Ravi Prakash

Professional with 7+ years of full time experience in Backend Software Development (Golang, Python, Java), MLOps, IOT across several product-based startups. Ability to hack around multiple technologies and build highly scalable, low latency, concurrent, distributed systems, RESTful services

9868576600, prakashravip1@gmail.com, Github, LinkedIn, Blog

TECHNICAL SKILLS

Programming Languages: Golang, Python, Java, C++, C

Databases: Cassandra, Bigtable, Neo4J, DynamoDB, Apache Pinot, MongoDB, ElasticSearch, Kafka, RabbitMQ, MySQL, Postgres, Bigquery, Redis, Aerospike

Frameworks and Libraries: Spring-Boot, Go-gin, go mux, gorm, Pandas, Django, Flask, FastAPI, Sanic, Keras,

Computer Vision, Asyncio

Cloud Platforms: AWS, Azure, GCP

Containerization and Orchestration: Docker, Kubernetes

Monitoring and Observability: Prometheus, Datadog, Grafana

APIs and Protocols: REST, gRPC, JSON, XML

EXPERIENCE

AiStrike, Pune — Senior Engineer

Dec 2024 - Present

- Architected and implemented enterprise-scale microservices using Spring Boot and OpenJDK 17, integrating PostgreSQL, Neo4j, and MongoDB databases in a Kubernetes environment, while ensuring high availability and scalability
- Led development of mission-critical security systems including a Risk Booster platform with advanced Neo4j query capabilities and an AI-powered Compliance Questionnaire system leveraging AWS S3
- Engineered automated AWS infrastructure solutions including multi-account onboarding and cross-regional monitoring systems, improving operational efficiency and resource management
- Implemented comprehensive security enhancements across Docker containers and developed a centralized vulnerability management platform integrating multiple scanning sources
- Built scalable threat intelligence systems including a domain reputation service managing 1M+ domains and an advanced correlation engine for security analysis
- Optimized system performance through Redis caching implementation and Kafka-based real-time data synchronization, significantly improving response times

Brightedge, Remote — Software Engineer 3

Aug 2023 - Dec 2024

- Collaborated with the VIP team to resolve customer support JIRA tickets and addressed a critical bug in PDF report generation.
- Contributed to the PBR team by maintaining and enhancing backend systems, improving recommendation accuracy from 92% to 98% for global and corporate accounts.
- Developed and optimized internal admin tools and backend systems, including a centralized data source for local businesses and a public API for accessing local business data.
- Enhanced backend API functionality for local SEO, integrating Google Business API to manage service areas and business insights.
- Achieved significant performance improvements, reducing script processing time from over 5 hours to under 45 minutes through database optimization and code efficiency.
- Implemented features for monitoring and syndicating business listings across Google, Apple, and Bing Maps, including holiday hours and review management.
- Designed and developed trial features for new businesses on the Brightedge Local business management platform, facilitating easier onboarding and feature access.

BatterySmart, Gurugram — SDE-3

Jan 2023 - June 2023

- Develop EV battery's IOT infrastructure by collecting IOT events from multiple battery vendors and storing them in a Apache Pinot realtime table. It is able to handle scale of over 70 qps and manage TBs of data storage and power real time analytics. Tech used Golang, Kafka, Avro, Apache Pinot, Prometheus, Grafana, Redis, Trino, Redash, Kubernetes
- Written a kafka producer to send Avro serialized electric battery IOT event data for a vendor. Tech used Python, Avro, Kafka, Kubernetes
- Developed an Insights service which captures IOT event for every EV Battery swap transaction and stores it in a Apache Pinot realtime table. Tech used Golang, Kafka, Pinot, Kubernetes
- Developed showing live EV battery location based on IOT data. Tech used: NodeJS, Redis, Kubernetes, Mysql, Redis Geolocation

magicpin, Gurugram — Tech Lead

Aug 2022 - Jan 2023

- Led Observability Initiatives
 - Deployed Prometheus in GKE for monitoring and integrated clients in Go, Java, and Python.
 - Optimized metric ingestion with relabeling filters.
- Developed Real-Time ML Prediction Architecture
 - Designed serving architecture for Google Vertex AI models using FastAPI and Bigtable.
 - Built pipelines for feature calculation, real-time predictions, and data sync between BigQuery and Bigtable.
 - Implemented Kubernetes jobs for model scoring and established CI/CD workflows for deployment.

Tekion Corp, Bengaluru — *ML Engineer*

July 2021 - Aug 2022

- Driving designing and development of distributed batch and streaming ETL processing pipeline, offline and online ML model serving architecture.
- Responsible for setting up a scalable, cheap, ACID compliant delta lake storage over S3 for storing processed raw facts known as features and batch prediction results of ML models, replacing postgres. <u>Blog</u>
- Responsible for development and maintenance of several internal python clients and tools packages used by ML Operations and data science team.
- Instrumental in bringing modern software development principles in all ML data pipeline batch jobs and ML model inference APIs.
- Optimized top k recommendation API latency by doing batch cosine similarity calculation and serializing high dimension feature vectors in parquet and storing it in an in-memory based fast storage.
- Created a bare SQL parser over pyarrow for ML data pipeline to facilitate direct read of raw data from delta lake warehouse. It supports query filtering over partitioned and non-partitioned columns, thereby reducing the need and cost of querying a managed query engine.
- Developed a scalable architecture for sharing duplicate data merge pipeline results using spark, delta lake, Kafka, and AWS Athena.
- Splitting sales chat app built using RASA framework into containerized (core and ACTION) microservice to facilitate communication among them using custom API routes and integrated our observability tools and dynamic model loading.
- POC for automating the manual task of vehicle data induction in Tekion from an OEM website using selenium, beautiful soup and Kafka.
- Added concurrency to ML serving APIs by converting all I/O bound tasks into asynchronous calls and using ASGI supporting web frameworks like FastAPI and adding parallelism using multiple uvicorn workers, integrated with open-source elastic stack observability. [opentelementry, elastic-apm, ElasticSearch, Kibana]

daloopa, Noida — Software Engineer III

Dec 2020 - June 2021

- Researched and implemented foreign language OCR detection, recognition, and translation for financial documents using open-source technologies, resulting in cost savings.
- Designed and developed a data pipeline for pattern-based search on financial documents utilizing machine learning, OCR, NLP, and heuristics.
- Automated manual tasks through AWS batch job automation.
- Enhanced system performance by improving latency of critical workflows via caching, concurrency, and

multiprocessing.

• Configured and managed a dockerized ECS Fargate pipeline for development and production environments.

magicpin, Gurugram — Software Engineer

June 2018 - Dec 2020

- Responsible for continuously evaluating recent technologies and driving architecture and design decisions on various critical backend systems and ownership to run and maintain cloud infrastructure.
- Took initiative for dataset collection, feature engineering, training using transfer learning, model inference using RESTful flask service on CNN based deep learning classification model for <u>selfie classification</u>.
- Migrated 21 TB data from 21 node Cassandra cluster to Cloud Bigtable cluster with redesigned schema ensuring row key entropy
- Migrated a tightly coupled java-MySQL monolith backend stack to decoupled, low latency, go service with denormalized MySQL schema thereby improving availability and reducing infra costs.
- Designed and developed a new low latency, highly available reverse-proxy layer between API layer and frontend for several backend systems which resulted in secure and robust backend systems.
- Designed and developed concurrent 5 layer trending pipeline which generates scores accumulated over days with decaying functions for various social entity in our app like posts, hashtags, merchants, users, friends and using worker-pool, pipeline, barrier design patterns
- Designed and Implemented Consumer and API Flow for projects like Merchant Rating calculation based on Bayes estimation , merchant complaints and review flow and Hashtag Feed and Campaign Feed and selfie and hashtag approval flows, groupbuys, merchant and user feed
- Designed and developed notification, sms, what sapp, IVR trigger logic along with an order cleanup service for Order Management Team

EDUCATION

NSIT, New Delhi — Bachelor in Engineering

2014 - 2018 Branch - Computer Engineering - CGPA 8

BGS International Public School, New Delhi – *Class 6 to 12*

2007 - 2014

12th CBSE Board - 93.8 %

10th CBSE Board - CGPA 10

COURSES AND SPECIALIZATIONS

- Programming with Google Go by University of California, Irvine
- Google Cloud Platform Fundamentals: Core Infrastructure
- DeepLearning.AI TensorFlow Developer
- Introduction to Machine Learning in Production
- <u>Go (Intermediate)</u>
- <u>Confluent Fundamentals Accreditation</u>