

Dipankar Banik

Lead Distributed Systems Engineer | Pleasanton, CA | dipankar.banik@gmail.com | (201) 736-5586 | dipankarbanik.com | linkedin.com/in/dipankarbanik | US Permanent Resident (Green Card), no sponsorship required

PROFESSIONAL SUMMARY

Lead Distributed Systems Engineer with 20+ years designing and operating large scale multi-tenant SaaS platforms on GCP and AWS. Deep expertise in Java 21, Spring Boot 3, event-driven architecture, distributed messaging (Apache Pulsar, Kafka), Kubernetes, Terraform, and relational and non-relational databases (Postgres, AlloyDB, Redis, BigQuery). Proven record of architecture leadership, mentoring, high-severity incident resolution, and delivery at scale. Currently applying **deterministic AI** and **agentic AI** to fraud detection and customer data platforms, combining rules driven, auditable decisioning with LLM powered workflows.

PROFESSIONAL EXPERIENCE

Lead Distributed Systems Engineer, Returns Platform | Narvar Inc., San Francisco, CA | March 2016 to Present

- Designed and owned the distributed backend of a multi-tenant Returns SaaS platform serving hundreds of global retail brands; architected autoscaling Java 21 / Spring Boot 3 microservices on GCP Kubernetes, including a full production cutover from AWS to GCP.
- Built multi-consumer Apache Pulsar event-driven architecture with shared subscriptions and dead letter queue policies for carrier event streams (UPS, FedEx, USPS); designed a reconciliation layer guaranteeing per-shipment event ordering under late, out-of-order, and duplicate delivery.
- Owned end-to-end infrastructure as code: Terraform across AWS and GCP, 79 Helm charts, ArgoCD GitOps, Ansible, and Packer image builds (300+ IaC commits).
- Managed data across Postgres 16, AlloyDB, Redis 7, and a 494-table BigQuery analytics surface with Avro and Protobuf schemas; led Python data pipeline tooling and SQL schema governance.
- Authored the returns rule engine (Drools, custom ANTLR DSL) running as 38 serverless Java Lambdas; delivered Stripe payment integration; built a Go webhook pipeline with Cloud Tasks scheduling; deep Kafka experience for inter-service streaming.
- Built **deterministic AI for fraud detection**: combined Drools rules based signal evaluation with agentic LLM workflows (MCP, RAG, vector databases) to produce auditable, repeatable fraud decisions on return fraud and abuse patterns across retailer traffic.
- Designed **deterministic AI workflows for the Customer Data Platform (CDP)** project: unified customer profile and event data across the post-purchase lifecycle, powering personalized experiences and predictive insights with rules driven guardrails around LLM outputs.
- Hardened platform security with OAuth2 / Okta, mTLS via Linkerd, HashiCorp Vault, and ModSecurity WAF; owned a Google Tink PII encryption library with Resilience4j retries used platform-wide.
- Mentored engineers, led code reviews and architecture decisions, ran on-call; integrated LLM and agentic AI capabilities (MCP, RAG, vector databases) for intelligent tracking and predictive delivery.

Applications Engineer, Platform and Infrastructure | Gilead Sciences, Foster City, CA | September 2014 to March 2016

- Architected migration of the FDA-validated Product Quality System from .NET to a distributed J2EE platform deployed across four global sites; built a cross-database data layer (Oracle, MSSQL), Spring MVC / Hibernate services, and LDAP authentication on Liferay.

Portal Lead and Distributed Systems Architect | Cisco, San Jose, CA | August 2010 to August 2014

- Led distributed platform engineering for Cisco GTMS uCRM (Salesforce integration, WebSphere Portal, Spring MVC) and CVCM QUAD, a Liferay platform serving 200+ tools with real-time event subscriptions and role-based search for Cisco's global user base.

Lead Platform Architect, Workflow and Data Systems | Kaiser Permanente, Pleasanton, CA | October 2009 to August 2010

- Built Kaiser's HR workflow platform on WebSphere, Spring, and Hibernate JPA with SSO via Oracle Access Manager and Spring Batch ETL from DB2 mainframe, processing high-volume HR records across a nationally distributed organization.

Lead Architect and Developer | MLB, Marsh USA, Guy Carpenter, Cravath, CWICO | New York, NY | October 2005 to October 2009

- Led distributed platform and data systems across finance, insurance, and sports, including the MLB Financial Desktop, MarshConnect global client portal (SOA, Hibernate, WebSphere), and real-time STP reporting.

Senior Programmer Analyst, J2EE Engineer | Marsh Inc., John Deere, Geologistics, Vigilos | 2001 to 2005

- Built enterprise Java / J2EE systems (EJB, WebLogic, Oracle) including the John Deere Asset Information System; early-career foundation in distributed Java systems and relational databases.

TECHNICAL SKILLS

Languages: Java 21, Python 3.12, Go 1.22, SQL, TypeScript, Kotlin

Distributed Systems: Apache Pulsar, Kafka, event-driven architecture, distributed locking, Quartz scheduling, Cloud Tasks, DLQ, exactly-once semantics, reconciliation at scale, REST, gRPC, Webhooks

Cloud and Infrastructure: GCP, AWS (S3, Lambda, VPC, ECS), Kubernetes, Docker, Terraform, Helm, ArgoCD, Linkerd, Ansible, Packer, CircleCI

Databases: Postgres, AlloyDB, Yugabyte, Redis, BigQuery, Oracle, DB2, MSSQL, Avro, Protobuf, Liquibase

Backend and Security: Spring Boot 3, Hibernate 6, Spring Batch, Drools, ANTLR, Resilience4j, JUnit 5, OAuth2, Okta, JWT, mTLS, HashiCorp Vault, Google Tink, ModSecurity WAF, Keycloak

AI and Data: deterministic AI, agentic AI, LLMs, RAG, vector databases, embeddings, agentic workflows, MCP, fraud detection AI, Customer Data Platform (CDP), MLOps

EDUCATION AND WORK AUTHORIZATION

Bachelor of Science in Physics | University of Pune, India | 1999

Work Authorization: US Permanent Resident (Green Card); no visa sponsorship required