

SARTHAK MAKHIJA

Principal Architect

@ sarthak.makhija@gmail.com

<https://www.linkedin.com/in/sarthak-makhija/> <https://tech-lessons.in>

Pune, India 411028



SUMMARY

Principal Architect with 15+ years of expertise driving the design of scalable distributed systems, search platforms, and storage architectures.

Currently leading the architecture of a petabyte-scale distributed search system at Caizin, focusing on high-volume document management, advanced tombstone stores, and significant cost optimization on top of Quickwit.

Demonstrated track record in storage internals, having previously led an engineering team to build a high-performance, strongly consistent Key/Value storage engine in Go.

Recognized industry contributor, validating core concepts for the book "*Patterns of Distributed Systems*" and publishing deep-dive articles on persistent memory.

Dedicated technical educator and open-source developer, building robust tools for caching and query execution, while leading hands-on workshops on LSM storage engines and advanced refactoring.

EXPERIENCE

Principal Architect

2024 - Present

Caizin

Pune, India (Remote)

A technology consultancy that provides software design and development services

- **Spearheading the architecture of a petabyte-scale distributed search system** built on top of Quickwit, strategically designed to optimize performance while drastically reducing infrastructure costs.
- **Driving open-source innovation** by authoring and maintaining relop, an in-memory SQL query engine written in Rust that features custom implementations for lexical analysis, parsing, logical planning, and execution.
- **Fostering an engineering culture** of excellence by conducting deep-dive, hands-on workshops on the internals of LSM-based storage engines and mentoring engineers on complex distributed system design.

Lead Consultant

11/2016 - 2024

ThoughtWorks

Pune, India (Remote)

A technology consultancy that provides software design and development services

- **Architected and built a strongly consistent, distributed key/value storage engine** from the ground up, implementing two-phase commit, partitioning, and consistent hashing to guarantee data reliability and high performance.
- **Spearheaded the 'Nudge' ecosystem on AWS Serverless**, streamlining core business workflows and driving significant infrastructure cost reductions.
- **Upskilled engineering teams** by designing and delivering **hands-on workshops** covering Go application profiling, **LSM-based storage architectures**, and **advanced refactoring** techniques for legacy systems.
- **Authored and maintained open-source tools** (including Clearcheck, CacheD, and Goselect) and contributed to industry literature by validating architectural patterns for the book "*Patterns of Distributed Systems*".
- **Published extensive technical content**, including a specialized series on Persistent Memory for the Kt. Academy blog and deep dives into storage engine internals on tech-lessons.in.

Senior Software Engineer

08/2014 - 10/2016

Citibank

Pune, India

A multinational investment bank and financial services corporation

- **Architected and developed** a microservices-based reporting application using Spring Boot, Spring REST, Hibernate, and AngularJS
- **Built the N99 payment processing system** using Spring MVC, Spring REST, and Spring Data JPA, integrating multiple payment gateways
- **Mentored the team** in adopting best practices through workshops on Java 8, Spring Boot, and robust testing methodologies

KEY ACHIEVEMENTS

🏆 Industry Publications

Recognized for validating various patterns for the book "*Patterns of Distributed Systems*" and authoring articles on Persistent Memory.

⚡ Open-source Contributions

Developed several open-source tools contributing to industry advancements and knowledge sharing.

SKILLS

Languages

Go · Rust · Java · Zig

Architecture & Core Internals

Distributed Systems · Storage Engines ·

Database Internals · Apache Lucene ·

Query Processing · Raft ·

Deterministic Simulation Testing

Engineering Practices

Advanced Refactoring ·

Legacy System Modernization ·

Clean Code Architecture ·

Performance Profiling

Infrastructure & Cloud

AWS · Kubernetes · Microservices

Data Technologies

PostgreSQL · Redis · Kafka

TRAINING / COURSES

Internals of Storage Engines

<https://github.com/SarthakMakhija/go-lsm-workshop>

Gamifying Refactoring

<https://gamifying-refactoring.github.io/>

EXPERIENCE

Application Developer 07/2008 - 07/2014

[TCS](#) Ahmedabad, India

A leading global IT services, consulting and business solutions organization

- **Engineered and scaled core VAT processing systems** deployed across multiple state governments.
- **Architected a strict XML validation framework** using Apache Betwixt to guarantee data integrity across critical tax applications.
- **Accelerated team velocity** by building custom internal developer tools, including an automated JSP Generator.
- **Designed a reusable state-management component** ('Interim Save') to optimize data persistence and user experience for complex, large-scale forms.

OPEN-SOURCE PROJECTS

relop (Rust) 2026

Pune, India (Remote)

An open-source, in-memory query engine featuring custom implementations for lexical analysis, parsing, logical planning, and SQL query execution.

(<https://github.com/SarthakMakhija/relop>).

Clearcheck (Rust) 2025

Pune, India (Remote)

A library that simplifies assertion statements in Rust, enabling chained assertions for improved readability and self-documenting test cases.

(<https://github.com/SarthakMakhija/clearcheck/>)

CacheD (Rust) 2025

Pune, India (Remote)

A high-performance in-memory cache inspired by Ristretto utilizing LFU (Least Frequently Used) eviction strategy. (<https://github.com/SarthakMakhija/cached/>).

Goselect (Go) 2024

Pune, India (Remote)

A library that provides a SQL-like select interface for files, supporting various scalar and aggregate functions. (<https://github.com/SarthakMakhija/goselect>).

EDUCATION

Bachelor of Technology 08/2004 - 05/2008

[Nirma University](#) Ahmedabad, India