Aditya Pandey

Linkedin: https://in.linkedin.com/in/aditya-kumar-pandey-cloud

Email : aditya141199@gmail.com
Mobile : +91 8368336244

Skills

- **Backend Engineering**: Java, Python, Go, C++, SQL, Shell/Bash, JVM Performance Tuning, Microservices Architecture
- **Cloud Platforms**: GCP (Google Cloud Platform), AWS (EKS, ECS), Terraform, Kubernetes, Docker, Linux/Unix
- **Site Reliability Engineering**: Monitoring & Alerting, Runbooks, Dashboards, High-Availability Systems, Fault Tolerance, Scalability
- **Datastores & Databases**: PostgreSQL (RDBMS), Redis, ClickHouse (Time-Series/Analytics), Snowflake, NoSQL
- **Observability & Tools**: Splunk, Datadog, Grafana, Prometheus, PagerDuty, REST/GraphQL APIs

Experience

- **Amagi Media Labs** — Bengaluru, India
  Senior Platform Engineer — *Dec 2024 - Present*
  - **GitOps & Kubernetes Platform**: Architected production **GitOps platform** using **ArgoCD** on **AWS EKS**, managing **50+ applications** across multi-environment clusters. Implemented dual-repository pattern with **Terraform** and **Helm**, achieving **100%** deployment automation.
  - **Advanced Autoscaling & Cost Optimization**: Designed **Karpenter v1.8.0** autoscaling with **15+ workload-specific node pools** and **spot instances** (90% savings), achieving **$30K+/month** compute cost reduction through just-in-time provisioning and right-sizing.
  - **Ephemeral GPU Inference Platform**: Architected ephemeral GPU inference platform using **Karpenter** to provision on-demand **GPU nodes** (G4dn, G6 instances with NVIDIA T4/L4) for AI model inference and video processing workloads. Implemented workload-specific node pools with taints/tolerations, enabling just-in-time GPU provisioning and automatic scale-down, reducing GPU infrastructure costs by **60%** through efficient resource utilization.
  - **Container Runtime Optimization**: Pioneered production deployment of **SOCI (Seekable OCI) snapshotter** on **Bottlerocket OS** with parallel image pull/unpack (20 concurrent downloads, 12 concurrent unpacks), reducing container startup by **50-70%** and workflow execution time by **40%**. Configured dedicated NVMe ephemeral storage for container runtime directories, optimizing I/O performance across all 15+ node pools.
  - **AI Platform & Orchestration**: Built **Model Context Protocol (MCP)** architecture with **5 production services** (RAG, Analytics, Search) integrated with **Weaviate** vector DB. Implemented dual orchestration with **Temporal** and **Argo Workflows** for AI agents and video processing pipelines.
  - **Data Platform & Analytics**: Scaled cloud-native platform to **1.2 TB/day** using **Apache Spark** on **Databricks**. Designed **ClickHouse** serving layer achieving **p95 < 800ms** on billion-row queries, reducing costs by **$20K/month**.
  - **Security & Observability**: Implemented **SLV operator** for secrets management, **IAM Roles for Service Accounts**, and comprehensive observability stack (**Prometheus**, **Grafana**, **Loki**, **KubeCost**). Maintained **99.9%** uptime with production on-call rotation.
  - **Technical Leadership & Documentation**: Led platform architecture decisions across multi-environment clusters, mentored engineers on Kubernetes and GitOps best practices. Documented system innovations (SOCI snapshotter, Karpenter autoscaling, MCP architecture) in comprehensive technical specifications, enabling knowledge transfer and team scalability.

- **Take Two Interactive Ltd.** — Bengaluru, India
  Data Engineer - AWS (Platform Engineering) — *Oct 2022 - Dec 2024*
  - **Cloud Infrastructure at Scale**: Operated and scaled distributed software across **AWS and GCP**, managing **1000+ ECS services**, **150+ ECR images**, **200+ S3 buckets**, and **50+ Lambda functions** using **Terraform** IaC. Built comprehensive monitoring, dashboards, and alerts for high-availability systems with **Docker** containerization, reducing manual operations by **85%**.
  - **Kubernetes & Multi-Cloud**: Deployed **Kubernetes (EKS)** clusters running microservices architecture for production applications. Built web services using **Python (Flask)** with **REST APIs**, integrated **PostgreSQL (RDBMS)** and **Redis (NoSQL)**, and handled authentication via **AWS Cognito** with **Okta SSO**. Reduced application downtime by **40%** and improved deployment speed by **50%**.
  - **GCP & API Management**: Built scalable token management system on **GCP (APIGEE)** for API gateway operations, designed to renew and generate tokens for proxies. Integrated **RabbitMQ** for messaging, reducing token-related failures by **70%** and increasing API availability by **25%**.
  - **Observability & Automation**: Built comprehensive observability infrastructure using **Splunk**, **Datadog**, and **Grafana/Prometheus**, correlating logs across multiple platforms (**IICS**, **AWS**, **Databricks**). Wrote runbooks, automated incident response (60% MTTR reduction), and migrated **175+ repositories** to **GitHub Actions**, eliminating **140 FTE hours** annually.

- **Amazon Web Services** — Bangalore, India
  Cloud Support Associate (Databases) — *Mar 2021 - Oct 2022*
  - **Database Reliability Engineering**: Built automation tools using **Python (Django)**, **React**, and **Java** for **RDS PostgreSQL (RDBMS)** upgrade operations, reducing maintenance downtime from **1 hour to <10 minutes**. Demonstrated strong debugging skills and code optimization for production database reliability.
  - **Backend Automation & Tooling**: Developed scalable backend services using **Python**, **Java**, and **Go** on Unix/Linux systems. Created Slack bot integrating **AWS Lambda** that eliminated **1000+ hours** of routine reporting tasks annually, winning internal hackathon for automation innovation.

Education

- **KLE Technological University** — Hubballi, India
  Bachelor of Engineering (B.E.), Computer Science and Engineering — *Aug 2017 - Jun 2021*