

DESMOND CHEONG

Email: desmond@desmondcheong.com · Website: desmondcheong.com · Github: github.com/desmondcheongzx

EXPERIENCE

Eventual

Software Engineer (5th engineer)

San Francisco, CA, USA

Jul 2024 – Present

- Designed a 10x faster parallel CSV reader that's on par with state-of-the-art, and outperforms major engines like Apache Spark, to remove data ingestion as a bottleneck
- Adding statistics and cost-based optimization to improve query planning

Databricks

Software Engineer (Photon + Geospatial teams)

San Francisco, CA, USA

Jul 2022 – Jun 2024

- Shipped the fastest known implementation of the H3 global grid indexing system, with up to 12x performance improvements over the widely-used open source library, and attracted over \$2M ARR in workload usage within a year of release
- Added Geospatial data types and algorithms to win multi-million dollar customer accounts attached to Spatial SQL support
- Owned operations for Geospatial engineering, which included creating and tracking usage dashboards, setting up alerts for correctness issues, designing rollback and feature-control mechanisms, and promoting usage via in-product messaging
- Spearheaded prototyping of automated statistics computation on data during table creation and incremental writes to improve Databricks' out-of-the-box query performance
- Delivered a performant and concurrency-safe solution for long-term storage of automated statistics
- Brought automated statistics to public-preview-readiness by resolving the main performance blockers and enhancing the feature to support most data types, including nested types
- Added a new semi-structured data type, unlocking customer migrations and future JSON-based workloads

Linux Foundation

Linux Kernel Intern

Remote

Jun 2021 – Aug 2021

- Improved the security of the mainline Linux kernel by analyzing crash logs, and correcting architecture designs, to fix dozens of memory corruption and potential denial-of-service errors
- Collaborated with maintainers to debug a wide range of Linux subsystems, including the graphics stack, the Bluetooth stack, core memory management, multiple file systems, and other device drivers

Intellex

Data Science Intern

Singapore

May 2020 – Aug 2020

Findify

Business Development Intern

Stockholm, Sweden

Jun 2019 – Jul 2019

LEADERSHIP

CSCI 1270: Database Management Systems

Head Teaching Assistant

Mar 2021 – Dec 2021

- Led 9 teaching assistants to create Brown's first integrated and fully-functional database system project from scratch, featuring cache management, B+ tree indexing, query processing, transactions, concurrency control, and crash recovery

PROJECTS

TRustDB

- Designed a highly performant database in Rust to manage high-cardinality time series data with boolean logic-based query optimizations and storage cost reduction through index compression with finite state transducers

CPython (Open Source Contributor)

- Enhanced the Python programming language to support cache invalidation for modules imported from zip archives

SKILLS

Performance & Internals Languages

SQL & DataFrame Engines, Apache Spark, Java Virtual Machine, Linux
C++, C, Rust, Go, Python, SQL, Scala, JavaScript, OCaml, Racket, Clojure

EDUCATION

Brown University

Sc.B. in Computer Science (magna cum laude), GPA 3.97/4.0

Sep 2018 – May 2022

Coursework: Distributed Computer Systems, Computer Networks, Database Management Systems, Topics in Database Management (Graduate), Operating Systems Laboratory