

Satheesh Gopalan

Data Engineer

✉ satheeshgopalan@gmail.com ☎ +91-7676251253 in [LinkedIn](#) 🐙 [GitHub](#) 📍 Bengaluru

SUMMARY

6+ years experienced, meticulous & result-oriented Data Engineer armed with a proven track record of analytical acumen in solving problems involving Big Data and deploying complex data pipelines. Possesses diverse experience in planning, executing and liaising with the key stakeholders to identify & resolve business problems for one of the largest CDNs and deliver excellent results.

KEY SKILLS

Python

PySpark

Hive

Spark

HBase

Sqoop

Flume

Kafka

Airflow

PostgreSQL

MySQL

Linux

Git

Perforce

Redshift

AWS Cloud

S3

EC2

EMR

Athena

Glue

PROFESSIONAL EXPERIENCE

DDC Specialist

Dec '18 - Present

Akamai Technologies

Bengaluru, IN

Distributed Data Collection Team

- Manage billing pipelines that handles the company revenues worth **over \$160 Million**
- Manage Hadoop clusters involved in **large data retrieval** requests of **over 100 GBs**
- Assess the workflow and increase the efficiency of log delivery service that process **over 200 TB of data** on daily basis
- Deploy and maintain reporting infrastructure to facilitate visual representation of **URL reports, Unique Visitors etc**
- Create and deploy **Scripts/Alerts/Grafana dashboards/Playbooks** to monitor & automate daily tasks
- Perform **Backlog analysis, Data Migration, capacity planning** and check real-time production reports
- Act as **Incident Representative** for the DDC Specialist Team for high severity incidents

Platform Operations Engineer

Oct '15 - Nov '18

Akamai Technologies

Bengaluru, IN

Network Operations Command Center

- Level 2 Engineer maintaining a platform of more than **200,000 machines** across distributed network

EDUCATION

Post Graduate Diploma in Data Science (CGPA : 3.7/4)

Oct '20 - Oct '21

IIIT Bangalore & upGrad

Bengaluru, IN

Key Projects :

- [ETL Pipeline for Nord Spar ATM Transactions Dataset](#) | Tech Stack : Sqoop, PySpark, EC2, S3, Redshift
- [Credit Analysis \(Exploratory data analysis. \)](#) | Tech Stack : Python, NumPy, matplotlib, Seaborn
- [Lead Scoring Based on Logistic Regression](#) | Tech Stack : Python, Pandas, scikit-learn

B.E. in Information Science (CGPA : 8.26/10)

Aug '11 - Jul '15

Nitte Meenakshi Institute of Technology

Bengaluru, IN

PUBLICATIONS

- [Risk based access control in cloud computing](#) | [IEEE Conference Publication](#) | [IEEE Xplore](#) , IEEE 2015

AWARDS & CERTIFICATIONS

- **Spot award** titled **Network Urgency-Persistence-Implementation** for fixing a DDC network wide issue
- **Akalades award** titled **Outstanding support for DDC!** for going above and beyond to resolve customer issue
- [Hackerrank Skill Certificate \(Python \)](#)
- [Programming for Everybody](#)