

Apurv Singh

Seattle, WA

| email.singh.apurv@gmail.com

| apurvsingh.com

Senior Software Developer — AI Systems, AWS Platforms & Distributed Infrastructure

Impact Snapshot

- Amazon L6 Senior SDE with 8+ years of experience across distributed systems, AWS platforms, encoding infrastructure, backend services, data systems, and AI platforms.
- Delivered 10x dynamic fleet scaling, improved encoding reliability from 95% to 99%, and reduced processing latency to approximately 1x content duration.
- Designed AWS compute, fleet management, observability, security remediation, automation, and AI-assisted developer tooling for high-scale production systems.
- Recognized for ownership, bias for action, mentoring, operational excellence, and deep technical problem solving across current-generation and next-generation platforms.

Professional Summary

Senior Software Developer with deep experience building large-scale distributed systems, AI-driven platforms, and cloud-native infrastructure at Amazon. Strong background in AWS system design, backend architecture, workflow orchestration, compute platforms, fleet management, security, data systems, and operational excellence.

Experienced in driving technical direction, simplifying ambiguous problems, mentoring engineers, and designing scalable systems that improve reliability, efficiency, security, and long-term maintainability.

Technical Leadership

- Drive technical direction for backend, compute, fleet, security, operational excellence, and AI platform initiatives across distributed systems.
- Translate ambiguous requirements into architecture, execution plans, tradeoff analysis, prototypes, and long-term platform roadmaps.
- Partner with cross-functional teams to align priorities, reduce technical debt, improve operational visibility, and scale platform adoption.
- Communicate complex architecture, workflow, fleet, security, AI tooling, and operational designs clearly across engineering and product stakeholders.
- Mentor engineers through onboarding, architecture reviews, code reviews, knowledge-sharing, integration testing, CloudFormation tasks, and operational troubleshooting.

Professional Experience

Amazon

Senior Software Development Engineer (L6), Prime Video Encoding

Seattle, WA

Sep 2022 – Present

- Drive architecture and technical direction for Prime Video's global encoding infrastructure, improving scalability, reliability, security, operational excellence, and long-term platform maintainability.
- Lead system-level design for encoding compute platforms, fleet management, deployment automation, capacity scaling, observability, and security remediation across distributed media workflows.
- Designed and delivered fleet auto-scaling capabilities to improve compute elasticity, reduce operational overhead, and support peak demand and large backfill workloads.
- Built operational tooling including fleet dashboards, HPC tracking, observability constructs, workflow health indicators, and alarms to improve incident response and infrastructure visibility.
- Designed distributed workflows supporting HEVC, AV1, H.264, Dolby Vision, and Dolby Atmos pipelines, optimizing for quality, cost, correctness, and playback reliability.
- Led flexible fleet migration from static EC2 capacity to elastic AWS Auto Scaling based infrastructure, improving utilization, reducing cost, and modernizing compute operations.

- Designed dynamic EBS-based storage provisioning to support per-job elasticity without relying on pre-provisioned infrastructure.
- Contributed to next-generation encoding platform initiatives, including workflow integration, async activity orchestration, communication libraries, and migration paths for current-generation systems.
- Drove security review remediation and certification efforts by identifying risks early, coordinating follow-ups, resolving blockers, and improving operational readiness before escalation.
- Built AI-assisted developer and operational tooling prototypes to accelerate troubleshooting, documentation, ticket handling, and engineering productivity.
- Partner with cross-functional engineering teams to synthesize requirements, evaluate design tradeoffs, raise risks early, and align execution around scalable platform improvements.
- Mentor engineers through onboarding, system walkthroughs, architecture discussions, code reviews, integration testing, CloudFormation tasks, and operational troubleshooting.

Selected Impact

- Reduced encoding latency to near real-time processing, approximately 1x content duration.
- Enabled 10x dynamic scaling during peak demand and large backfill workloads.
- Improved encoding reliability from 95% to 99% success rate.
- Delivered cost savings through flexible fleet migration, fleet auto-scaling, and improved compute utilization.
- Improved operational health through dashboards, observability tooling, better alarms, documentation, and faster incident response.
- Earned trust as a senior technical owner for design reviews, compute platform decisions, fleet operations, security issues, and operational excellence.

Amazon

Software Development Engineer II (L5), Amazon Appstore / Alexa

Seattle, WA

Jun 2019 – Sep 2022

- Architected and built a centralized privacy and compliance platform supporting GDPR, COPPA, and CCPA across multiple Amazon services.
- Designed reusable connector-based backend architecture enabling teams to onboard heterogeneous data sources without building custom integrations.
- Synthesized requirements from partner teams and translated them into scalable platform capabilities, APIs, workflows, and data contracts.
- Built distributed data ingestion, validation, and processing pipelines using Spark, EMR, S3, Athena, Redshift, and AWS-native services.
- Designed backend services for compliance orchestration, policy enforcement, auditability, workflow reliability, and operational visibility.
- Identified technical debt in repeated one-off compliance integrations and replaced them with reusable platform abstractions.
- Improved system scalability through idempotent processing, retry-safe workflows, fault isolation, and optimized data access patterns.
- Built Redshift-backed analytics layers powering Tableau and Superset dashboards for compliance visibility and operational decision-making.
- Designed and built multi-tenant platform capabilities enabling multiple teams to onboard quickly through reusable backend abstractions.
- Led certification and operational readiness efforts by coordinating risks, follow-ups, technical milestones, and partner onboarding.

Selected Impact

- Enabled scalable privacy compliance across multiple Amazon business units.
- Reduced onboarding complexity for new tenants and data sources through reusable connector architecture.
- Launched multi-tenant platform capabilities adopted by multiple Amazon teams.
- Contributed to US Patent 11,714,918 B2 for reusable data privacy compliance systems.

Amazon

Software Development Engineer I (L4), Appstore Developer Portal / Data Analytics

Bangalore, India

Aug 2017 – Jun 2019

- Built backend data platforms for multi-tenant sales, earnings, developer analytics, and reporting workloads across Amazon Appstore.
- Designed scalable ETL pipelines using Spark, Hive, EMR, and S3-based data lakes to support analytics and business reporting.
- Developed Redshift-backed analytics models optimized for query performance, correctness, and storage efficiency.

- Built OLAP-based analytics systems using Apache Kylin to support low-latency aggregated reporting over large datasets.
- Delivered automation for DSAR and GDPR-related workflows by understanding evolving requirements, integrating with partner systems, and shipping production-ready solutions under tight timelines.
- Led development and delivery of key internal initiatives including Valkyrie and Bullseye, coordinating with cross-time-zone stakeholders to clarify requirements and keep execution on track.
- Improved operational health during on-call by deep-diving recurring issues, identifying root causes, and reducing ticket volume from approximately 70–80 to 10–15.
- Helped partner teams understand analytics system flows, edge cases, and integration behavior through detailed technical guidance and troubleshooting support.

Selected Impact

- Delivered high-quality backend and automation projects with strong ownership despite ambiguous and evolving requirements.
- Reduced operational load by resolving recurring on-call issues and improving system stability.
- Earned trust across partner teams as a go-to engineer for analytics systems, integrations, and production troubleshooting.
- Successfully launched cross-team initiatives such as Bullseye on time through clear communication and fast issue resolution.

Amazon

Software Development Engineer Intern, Amazon Appstore Analytics

Bangalore, India
Feb 2017 – Aug 2017

- Built abstraction layers over Elasticsearch to simplify complex analytical queries.
- Worked with AWS databases, data warehousing systems, and production analytics infrastructure.

Indian Institute of Remote Sensing (IIRS), ISRO

Research Intern

Nagpur, India
May 2016 – Dec 2016

- Developed web-based workflows for satellite image processing systems.
- Worked on backend processing pipelines, image processing algorithms, and remote sensing data systems.

Selected AI Projects

AI Research & Decision Systems Platform

- Designed and built AI-powered backend systems using RAG pipelines, embeddings, vector search, and LLM-based reasoning.
- Defined technical architecture for multi-stage retrieval, ranking, evaluation, and response-generation workflows.
- Built orchestration pipelines combining structured data, vector databases, external sources, ranking models, and AI-generated responses.
- Created roadmap for improving answer quality through hybrid search, citation-aware ranking, evaluation loops, and structured knowledge sources.
- Identified gaps in semantic retrieval quality and designed improved ranking, validation, and correctness layers.
- Built evaluation systems to measure retrieval accuracy, response quality, ranking performance, and system reliability.

Skills

Technical Leadership: Technical roadmap, architecture reviews, stakeholder alignment, system design, tradeoff analysis, platform strategy, mentoring

Systems & Architecture: Distributed systems, backend architecture, workflow orchestration, scalability, reliability engineering, operational excellence

Cloud & Infrastructure: AWS EC2, Auto Scaling, ECS, EKS, Lambda, S3, DynamoDB, RDS, SQS, EMR, Athena, Redshift, EBS, Docker, Helm, CloudFormation

AI Systems: RAG pipelines, embeddings, vector search, OpenAI APIs, inference orchestration, ranking, evaluation systems, AI-assisted developer tools

Data: PostgreSQL, Redshift, DynamoDB, Spark, Hive, Athena, vector databases

Languages: Java, Python, TypeScript, JavaScript

Frameworks & Tools: Node.js, Express, React, Temporal, ProseMirror, Streamlit

Education

Bachelor of Engineering

Nagpur University, Nagpur, India

CGPA: 8.1

2013 – 2017

Honors & Awards

- ACM International Collegiate Programming Contest (ICPC) — Amritapuri Onsite, Dec 2016.
- Project Inter-College Competition, GlobalLogic — Runner Up, Feb 2016.