GAUTAM RAJU

+1-857-370-2790 | raju.ga@northeastern.edu | LinkedIn | GitHub

PROFESSIONAL SUMMARY

Site Reliability Engineer with 3+ years building and operating large-scale distributed systems, specializing in automation, observability, and incident management. Expertise in Go, Python, Kubernetes, and cloud infrastructure with proven track record eliminating toil and improving system reliability. Completing MS in Computer Science (Apr 2026).

EDUCATION

Northeastern University   **Sep 2024 - Apr 2026**

Master of Science, Computer Science GPA: 3.7/4.0
  • **Coursework:** Distributed Systems, Machine Learning Operations, Cloud Computing, Algorithms

Visvesvaraya Technological University   **Aug 2016 - Jul 2020**

Bachelor of Engineering, Computer Science GPA: 8.05/10.0

WORK EXPERIENCE

Ancestry   **Jun 2025 - Dec 2025**

Software Engineer Cloud Platform (EKS Team)
  • Managed production on-call rotation for EKS infrastructure maintaining 99.95% uptime across 200+ services supporting 10+ teams
  • Built observability infrastructure using Prometheus and Grafana tracking SLIs and system health, reducing MTTR by 30%
  • Eliminated toil through automation using Go and Python reducing manual deployment effort by 20 engineer-hours weekly
  • Implemented monitoring dashboards for capacity planning tracking resource utilization across 500+ cloud instances
  • Collaborated with engineering teams on incident response executing blameless postmortems and implementing preventive measures

Apple Inc.   **Sep 2021 - Jul 2024**

Software Engineer
  • Operated distributed systems processing 100M+ monthly transactions with 99.97% uptime through effective SLO tracking
  • Participated in on-call rotation managing incident response reducing customer-impacting incidents by 25%
  • Built monitoring infrastructure with Prometheus and CloudWatch tracking system health across 50+ microservices
  • Automated infrastructure operations using Python and CI/CD pipelines reducing deployment time by 65%
  • Designed reliable microservices with API Gateway managing traffic distribution and fault-tolerant load balancing

PROJECTS

Infra-Archaeology-MCP
  • Built Python MCP server integrating CloudTrail, Cost Explorer, and Terraform state for cloud resource lineage analysis
  • Implemented automated compliance auditing cross-referencing CloudTrail events, Terraform configurations, and Git history
  • Designed dependency mapping tracking 500+ AWS resources preventing production incidents through cost wastage detection
  • Developed Claude-powered natural language interface reducing infrastructure investigation time from 2 hours to 5 seconds

EmailSender.AI
  • Built Python automation system with GitHub Actions CI/CD achieving 99% uptime through comprehensive error handling
  • Implemented health monitoring and alerting tracking system availability, API latency, and failure rates
  • Designed observability framework with structured logging and metrics collection for debugging and performance analysis
  • Automated OAuth 2.0 token refresh workflows eliminating manual intervention and reducing operational toil

GPU Profiler
  • Developed Python framework for GPU performance profiling analyzing memory usage and kernel-level metrics
  • Implemented automated monitoring tracking resource utilization, thermal throttling, and performance bottlenecks
  • Built optimization recommendations reducing training time by 40% through batch size and configuration analysis
  • Designed time-series metrics collection with visualization enabling capacity planning and performance trending

TECHNICAL SKILLS

  • **Languages**: Python, Go, Java, Bash, SQL, JavaScript, C++ (coursework)
  • **SRE Practices**: Production On-Call, Incident Management, Blameless Postmortems, SLO/SLI Tracking, Toil Reduction, Planning
  • **Monitoring & Observability**: Prometheus, Grafana, DataDog, CloudWatch, ELK Stack, Logs, Metrics, APM, Distributed Tracing
  • **Cloud & Infrastructure**: AWS (EC2, EKS, S3, Lambda, CloudWatch, VPC), GCP, Kubernetes, Docker, Linux, System Performance
  • **Automation & IaC**: Terraform, Ansible, Infrastructure as Code, CI/CD, Jenkins, GitHub Actions
  • **Distributed Systems**: Microservices, Load Balancing, Fault Tolerance, Scalability, High Availability
  • **AI-Assisted Development**: GitHub Copilot, Claude AI, ChatGPT, AI-powered debugging and automation
  • **Data Structures & Algorithms**: Graph Traversal, System Design, Complexity Analysis, Problem Solving
  • **Databases**: PostgreSQL, MySQL, MongoDB, Redis, Query Optimization
  • **Tools**: Git, pytest, Docker, Postman, PagerDuty (familiar), Linux Debugging, Performance Profiling