

Andy Pavlo

Carnegie Mellon University
Department of Computer Science
Gates-Hillman Center 9019
Pittsburgh, PA 15213-3891 USA

E-mail: pavlo@cs.cmu.edu
Web: <https://cs.cmu.edu/~pavlo/>
Bluesky: @andypavlo.bsky.social
ORCID: 0000-0001-6040-6991

Education

Ph.D., Computer Science Brown University	Summer 2013 <i>Providence, RI USA</i>
M.Sc., Computer Science Brown University	Spring 2009 <i>Providence, RI USA</i>
M.Sc., Computer Science Rochester Institute of Technology	Spring 2006 <i>Rochester, NY USA</i>
B.Sc. with Honors, Computer Science Rochester Institute of Technology	Spring 2005 <i>Rochester, NY USA</i>

Appointments

Associate Professor with Indefinite Tenure Carnegie Mellon University, Department of Computer Science	July 2022 – present <i>Pittsburgh, PA USA</i>
Associate Professor Carnegie Mellon University, Department of Computer Science	July 2019 – June 2022 <i>Pittsburgh, PA USA</i>
Assistant Professor Carnegie Mellon University, Department of Computer Science	September 2013 – June 2019 <i>Pittsburgh, PA USA</i>
Courtesy Associate Professor Carnegie Mellon University, Department of Electrical & Computer Engineering	January 2014 – present <i>Pittsburgh, PA USA</i>

Professional Experience

VP of Database Research ClickHouse, Inc.	August 2026 – Current <i>San Francisco, CA USA</i>
Co-Founder and CEO SO-YOU-DONT-HAVE-TO INCORPORATED'); DROP TABLE companies;	February 2025 – June 2026 <i>Pittsburgh, PA USA</i>
Advisory Board Voltron Data, Inc.	January 2021 – September 2025 <i>San Francisco, CA USA</i>
Co-Founder and CEO OtterTune, Inc.	May 2020 – August 2024 <i>Pittsburgh, PA USA</i>
Advisory Board Fauna, Inc.	April 2019 – April 2021 <i>San Francisco, CA USA</i>
Advisory Board Splice Machine, Inc.	November 2018 – August 2021 <i>San Francisco, CA USA</i>
Expert Witness Much Shelist, P.C.	May 2014 – May 2016 <i>Chicago, IL USA</i>
Research Assistant Brown University, Department of Computer Science	September 2007 – August 2013 <i>Providence, RI USA</i>

Non-Testifying Technical Expert

Choate, Hall & Stewart, LLP.

June 2009 – November 2011

Boston, MA USA

Research Intern

Vertica, Inc.

June 2008 – August 2008

Billerica, MA USA

Systems Programmer – Condor Project

University of Wisconsin–Madison, Department of Computer Sciences

September 2005 – August 2007

Madison, WI USA

Software Developer

PAID, Inc.

March 1999 – December 2004

Owings Mills, MD USA

Teaching Experience

- **15-445/645 – Intro to Database Systems** (Carnegie Mellon University, Spring 2026)
<https://15445.courses.cs.cmu.edu/spring2026/> 
- **15-445/645 – Intro to Database Systems** (Carnegie Mellon University, Fall 2025)
<https://15445.courses.cs.cmu.edu/fall2025/> 
- **15-799 – Special Topics: Query Optimization** (Carnegie Mellon University, Spring 2025)
<https://15799.courses.cs.cmu.edu/spring2025/> 
- **15-445/645 – Intro to Database Systems** (Carnegie Mellon University, Fall 2024)
<https://15445.courses.cs.cmu.edu/fall2024/> 
- **15-721 – Advanced Database Systems** (Carnegie Mellon University, Spring 2024)
<https://15721.courses.cs.cmu.edu/spring2024/> 
- **15-445/645 – Intro to Database Systems** (Carnegie Mellon University, Fall 2023)
<https://15445.courses.cs.cmu.edu/fall2023/> 
- **15-721 – Advanced Database Systems** (Carnegie Mellon University, Spring 2023)
<https://15721.courses.cs.cmu.edu/spring2023/> 
- **15-445/645 – Intro to Database Systems** (Carnegie Mellon University, Fall 2022)
<https://15445.courses.cs.cmu.edu/fall2022/> 
- **15-799 – Special Topics: Self-Driving Database Systems** (Carnegie Mellon University, Spring 2022)
<https://15799.courses.cs.cmu.edu/spring2022/> 
- **15-445/645 – Intro to Database Systems** (Carnegie Mellon University, Fall 2020)
<https://15445.courses.cs.cmu.edu/fall2020/> 
- **15-721 – Advanced Database Systems** (Carnegie Mellon University, Spring 2020)
<https://15721.courses.cs.cmu.edu/spring2020/> 
- **15-445/645 – Intro to Database Systems** (Carnegie Mellon University, Fall 2019)
<https://15445.courses.cs.cmu.edu/fall2019/> 
- **15-721 – Advanced Database Systems** (Carnegie Mellon University, Spring 2019)
<https://15721.courses.cs.cmu.edu/spring2019/> 
- **15-445/645 – Intro to Database Systems** (Carnegie Mellon University, Fall 2018)
<https://15445.courses.cs.cmu.edu/fall2018/> 
- **15-721 – Advanced Database Systems** (Carnegie Mellon University, Spring 2018)
<https://15721.courses.cs.cmu.edu/spring2018/> 
- **15-445/645 – Intro to Database Systems** (Carnegie Mellon University, Fall 2017)
<https://15445.courses.cs.cmu.edu/fall2017/> 
- **15-721 – Advanced Database Systems** (Carnegie Mellon University, Spring 2017)
<https://15721.courses.cs.cmu.edu/spring2017/> 
- **15-415/615 – Database Applications** (Carnegie Mellon University, Fall 2016)
<https://15415.courses.cs.cmu.edu/fall2016/> 
- **15-721 – Advanced Database Systems** (Carnegie Mellon University, Spring 2016)
<https://15721.courses.cs.cmu.edu/spring2016/> 
- **15-415/615 – Database Applications** (Carnegie Mellon University, Fall 2015)
<https://15415.courses.cs.cmu.edu/fall2015/> 

- **15-415/615 – Database Applications** (Carnegie Mellon University, Spring 2015)
<https://15415.courses.cs.cmu.edu/spring2015> 
- **15-415/615 – Database Applications** (Carnegie Mellon University, Spring 2014)
<https://15415.courses.cs.cmu.edu/spring2014> 
- **15-799 – Special Topics: Big Data Systems** (Carnegie Mellon University, Fall 2013)
<https://15799.courses.cs.cmu.edu/fall2013/> 
- **CSC2270 – NewSQL – The Course** (Brown University, Spring 2012)
<http://cs.brown.edu/courses/cs227/archives/2012/> 
- **CSC2270 – Non-standard, Web-scale Databases** (Brown University, Spring 2011)
<http://cs.brown.edu/courses/cs227/archives/2011/> 

Mentoring

PhD Students

- Yuchen Liang (Carnegie Mellon University, 2025 – present)
- Sam Arch (Carnegie Mellon University, 2022 – present)
- William Zhang (Carnegie Mellon University, 2021 – 2026)
- Wan Shen Lim (Carnegie Mellon University, 2020 – 2026)
- Matt Butrovich (Carnegie Mellon University, 2018 – 2024)
- Lin Ma (Carnegie Mellon University, 2015 – 2021)
- Prashanth Menon (Carnegie Mellon University, 2015 – 2021)
- Dana Van Aken (Carnegie Mellon University, 2014 – 2021)
- Huanchen Zhang (Carnegie Mellon University, 2014 – 2019)
- Joy Arulraj (Carnegie Mellon University, 2013 – 2018)
- Xiangyao Yu (Massachusetts Institute of Technology, 2013 – 2017)

Post-Docs

- Lin Ma (Carnegie Mellon University, 2021 – 2022)
- Andrew Crotty (Carnegie Mellon University, 2021 – 2022)

Master's Students

- Abi Kim (Carnegie Mellon University, B.S., M.Sc. 2023)
- Mike Xu (Carnegie Mellon University, B.S., M.Sc. 2023)
- Deepayan Patra (Carnegie Mellon University, B.S., M.Sc. 2022)
- Tanuj Nayak (Carnegie Mellon University, B.S., M.Sc. 2021)
- Amadou Ngom (Carnegie Mellon University, B.S., M.Sc. 2020)
- Lily Liu (Carnegie Mellon University, M.Sc. 2019)
- Gustavo Angulo (Carnegie Mellon University, B.S., M.Sc. 2019)
- Tianyu Li (Carnegie Mellon University, B.S., M.Sc. 2019)
- Sivaprasad Sudhir (Carnegie Mellon University, M.Sc. 2018)
- Prashasthi Prabhakar (Carnegie Mellon University, M.Sc. 2017)
- Patrick Huang (Carnegie Mellon University, M.Sc. 2017)
- Matt Perron (Carnegie Mellon University, M.Sc. 2016)
- Thomas Marshall (Carnegie Mellon University, M.Sc. 2015)
- Atreyee Maiti (Carnegie Mellon University, M.Sc. 2014)
- Xin Jia (Brown University, M.Sc. 2013)
- Yang Zou (Brown University, M.Sc. 2012)
- Yang Lu (Brown University, M.Sc. 2012)
- Visawee Angkanawaraphan (Brown University, M.Sc. 2011)
- Zhe Zhang (Brown University, M.Sc. 2010)
- Ning Shi (Brown University, M.Sc. 2010)

Awards & Honors

- ACM SIGMOD Distinguished Program Committee Member (Summer 2026)

- IEEE TCDE Ramez Elmasri Outstanding Database Education Award (Spring 2026)
- VLDB Early Career Award (Summer 2021)
- ACM SIGMOD Distinguished Program Committee Member (Summer 2019)
- NSF CAREER (Winter 2019)
- VLDB Distinguished Associate Editor (Summer 2018)
- ACM SIGMOD Best Paper Award (Summer 2018)
- ACM SIGMOD Distinguished Program Committee Member (Summer 2018)
- CMU Joel & Ruth Spira Teaching Award (Spring 2018)
- Sloan Research Fellowship (Winter 2018)
- Google Faculty Research Award (Winter 2018)
- ACM SIGMOD Distinguished Program Committee Member (Summer 2017)
- Facebook Faculty Research Award (Summer 2015)
- Google Faculty Research Award (Summer 2015)
- ACM SIGMOD Jim Gray Best Dissertation Award (Summer 2014)
- Amazon Research Grant (Winter 2012)
- ACM SIGMOD Student Travel Grant (Spring 2010)
- Class Clown for Mt. Hebron High School (Spring 1999)

Service

To the University

- CSD 5th Year Masters Program (2025 – 2026)
- CSD PhD Admissions Committee (2025)
- CSD Hiring Committee (2022)
- SCS Undergraduate Curriculum Review Committee (2020 – present)
- CSD Hiring Committee (2018)
- CSD MS Admissions Committee (2017)
- CSD PhD Admissions Committee (2016)
- CSD PhD Admissions Committee (2015)
- CSD PhD Admissions Committee (2014)

To the Profession

- Associate Editor (VLDB 2027)
- Program Committee (SIGMOD 2027)
- Associate Editor (VLDB 2026)
- Program Committee (SIGMOD 2026)
- Program Committee (CIDR 2021 – present)
- Editorial Board for Foundations and Trends in Databases (2020 – 2022)
- Program Committee (CIDR 2020)
- Associate Editor (VLDB 2020)
- Area Chair (SIGMOD 2020)
- Program Committee (CIDR 2019)
- Program Committee (SIGMOD 2019)
- Program Committee (VLDB 2019)
- Associate Editor (VLDB 2018)
- Program Committee (SIGMOD 2018)
- Co-Chair (FADS@VLDB 2017)
- Program Committee / Demo Co-Chair (SIGMOD 2017)
- Program Committee (CIDR 2017)
- Program Committee Area Chair (ICDE 2016)
- Program Committee / PhD Workshop (VLDB 2016)
- Co-Chair (IMDM 2016)
- Program Committee (SIGMOD 2016)
- Steering Committee / Program Committee (HPTS 2015)
- Co-Chair (IMDM 2015)

- Program Committee / Demo Committee / Dissertation Award Committee (SIGMOD 2015)
- Program Committee / Demo Committee (VLDB 2015)
- Co-Chair (IMDM 2014)
- Program Committee (SIGMOD 2014)
- Program Committee (SoCC 2014)
- Program Committee (CloudDP 2014)
- Program Committee (VLDB 2014)
- External Reviewer for LNCS Transactions on Large-Scale Data- and Knowledge-Centered Systems (2013)
- External Reviewer for Springer Journal of Distributed and Parallel Databases (2013)
- External Reviewer for IBM Journal of Research and Development (2012)
- Demo Judge (SIGMOD 2010)
- Demo Judge (SIGMOD 2009)

Publications

Refereed

- [1] William Zhang, Wan Shen Lim, and Andrew Pavlo. “This is Going to Sound Crazy, But What If We Used Large Language Models to Boost Automatic Database Tuning Algorithms By Leveraging Prior History? We Will Find Better Configurations More Quickly Than Retraining From Scratch!” In: *Proc. ACM Manag. Data* 4.1 (Apr. 2026). DOI: 10.1145/3786704. URL: <https://db.cs.cmu.edu/papers/2026/zhang-sigmod2026.pdf>.
- [2] Xinyu Zeng, Ruijun Meng, Martin Prammer, Wes McKinney, Jignesh M. Patel, Andrew Pavlo, and Huanchen Zhang. “F3: The Open-Source Data File Format for the Future”. In: *Proc. ACM Manag. Data* 3.4 (Sept. 2025). DOI: 10.1145/3749163. URL: <https://db.cs.cmu.edu/papers/2025/zeng-sigmod2025.pdf>.
- [3] Matthew Butrovich, Samuel Arch, Wan Shen Lim, William Zhang, Jignesh M Patel, and Andrew Pavlo. “BPF-DB: A Kernel-Embedded Transactional Database Management System For eBPF Applications”. In: *Proc. ACM Manag. Data* 3.3 (June 2025). DOI: 10.1145/3725272. URL: <https://db.cs.cmu.edu/papers/2025/butrovich-sigmod2025.pdf>.
- [4] Hyoungjoo Kim, Yiwei Zhao, Andrew Pavlo, and Phillip B. Gibbons. “No Cap, This Memory Slaps: Breaking Through the Memory Wall of Transactional Database Systems with Processing-in-Memory”. In: *Proc. VLDB Endow.* 18.11 (2025), pp. 4241–4254. DOI: 10.14778/3749646.3749690. URL: <https://www.vldb.org/pvldb/vol18/p4241-kim.pdf>.
- [5] Abigale Kim, Marco Slot, David G. Andersen, and Andrew Pavlo. “Anarchy in the Database: A Survey and Evaluation of Database Management System Extensibility”. In: *Proc. VLDB Endow.* 18.6 (2025), pp. 1962–1976. DOI: 10.14778/3725688.3725719. URL: <https://www.vldb.org/pvldb/vol18/p1962-kim.pdf>.
- [6] Christos Laspias, Andrew Pavlo, and Jignesh M Patel. “A Hot Take on the Intel Analytics Accelerator for Database Management Systems”. In: *ADMS@VLDB.* 2025. URL: <https://db.cs.cmu.edu/papers/2025/laspas-adms2025.pdf>.
- [7] Martin Prammer, Xinyu Zeng, Ruijun Meng, Wes McKinney, Huanchen Zhang, Andrew Pavlo, and Jignesh Patel. “Towards Functional Decomposition of Storage Formats”. In: *CIDR 2025, Conference on Innovative Data Systems Research.* 2025. URL: <https://db.cs.cmu.edu/papers/2025/p19-prammer.pdf>.
- [8] Michael Stonebraker and Andrew Pavlo. “What Goes Around Comes Around... And Around...” In: *SIGMOD Record* 53.2 (June 2024), pp. 21–37. DOI: 10.1145/3685980.3685984. URL: <https://db.cs.cmu.edu/papers/2024/whatgoesaround-sigmodrec2024.pdf>.
- [9] Samuel I Arch, Todd C. Mowry, Jignesh M. Patel, and Andrew Pavlo. “The Key to Effective UDF Optimization: Before Inlining, First Perform Outlining”. In: *Proc. VLDB Endow.* 18.1 (2024), pp. 1–13. DOI: 10.14778/3696435.3696436. URL: <https://www.vldb.org/pvldb/vol18/p1-arch.pdf>.
- [10] William Zhang, Wan Shen Lim, Matthew Butrovich, and Andrew Pavlo. “The Holon Approach for Simultaneously Tuning Multiple Components in a Self-Driving Database Management System with Machine Learning via Synthesized Proto-Actions”. In: *Proc. VLDB Endow.* 17.11 (2024), pp. 3373–3387. DOI: 10.14778/3681954.3682007. URL: <https://www.vldb.org/pvldb/vol17/p3373-zhang.pdf>.

- [11] Wan Shen Lim, Lin Ma, William Zhang, Matthew Butrovich, Samuel I Arch, and Andrew Pavlo. “Hit the Gym: Accelerating Query Execution to Efficiently Bootstrap Behavior Models for Self-Driving Database Management Systems”. In: *Proc. VLDB Endow.* 17.11 (2024), pp. 3680–3693. doi: 10.14778/3681954.3682030. URL: <https://www.vldb.org/pvldb/vol17/p3680-lim.pdf>.
- [12] Xinyu Zeng, Ruijun Meng, Andrew Pavlo, Wes McKinney, and Huanchen Zhang. “NULLS!: Revisiting Null Representation in Modern Columnar Formats”. In: *Proceedings of the 20th International Workshop on Data Management on New Hardware*. DaMoN ’24. 2024. doi: 10.1145/3662010.3663452. URL: <https://db.cs.cmu.edu/papers/2024/zeng-damon24.pdf>.
- [13] Kai Franz, Samuel I Arch, Denis Hirn, Torsten Grust, Todd Mowry, and Andrew Pavlo. “Dear User-Defined Functions, Inlining isn’t working out so great for us. Let’s try batching to make our relationship work. Sincerely, SQL”. In: *CIDR 2024, Conference on Innovative Data Systems Research*. 2024. URL: <https://db.cs.cmu.edu/papers/2024/p13-franz.pdf>.
- [14] Xinyu Zeng, Yulong Hui, Jiahong Shen, Andrew Pavlo, Wes McKinney, and Huanchen Zhang. “An Empirical Evaluation of Columnar Storage Formats”. In: *Proc. VLDB Endow.* 17.2 (2023), pp. 148–161. doi: 10.14778/3626292.3626298. URL: <https://www.vldb.org/pvldb/vol17/p148-zeng.pdf>.
- [15] Matthew Butrovich, Karthik Ramanathan, John Rollinson, Wan Shen Lim, William Zhang, Justine Sherry, and Andrew Pavlo. “Tigger: A Database Proxy That Bounces With User-Bypass”. In: *Proc. VLDB Endow.* 16.11 (2023), pp. 3335–3348. doi: 10.14778/3611479.3611530. URL: <https://www.vldb.org/pvldb/vol16/p3335-butrovich.pdf>.
- [16] Wan Shen Lim, Matthew Butrovich, William Zhang, Andrew Crotty, Lin Ma, Peijing Xu, Johannes Gehrke, and Andrew Pavlo. “Database Gyms”. In: *CIDR 2023, Conference on Innovative Data Systems Research*. 2023. URL: <https://db.cs.cmu.edu/papers/2023/p27-lim.pdf>.
- [17] Yu Xia, Xiangyao Yu, Matthew Butrovich, Andrew Pavlo, and Srinivas Devadas. “Litmus: Towards a Practical Database Management System with Verifiable ACID Properties and Transaction Correctness”. In: *Proceedings of the 2022 International Conference on Management of Data*. SIGMOD ’22. 2022, pp. 1478–1492. doi: 10.1145/3514221.3517851. URL: <https://db.cs.cmu.edu/papers/2022/yu-litmusdb-sigmod2022.pdf>.
- [18] Matthew Butrovich, Wan Shen Lim, Lin Ma, John Rollinson, William Zhang, Yu Xia, and Andrew Pavlo. “Tastes Great! Less Filling! High Performance and Accurate Training Data Collection for Self-Driving Database Management Systems”. In: *Proceedings of the 2022 International Conference on Management of Data*. SIGMOD/PODS ’22. 2022, pp. 617–630. doi: 10.1145/3514221.3517845.
- [19] Andrew Crotty, Viktor Leis, and Andrew Pavlo. “Are You Sure You Want to Use MMAP in Your Database Management System?” In: *CIDR 2022, Conference on Innovative Data Systems Research*. 2022. URL: <https://db.cs.cmu.edu/papers/2022/cidr2022-p13-crotty.pdf>.
- [20] Dana Van Aken, Dongsheng Yang, Sebastien Brillard, Ari Fiorino, Bohan Zhang, Christian Billian, and Andrew Pavlo. “An Inquiry into Machine Learning-based Automatic Configuration Tuning Services on Real-World Database Management Systems”. In: *Proc. VLDB Endow.* 14.7 (2021), pp. 1241–1253. URL: <https://db.cs.cmu.edu/papers/2021/p1241-aken.pdf>.
- [21] Lin Ma, William Zhang, Jie Jiao, Wuwen Wang, Matthew Butrovich, Wan Shen Lim, Prashanth Menon, and Andrew Pavlo. “MB2: Decomposed Behavior Modeling for Self-Driving Database Management Systems”. In: *Proceedings of the 2021 International Conference on Management of Data*. SIGMOD ’21. 2021, pp. 1248–1261. doi: 10.1145/3448016.3457276. URL: <https://db.cs.cmu.edu/papers/2021/ma-sigmod2021.pdf>.
- [22] Amadou Ngom, Prashanth Menon, Matthew Butrovich, Lin Ma, Wan Shen Lim, Todd C. Mowry, and Andrew Pavlo. “Filter Representation in Vectorized Query Execution”. In: *Proceedings of the 17th International Workshop on Data Management on New Hardware (DaMoN 2021)*. DAMON’21. 2021, 6:1–6:7. doi: 10.1145/3465998.3466009. URL: <https://db.cs.cmu.edu/papers/2021/ngom-damon2021.pdf>.
- [23] Xinjing Zhou, Joy Arulraj, Andrew Pavlo, and David Cohen. “Spitfire: A Three-Tier Buffer Manager for Volatile and Non-Volatile Memory”. In: *Proceedings of the 2021 International Conference on Management of Data*. SIGMOD ’21. 2021, pp. 2195–2207. doi: 10.1145/3448016.3452819. URL: <https://db.cs.cmu.edu/papers/2021/zhou-sigmod2021.pdf>.

- [24] Ling Zhang, Matthew Butrovich, Tianyu Li, Andrew Pavlo, Yash Nannapaneni, John Rollinson, Huanchen Zhang, Ambarish Balakumar, Daniel Biales, Ziqi Dong, Emmanuel J Eppinger, Jordi E Gonzalez, Wan Shen Lim, Jianqiao Liu, Lin Ma, Prashanth Menon, Soumil Mukherjee, Tanuj Nayak, Amadou Ngom, Dong Niu, Deepayan Patra, Poojita Raj, Stephanie Wang, Wuwen Wang, Yao Yu, and William Zhang. “Everything is a Transaction: Unifying Logical Concurrency Control and Physical Data Structure Maintenance in Database Management Systems”. In: *CIDR 2021, Conference on Innovative Data Systems Research*. 2021. URL: https://db.cs.cmu.edu/papers/2021/cidr2021_paper06.pdf.
- [25] Tianyu Li, Matthew Butrovich, Amadou Ngom, Wan Shen Lim, Wes McKinney, and Andrew Pavlo. “Mainlining Databases: Supporting Fast Transactional Workloads on Universal Columnar Data File Formats”. In: *Proc. VLDB Endow.* 14.4 (2020), pp. 534–546. URL: <https://db.cs.cmu.edu/papers/2020/p534-li.pdf>.
- [26] Yu Xia, Xiangyao Yu, Andrew Pavlo, and Srinivas Devadas. “Taurus: Lightweight Parallel Logging for In-Memory Database Management Systems”. In: *Proc. VLDB Endow.* 14.2 (2020), pp. 189–201. URL: <https://db.cs.cmu.edu/papers/2020/p189-xia.pdf>.
- [27] Prashanth Menon, Amadou Ngom, and Andrew Pavlo Lin Ma Todd C. Mowry. “Permutable Compiled Queries: Dynamically Adapting Compiled Queries without Recompiling”. In: *Proc. VLDB Endow.* 14.2 (2020), pp. 101–113. URL: <https://db.cs.cmu.edu/papers/2020/p101-menon.pdf>.
- [28] Huanchen Zhang, Xiaoxuan Liu, David G. Andersen, Michael Kaminsky, Kimberly Keeton, and Andrew Pavlo. “Order-Preserving Key Compression for In-Memory Search Trees”. In: *Proceedings of the 2020 International Conference on Management of Data*. 2020, pp. 1601–1615. URL: <https://db.cs.cmu.edu/papers/2020/zhang-sigmod2020.pdf>.
- [29] Yihan Sun, Guy E. Blelloch, Wan Shen Lim, and Andrew Pavlo. “On Supporting Efficient Snapshot Isolation for Hybrid Workloads with Multi-Versioned Indexes”. In: *Proc. VLDB Endow.* 13 (2 Oct. 2019), pp. 221–225. URL: <https://db.cs.cmu.edu/papers/2019/p211-sun.pdf>.
- [30] Yangjun Sheng, Anthony Tomasic, Tieying Zhang, and Andrew Pavlo. “Scheduling OLTP transactions via learned abort prediction”. In: *Proceedings of the Second International Workshop on Exploiting Artificial Intelligence Techniques for Data Management, aiDM@SIGMOD 2019*. 2019, 1:1–1:8. URL: <https://db.cs.cmu.edu/papers/2019/a1-sheng.pdf>.
- [31] Timo Kersten, Viktor Leis, Alfons Kemper, Thomas Neumann, Andrew Pavlo, and Peter Boncz. “Everything You Always Wanted to Know About Compiled and Vectorized Queries But Were Afraid to Ask”. In: *Proc. VLDB Endow.* 11 (13 Sept. 2018), pp. 2209–2222. URL: <https://db.cs.cmu.edu/papers/2018/p2209-kersten.pdf>.
- [32] Xiangyao Yu, Yu Xia, Andrew Pavlo, Daniel Sanchez, Larry Rudolph, and Srinivas Devadas. “Sundial: Harmonizing Concurrency Control and Caching in a Distributed OLTP Database Management System”. In: *Proc. VLDB Endow.* 11.10 (June 2018), pp. 1289–1302. URL: <https://db.cs.cmu.edu/papers/2018/p1289-yu.pdf>.
- [33] Ziqi Wang, Andrew Pavlo, Hyeontaek Lim, Viktor Leis, Huanchen Zhang, Michael Kaminsky, and David G. Andersen. “Building a Bw-Tree Takes More Than Just Buzz Words”. In: *Proceedings of the 2018 ACM International Conference on Management of Data*. SIGMOD ’18. 2018, pp. 473–488. URL: <https://db.cs.cmu.edu/papers/2018/mod342-wangA.pdf>.
- [34] Lin Ma, Dana Van Aken, Ahmed Hefny, Gustavo Mezerhane, Andrew Pavlo, and Geoffrey J Gordon. “Query-based Workload Forecasting for Self-Driving Database Management Systems”. In: *Proceedings of the 2018 ACM International Conference on Management of Data*. SIGMOD ’18. 2018, pp. 631–645. doi: 10.1145/3183713.3196908. URL: <https://db.cs.cmu.edu/papers/2018/mod435-maA.pdf>.
- [35] Huanchen Zhang, David G. Andersen, Michael Kaminsky, Andrew Pavlo, Hyeontaek Lim, Viktor Leis, and Kimberly Keeton. “SuRF: Practical Range Query Filtering with Fast Succinct Tries”. In: *Proceedings of the 2018 ACM International Conference on Management of Data*. SIGMOD ’18. 2018, pp. 323–336. URL: <https://db.cs.cmu.edu/papers/2018/mod601-zhangA-hm.pdf>.
- [36] Prashanth Menon, Todd C. Mowry, and Andrew Pavlo. “Relaxed Operator Fusion for In-Memory Databases: Making Compilation, Vectorization, and Prefetching Work Together At Last”. In: *Proc. VLDB Endow.* 11.1 (Sept. 2017), pp. 1–13. URL: <https://db.cs.cmu.edu/papers/2017/p1-menon.pdf>.
- [37] Yingjun Wu, Joy Arulraj, Jiexi Lin, Ran Xian, and Andrew Pavlo. “An Empirical Evaluation of In-Memory Multi-Version Concurrency Control”. In: *Proc. VLDB Endow.* 10.7 (Mar. 2017), pp. 781–792. URL: <https://db.cs.cmu.edu/papers/2017/p781-wu.pdf>.

- [38] Rachael Harding, Dana Van Aken, Andrew Pavlo, and Michael Stonebraker. “An Evaluation of Distributed Concurrency Control”. In: *Proc. VLDB Endow.* 10.5 (Jan. 2017), pp. 553–564. URL: <https://db.cs.cmu.edu/papers/2017/p553-harding.pdf>.
- [39] Dana Van Aken, Andrew Pavlo, Geoffrey J. Gordon, and Bohan Zhang. “Automatic Database Management System Tuning Through Large-scale Machine Learning”. In: *Proceedings of the 2017 ACM International Conference on Management of Data*. SIGMOD ’17. 2017, pp. 1009–1024. DOI: 10.1145/3035918.3064029. URL: <https://db.cs.cmu.edu/papers/2017/p1009-van-aken.pdf>.
- [40] Lianghong Xu, Andrew Pavlo, Sudipta Sengupta, and Gregory R. Ganger. “Online Deduplication for Databases”. In: *Proceedings of the 2017 ACM International Conference on Management of Data*. SIGMOD ’17. 2017, pp. 1355–1368. URL: <https://db.cs.cmu.edu/papers/2017/p1355-xu.pdf>.
- [41] Andrew Pavlo, Gustavo Angulo, Joy Arulraj, Haibin Lin, Jiexi Lin, Lin Ma, Prashanth Menon, Todd Mowry, Matthew Perron, Ian Quah, Siddharth Santurkar, Anthony Tomasic, Skye Toor, Dana Van Aken, Ziqi Wang, Yingjun Wu, Ran Xian, and Tieying Zhang. “Self-Driving Database Management Systems”. In: *CIDR 2017, Conference on Innovative Data Systems Research*. 2017. URL: <https://db.cs.cmu.edu/papers/2017/p42-pavlo-cidr17.pdf>.
- [42] Joy Arulraj, Matthew Perron, and Andrew Pavlo. “Write-Behind Logging”. In: *Proc. VLDB Endow.* 10.4 (Dec. 2016), pp. 337–348. URL: <https://db.cs.cmu.edu/papers/2016/p337-arulraj.pdf>.
- [43] Marco Serafini, Rebecca Taft, Aaron J Elmore, Andrew Pavlo, Ashraf Aboulmaga, and Michael Stonebraker. “Clay: Fine-Grained Adaptive Partitioning for General Database Schemas”. In: *Proc. VLDB Endow.* 10.4 (Dec. 2016), pp. 445–456. URL: <http://hstore.cs.brown.edu/papers/hstore-clay.pdf>.
- [44] Lin Ma, Joy Arulraj, Sam Zhao, Andrew Pavlo, Subramanya R. Dulloor, Michael J. Giardino, Jeff Parkhurst, Jason L. Gardner, Kshitij Doshi, and Stanley Zdonik. “Larger-than-memory Data Management on Modern Storage Hardware for In-memory OLTP Database Systems”. In: *Proceedings of the 12th International Workshop on Data Management on New Hardware*. DaMoN ’16. 2016, 9:1–9:7. URL: <https://db.cs.cmu.edu/papers/2016/ma-damon2016.pdf>.
- [45] Joy Arulraj, Andrew Pavlo, and Prashanth Menon. “Bridging the Archipelago Between Row-Stores and Column-Stores for Hybrid Workloads”. In: *Proceedings of the 2016 International Conference on Management of Data*. SIGMOD ’16. 2016, pp. 583–598. DOI: 10.1145/2882903.2915231. URL: <https://db.cs.cmu.edu/papers/2016/arulraj-sigmod2016.pdf>.
- [46] Huanchen Zhang, David G. Andersen, Andrew Pavlo, Michael Kaminsky, Lin Ma, and Rui Shen. “Reducing the Storage Overhead of Main-Memory OLTP Databases with Hybrid Indexes”. In: *Proceedings of the 2016 International Conference on Management of Data*. SIGMOD ’16. 2016, pp. 1567–1581. URL: <https://db.cs.cmu.edu/papers/2016/zhang-sigmod2016.pdf>.
- [47] Xiangyao Yu, Andrew Pavlo, Daniel Sanchez, and Srinivas Devadas. “TicToc: Time Traveling Optimistic Concurrency Control”. In: *Proceedings of the 2016 International Conference on Management of Data*. SIGMOD ’16. 2016, pp. 1629–1642. DOI: 10.1145/2882903.2882935. URL: <https://db.cs.cmu.edu/papers/2016/yu-sigmod2016.pdf>.
- [48] John Meehan, Nesime Tatbul, Stan Zdonik, Cansu Aslantas, Ugur Çetintemel, Jiang Du, Samuel Madden, David Maier, Andrew Pavlo, Michael Stonebraker, Kristin Tufte, and Hao Wang. “S-Store: Streaming Meets Transaction Processing”. In: *PVLDB* 8.13 (2015), pp. 2134–2145. URL: <https://db.cs.cmu.edu/papers/2015/p2134-meehan.pdf>.
- [49] Lianghong Xu, Andrew Pavlo, Sudipta Sengupta, Jin Li, and Gregory R. Ganger. “Reducing Replication Bandwidth for Distributed Document Databases”. In: *Proceedings of the Sixth ACM Symposium on Cloud Computing*. SoCC ’15. 2015, pp. 222–235. URL: <https://db.cs.cmu.edu/papers/2015/socc15-sdedup.pdf>.
- [50] Aaron J. Elmore, Vaibhav Arora, Rebecca Taft, Andrew Pavlo, Divyakant Agrawal, and Amr El Abbadi. “Squall: Fine-Grained Live Reconfiguration for Partitioned Main Memory Databases”. In: *Proceedings of SIGMOD*. SIGMOD ’15. 2015, pp. 299–313. URL: <https://db.cs.cmu.edu/papers/2015/p299-elmore.pdf>.
- [51] Joy Arulraj, Andrew Pavlo, and Subramanya Dulloor. “Let’s Talk About Storage & Recovery Methods for Non-Volatile Memory Database Systems”. In: *Proceedings of the 2015 International Conference on Management of Data*. SIGMOD ’15. 2015, pp. 707–722. URL: <https://db.cs.cmu.edu/papers/2015/p707-arulraj.pdf>.

- [52] Xiangyao Yu, George Bezerra, Andrew Pavlo, Srinivas Devadas, and Michael Stonebraker. “Staring into the Abyss: An Evaluation of Concurrency Control with One Thousand Cores”. In: *PVLDB* 8.3 (2014), pp. 209–220. URL: <https://www.vldb.org/pvldb/vol8/p209-yu.pdf>.
- [53] Rebecca Taft, Essam Mansour, Marco Serafini, Jennie Duggan, Aaron J. Elmore, Ashraf Aboulmaga, Andrew Pavlo, and Michael Stonebraker. “E-Store: Fine-Grained Elastic Partitioning for Distributed Transaction Processing”. In: *PVLDB* 8.3 (2014), pp. 245–256. URL: <https://hstore.cs.brown.edu/papers/hstore-elastic.pdf>.
- [54] Justin DeBrabant, Joy Arulraj, Andrew Pavlo, Michael Stonebraker, Stanley B. Zdonik, and Subramanya Dullloor. “A Prolegomenon on OLTP Database Systems for Non-Volatile Memory”. In: *ADMS @ VLDB*. 2014, pp. 57–63. URL: <http://hstore.cs.brown.edu/papers/hstore-nvm.pdf>.
- [55] Djellel Eddine Difallah, Andrew Pavlo, Carlo Curino, and Philippe Cudré-Mauroux. “OLTP-Bench: An Extensible Testbed for Benchmarking Relational Databases”. In: *PVLDB* 7.4 (2013), pp. 277–288. URL: <https://www.cs.cmu.edu/~pavlo/static/papers/oltpbench-vldb.pdf>.
- [56] Justin DeBrabant, Andrew Pavlo, Stephen Tu, Michael Stonebraker, and Stanley B. Zdonik. “Anti-Caching: A New Approach to Database Management System Architecture”. In: *PVLDB* 6.14 (2013), pp. 1942–1953. URL: <http://hstore.cs.brown.edu/papers/hstore-anticaching.pdf>.
- [57] Carlo A. Curino, Djellel E. Difallah, Andrew Pavlo, and Philippe Cudre-Mauroux. “Benchmarking OLTP/Web Databases in the Cloud: The OLTP-bench Framework”. In: *CloudDB*. 2012, pp. 17–20. URL: <http://www.cs.cmu.edu/~pavlo/static/papers/oltpbench.pdf>.
- [58] Andrew Pavlo, Carlo Curino, and Stanley Zdonik. “Skew-Aware Automatic Database Partitioning in Shared-Nothing, Parallel OLTP Systems”. In: *SIGMOD*. 2012, pp. 61–72. URL: <https://hstore.cs.brown.edu/papers/hstore-partitioning.pdf>.
- [59] Andrew Pavlo, Evan P.C. Jones, and Stan Zdonik. “On Predictive Modeling for Optimizing Transaction Execution in Parallel OLTP Systems”. In: *Proc. VLDB Endow.* 5 (2 Oct. 2011), pp. 85–96. URL: <https://hstore.cs.brown.edu/papers/hstore-markov.pdf>.
- [60] 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: *Proceedings of SIGMOD*. 2009, pp. 165–178. URL: <https://www.cs.cmu.edu/~pavlo/papers/benchmarks-sigmod09.pdf>.
- [61] Ewa Deelman, Miron Livny, Gaurang Mehta, Andy 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: *HPC and Grids in Action*. Vol. 16. Advances in Parallel Computing. 2008, pp. 56–74. URL: https://www.cs.cmu.edu/~pavlo/static/papers/cetraro_final.pdf.
- [62] 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*. 2006, pp. 263–273. URL: <https://www.cs.cmu.edu/~pavlo/static/papers/nmi-lisa2006.pdf>.

Books

- [1] Joy Arulraj and Andrew Pavlo. *Non-Volatile Memory Database Management Systems*. Synthesis Lectures on Data Management. Morgan & Claypool Publishers, 2019. doi: 10.2200/S00891ED1V01Y201812DTM055.

Book Chapters

- [1] Andy Pavlo. “Making Databases Work”. In: ed. by Michael L. Brodie. Association for Computing Machinery and Morgan & Claypool, 2019. Chap. In-memory, Horizontal, and Transactional: The H-store OLTP DBMS Project, pp. 245–251.
- [2] Franz Faerber, Alfons Kemper, Per-Åke Larson, Justin J. Levandoski, Thomas Neumann, and Andrew Pavlo. “Main Memory Database Systems”. In: *Foundations and Trends in Databases* 8.1-2 (2017), pp. 1–130.

Demonstrations / Tutorials

- [1] Patrick Wang, Wan Shen Lim, William Zhang, Samuel Arch, and Andrew Pavlo. “Automated Database Tuning vs. Human-Based Tuning in a Simulated Stressful Work Environment: A Demonstration of the Database Gym”. In: *Companion of the 2025 International Conference on Management of Data*. SIGMOD/PODS ’25. 2025, pp. 247–250. DOI: 10.1145/3722212.3725083.
- [2] Bohan Zhang, Dana Van Aken, Justin Wang, Tao Dai, Shuli Jiang, Jacky Lao, Siyuan Sheng, Andrew Pavlo, and Geoffrey J. Gordon. “A Demonstration of the OtterTune Automatic Database Management System Tuning Service”. In: *PVLDB* 11.12 (2018), pp. 1910–1913.
- [3] Joy Arulraj and Andrew Pavlo. “How to Build a Non-Volatile Memory Database Management System”. In: *Proceedings of the 2017 ACM International Conference on Management of Data*. SIGMOD ’17. 2017, pp. 1753–1758.
- [4] Dana Van Aken, Djellel Eddine Difallah, Andrew Pavlo, Carlo Curino, and Philippe Cudre-Mauroux. “BenchPress: Dynamic Workload Control in the OLTP-Bench Testbed”. In: *Proceedings of SIGMOD*. 2015, pp. 1069–1073.
- [5] Ugur Çetintemel, Jiang Du, Tim Kraska, Samuel Madden, David Maier, John Meehan, Andrew Pavlo, Michael Stonebraker, Erik Sutherland, Nesime Tatbul, Kristin Tufte, Hao Wang, and Stanley B. Zdonik. “S-Store: A Streaming NewSQL System for Big Velocity Applications”. In: *PVLDB* 7.13 (2014), pp. 1633–1636.
- [6] 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”. In: *Proc. VLDB Endow.* 1.2 (2008), pp. 1496–1499.
- [7] Christopher Homan, Andrew Pavlo, and Jonathan Schull. “Smoother Transitions Between Breadth-First-Spanning-Tree-Based Drawings”. In: *14th International Symposium on Graph Drawing*. Vol. 4372. Lecture Notes in Computer Science. 2006, pp. 442–445.

Miscellaneous Writings

- [1] Samuel Arch, Yuchen Liu, Todd C. Mowry, Jignesh M. Patel, and Andrew Pavlo. “Partial UDF Inlining”. In: *SIGMOD Rec.* 55.1 (Apr. 2026), pp. 74–83. DOI: 10.1145/3810900.3810914. URL: <https://db.cs.cmu.edu/papers/2026/arch-sigmodrec2026.pdf>.
- [2] Huanchen Zhang, Hyeontaek Lim, Viktor Leis, David G. Andersen, Michael Kaminsky, Kimberly Keeton, and Andrew Pavlo. “Succinct Range Filters”. In: *Commun. ACM* 64.4 (Mar. 2021), pp. 166–173. URL: <https://doi.org/10.1145/3450262>.
- [3] Andrew Pavlo, Matthew Butrovich, Lin Ma, Prashanth Menon, Wan Shen Lim, Dana Van Aken, and William Zhang. “Make Your Database System Dream of Electric Sheep: Towards Self-Driving Operation”. In: *Proc. VLDB Endow.* 14.12 (2021), pp. 3211–3221. URL: <http://www.vldb.org/pvldb/vol14/p3211-pavlo.pdf>.
- [4] Huanchen Zhang, Hyeontaek Lim, Viktor Leis, David G. Andersen, Michael Kaminsky, Kimberly Keeton, and Andrew Pavlo. “Succinct Range Filters”. In: *ACM Trans. Database Syst.* 45.2 (2020), 5:1–5:31.
- [5] Andrew Pavlo, Matthew Butrovich, Ananya Joshi, Lin Ma, Prashanth Menon, Dana Van Aken, Lisa Lee, and Ruslan Salakhutdinov. “External vs. Internal: An Essay on Machine Learning Agents for Autonomous Database Management Systems”. In: *IEEE Data Engineering Bulletin* (June 2019), pp. 32–46.
- [6] Huanchen Zhang, Hyeontaek Lim, Viktor Leis, David G. Andersen, Michael Kaminsky, Kimberly Keeton, and Andrew Pavlo. “Succinct Range Filters”. In: *SIGMOD Record* 48.1 (Mar. 2019), pp. 78–85.
- [7] Dana Van Aken, Andrew Pavlo, and Geoffrey J. Gordon. *Amazon AWS AI Blog – Tuning Your DBMS Automatically with Machine Learning*. <https://aws.amazon.com/blogs/ai/tuning-your-dbms-automatically-with-machine-learning/>. June 2017.
- [8] Andrew Pavlo. “What Are We Doing With Our Lives?: Nobody Cares About Our Concurrency Control Research”. In: *Proceedings of the 2017 ACM International Conference on Management of Data*. SIGMOD ’17. 2017, p. 3.
- [9] Peter Bailis, Camille Fournier, Joy Arulraj, and Andrew Pavlo. “Research for Practice: Distributed Consensus and Implications of NVM on Database Management Systems”. In: *Commun. ACM* 59.11 (Oct. 2016), pp. 52–55. URL: <https://db.cs.cmu.edu/papers/2016/cacm-nvm2016.pdf>.

- [10] Andrew Pavlo and Matthew Aslett. “What’s Really New with NewSQL?” In: *SIGMOD Rec.* 45.2 (June 2016), pp. 45–55. URL: <https://db.cs.cmu.edu/papers/2016/pavlo-newsql-sigmodrec2016.pdf>.
- [11] Andrew Pavlo. “Emerging Hardware Trends in Large-Scale Transaction Processing”. In: *IEEE Internet Computing* 19.3 (May 2015), pp. 68–71. URL: https://www.cs.cmu.edu/~pavlo/static/papers/IC_20150501_May_2015.pdf.
- [12] Michael Stonebraker, Andrew Pavlo, Rebecca Taft, and Michael L. Brodie. “Enterprise Database Applications and the Cloud: A Difficult Road Ahead”. In: *2014 IEEE International Conference on Cloud Engineering*. 2014, pp. 1–6.
- [13] Andrew Pavlo. “Not Your Traditional Data Management - HPTS Conference Report”. In: *USENIX ;login:* (Feb. 2012), pp. 76–78. URL: <https://www.usenix.org/publications/login/february-2012/14th-international-workshop-high-performance%20-transaction-systems>.
- [14] Xin Jia, Andrew Pavlo, and Stanley B. Zdonik. “Tastes Great, Less Filling: Low-Impact OLAP MapReduce Queries on High-Performance OLTP Systems”. In: *TinyToCS 1* (2012). URL: <http://tinytocs.org/vol1/papers/tinytocs-v1-jia.pdf>.
- [15] Andrew Pavlo and Ning Shi. “Graffiti Networks: A Subversive, Internet-Scale File Sharing Model”. In: *CoRR* abs/1101.0350 (2011). URL: <http://arxiv.org/abs/1101.0350>.
- [16] Michael Stonebraker, Daniel Abadi, David J. DeWitt, Sam Madden, Erik Paulson, Andrew Pavlo, and Alexander Rasin. “MapReduce and Parallel DBMSs: Friends or Foes?” In: *Communications of the ACM* 53.1 (2010), pp. 64–71.
- [17] Andrew Pavlo, Christopher Homan, and Jonathan Schull. “A parent-centered radial layout algorithm for interactive graph visualization and animation”. In: *CoRR* abs/cs/0606007 (2006). URL: <http://arxiv.org/abs/cs/0606007>.

Synergistic Activities

Open-Source Software:

- optd Query Optimizer (<https://github.com/cmu-db/optd>)
- BenchBase Database Benchmark Suite (<https://github.com/cmu-db/benchbase>)
- NoisePage Database Management System (<https://noisepage.com>)
- Peloton Database Management System (<https://db.cs.cmu.edu/peloton>)
- OtterTune (<https://db.cs.cmu.edu/projects/ottertune/>)
- Database of Databases (<https://dbdb.io>)
- H-Store Database Management System (<https://hstore.cs.brown.edu>)

Database Seminar Series:

- **Whatever Andy Found Interesting While Out At Sea** (Fall 2026)
<https://db.cs.cmu.edu/seminars/fall2026> 
- **PostgreSQL vs. The World Seminar Series** (Spring 2026)
<https://db.cs.cmu.edu/seminars/spring2026> 
- **Future Data Systems Seminar Series** (Fall 2025)
<https://db.cs.cmu.edu/seminars/fall2025> 
- **SQL or Death? Seminar Series** (Spring 2025)
<https://db.cs.cmu.edu/seminar2025> 
- **Database Building Blocks Technical Talks** (Fall 2024)
<https://db.cs.cmu.edu/seminar2024> 
- **ML $\rightleftharpoons$ DB Technical Talks** (Fall 2023)
<https://db.cs.cmu.edu/seminar2023> 
- **¡Databases! - A Database Seminar Series** (Fall 2022)
<https://db.cs.cmu.edu/seminar2022> 
- **Vaccination Database Tech Talks – Booster** (Spring 2022)
<https://db.cs.cmu.edu/seminar2022-booster> 
- **Vaccination Database Tech Talks – Second Dose** (Fall 2021)
<https://db.cs.cmu.edu/seminar2021-dose2> 
- **Vaccination Database Tech Talks – First Dose** (Spring 2021)
<https://db.cs.cmu.edu/seminar2021> 

- **Quarantine Database Tech Talks** (Fall 2020)
<https://db.cs.cmu.edu/seminar2020> 
- **Hardware-Accelerated Database Lectures** (Fall 2018)
<https://db.cs.cmu.edu/seminar2018> 
- **Time Series Database Lectures** (Fall 2017)
<https://db.cs.cmu.edu/seminar2017> 
- **Databaseology Lectures** (Fall 2015)
<https://db.cs.cmu.edu/seminar2015> 
- **Seven Databases in Seven Weeks** (Fall 2014)
<https://db.cs.cmu.edu/seminar2014> 

Internship Programs:

- Summer 2025 (<https://db.cs.cmu.edu/internship2025>)
- Summer 2021 (<https://db.cs.cmu.edu/internship2021>)
- Summer 2019 (<https://db.cs.cmu.edu/internship2019>)
- Summer 2018 (<https://db.cs.cmu.edu/internship2018>)
- Summer 2017 (<https://db.cs.cmu.edu/internship2017>)
- Summer 2017 Pennsylvania Governor's School for the Sciences (Pre-College)

Talks

- **Databases: The Final Boss of Agents**
Snowflake Tech Talk (July 2026)
Google Data Intelligence Academic Seminar (July 2026)
Supporting Our AI Overlords @ CAIS Keynote (May 2026)
Percona Live Keynote (May 2026)
Greater Chicago Area Systems Research Workshop (May 2026)
- **ChatGPT Ain't Got \$%& On Me! The Future of Automated Database Tuning**
OpenAI (January 2026)
P99 Conf Keynote (October 2025)
MDI Summit 2025 (September 2025)
South Bay Systems Meetup (August 2025)
Google DANI Academic Seminar (August 2025)
- **The Next Chapter in the Sordid Love/Hate Relationship Between DBs and OSES**
P99 Conf Keynote (October 2024)
Google DANI Academic Seminar (August 2024)
- **What Goes Around Comes Around... And Around...**
CWI Lectures & Dijkstra Fellowship (November 2024)
SF Systems Meetup (October 2024)
University of California – Berkeley Database Group (October 2024)
PGConf NYC 2023 Keynote (October 2023)
- **Why Machine Learning for Automatically Optimizing Databases Doesn't Work**
Databricks (July 2023)
Silicon Valley PostgreSQL Meetup (July 2023)
7 on the Beach 2023 Keynote (May 2023)
Percona Live (May 2023)
PGDay Chicago (April 2023)
BTW 2023 Keynote (March 2023)
- **The Building Blocks for Self-Driving PostgreSQL**
Citus Con 2022 Keynote (April 2022)
- **Do We Still Need People to Write Database Systems?**
Dremio Subsurface 2022 (March 2022)
OSACON 2021 (November 2021)

- **The OtterTune Database Configuration Auto-Tuning Service**
Postgres Vision 2021 (July 2021)
Percona Live (May 2021)
University of Iowa (December 2020)
University of Wisconsin–Madison (November 2020)
National Database Conference of China (October 2020)
- **Self-Driving Databases & My Pregnant Wife: The Hard Parts**
Cornell University (September 2019)
AIDB @ VLDB (August 2019)
Yelp (August 2019)
Oracle (August 2019)
Rockset (August 2019)
Snowflake (August 2019)
Two Sigma (August 2019)
MongoDB (August 2019)
CockroachDB (August 2019)
- **Make Your Database Dream of Electric Sheep: Designing for Autonomous Operation**
Microsoft Research Faculty Summit (August 2018)
Percona Live (April 2018)
Amazon AWS (March 2018)
- **What Non-Volatile Memory Means for the Future of Database Systems**
Two Sigma (July 2017)
MongoDB (July 2017)
- **What Are We Doing With Our Lives?: Nobody Cares About Our Concurrency Control Research**
SIGMOD Keynote (May 2017)
- **How to Build a Non-Volatile Memory Database Management System**
SIGMOD Tutorial (May 2017)
- **Self-Driving Database Management System**
CIDR / Google / DataBricks (January 2017)
- **Peloton: The Self-Driving Database Management System**
DataEngConf (NYC) (November 2016)
PDL Visit Day (May 2016)
- **What Non-Volatile Memory Means for the Future of Database Systems**
In-Memory Computing Summit (May 2016)
- **Andy Pavlo’s Mostly Professional and Collected Presentation about Getting Down & Dirty in OLTP Databases with Intel’s NVM SDV**
ISTC for Big Data Retreat (September 2015)
- **Reducing Replication Bandwidth for Distributed Document Databases**
MongoDB (July 2015)
- **Moving the Abyss: Database Management on Future 1000-core Processors**
NSF Workshop on Exploiting Parallelism and Scalability (June 2015)
- **I Don’t Want to be the Mitt Romney of Databases**
University of Waterloo (November 2015)
IBM Research (September 2015)
University of Michigan (April 2015) *Two Sigma* (July 2015)
- **Database Systems Do Not Scale to 1000 CPU Cores and Other Tales of the Macabre**
Qatar Computing Research Institute (December 2014)
CMU Qatar (December 2014)
New England Database Summit (December 2014)

- **Let's Talk About Storage & Recovery Methods for Non-Volatile Memory OLTP Database Systems**
ISTC for Big Data Retreat (August 2014)
- **H-Store Retrospective**
SIGMOD Jim Gray Doctoral Dissertation Award (June 2014)
- **Cache Rules Everything Around Me**
University of Pittsburgh (February 2014)
Stanford University (January 2014)
Univ. of California, Berkeley (January 2014)
Univ. of California, Santa Cruz (January 2014)
- **OLTP on the NVM SDV: YMMV**
ISTC for Big Data Retreat (December 2013)
- **What's Really New With NewSQL**
Northwest Database Society (October 2013)
Google Kirkland (October 2013)
- **Random Prognostications: An Evening with Andy Pavlo**
HPTS 2013 (September 2013)
- **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)
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 (February 2012)
- **Magical Parallel OLTP Databases**
Massachusetts Institute of Technology (November 2011)
- **Life After the Stonebraker Stores**
HPTS (October 2011)
- **MapReduce and Parallel DBMSs: Together At Last**
Brown IPP Symposium on Cloud Computing (May 2010)
New England Database Summit 2010 (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)

Panels

- **How Should Databases Be Taught in the Future?** (Moderator)
ICDE (May 2026)
- **Magical Machine Learning for Database Tuning** (Moderator)
VLDB (August 2021)
Video: <https://youtu.be/Z-R8q-LakW8>
- **Junior Faculty Panel** (Participant)
CMU/MIT Rising Stars Workshop (November 2016)
- **Computing Futures** (Participant)
CMU CSD50 Celebration (October 2015)
- **Can Big Data Platforms EVER Deliver Real-time Analytics?** (Participant)
BIRTE @ VLDB (August 2015)
- **Mobility and Big Data: Adding an M to the Vs, or not?** (Participant)
IEEE International Conference on Mobile Data Management (June 2015)