

Tom Kurian

tomkurian77@gmail.com | +44 7824075045 | Bristol, United Kingdom

github.com/tom | [linkedin/tom](https://linkedin.com/in/tom)

EDUCATION AND HONORS

University of Bristol, School of Computer Science, Electrical and Electronic Engineering, and Engineering Maths Bristol, UK
MSc Data Science Sep 2021 - Sep 2022

- Relevant Coursework: Machine Learning, Deep Learning, Large Scala Data Engineering, Probability & Statistics, Cloud Computing, Visual and Textual Analytics

Mahatma Gandhi University, Mar Athanasius College of Engineering Kothamangalam, India
Bachelor of Technology with Honours (Electronics and Communication Engineering) Aug 2014 - May 2018

- Cumulative GPA: 8.66/10.0
- Academic Topper in third, fifth, sixth, and eighth semester

PROFESSIONAL EXPERIENCE

Tata Consultancy Services Chennai, India

Data Engineer Aug 2018 – Aug 2021

- Designed and implemented the organization's real-time and batch data pipeline solutions to process semi-structured and structured data from many data sources such as Kafka, S3, HDFS, Relational Databases using PySpark, PyFlink, and Elastic Map Reduce with various data sinks.
- Designed and implemented a real-time change data capture application in PyFlink and PostgreSQL to sync the employee learning progress, which further reduced the incoming tickets based on the sync failures.
- Remodelled the existing PySpark jobs with Spark deployment best practices and concurrency, which reduced the running time by half and an annual cost reduction of 75 percent.
- Designed and built the organization's data lake in AWS S3 and used AWS Athena to run ad-hoc queries on top of it and visualised the results in Apache Superset.
- Migrated informatica based ETL jobs to PySpark using AWS EMR, this reduced the operation cost to a certain extent.
- Automated the periodic ETL jobs using Jenkins and Boto3, which reduced a considerable amount of manual workload.
- Crafted a serverless automated reporting solution for summary reports using AWS Lambda, AWS Glue, AWS Simple Email Service (SES), and AWS S3.
- Part of the talent development team in Tata Consultancy Services and handled developing key performance features based on employee learning progress data.
- Implemented a containerised real-time ticket classification application using Flask, Sklearn, and Docker and deployed in AWS EKS, which helped the operations team to classify the incoming tickets in real-time.

Tata Consultancy Services Kochi, India

Data Engineering Intern Dec 2017 – Jun 2018

- Designed and implemented a batch data processing application using PySpark, and Hive in Cloudera Environment for an FMCG client. Each job is scheduled using Oozie schedulers.
- Implemented several Tableau Dashboards to display various customer KPIs.

SKILLS

- Programming Languages: Python, Scala, Java, C, Bash
- Big Data Frameworks: Apache Spark, PyFlink, Apache Hive, Apache Kafka, Apache Hudi, Apache Sqoop
- Databases and Data Warehouses: PostgreSQL, MySQL, Cassandra, Amazon Redshift
- Visualization Tools: Tableau, Apache Superset, AWS QuickSight
- Machine Learning Frameworks: scikit-learn, TensorFlow 2.0, AWS SageMaker
- Backend Frameworks: Flask, Python Fast API
- DevOps Tools: Docker, Jenkins, Terraform, Kubernetes
- Version Control: GitHub, GitLab
- AWS Services: Elastic Map Reduce, Athena, Redshift, Lambda, Glue, S3, Presto in EMR.

AWARDS AND ACHEIVEMENTS

Innovation Pride Award

April 2020

Tata Consultancy Services

- Secured first prize in inter unit big data hackathon conducted within the organization.
- The theme of the hackathon was to create a cost-effective data pipeline to read data from MongoDB, apply the required transformations, and write to S3 data lake. The entire job should be automated using a suitable scheduler.
- Tools: PySpark, MongoDB, S3, Apache Hudi, AWS Athena, Jenkins

Innovation Pride Award

February 2020

Tata Consultancy Services

- Secured runner up position in inter unit hackathon conducted within the organization.
- The theme of the hackathon was to create an attrition prediction system to help the HR team.
- Tools: XgBoost, FastAPI, ReactJS

Applause for Team Award

May 2020

Tata Consultancy Services

- In appreciation for the contribution given for the organization's campus placement drive program.
- Delivered workshops and online sessions for explaining the recruitment process for various colleges.