Andy Huynh

PhD Candidate in Computer Science

huynh.n.andy@gmail.com · ndhuynh.com

Research Interest

ML for systems, systems for ML, Tuning Database Systems, Robust Data Systems

Education

Ph.D. Computer Science 2017 - Expected Summer 2025
Boston University with Manos Athanassoulis

B.Eng Computer Engineering *Magna Cum Laude* with *Distinction* 2014 - 2017
University of Minnesota, Twin Cities

Honors and Awards

IBM Ph.D. Student Fellowship *2020*

Dean's List 6 Semester at University of Minnesota *2014 - 2017*

Professional Experience

Software Engineering Intern at Meta (formerly Facebook Inc.) 06/2022 - 02/2023

Research Intern at NetApp in Advanced Technology Group 06/2019 - 09/2019

Machine Learning Research Intern at Bose in Automotive Group 06/2018 - 09/2018

Firmware Engineering Intern at Medtronic 06/2017 - 09/2017

Publications

1. *AXE: A Task Decomposition Approach to Learned LSM Tuning*
   Huynh A, Saha A, Chaudhari H.A., Athanassoulis M.
   Under Submission, 2025

2. *Benchmarking Learned and LSM Indexes for Data Sortedness*
   Raman A, **Huynh A**, Lu J, Athanassoulis M
   Proceedings of the Tenth International Workshop on Testing Database Systems, 2024

3. *Towards Flexibility and Robustness of LSM Trees*
   Huynh A, Chaudhari H.A., Terzi E, Athanassoulis M.
   The VLDB Journal, 2023

4. *Endure: A Robust Tuning Paradigm for LSM Trees Under Workload Uncertainty*
   Huynh A, Chaudhari H.A., Terzi E, Athanassoulis M.
   Proceedings of the VLDB Endowment, 15, 8 (April 2022), 1605–1618.

5. *Modeling of Swine Diaphragmatic Tissue Under Uniaxial Loading*
   Huynh A, Molina Espinosa M, Lobo Fenoglietto F, Singal A, Iaizzo P.
   ASME Journal of Medical Devices, 9(3), 3-3. 2015.

Presentations and Posters

North East Database Day, Main Track Presentation, March 2023
RedHat Research Days, Presentation, Feb 2023
"Endure: A Robust Tuning Paradigm for LSM Trees Under Workload Uncertainty" VLDB, Presentation, September 2022
RocksDB Internal Team, Presentation, August 2022
PingCAP Community Meetup, Presentation, May 2022
RedHat Greater New England Research Interest Group Meeting, Presentation, May 2021
North East Database Day 2020, Poster

Service

Organizations

SIGMOD 2024 - External Reviewer
IEEE Big Data 2023 - External Reviewer
SIGMOD 2022 - Reproducibility Reviewer
SIGMOD 2023 - External Reviewer

Department

2023 PhD Admits Organizing Committee
2022 MiDAS Seminar Organizer

Teaching

CS 460: Introduction to Database Systems	Fall 2020
CS 591A: Data Systems Architecture	Spring 2020
CS 591P: Object-Oriented Programming in Java	Fall 2018, Spring, Summer, Fall 2019
CS 112: Introduction to Computer Science II	Spring 2018, Summer 2018
CS 111: Introduction to Computer Science I	Fall 2017
PHYS 1302W: Introductory Physics for Science and Engineering II	Spring 2015

Main Work Descriptions

Learning to Tune Data Systems When a systems designer is given a workload or application and asked to optimize a database for the best possible performance, they often do not have the tools to help recommend the best configuration. Rather, they use random testing and past experiences to tune their database. I am exploring machine learning methods to help solve the tuning problem; I am particularly interested in the most efficient way to learn from physical database executions. Systems are unique in that to gather example data, we must pay the price of physically waiting for an executed workload to be finished. My research aims to learn efficiently from past examples or to learn from surrogate cost models and translate this to high-performing configurations.

Robust LSM Trees: I am implementing a new robust tuning paradigm for LSM Trees. By framing the tuning problem as an optimization problem that takes into consideration uncertainties in the input factors, we can find a design that, when deployed, is robust to changes in the expected workload and resources. I utilize Python to create an optimization framework that solves the modeled problem, then pipes the design decisions into a C++ framework that exposes tuning knobs of RocksDB. An instance of the database is deployed, and we can test this on randomized or real world workloads.

Mentoring

Suhruth Vuppala - Undergraduate Student - Boston University	2024
Lucas Yoon - Undergraduate Student - Boston University	2024
Anwesha Saha - Master Student - Boston University - Now PhD	2023–2024
Jida Li - Undergraduate Student - Boston University	Spring 2023
Caterina Caravaggio - Masters Student - Universit‘a di Bologna	Summer 2020