

HPTS 2024

~20 Students in ~20 mins

Students @ HPTS'24

Robert Bayer, IT University of Copenhagen

Achilles Benetopoulos, UC Santa Cruz

Matthew Butrovich, Carnegie Mellon University

Audrey Cheng, UC Berkeley

David Chu, UC Berkeley

Chris Douglas, UC Berkeley

Shahrzad Haji amin Shirazi, UC Riverside

Shadaj Laddad, UC Berkeley

Tianyu Li, MIT

Mehnaz Tabassum Mahin, UC Riverside

Hamish Nicholson, EPFL

Liana Patel, Stanford

Conor Power, UC Berkeley

Ties Robroek, IT University of Copenhagen

Viktor Sanca, EPFL

Daniel ten Wolde, CWI

Bobbi Yogatama, UW - Madison

Geoffrey Yu, MIT

Xinjing Zhou, MIT

Tobias Ziegler, TU Darmstadt

Robert Bayer, IT University of Copenhagen

- 2nd year PhD student, supervised by Pinar Tözün
- Machine learning at the edge
 - Image analysis on satellites - Gong show talk
 - Performance analysis
 - Resource-allocation
- Very amateur triathlete and pizza lover



Achilles Benetopoulos, UC Santa Cruz

- Fourth year PhD student, advised by Peter Alvaro
- Research focus on novel programming paradigms for distributed systems
 - Currently working on magpie, a distributed single-level store and runtime for HTAP workloads
- Interests: cooking, reading, travelling



Matt Butrovich (CMU)

- Escaped from [Andy Pavlo](#) in May
- [Thesis](#) explored DBMS and OS co-design, primarily using [eBPF](#)
- Joined Apple in August
 - Open source work on [Spark](#), [DataFusion](#), and [DataFusion Comet](#)
- Find us hiking (just spent 3 days in Big Sur) or cheering on the Dodgers at road games



New River Gorge National Park



Dodgers @ PNC Park

Audrey Cheng (UC Berkeley)



5th year PhD student, advised by Ion Stoica and Natacha Crooks

Working on **optimizing performance for modern transaction processing**

- Towards optimal transaction scheduling (VLDB' 24)
- Mammoth transactions on graph data (VLDB' 24)
- Optimizing caching for transactions (OSDI '23)
- TAOBench: open-sourcing Meta's transactional workloads (VLDB'22)
- RAMP-TAO: read-atomic transactions for Meta (VLDB '21)

Snorkeling enthusiast



David Chu (UC Berkeley)



- 5th year PhD student, advised by Natacha Crooks & Joe Hellerstein
- Working on **automatically rewriting distributed protocols** such as Paxos or PBFT (or any cloud program) so they can scale up
 - Optimizing Distributed Protocols with Query Rewrites (SIGMOD '24)
 - Bigger, not Badder: Safely Scaling BFT Protocols (PaPoC '24)
- My fiance's birthday was yesterday, so we went to LA. I will be here later today at the gong show!



Chris Douglas, UC Berkeley

- Third year PhD student advised by Joe Hellerstein
- Previously, worked on distributed storage/resource management infrastructure for large scale analytics at Yahoo!, Microsoft, and Apple.
- Interests include incremental computation and concurrency control. Storage-only log-structured tables (e.g., Iceberg)



Shahrzad Haji amin Shirazi, UC Riverside

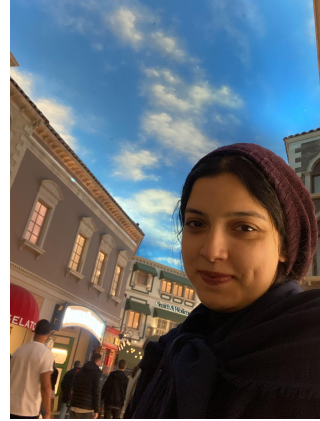
- 5th year PhD student, advised by Vassilis Tsotras & Michael Carey.
- Have been working on **query optimization techniques** for Big Data in **Apache AsterixDB**.



- In today's fast-paced world:
 - asking someone to "Come and run your query to get a result" is ***boring!***
 - the BAD extension of AsterixDB is ***not boring!***

(**B**ig **A**ctive **D**ata System)

- It is Ready to cope with the today's big data challenges.
- Interests:
 - Traveling
 - Escape rooms (Already escaped from all the ones within a 41-mile radius of my house!:))



Shadaj Laddad, UC Berkeley

- Final (!) year PhD student, advised by Alvin Cheung & Joe Hellerstein
- Working on **programming models** for building *correct, high performance* distributed systems with ease
 - What if quorum in Paxos was a library function?
- Intersection of programming languages, distributed systems, formal methods, and (of course) databases
 - Suki: Choreographed Distributed Dataflow in Rust (CP '24)
 - Keep CALM and CRDT On (VLDB '23)
 - Katara: Synthesizing CRDTs with Verified Lifting (OOPSLA '22)
- Interests: pickleball, k-pop, theme park history
- David is my ride, I will be here soon!



Tianyu Li, MIT

- Final year PhD student advised by Sam Madden
- I build modern abstractions for the modern cloud
 - Durable Execution and Resilient Cloud Programming (Main line of my thesis work, in SIGMOD 21',23', CIDR 24', and more)
 - Virtualized cloud data management (HPTS Talk! Jointly with Geoffrey Yu, in VLDB 23', 24')



● ~~HPTS ruined my music career~~



Mehnaz Tabassum Mahin, UC Riverside

- 5th year PhD student, advised by Vassilis Tsotras and Michael Carey
- Developed a mixed workload benchmark for NoSQL
 - CH3 (IEEE Big Data '22)
 - <https://github.com/couchbaselabs/ch3>
- Worked on analyzing large-scale spatio-temporal data
 - Dwell Regions: Generalized Stay Regions (ACM TSAS '23)
- Working on enhancing cost-based optimizer (CBO)
 - Apache AsterixDB and NoSQL platforms
- Hobbies: collecting stamps, any good reads!



Hamish Nicholson, EPFL

- 4th year PhD advised by Natassa Ailamaki
- Formerly SWE @ Ocident
- Working on data movement on modern hardware

- When I'm not in the office, I'm in the mountains

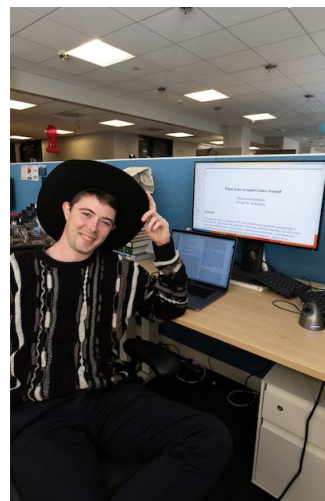


Liana Patel, Stanford

- Fourth year PhD Student, co-advised by Matei Zaharia & Carlos Guestrin
- Research Interests: data systems for enabling knowledge intensive AI apps
 - ACORN: an index for combining vector search + relational queries [SIGMOD '24]
 - LOTUS: a declarative programming model & query engine for AI-based data processing
 - <https://github.com/TAG-Research/lotus>
 - TAG: a paradigm for serving NL queries over databases
 - --> gong show talk later!
- Hobbies/Interests: squash, swimming, pickleball, reading

Conor Power, UC Berkeley

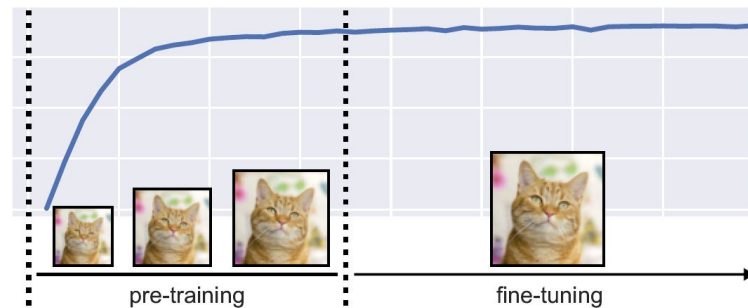
- 4th year PhD student advised by Joe Hellerstein
- Former senior software engineer on Cosmos and Azure Storage at Microsoft
- Research on the role of algebraic properties in distributed data systems:
 - How exactly do the CALM Theorem and CRDTs relate to one-another?
 - What optimizations do properties like commutativity, distributivity, invertibility, and idempotence offer?
 - Tooling to help developers identify properties of their code and build more efficient distributed systems.
- Love board games and reality tv
- Co-host of the Thinking About Computers Podcast
youtube.com/@tatacpodcast
- Co-organizer of SF Systems Club lu.ma/sf-distsys



Ties Robroek, IT University of Copenhagen

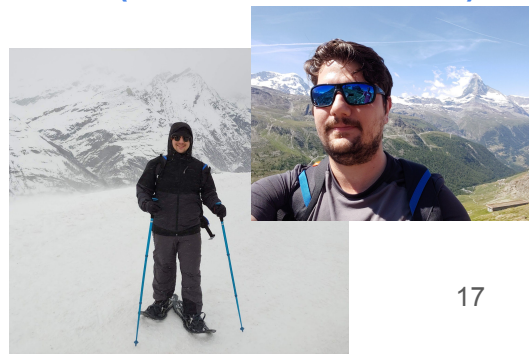
- Just handed in PhD thesis under Pinar Tözün
- Training more models with less resources
- Focus on the data pipeline

- Bouldering and triathlon enthusiast



Viktor Sanca, EPFL

- Natassa Ailamaki's 30th PhD graduate (a few weeks ago)
- Thesis marries approximate analytics & modern systems
 - Efficient ad-hoc sampling for scale-up systems + new applications
[DAMON'22, SIGMOD'23, SIGMOD Record'24, SIGMOD Research Highlight Award]
 - Speculative execution for analytics for better || resource use
[CIDR'21, AIDB'23] - Just like HW one, avoid waiting for results: speculate-repair
 - Relational model++ and operators for ML-powered analytics - **Gong show talk!**
[ICDE'23, TKDE'24, ICDE'24, DAMON'24] - vector-relational systems and optimizations (before it was cool?)
 - Impact of system and hardware heterogeneity [ADMS'23, TPC-TC'23]
- EPFL→Oracle, Bay Area (soon) - hints & tips appreciated!
- Switzerland + hiking is always good option (any season)

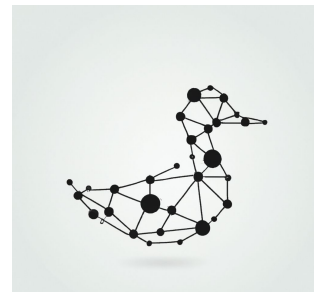


Daniël ten Wolde (Database Architectures, CWI)



- 2nd year PhD student, advised by Peter Boncz
- Integrating **Graph Processing** in Relational Query Engines

- Developing **DuckPGQ** (CIDR '23 & VLDB demo '24)
 - Graph extension for DuckDB
- Support SQL/PGQ (SQL:2023)
- Path-finding operator
 - novel parallelism model
- Adaptive Factorized Query Processing
- Worst-Case Optimal Join Algorithms



- Avid sport watcher (Going to World Championship road cycling next week)

[DuckPGQ](#)
[documentation:](#)



Bobbi Yogatama, UW - Madison

- Final-year PhD student, advised by Xiangyao Yu.
- Working on **accelerating database using GPUs.**
 - SIGMOD'22, VLDB'22, DaMoN'23, VLDB'24
 - Check out **my talk on Tuesday!**
- Interest: I do competitive Ballroom Dancing



Geoffrey Yu, MIT



- 5th year PhD student, advised by Tim Kraska
- Working on **virtualizing** and **automatically designing** cloud data infrastructures [VLDB '24 and '23]
 - Talk on Tuesday w/ Tianyu!
- Previous work on key value stores [VLDB '22], GPU performance prediction for DNN training [ATC '21]
- Hobbies: Photography, hiking, ice skating, collecting vinyl

Xinjing Zhou, MIT

- 4th year PhD student, advised by Michael Stonebraker
- Working on co-designing DBMS with OS using virtualization
 - Talk this afternoon
- Previously worked on transactions [VLDB 22], indexing / buffer management on tiered memory [VLDB 19, CIDR 23, SIGMOD 21].

- Hiking, soccer watcher



Tobias Ziegler, TU Darmstadt

Postdoctoral Researcher

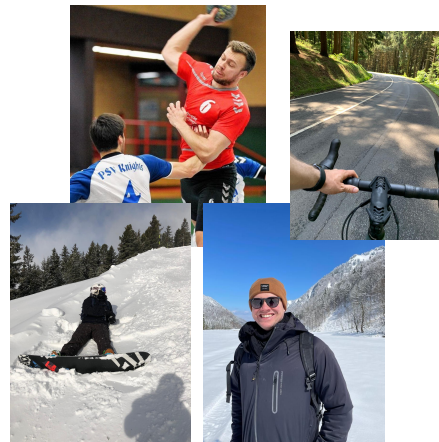
TU Darmstadt → soon at TU Munich

Research Focus:

- **Thesis:** Scalable OLTP over Fast Networks (RDMA)
- **Current Work:** Cloud DBMS/Services - Tackling “missing” primitives for the cloud
e.g., journaling/log service
-

Interests:

- Coffee enthusiast
- “Security through obscurity
- Like all kinds of sports



Now go and mingle! ...

Happy HPTS'24