

Pradyumna Prakash Khawas

(347) 680-5573 | pradyumna.khawas@gmail.com
www.linkedin.com/in/pradyumna-khawas | therealppk.github.io

Education

North Carolina State University, Raleigh, United States

August 2022 - May 2024

Master of Computer Science

GPA - 3.85/4.0

Courses: Design and Analysis of Algorithms, Operating Systems, Cloud Computing

Vellore Institute of Technology, Chennai, India

June 2016 - May 2020

Bachelor of Technology, Computer Science and Engineering

GPA - 9.01/10.0

Courses: Data Structures and Algorithms, Operating Systems, Parallel and Distributed Computing

Experience

Data Engineer Intern - Cloudflare | Remote, US

June 2023 - August 2023

- Built an explainable times series forecasting model to predict the CPU demand for servers on Cloudflare's global edge network while having a similar performance to the current forecasting model.
- Analyzed sales and marketing data to understand its implications and effects on CPU utilization across the network. Performed Exploratory Data Analysis to understand the data and used visualizations to explain insights to stakeholders.

Software Engineer - WorkIndia | Bangalore, India

July 2020 - June 2022

- **Mentored interns** and guided them to develop applications independently and take ownership of projects.
- Awarded the **Best Practices and Center of Excellence Award** at WorkIndia, 2022 Q1, for pushing the team to continuously learn and choose the right strategies through tech sessions and code reviews.
- Handled cloud admin responsibilities for various projects, including IAM, infrastructure and network planning and provisioning (using **Infrastructure-as-Code** tools like **Terraform** and **Ansible**), and cost analysis.

Data Team

- Built a robust **data lake and warehouse** using **Django (Python)** for the metadata service, **Apache Kafka** for pipelines, **Apache Spark** for processing, Amazon Athena as the query engine, and Tableau and Metabase for the dashboards.
- **Reduced data costs to two-thirds** by improving the efficiency of Spark processes by adding appropriate configurations, optimizing cron mechanisms, and right-sizing infrastructure.
- Built a **highly available, high-traffic API** capable of capturing **~126 million daily app events**. Implemented the service in **Go** using the **Gin web framework**.
- Collaborated with marketing, revenue, and product teams to solve complex data requirements using **ELT / ETL** approaches.

Platform Team

- Designed and developed a **unified communication platform** using **Django (Python)** for the service, **PostgreSQL** for storing metadata, and **Apache Cassandra** for storing all the communication data.
- Scaled the platform to deliver **~ 10 million communications per day** by incorporating **complex distributed strategies** and revamping the infrastructure. Utilized **Apache Kafka** for the communication queues, migrated the queues from Amazon EC2 to Kubernetes for faster scaling-in and easier orchestration, and used **Apache Zookeeper** to coordinate clusters.
- Managed stakeholders' requirements, assessed the current backlog and planned future tasks in an **agile environment**.

Software Engineer Intern - WorkIndia | Bangalore, India

Jan 2020 - June 2020

- Created a **data lake** to stream data from various databases to the warehouse for further analytics.
- Built multiple **data pipelines** using **Apache Kafka**, capable of handling **~ 1 million messages per minute**.

Software Engineer Intern - Kristal.AI | Bangalore, India

May 2019 - June 2019

- Migrated the website search functionality from legacy **Apache Solr** to SolrCloud, thus improving reliability and performance.
- Integrated the search data with data models in **Spring Boot (Java)** to enable CRUD operations.
- Proactively took the initiative to refine the search results by tweaking filters and tokenizers.

Skills

Languages - Python, Java, Ruby, Go, SQL, HTML/CSS, JavaScript **Frameworks** - Django, Spring Boot, Ruby on Rails, Gin

Platform and Tools - Apache Spark, Kubernetes, Docker, Apache Kafka, Apache Zookeeper, Ansible, Git

Cloud Platforms - Amazon Web Services (EC2, S3, EKS, RDS, ECR, Athena, Glue, ElastiCache, CloudFormation, IAM, VPC)

Databases - MySQL, PostgreSQL, Apache Cassandra, Redis, Apache Solr, Elasticsearch, PrestoDB

Monitoring - Prometheus, Grafana, Zabbix, New Relic, PagerDuty

Activities and Achievements

- Secured the prize for the best project in the Social Impact track at **DiamondHacks'24** hosted by WICS at NC State University.
- Attended the All Things Open (ATO) conferences in 2022 and 2023 to gain exposure and keep up with cutting-edge tech.
- Secured the first prize in **HackNITJ 2019**, a hackathon conducted by NIT Jalandhar.
- Received School Honor for securing the highest score in Computer Science in AISCCE-2016.