Arpan Sahu

Software Engineer

Delhi, India | arpansahu@zohomail.in | +917828007426 | arpansahu.space | arpansahu | arpansahu | arpansahu

Profile

Specializes in designing robust, cloud-native architectures, integrating complex third-party services, and managing large-scale workloads with millions of concurrent requests. Proficient in CI/CD pipelines, automation, and infrastructure optimization to deliver efficient, fault-tolerant deployments. Passionate about transforming complex workflows into impactful, real-world solutions at enterprise scale.

Skills

C | C++ | Python | Java | Django | Flask | Fast API | React | DevOps | Docker | Kubernetes | Jenkins | Github | AWS | GraphQL | CSS | HTML | Bootstrap | PostgreSQL | MySQL | MongoDB | Nginx | Redis | Selenium | Javascript | Celery | RabbitMQ | Apache Kafka | Scrapy | Websockets | Payments Integration | Amazon Web Services | Data Structures & Algorithm | CI/CD | Taking Ownership | ElasticSearch | NodeJS | Go | Spring Boot | SQLAlchemy | Hibernate | Sentry | Datadog | Wireshark | Prometheus | Graphana | OpenTelementry | Jaeger | Paypal | Airflow | RedShift | RAG | Lachang | gRPC

Work Experience

Software engineer 2, *Microsoft* 03/2025 – present | Noida, India

- Engineered hyperscale orchestration systems powering Microsoft's data collection and analytics pipelines, enabling petabyte-scale processing and 100M+ requests per minute with high reliability.
- Optimized distributed infrastructure for performance, fault tolerance, and automation, reducing operational overhead while ensuring SLA compliance and global availability.
- that empower downstream data teams to efficiently collect, transform, and analyze massive datasets, serving as the foundation for Microsoft's data-driven insights and decision-making.

Software Engineer, *Trellix/Fireeye-McAfee* 07/2022 – 03/2025 | Bengaluru, India

- **Developed a SaaS platform for Extended Detection** and Response (XDR), migrating customer data from individually owned, on-premises EC2 instances to a centralized architecture leveraging Kubernetes clusters and a unified database. This transition reduced infrastructure management overhead, enabled scalable, pay-as-you-go service offerings, optimized resource utilization, and facilitated modular cybersecurity services.
- **Enhanced the SaaS platform's (SOAR - security orchestration[conterisation] and response system) efficiency & Architected Integrations** by building a robust CI/CD pipeline, automating testing, deployment, and scaling processes, reducing deployment times by 40% and increasing system reliability and responsiveness.- Architected integrations with Tenable, Microsoft Azure, VirusTotal, ServiceNow, and Trellix/McAfee, improving platform efficiency by 25% and enabling faster and more effective threat response.These integrations handled 3 million RPM and supported 500,000 concurrent playbook(REST APIs Integrations) executions, improving platform efficiency by 25% and enabling faster, more effective threat responses.
- **Engineered a high-performance data segregation engine** to identify customer-specific threat impacts across billions of records. By transitioning from traditional database queries to Redis key-value pairs and ultimately optimizing to a bit-string representation, achieved a ~99.88% reduction in memory usage and improved query speeds by over 100x, enabling real-time threat analysis and decision-making at scale. **Made Real time Streaming supported in an analytical system utilising Kafka, worked with ACL and other kafka cluster management too along with Zookeeper.**
- **Built a Search Microservice**, consolidating the search functionalities of three existing Trellix products into a single interface, and implemented Elasticsearch for additional search capabilities. This innovation improved data retrieval speed by 80% and simplified the user experience by eliminating the need for multiple search portals.

- **Implemented a dynamic rule engine** that enables the SaaS platform to define, manage, and execute complex cybersecurity rulesets stored in the database. By eliminating the need for hardcoded logic, this solution streamlined the creation and deployment of new threat detection rules, reducing response times and enhancing agility in addressing emerging threats in a rapidly evolving cybersecurity landscape.
- **Worked on developing the Trellix IAM microservice** to enable seamless authentication across internal cybersecurity platforms, significantly reducing human involvement in the authentication process.
- **Leveraged Generative AI with LangChain, RAG, and gRPC** to automate the generation of SOAR marketplace apps, reducing manual effort through intelligent code generation, playbook creation, and seamless microservice communication.

Associate Software Engineer, *Trellix/Fireeye-McAfee* 07/2022 – 01/2023 | Pune, India
Got merged with McAfee to form Trellix

Software Engineer, *Gammastack* 02/2021 – 06/2022 | Indore, India
- **Delivered high-quality**, maintainable code by implementing best practices, automated testing, and code reviews, reducing bug resolution time by 15% and release cycles by 30%.
- **Collaborated with cross-functional stakeholders** to define requirements and optimize implementation strategies, achieving a 25% decrease in project turnaround time.
- **Designed and implemented a scalable betting platform for Betfaro**, supporting 50,000+ concurrent users during peak match times, processing over $1M in daily transactions, and leveraging the **New Double Bet model** to drive a 60% increase in client profitability and a 40% surge in website traffic.
- **Optimized data processing workflows** for KodeKloud by 20% through the development of scalable features leveraging MongoDB, Flask, and AngularJS.
- **Designed and optimized database schemas** to support **high-volume transactions**, ensuring scalability and efficiency for Betfaro's betting platform.
- **Integrated Payment Gateway:** Integrated subscription based and order based payment system with Crypto as well as Real Currency based gateways like Paypal and Razorpay.

Education

Bachelors of Technology (Computer Science), 08/2017 – 07/2021 | Bhopal, India
Lakshmi Narain College of Technology & Excellence
CGPA: 8.55

Higher Secondary, *Shir Ram Adarsh Higher Secondary* 06/2015 – 06/2016 | Seoni, India
Percentage: 70.40

High Secondary, *Bhonsala Military School* 05/2013 – 05/2014 | Nagpur, India
Percentage: 88.60

Projects

Scrape Optimus, *A Versatile Web Scraping Solution* 09/2024 – present
Checkout at **Live**
Technologies Used: PYTHON | DJANGO | JQUERY | HTML | CSS | JQUERY | BOOTSTRAP | JAVASCRIPT | REDIS | GITHUB | DOCKER | HARBOR | KUBERNETES | JENKINS | | NGINX | MINIO | UBUNTU | MAILJET | SENTRY | RAZORPAY | PAYPAL | OPENVPN | KAFKA | RABBITMQ | GEOCODIFY | RANCHER | POSTGRESQL
- **Conceptualized, designed, and self-managed** the end-to-end deployment of a versatile web scraping solution on VPS, overseeing every stage from CI/CD automation to production deployment, payment integration, and feature development.
- **Launched with a freemium model**, achieving early traction and user engagement within the first week.
- **Integrated robust scraping capabilities** by connecting with over 200 proxies and 50 VPN providers, enabling accurate data collection across 1,000+ websites.
- **Implemented automated CI/CD pipelines**, reducing deployment time by 35% and supporting rapid updates, bug fixes, and scalability.
- **Built a scalable payment system** for premium users, establishing a revenue stream and enabling seamless user upgrades.

Talent Bridged, *Bridging Talent with Opportunity* 01/2022 – present
Checkout **Live** .
Technologies Used: PYTHON | DJANGO | JQUERY | HTML | CSS | BOOTSTRAP | SCRAPY | JAVASCRIPT | POSTGRES | NGINX | SCRAPE OPTIMUS DOCKER | MAILJET | SENTRY | KUBERNETES | DOCKER | JENKINS | MINIO | HARBOR | PORTAINER | UBUNTU | RACNHER | REDIS | KAFKA | RABBITMQ | ELASTICSEARCH
- **Aggregated job listings from multiple sources** into a single platform, streamlining the search experience and reducing job search time for users by 40%.
- **Designed and implemented skill-based categorization**, enhancing candidate-job matching accuracy by 50% to provide more relevant opportunities.
- **Developed an automated web scraping pipeline** to ensure up-to-date job data with a 90% accuracy rate, reducing manual data updates by 80% and significantly improving data freshness.
- **Integrated a Kibana-like board system to track application** processes, enabling real-time monitoring and management, which improves operational oversight and decision-making for administrators.

- **Deployed and managed the platform on a VPS**, handling all aspects of setup, scaling, and maintenance to support stable, user-friendly access.

Razorpay IPN Django Handler, *OpenSource Python Library* 10/2024 – 11/2024
Checkout **Github** or **Install Via**
Technologies Used: PYTHON | **DJANGO** | **PYPI** | **GITHUB** | **GITHUB ACTIONS**

- **Released a Django library** for managing Razorpay IPN webhook events, designed to streamline payment, order, and subscription notifications for developers.
- **Anticipated to reduce manual handling** of payment notifications by up to 90%, improving response times and reliability for projects using Razorpay.
- **Early traction**: Gained initial interest on GitHub, with a positive response from developers for ease of integration and signal-based event tracking capabilities.

Great Chat, *Django Chat Project* 03/2024 – present
Checkout **Github** or **Live**
Technologies Used: PYTHON | **DJANGO** | **HTML** | **CSS** | **JQUERY** | **BOOTSTRAP** | **JAVASCRIPT** | **REDIS | POSTGRES** | **GITHUB | DOCKER | KUBERNETES** | **JENKINS | DJANGO CHANNELS | WEB SOCKETS | HTMX**

- **Concept Project** demonstrating expertise with Django Channels for real-time chat capabilities.
- **Features**: Private, group, and global chat functionalities with online/offline notifications and profile management.
- **Deployment**: Hosted on Ubuntu VPS with Docker and Kubernetes, CI/CD-enabled for continuous integration.

For a comprehensive portfolio of 12+ live projects,
I have built and deployed, visit: arpansahu.me/projects.

- **Deployed and managed a cost-efficient VPS-based infrastructure using Hostinger,** hosting PostgreSQL, Redis, Docker (with Harbor), Kubernetes, Jenkins, Kafka, RabbitMQ, and Nginx, secured with SSL for all services and optimized for scalability across multiple personal projects.