Daniel J. Abadi

Office: Iribe Center for Computer Science and Engineering 5156
8125 Paint Branch Drive
College Park, MD 20742

abadi@umd.edu

www.cs.umd.edu/~abadi/

BIOGRAPHICAL INFORMATION

Name: Daniel Abadi

Rank: Darnell-Kanal Professor of Computer Science at University of Maryland, College Park

Education:

Massachusetts Institute of Technology, Cambridge, MA

PhD in Electrical Engineering and Computer Science, February 2008

Cambridge University, Cambridge, England

M.Phil. in Computer Speech Text and Internet Technology, July 2003

Brandeis University, Waltham, MA

B.S. in Computer Science (double-majored in Neuroscience), May 2002

GPA: 3.98 (out of 4.0)

Class Rank: 1 out of 816

Summary of professional career: Daniel Abadi's primary areas of research include database system architecture, implementation, and scalability. He spent 9 years on the computer science faculty at Yale University (2008-2017) before moving to his current position at UMD. He is best known for contributions to the following projects:

[Notable Completed Projects]

HadoopDB is an extremely scalable parallel database system designed specifically for doing complex analytics over structured (relational) and unstructured data. It is architected to achieve extremely high performance and efficiency, while still yielding robust fault tolerance and query flexibility. HadoopDB is designed to run on a shared-nothing cluster of commodity machines, or in the cloud, and won the VLDB "Test of Time" award in 2019. This project was commercialized by Hadapt, which was acquired in 2014 by Teradata.

C-Store used a novel column-oriented database design to outperform traditional database systems by two orders of magnitude on read-mostly workloads. Abadi's work on the central query execution engine for C-Store resulted in the 2008 SIGMOD Jim Gray Doctoral Dissertation Award and the VLDB 10-Year Best Paper Award. The C-Store technology has been used (through the Vertica commercialization) by thousands of businesses across the world, including Facebook (before Meta), Uber, Wayfair, Intuit, New York Times, edX, and Chevron.

H-Store investigated how modern architectural and application shifts affect the performance of database systems and proposed a new architecture for high throughput, in-memory transaction processing. This work was commercialized by VoltDB which has raised over \$41 million in venture capital to this date.

Calvin is a distributed database system designed for scalable transactional (OLTP) processing. It uses deadlock avoidance techniques and new concurrency control schemes in order to remove all nondeterminism from typical OLTP workloads, which enables active replication of database systems with no synchronization overhead whatsoever. Further, Calvin eliminates the requirement for two-phase commit (2PC) for any kind of distributed transaction (removing 2PC significantly improves scalability). Calvin is therefore able to achieve much higher transactional throughput than modern distributed database systems.

Aurora and Borealis were large multi-university projects (Brandeis, Brown, MIT) focused on data stream management. They were among the earliest systems that supported large scale continuous and historical queries over streaming data.

Professional honors and recognition:

CIDR Test of Time Award (2025)

ACM Fellow (2020)

VLDB Test of Time Award (2019)

Dissertation advisor of 2020 SIGMOD Dissertation Award Winner (Jose Faleiro)

VLDB 10 Year Best Paper Award (2015)
Dissertation advisor of 2014 SIGMOD Dissertation Award Winner (Alexander Thomson)
2013-2014 Yale Provost's Teaching Prize
VLDB Early Career Researcher Award (2013)
Sloan Fellowship (2011)
IBM Scalable Data Analytics for a Smarter Planet Innovation award (2011)
National Science Foundation (NSF) Career Award (2009)
SIGMOD Jim Gray Doctoral Dissertation Award (2008)
VLDB Best Paper Award (2007)
Conference on Innovative Data Systems Research (CIDR) Best Poster Presentation Award (2005)

TEACHING:

CMSC 424 (Database Design)
 CMSC 624 (Database System Architecture and Implementation)
 CMSC 828N (Database System Architecture and Implementation)

PUBLICATIONS (each publication listed is a fully refereed research article):

- [54] **Are Database System Researchers Making Correct Assumptions about Transaction Workloads?** Cuong D. T. Nguyen, Kevin Chen, Christopher DeCarolis, Daniel J. Abadi. *In Proc. ACM Manag. Data* 3(3): 131:1-131:26 (2025).
- [53] **AutoComp: Automated Data Compaction for Log-Structured Tables in Data Lakes.** Anja Gruenheid, Jesús Camacho-Rodríguez, Carlo Curino, Raghu Ramakrishnan, Stanislav Pak, Sumedh Sakdeo, Lenisha Gandhi, Sandeep K. Singhal, Pooja Nilangekar, Daniel J. Abadi. *SIGMOD Conference Companion 2025*: 404-417.
- [52] **Bullion: A Column Store for Machine Learning.** Gang Liao, Ye Liu, Jianjun Chen, Daniel J. Abadi. *Proceedings of 15th Conference on Innovative Data Systems Research (CIDR)*, 2025.
- [51] **FileScale: Fast and Elastic Metadata Management for Distributed File Systems.** Gang Liao and Daniel J. Abadi. *In Proceedings of SoCC 2023*.
- [50] **Detock: High Performance Multi-region Transactions at Scale** Cuong D. T. Nguyen, Johann K. Miller, Daniel J. Abadi. *In Proceedings of SIGMOD*, 2023.
- [49] **C5: Cloned Concurrency Control That Always Keeps Up** Jeffrey Helt, Abhinav Sharma, Daniel J. Abadi, Wyatt Lloyd, Jose M. Faleiro. *PVLDB* 16(1), 2022.
- [48] **BullFrog: Online Schema Evolution via Lazy Evaluation** Souvik Bhattacharjee, Gang Liao, Michael Hicks, and Daniel J. Abadi. *In Proceedings of SIGMOD*, 2021.
- [47] **AnyLog: a Grand Unification of the Internet of Things** Daniel J. Abadi, Owen Arden, Faisal Nawab, and Moshe Shadmon. *In Proceedings of CIDR*, 2020.
- [46] **Integration of Large-Scale Data Processing Systems and Traditional Parallel Database Technology** Azza Abouzied, Daniel J. Abadi, Kamil Bajda-Pawlikowski, and Avi Silberschatz. *PVLDB* 12(12), 2019.
- [45] **SLOG: Serializable, Low-latency, Geo-replicated Transactions** Kun Ren, Dennis Li, and Daniel J. Abadi. *PVLDB* 12(11), 2019.
- [44] **The FuzzyLog: A Partially Ordered Shared Log** Lockerman, J.; Faleiro, J. M.; Kim, J.; Sankaran, S.; Abadi, D. J.; Aspnes, J.; Sen, S.; and Balakrishnan, M. *In OSDI*, 2018.
- [43] **An Overview of Deterministic Database Systems.** Abadi, D. J.; and Faleiro, J. M. *In CACM*, 61(9), September 2018
- [42] **High Performance Transactions via Early Write Visibility** Jose M. Faleiro, Daniel Abadi, and Joseph M. Hellerstein. *In PVLDB* 10(5): 613-624, 2017.
- [41] **Latch-free Synchronization in Database Systems: Silver Bullet or Fool's Gold?** Jose M. Faleiro and Daniel J. Abadi. *In Proceedings of CIDR*, 2017.
- [40] **Scalable Pattern Matching over Compressed Graphs via Dedensification** Antonio Maccioni and Daniel J. Abadi. *In Proceedings of KDD*, 2016.
- [39] **Automatic Generation of Normalized Relational Schemas from Nested Key-Value Data** Michael DiScala and Daniel J. Abadi. *In Proceedings of SIGMOD*, 2016.
- [38] **Design Principles for Scaling Multi-core OLTP Under High Contention** Kun Ren, Jose Faleiro, and Daniel J. Abadi. *In Proceedings of SIGMOD*, 2016.
- [37] **Low-Overhead Asynchronous Checkpointing in Main-Memory Database Systems** Kun Ren, Thaddeus Diamond, Daniel J. Abadi, and Alexander Thomson. *In Proceedings of SIGMOD*, 2016.
- [36] **LEOPARD: Lightweight Edge-Oriented Partitioning and Replication for Dynamic Graphs** Jiewen Huang and Daniel J. Abadi. *In PVLDB* 9(7): 540-551, 2016.

- [35] **CalvinFS: A Scalable File System Built on Top of a Distributed DBMS**
Alexander Thomson and Daniel J. Abadi. Proceedings of FAST, 2015.
- [34] **Rethinking serializable multiversion concurrency control**
Jose Faleiro and Daniel J. Abadi. In PVLDB 8(10), 2015.
- [33] **VLL: A Lock Manager Redesign For Main Memory Database Systems**
Kun Ren, Alexander Thomson, and Daniel J. Abadi. In VLDB Journal 24(5): 681-705, 2015.
- [32] **FIT: A Distributed Database Performance Tradeoff**
Jose M. Faleiro, Daniel J. Abadi. IEEE Data Engineering Bull. 38(1): 10-17, 2015.
- [31] **Lazy Evaluation of Transactions in Database Systems**
Jose Faleiro, Alexander Thomson, and Daniel J. Abadi. In Proceedings of SIGMOD, 2014.
- [30] **Sinew: a SQL System for Unified Analytics of Multi-structured Data**
Daniel Tahara, Thaddeus Diamond, and Daniel J. Abadi. In Proceedings of SIGMOD, 2014.
- [29] **An Evaluation of the Advantages and Disadvantages of Deterministic Database Systems**
Kun Ren, Alexander Thomson, and Daniel J. Abadi. In PVLDB 7(10): 821-832, 2014.
- [28] **Fast Distributed Transactions and Strongly Consistent Replication for OLTP Database Systems**
Alexander Thomson, Thaddeus Diamond, Shu-Chun Weng, Kun Ren, Philip Shao, and Daniel J. Abadi. In ACM TODS 39(2): 11, 2014.
- [27] **Query Optimization of Distributed Pattern Matching**
Jiewen Huang, Kartik Venkatraman, and Daniel J. Abadi. In Proceedings of ICDE, 2014.
- [26] **Modularity and Scalability in Calvin.**
Alexander Thomson, Daniel J. Abadi. IEEE Data Eng. Bull. 36(2): 48-55 (2013)
- [25] **Invisible loading: access-driven data transfer from raw files into database systems.**
Azza Abouzied, Daniel J. Abadi, Avi Silberschatz. EDBT 2013: 1-10
- [24] **Lightweight Locking for Main Memory Database Systems**
Kun Ren, Alexander Thomson, and Daniel J. Abadi. PVLDB, 2013.
- [23] **Consistency Tradeoffs in Modern Distributed Database System Design: CAP is Only Part of the Story**
Daniel J. Abadi. In IEEE Computer, 45(2), 2012.
- [22] **Calvin: Fast Distributed Transactions for Partitioned Database Systems**
Alexander Thomson, Thaddeus Diamond, Shu-Chun Weng, Kun Ren, Philip Shao, and Daniel J. Abadi. In Proceedings of SIGMOD, 2012.
- [21] **Scalable SPARQL Querying of Large RDF Graphs**
Jiewen Huang, Daniel J. Abadi, and Kun Ren. PVLDB 4(21), 2011.
- [20] **Efficient Processing of Data Warehousing Queries in a Split Execution Environment**
Kamil Bajda-Pawlikowski, Daniel J. Abadi, Avi Silberschatz, and Erik Paulson. In Proceedings of SIGMOD, 2011.
- [19] **The Case for Determinism in Database Systems**
Alexander Thomson and Daniel J. Abadi. PVLDB 3(1), 2010.
- [18] **Low Overhead Concurrency Control for Partitioned Main Memory Databases**
Evan P. C. Jones, Daniel J. Abadi, and Samuel Madden. In Proceedings of SIGMOD, 2010.
- [17] **MapReduce and Parallel DBMSs: Friends or Foes?**
Michael Stonebraker, Daniel J. Abadi, David J. DeWitt, Samuel Madden, Erik Paulson, Andrew Pavlo, and Alexander Rasin. In CACM, 53(1), 2010.
- [16] **HadoopDB: An Architectural Hybrid of MapReduce and DBMS Technologies for Analytical Workloads**
Azza Abouzeid, Kamil Bajda-Pawlikowski, Daniel J. Abadi, Avi Silberschatz, and Alexander Rasin. PVLDB, 2(1), 2009. Test of Time Award.
- [15] **A Comparison of Approaches to Large Scale Data Analysis**
Andrew Pavlo, Erik Paulson, Alexander Rasin, Daniel J. Abadi, David J. DeWitt, Samuel R. Madden, and Michael Stonebraker. In Proceedings of SIGMOD, 2009.
- [14] **Data Management in the Cloud: Limitations and Opportunities**
Daniel J. Abadi. In IEEE Data Engineering Bulletin, 32(1), 2009.
- [13] **SW-Store: A Vertically Partitioned DBMS for Semantic Web Data Management**
Daniel J. Abadi, Adam Marcus, Samuel R. Madden, and Kate Hollenbach. In VLDB Journal, 18(2), 2009.
- [12] **Column-Stores vs. Row-Stores: How Different Are They Really?**
Daniel J. Abadi, Samuel R. Madden, and Nabil Hachem. In Proceedings of SIGMOD, 2008.
- [11] **OLTP Through the Looking Glass, And What We Found There**
Stavros Harizopoulos, Daniel J. Abadi, Samuel R. Madden, and Michael Stonebraker. In Proceedings of SIGMOD, 2008.

[10] Scalable Semantic Web Data Management Using Vertical Partitioning

Daniel J. Abadi, Adam Marcus, Samuel R. Madden, and Kate Hollenbach. *Proceedings of VLDB*, 2007. **Best Paper Award.**

[9] The End of an Architectural Era (It's Time for a Complete Rewrite)

Michael Stonebraker, Samuel R. Madden, Daniel J. Abadi, Stavros Harizopoulos, Nabil Hachem, and Pat Helland. *Proceedings of VLDB*, 2007.

[8] Column Stores for Wide and Sparse Data

Daniel J. Abadi. In *Proceedings of CIDR*, 2007.

[7] Materialization Strategies in a Column-Oriented DBMS

Daniel J. Abadi, Daniel S. Myers, David J. DeWitt, and Samuel R. Madden. In *Proceedings of ICDE*, 2007.

[6] Performance Tradeoffs in Read-Optimized Databases

Stavros Harizopoulos, Velen Liang, Daniel J. Abadi, and Samuel R. Madden. In *Proceedings of VLDB*, 2006.

[5] Integrating Compression and Execution in Column-Oriented Database Systems

Daniel J. Abadi, Samuel R. Madden, and Miguel C. Ferreira. *Proceedings of SIGMOD*, 2006.

[4] REED: Robust, Efficient Filtering and Event Detection in Sensor Networks

Daniel Abadi, Samuel Madden, and Wolfgang Lindner. *Proceedings of VLDB*, 2005.

[3] C-Store: A Column Oriented DBMS

Mike Stonebraker, Daniel Abadi, Adam Batkin, Xuedong Chen, Mitch Cherniack, Miguel Ferreira, Edmond Lau, Amerson Lin, Samuel Madden, Elizabeth O'Neil, Pat O'Neil, Alex Rasin, Nga Tran and Stan Zdonik. *Proceedings of VLDB*, 2005. **10 Year Best Paper Award.**

[2] The Design of the Borealis Stream Processing Engine

Daniel J. Abadi, Yanif Ahmad, Magdalena Balazinska, Ugur Cetintemel, Mitch Cherniack, Jeong-Hyon Hwang, Wolfgang Lindner, Anurag S. Maskey, Alexander Rasin, Esther Ryzkina, Nesime Tatbul, Ying Xing and Stan Zdonik. *Proceedings of the Conference for Innovative Database Research (CIDR)*, 2005.

[1] Aurora: A New Model and Architecture for Data Stream Management

Daniel Abadi, Don Carney, Ugur Cetintemel, Mitch Cherniack, Christian Convey, Sangdon Lee, Michael Stonebraker, Nesime Tatbul, and Stan Zdonik. *VLDB Journal*, 12(2): 120-139, 2003.

PATENTS

US Patent 10,868,865 System and apparatus to manage data using a peer-to-peer network and the blockchain (2020)

US Patent 9,934,276 Systems and methods for fault tolerant, adaptive execution of arbitrary queries at low latency (2018)

US Patent 9,495,427 Processing of data using a database system in communication with a data processing framework (2016)

US Patent 9,471,711 Schema-less access to stored data (2016)

US Patent 9,336,263 Data loading systems and methods (2016)

US Patent 8,935,232 Query execution systems and methods (2015)

US Patent 8,886,631 Query execution systems and methods (2014)

US Patent 8,700,563 Deterministic database systems (2014)

DISTINGUISHED TALKS

“Integration of Large-Scale Data Processing Systems and Traditional Parallel Database Technology” VLDB Test of Time Award **Plenary Talk**, VLDB 2019

“C-Store: Looking Back and Looking Forward” VLDB 10 Year Best Paper Award **Plenary Talk**, VLDB 2015

“From HadoopDB to Hadapt: A Case Study of Transitioning a VLDB paper into Real World Deployments” VLDB Early Career Research Award **Plenary Talk**, VLDB 2013

“Data loading challenges when transforming Hadoop into a parallel database system”. **Keynote** at EDBT 2013

“Tradeoffs Between Parallel Database Systems, Hadoop, and HadoopDB as Platforms for Petabyte-Scale Analysis”. **Keynote** at SSDBM 2010

“Query Execution in Column-Oriented DBMSs” ACM SIGMOD Jim Gray Award **Plenary Talk**, SIGMOD 2009

SOFTWARE

Calvin (<https://github.com/yaledb/calvin>): A scalable distributed transactional database system.
HadoopDB (<https://sourceforge.net/projects/hadoopdb/>): A parallel database management system.
H-Store (<http://hstore.cs.brown.edu/>): A high-throughput main memory database management system.
C-Store (<http://db.csail.mit.edu/projects/cstore>): A column-oriented database management system.
Borealis (<http://www.cs.brown.edu/research/borealis/public>): A distributed data stream management system.

SERVICE

SIGMOD 2025 Associate Editor **ICDE 2022** Associate Editor **CIDR 2019** Program Committee **SIGMOD 2018** Group Leader **CIDR 2017** Program Committee, **VLDB 2016** Associate Editor, **VLDB 2015** Review Board, **SIGMOD 2014** Program Committee, **CIDR 2013** Program Committee, **SIGMOD 2011** Program Committee, **CIDR 2011** Program Committee, **New England Database Summit 2010** Program Chair, **VLDB 2010** Program Committee, **SIGMOD 2010** Program Committee, **VLDB 2009** Program Committee, **SIGMOD 2009** Local Arrangements Chair, **SIGMOD IDAR 2009** Program Committee, **EDBT 2009 Demo Track** Program Committee, **Winston Churchill Foundation 2008** Selection Committee, **VLDB 2008** Program Committee, **ESWC 2008 Ph.D. Symposium** Program Committee.