

Dwai Banerjee

Engineering Leader

✉ dwaibanerjee78@gmail.com

📍 Bengaluru, India

🌐 dbanerj

🗣️ dwaiba

Professional Summary

Engineering leader with 24+ years across startup and enterprise platforms, delivering reliable products with practical, customer-centered execution. Built and guided cross-country teams with clear ownership growth. Worldwide footprint across India, the European Union, UK, South Africa, Southeast Asia, and the Americas.

Experience

cloudgear.io, Director

Own platform architecture across GenAI systems and GPU-accelerated inference platforms, delivering enterprise-grade AI backend and deterministic code generation frameworks.

- Architected and developed a private tri-repo (App / Infra / Deploy) AI backend framework using deterministic proto-first code generation, preventing contract-to-production drift via deterministic generation and policy gates across service, migration, and runtime generation.
- Embedded reusable distributed systems primitives including multi-tier caching (in-memory, distributed, write-through, and write-behind), event-driven streaming (pub/sub fanout, backpressure control, idempotent consumers), outbox reliability patterns, and zero-trust auth chains across 150+ auto-generated services.
- Scaled production rollout across 6+ enterprise customers in fintech, healthcare, and digital commerce, delivering 400+ migrations with observability, resilience, and tenancy patterns wired by construction.
- Built a GCP-based lakehouse and MLOps backbone processing 2.5B+ events per month, enabling distributed model training, hyperparameter tuning, and sub-150ms p95 inference (service-side).
- Reduced deployment cycle time from 3-4 weeks to under 5 days and improved GPU utilization by approximately 35 percent through inference optimization, workload batching, and promotion-gated CI/CD pipelines.

India

Jan 2024 – present

2 years 3 months

AWS, Senior Solutions Architect

Field leadership across GenAI, HPC, and high-growth startup ecosystems.

- Led Amazon Bedrock private previews and GenAI enablement for 15+ high-growth startups, helping teams move from experimentation to production-grade RAG and inference architectures.
- Designed high-throughput, low-latency inference systems using SageMaker and optimized vector retrieval, reducing response latency by approximately 40 percent and supporting 8K+ requests/sec peak.
- Architected GPU-accelerated GenAI and HPC workloads spanning distributed training and model tuning, improving GPU utilization by approximately 30 percent and accelerating prototype-to-production timelines by approximately 50 percent.
- Operationalized secure AI workloads with production MLOps, governance guardrails, disaster recovery patterns, and scalable multi-account architectures.

India

Apr 2023 – Jan 2024

10 months

<p>Citrix Systems, Senior Principal Software Engineer Principal leadership on Citrix DaaS (formerly Citrix Virtual Apps and Desktops service) engineering across AWS, Azure, and GCP.</p> <ul style="list-style-type: none"> • Architected and developed zero-trust, cloud-native microservices and control-plane systems in Go for a multi-product portfolio generating approximately \$2B in annual revenue, enforcing strong identity, service-to-service auth, and policy-driven runtime isolation. • Delivered frameworks using Terraform Enterprise and policy-as-code (OPA/Gatekeeper), embedding governance, drift detection, and FedRAMP-aligned controls directly into CI/CD and infrastructure workflows while mentoring distributed engineering teams. • Supported FedRAMP certification tracks aligned to NIST 800-53 controls at High impact levels. 	<p style="text-align: right;">India Oct 2020 – Mar 2023 2 years 6 months</p>
<p>Philips Healthcare, Principal Architect Led architecture and development for Health Suite on-Premise and hybrid healthcare platform rollout.</p> <ul style="list-style-type: none"> • Implemented HL7/FHIR/OIDC core services and secure GitOps/CI pipelines across multi-hypervisor environments. • Defined multi-GPU and compute benchmark programs for CT, insights, and radiation oncology workloads. 	<p style="text-align: right;">India Dec 2018 – May 2020 1 year 6 months</p>
<p>HCL Technologies, Senior Solution Director European solution leadership across verticals.</p> <ul style="list-style-type: none"> • Led cross-country teams delivering reusable cloud-native and Kubernetes/OpenShift stacks for BFSI and digital banking programs. • Built HPC/GPU compute fabrics for clinical and automotive simulation workloads. 	<p style="text-align: right;">The Netherlands & EU May 2017 – Nov 2018 1 year 7 months</p>
<p>Microsoft, Cloud Solution Architect Built Azure open-source HPC/GPU automation and customer enablement programs for NVIDIA virtual GPU solutions.</p> <ul style="list-style-type: none"> • Scaled high-fidelity GPU fabrics for upstream exploration and production and for seakeeping and wave-interaction simulations for ship builders. 	<p style="text-align: right;">The Netherlands and the USA Jan 2016 – Apr 2017 1 year 4 months</p>
<p>Exact Software, Principal, Solution Architecture Led ERP modernization prototypes and architecture roadmap recommendations for board-level planning.</p> <ul style="list-style-type: none"> • Drove componentization of legacy stacks into cloud-ready services and internal rapid prototyping tracks. 	<p style="text-align: right;">The Netherlands Aug 2015 – Nov 2015 4 months</p>
<p>Enable-U, Senior Consultant Scaled integration and governance solutions across Benelux, EU institutions, and Africa.</p> <ul style="list-style-type: none"> • Productized middleware autodiscovery and governance offerings and led API/SOA security programs in multi-country deployments. 	<p style="text-align: right;">The Netherlands, Africa Feb 2013 – July 2015 2 years 6 months</p>
<p>Canon Europa, Domain Solution Architect Led Canon Unity integration and data transformation across EMEA.</p> <ul style="list-style-type: none"> • Platform operated at 1.5M+ real-time transactions per month with nearly 2x ETL load; built integration/data CoE and vendor delivery model. 	<p style="text-align: right;">The Netherlands Oct 2011 – Jan 2013 1 year 4 months</p>
<p>TMNS, Technical Architect Product architecture and enterprise integration consulting across EU clients. TMNS was acquired by Devoteam.</p> <ul style="list-style-type: none"> • Built CEP product for an EU startup and improved Swisscom transaction throughput by 35 percent; introduced hybrid cloud self-service patterns. 	<p style="text-align: right;">The Netherlands, Switzerland Apr 2009 – Sept 2011 2 years 6 months</p>

<p>HCL Technologies, Technical Architect Early architecture leadership in SOA governance, platform virtualization, and Cisco AON solution labs.</p> <ul style="list-style-type: none"> Delivered large client programs across banking, telecom, and consumer sectors with reusable integration patterns. 	<p>India, UK, Malaysia June 2007 – Feb 2009 1 year 9 months</p>
<p>Oracle, Senior Member Technical Staff Core contributor to Oracle SOA Suite and Fusion Middleware integration engines.</p> <ul style="list-style-type: none"> Implemented standards-aligned service integration and HA topologies for enterprise deployments. 	<p>Bengaluru, India Apr 2006 – June 2007 1 year 3 months</p>
<p>IBM, Engineer Delivered WebSphere/MQ integration, DR, and security architecture programs across three continents.</p> <ul style="list-style-type: none"> Supported public sector, banking, and global B2B/EDI modernization initiatives. 	<p>India, Switzerland, UK, Barbados Oct 2003 – Mar 2006 2 years 6 months</p>
<p>Various, C++/Web developer Built distributed logistics and messaging systems in C++/EJB stacks for UK and Japan customers.</p>	<p>India Jan 2001 – Oct 2003 2 years 10 months</p>

Education

<p>National Institute of Technology, Durgapur, Engineering</p>	<p>Durgapur, India 1997 – 2001</p>
---	---

Certificates

<p>TOGAF 9 Certified Credential ID 32021</p>	<p>2010</p>
--	-------------

Skills

- Distributed Systems and Microservices**
- HPC, GPU, and MLOps**
- GenAI Systems and Model Tuning**
- Data Architecture, Analytics, and Governance**
- Leadership and Scalable Delivery**