Sahil Parekh

+1 445-256-2439 | sahilparekh08@gmail.com | linkedin.com/in/parekh-sahil | sahilparekh08.github.io

Education

University of Pennsylvania, School of Engineering and Applied Science **Aug 2023 – May 2025**
Master of Science in Engineering, Computer and Information Science (Concentration: Systems) CGPA: 3.83/4.00
Courses: Software Systems, Internet and Web Systems, Foundations of Streaming Computations, Analysis of Algorithms, ML

University of Mumbai, Sardar Patel Institute of Technology **Aug 2016 – May 2020**
Bachelor of Engineering, Information Technology CGPA: 9.53/10.00
Courses: Distributed Systems, Operating Systems, Advanced Databases, Advanced DSA, Computer Networks, Cloud Computing

Skills

- **Languages:** Java, C, C++, Python, SQL, bash, JavaScript, HTML, CSS, R, NoSQL, Rust
- **Frameworks and Libraries:** Spring, JUnit, Mockito, gRPC, Protobuf, PyTorch, Node.js, Express.js, Apache Kafka, Apache Spark, Apache Avro, Apache Lucene, Hadoop, Pandas, scikit-learn, jQuery
- **Technologies:** Jenkins, MySQL, Oracle DB, MongoDB, EhCache, Docker, Kubernetes, AWS, Oracle BDB, FIX Engine, REST
- **Tools and OS:** Git, Gradle, Maven, GDB, Control-M, Dynatrace, Postman, Pentaho, Jira, Linux, Unix, LITMUS-RT, Windows
- **Certifications:** AWS Solutions Architect - Associate (*Amazon*), Applied Algorithms and Data Structures (*Morgan Stanley*)

Work Experience

University of Pennsylvania, Research Assistant **Jun 2024 – Aug 2024**
- Built EDF-based real-time scheduling algorithms for Linux Kernel in C on top of LITMUS-RT for multi-core systems
- Leveraged Intel-CAT and MemGuard to improve schedulability of real-time tasks on multi-core platforms

JP Morgan Chase & Co., Software Engineer II **Feb 2023 – Jul 2023**
- Redesigned a monolithic application into microservices deployed using Docker, boosting throughput by 250%
- Led an Agile team of 3 to develop a micro-batch processing service processing ∼1.4m messages per day per server
- Created Kafka and REST endpoints using Spring MVC for trade report correction, reducing over-reporting by 45%

JP Morgan Chase & Co., Software Engineer **Aug 2020 – Jan 2023**
- Conceptualized a low-latency real-time stream processing framework in Spring, achieving a 60ms latency per Avro payload
- Reduced peak query execution time on Oracle DB by 60% and attained a 90% cache-hit rate using EhCache
- Formulated an exception-management service to aid telemetry, helping decrease trade exceptions by 75% over 3 months

JP Morgan Chase & Co., Software Engineer Intern **Jan 2020 - Jul 2020**
- Established CI/CD with Jenkins, automation with Control-M and monitoring with Dynatrace, saving 9 developer hours per week
- Benchmarked the throughput of Kafka, TransactTools FIX Engine, and Oracle BDB on a 3-pod replica set on Kubernetes

Projects

Carbon-Aware Scheduling for Data Centers *(Master's Thesis at University of Pennsylvania)* **Aug 2024 – Present**
- Designing scheduling algorithms and resource allocation strategies for workloads in data centers to minimize carbon emissions
- Analyzing approaches to model microservice DAGs and predict green energy availability and power consumption

Crawl and Order, a Distributed Search Engine (Description) (GitHub) **Aug 2024 – Dec 2024**
- Built a scalable, distributed computing framework to crawl, index and PageRank 100k web pages on the internet
- Constructed an inverted indexed webpage content and stored it on a distributed fault-tolerant Key-Value store in Protobuf format

PennCloud, a Distributed Google Suite Replica (Description) (GitHub) **Jan 2024 – May 2024**
- Architected a replicated, durable and causally consistent distributed storage system in C++ with gRPC for IPC
- implemented health status monitoring and load balancing algorithms to handle scaling of servers in under 1.5s

Spotify Trending Song Prediction (Description) (GitHub) **Jan 2024 – May 2024**
- Conducted EDA using Apache Spark and designed DNN and Decision Tree-based models, both achieving > 85% accuracy

Kafka Streams Application (Description) (GitHub) **Jan 2024 – Mar 2024**
- Automated Kafka cluster setup using bash scripts and developed stream processing pipelines using Kafka Streams API

Algorithmic Strategies for Competitive Games (Description) **Aug 2023 – Dec 2023**
- Implemented path-finding, ray-tracing, and multi-threaded heuristic algorithms, achieving 260ms latency per gameplay move

Learning Management System *(with Tata Institute of Social Sciences)* (Description) **Dec 2020 – Aug 2021**
- Developed backend APIs with Node.js and Express.js, optimizing for a request latency of 150ms
- Configured AWS VPC, DynamoDB, S3, EC2, and CloudWatch to handle 100k concurrent connections

Securities References Data Management System *(with Morgan Stanley)* **Jan 2018 – Apr 2018**
- Designed an MVC web app using AngularJS, Node.js and MongoDB to display real-time stock trends, updating every 30s

Publication

An Approach to Reducing Uncertainty Problem in Network intrusion Detection Systems (Paper)
- Collected network packets using Wireshark and implemented a Genetic Algorithm in Python and R, reducing feature space by 62%
- Trained a Deep Neural Network, Depth-limited Random Forest, and KNN models, achieving an F-1 score above 0.9