Matthew Perron

mperron@mit.edu
+1-(802)-238-3141
matthew-perron.com

Expertise

Cloud Data Systems

My research makes data systems easier for non-expert users to deploy and use. My work focuses on how analytical database systems can use elasticity to meet user performance goals in dynamic environments while maintaining low cost. I have a strong record of finding and implementing creative solutions to hard problems in cloud analytical database systems.

Education

Ph.D. Computer Science, Massachusetts Institute of Technology EECS — 2024
Thesis: Elasticity In Cloud Analytical Database Management Systems
Advisor: Samuel Madden

M.S. Computer Science, Carnegie Mellon University, CSD — 2017

B.S. Computer Science, Rochester Institute of Technology — 2013
Minors: Japanese Language/Japanese Language and Culture
Study Abroad: Sophia University, Tokyo, Japan — 2010-2011

Experience

Senior Researcher, Microsoft Research Data Systems Group — 2024-present

Research Assistant, MIT CSAIL — 2017–2024
Advisor: Samuel Madden
Coauthors: David DeWitt, Raul Castro Fernandez, Michael Cafarella, Tim Kraska, Michael Stonebraker
Research Topics: Cloud Analytical Databases, Query Reoptimization, Cloud Database Benchmarking
Publication Venues: SIGMOD, VLDB, ICDE

- Currently researching buffer pool management in elastic analytical database systems to meet end user performance and cost goals.
- Designed and built elastic scaling systems for analytical database system to minimize cost while retaining elasticity. Designed and built purpose built shuffling mechanism. Published in SIGMOD 2024.
- Designed and built query execution engine atop cloud functions. Published results in SIGMOD 2020.
- Investigated and built research prototype re-optimization mechanism in PostgresSQL, demonstrating benefits over improved cardinality estimation alone. Published in ICDE 2019.
- Worked collaboratively to systematically benchmark several cloud databases. Published results in VLDB 2019.
- Met weekly with two undergraduate research assistants, guided and advised their projects.

Intern, Amazon Redshift — Summer 2020 and 2021
Supervised by Ippokratis Pandis

- Researched techniques to extending Redshift with elastic stateless compute to reduce query latency.
- Received patent 18/171,245 for this work.

Research Intern, Microsoft Research, DMX Group — Summer 2018
Supervised by Srikanth Kandula, Surajit Chaudhuri

- Researched techniques for cardinality estimation using machine learning.

Graduate Technical Intern, Intel Labs, Infrastructure Research Lab — January-August 2017

Engineer, SoftBank Corp, Middleware Platform Division — 2013-2015

Honors & Awards

NSF GRFP Honorable Mention — 2017

Irwin Mark Jacobs (1957) and Joan Klein Jacobs Presidential Fellow 2017-2018

Other Activities

President, MIT Rowing Club 2020-2022
- Responsible for the period during and after COVID pandemic, growing the club to 100 members.
- Planned season details, arranged coaches, and coach payments.
- Recruited club officers, organized regular officer meetings.

Member, CSAIL Researcher Council 2021-2023
- Participated in regular meetings with the lab director.
- Brought concerns from students to the attention of lab administration.

Mentor, MIT Graduate Application Assistance Program 2020
- Advised two students from underrepresented minorities in revising PhD application materials over several meetings.
- Both students now enrolled in PhD programs.

New Member Manager, MIT Rowing Club 2022-2023
- Responsible for communicating with prospective club members.

Teaching Assistant, MIT 6.S080: Software Systems for Data Science Fall 2019
- Ran office hours once weekly for a class of 74 students.
- Designed and graded course assignments
- Gave a lecture

Teaching Assistant, MIT Brave Behind Bars Summer 2023
- Ran twice weekly sessions with small groups of incarcerated students, reinforcing course material.

Talks

Cockroach Labs	January 2020
University of Chicago ChiData Group	April 2020
Databricks	June 2020
IBM	Sept 2020

Publications

[1] **Matthew Perron**, Raul Castro Fernandez, David DeWitt, Michael Cafarella, and Samuel Madden. Cackle: Analytical Workload Cost and Performance Stability With Elastic Pools. *Proc. ACM Manag. Data*, 1(4), Dec 2023

[2] **Matthew Perron**, Raul Castro Fernandez, David DeWitt, and Samuel Madden. Starling: A Scalable Query Engine on Cloud Functions. In *Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data*, SIGMOD '20, page 131–141, New York, NY, USA, 2020. Association for Computing Machinery

[3] Junjay Tan, Thanaa Ghanem, **Matthew Perron**, Xiangyao Yu, Michael Stonebraker, David DeWitt, Marco Serafini, Ashraf Aboulnaga, and Tim Kraska. Choosing a Cloud DBMS: Architectures and Tradeoffs. *Proceedings of the VLDB Endowment*, 12(12):2170–2182, 2019

[4] **Matthew Perron**, Zeyuan Shang, Tim Kraska, and Michael Stonebraker. How I Learned to Stop Worrying and Love Re-optimization. In *2019 IEEE 35th International Conference on Data Engineering (ICDE)*, pages 1758–1761. IEEE, 2019

[5] Andrew Pavlo, Gustavo Angulo, Joy Arulraj, Haibin Lin, Jiexi Lin, Lin Ma, Prashanth Menon, Todd C Mowry, **Matthew Perron**, Ian Quah, et al. Self-Driving Database Management Systems. In *CIDR*, volume 4, page 1, 2017

[6] Joy Arulraj, **Matthew Perron**, and Andrew Pavlo. Write-Behind Logging. *Proceedings of the VLDB Endowment*, 10(4):337–348, 2016

PATENTS

[7] Ippokratis Pandis and **Matthew James Perron**. Selecting Between Hydration-Based Scanning and Stateless Scale-out Scanning to Improve Query Performance, June 22 2023. US Patent App. 18/171,245

Samuel Madden
Professor
MIT CSAIL
32-G938
32 Vassar Street
Cambridge, MA 02139
madden@csail.mit.edu

David DeWitt
Adjunct Professor
MIT CSAIL
32-G428
32 Vassar Street
Cambridge, MA 02139
david.dewitt@outlook.com

Raul Castro Fernandez
Assistant Professor
Department of Computer Science
University of Chicago
Crerar 245
5730 S Ellis Ave
Chicago, IL 60637
raulcf@uchicago.edu

Michael Cafarella
Principal Research Scientist
MIT CSAIL
32-G924
32 Vassar Street
Cambridge, MA 02139
michjc@csail.mit.edu