

TANUJ GUPTA

P: (831) 266-5441 | tanuj.professional@gmail.com | [WEBSITE](#) | [MEDIUM](#) | [LINKEDIN](#) | [LEETCODE](#) | [GITHUB](#) | [GITHUB2](#)

EDUCATION

University Of California, Santa Cruz | Research Assistant

Masters in Computer Science (GPA 4.0+/4.0)

Coursework: Distributed Systems, Design, and Implementation of Databases, Deep Learning, High-Performance Computing

Santa Cruz, California

September 2021 - June 2023

Birla Institute of Technology and Science, Pilani

Bachelors in Computer Science

Hyderabad, India

August 2015 - July 2019

EXPERIENCE

[MetaJungle LLC](#)

Missouri

Team lead / Founding Software Engineer

May 2023 - Present

- Leading a team of 9 engineers to develop a Web3 portal using **Next.js**, **Node.js**, **Express.js**, **GraphQL**, and **AWS** featuring wallet, creator, NFT, and collection pages with integrated social media, aggregating data from blockchains and the internet.
- Designed Blockchain Index using **Cassandra**, **MongoDB**, **Hadoop**, **Ethereum**, and **Redis** for efficient storage and querying of 20 million blockchain transactions with more than 1 billion Nft and ERC20 Transfers.
- Built distributed data pipelines for aggregating NFT data from blockchains and the internet at the scale of a million daily operations using **Airflow**, **Spark**, and **NodeJS**.
- Implemented **AWS** Based cloud infrastructure as code using **Terraform** with services like **ECS**, **ELB**, **ASG**, etc
- Developed Chat Service for all pages on Metajungle using **WebSockets**, load balancers, change streams, and **Kafka** queue.
- Developed an **advanced (RAG)** system to dynamically enhance NFT data generation, significantly improving the accuracy and depth of information available to customers for making informed decisions

[Meta \(Facebook\)](#)

Menlo Park, California

Software Engineer Internship

June 2022 - September 2022

Tech stack: Javascript, **React.js**, **XHP**, **GraphQL**, **XControllers**, **Hack PHP**, **MySQL**, **Scuba**, **Presto**, **ODS(Time series)**, **Data Swarm**.

- Developed two end-to-end applications in **React.js** and **Hack PHP** to carry the Profile to Profile+ transition on the Facebook app.
- Build data pipelines using the **Data Swarm** framework to create smaller aggregated tables to reduce GraphQL response time.

[Airtel X Labs](#)

Gurgaon, India

Senior Software Engineer

May 2020 - August 2021

- Led development and maintenance of FTTH(Fiber To The Home) project spanning three microservices using **Java Spring Boot**.
- Streamlined communication between microservices by introducing **Kafka**-based microservice orchestration,

[OYO Group](#)

Gurgaon, India

Data Engineer

June 2019 - May 2020

- Developed big data pipelines, batch and real-time, to sync data from various micro services into our hive-based data lake for analysis using **Spark**, **Airflow**, **Presto**, **Kafka**, **Kafka Streams**, **S3**, **Hadoop**, and **Hive**.
- Optimized data lake UPSERT efficiency using **Apache Hudi** and **AWS S3**-based Bloom Filters for reduced SLA.
- Developed APIs and wrappers in **Java Spring Boot** enabling seamless data migration between sources for developers.
- Migrated **Kafka** and **Airflow** clusters to **Kubernetes** to optimize AWS resource use by 20% and cost by \$40 a day.

[HDFC Bank](#)

Gurgaon, India

Software Engineer Intership

January 2019 - June 2019

- At HDFC Life's ADCT team, developed features to streamline life insurance processes with **React.js** and **Node.js**

RESEARCH / PROJECTS

Large-Scale Graph Processing with Actors | C++(**OpenSHMEM** and **HCLib**) | [Github](#) | [link](#)

Jan 2023 - June 2023

Designed and implemented an actor-based graph library kernel for distributed graph construction and generation using **HCLIB** library written in **OpenMP** and C++. implemented distributed Breadth-First Search (BFS) algorithms using **AGL**.

Rocks DB Embedded key-value database(Developed by Meta) | C++(**Bloom filters**) | [Github](#)

Jan 2022 - March 2022

Proposed a probabilistic constant time solution to reduce high read latency for high-read/low-write data in LSM Tree based Rocks DB, an embedded key-value store. We solve this problem using probabilistic data structures(**Bloom filters**).

SKILLS

Distributed Systems, Node.js, React.js, Next.js, Spring Boot, GraphQL, REST APIs, Design Patterns, Multi-Threading, HTML, JavaScript, CSS, Tailwind CSS, AWS, Kafka, Docker, Kubernetes, MongoDB, Cassandra, PostgreSQL, Redis, ETL, Airflow, Apache Spark, Hadoop, Presto, Hive, LangChain, Transformers, RAG Systems, Pinecone, Solidity, Ethereum, EVM, Geth Nodes