

# Kartik Batheja

## Senior Data Engineer

Bangalore | 09999691088 | [kartikbatheja@gmail.com](mailto:kartikbatheja@gmail.com)

### Summary

Seasoned Big Data Engineer with 7 years of extensive experience designing, coding and testing data solutions in development and production environments. Proficient in working with diverse technologies such as Spark, Azure data factory and big data technologies. Motivated to stay informed of latest industry innovations to provide users with most up-to-date solutions.

### Core Competencies

- Maersk central datalake engineer
- Managing different pattern sources
- Spark job designer
- Performance management
- Troubleshooting
- Analyzing data
- Worked in different environments.
- Client Service
- Proficient in cloud computing.
- Willing to work as individual contributor as well as team player.
- Self motivator
- Critical thinking for complex problem.
- Problem solving capabilities.
- Lead by setting example.

### Work experience

Sr. Data Engineer Dec, 2019 - Present

Maersk GSC

- Extensively worked on Azure cloud platform to create a datalake having data for numerous sources.
- Collected requirements and designed pipelines to ingest data from different sources and publishing the data as is in raw layer.
- Design strategies to capture historical loads and ingest incremental data.
- Design Spark jobs to parse and cleanse the data on metadata provided by the source.
- Design partitioning and purge strategy on cleanse layer of datalake.
- Implement business logic over cleansed data to be used for different platforms.
- Contributed to internal activities for overall process improvements, efficiencies and innovation.
- Closely collaborated with project members to identify and quickly address problems.

Big Data Engineer Feb, 2019 - Dec, 2019

Mediakind Pvt. Ltd.

- Development and maintenance of the data platform utilized by all platforms teams.
- Development and optimiziing of highly scala recommendation jobs over Spark with Scala.
- Developed, implemented and supported the analytics engine using Spark jobs, Kafka and different tools.
- Extensively worked on ELK stack to get the realtime stats.
- Analyzing data to debug specific issue pertaining to spark jobs.

Application Developer July, 2017 - Feb, 2019

Alphaserve Technologies

- Worked closely with other SOC analysts and infrastructure specialists to deliver high availability solutions for developing a SIEM tool from scratch using several open source frameworks in a phased manner.

### Skills

Spark 2.4

Spark jobs, Spark SQL, Spark Streaming

Databases

MySQL, Azure SQL Database

Languages

Scala, Python

Applications

Azure Cloud Platform - Data Factory, Datalake, EventHub, Function App, Logic App

ELK Stack - ElasticSearch, Logstash, Kibana

Databricks with Azure

- Setting up the environment for Hadoop tools and services over a cluster.
- Implementing Active directory authentication using NGINX and Zuul proxy.
- Created the final product using tools like ElasticSearch, Logstash and Kibana.
- Ingesting streaming data from different source systems to a central system using Nifi and Kafka services
- Parsing this continuous data to a common format using Spark Streaming.
- Storing high frequency data to ElasticSearch in indexes and implement data analysis over Kibana.

Sr. Software Engineer

June, 2014 - July, 2017

Infosys Limited

- Developed web application to view, edit and add checks and documents used by large US Bank.
- Delivered and maintained scalable system architecture to support high-availability Internet site with various internal applications.

## Education

---

Bachelors in Technology

2010 - 2014

GGSIU