

TANUJ GUPTA

P: (831) 266-5441 | tanujinusa@gmail.com | [LINKEDIN](#) | [LEETCODE](#) | [GITHUB](#) | [GITHUB2](#) | [Website](#)

EDUCATION

The University Of California, Santa Cruz | Research Assistant

Masters in Computer Science (GPA 4.0+/4.0)

Santa Cruz, California

Sept 2021 - June 2023

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

Birla Institute of Technology and Science, Pilani

Bachelor of Engineering in Computer Science

Hyderabad, India

Aug 2015 - July 2019

EXPERIENCE

MetaJungle LLC

Team lead / Founding Software Engineer

Missouri

2023 - Present

Tech Stack: Next JS, Node JS, Web3.JS, Solidity, Hadoop, MongoDB, Cassandra, Airflow, AWS ECS, Elastic Search, Jenkins

- Leading a team of 7 engineers to develop a groundbreaking one-stop portal using **Next JS**, **Express JS**, and **AWS** where people can interact with their NFTs by aggregating data from various marketplaces, 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

Meta (Facebook)

Software Engineer Internship

Menlo Park, California

2022

Tech stack: 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

Senior Software Engineer

Gurgaon, India

2020 - 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 Rooms_

Data Engineer

Gurgaon, India

2019 - 2020

Tech Stack: Java, Java Spring Boot, Spark, Kafka, Spark, Presto, Hive, AWS S3, Hadoop, and Airflow

- 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**, **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.

ACADEMIC 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 an OpenMP and C++. implemented distributed Breadth-First Search (BFS) algorithms using AGL.

Rocks DB Embedded key-value database(Developed by Meta) | 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 Rocks DB, an embedded key-value store. We solve this problem using probabilistic data structures(Bloom).

SKILLS

Databases: PostgreSQL(Debezium) | MongoDB | Oracle | MySQL | TAO

API: REST | GraphQL

DevOps: Docker | Kubernetes | Jenkins | Git | Mercurial

Languages/Frameworks: Java | Python | C++ | Spring Boot | Node JS | React JS | Hack PHP

Big Data: Hadoop | Presto | Scuba | Hive | Apache Spark | Kafka | AWS | DataSwarm | Zookeeper

OPL Stack: LangChain | OpenAI | Streamlit | LangChain | Pinecone