

# Srinivas Reddy Alluri

Singapore

[srinivas96alluri@gmail.com](mailto:srinivas96alluri@gmail.com)

+65 82255712

[linkedin.com/in/mrsrinivas](https://www.linkedin.com/in/mrsrinivas)

<https://mrsrinivas.com>

## Summary

- Lead Engineer with a decade of experience designing and developing big data applications and platforms in the cloud and on-prem.
- Hands-on knowledge of tools like Apache Spark, Apache Kafka, AWS, Presto, Airflow, Scala, and Java.
- Built petabyte(PB) scale multi-tenant cloud data platforms and services iteratively.
- Acquired new skills based on project needs with self-learning via online and peers.
- Ranked all-time top 11 answerers in StackOverflow for Apache Spark, Apache Hadoop, and Spark SQL tagged questions.

## Experience

### **Cloud Data/ML/Backend Engineer | Experimentation Platform**

Grab

Oct 2019 - Present (3 years +)

My day-to-day work in GrabX(experimentation platform) team,

- Leads a team of data and backend engineers in developing and operationalizing (data) products/features (scale per day: ~ 80 TB size, 170 B records)
- Creates RFC(proposal) documents from an idea/problem, then gets the team/department level alignment for project execution and rollout from various stakeholders.
- Mentor team members with code/document review in regular sync-ups.

Achievements:-

- Conceptualized a multi-tenant data platform for Grab subsidiaries to leverage the GrabX platform in an iterative agile manner.
- Developed modular RBAC user authorization solution for GrabX multi-tenancy.
- Successfully migrated cloud data lake (~ 2.5 PB) from AWS to Azure with minimal interruption to downstream users.
- Improved the data pipeline cost efficiency by 50% with pragmatic cluster scaling rules and jobs performance tuning with better compressions.

### **Data Platform Engineer/Solution Arch - Apache Kafka | Enterprise Data Security**

Accenture

Jun 2018 - Oct 2019 (1 year 5 months)

As a Data Solutions-Architect/Engineer for DBS Bank ADA datalake platform,

- Worked in the Enterprise Data Security Framework(EDSF) team to provide cell-level security enforcement on datasets in the data lake platform with BlueTalon and Protegrity.
- Solution Architecture the data pipelines in the platform using tools like Airflow, Alluxio, Spark, YARN, Hive, Object store, and Kafka.
- Worked in the Enterprise Streaming Platform team to setup Kafka/Connect/SchemaRegistry in data lake then integrate with metadata system and kerberos authentication with Active directory.

- Publish best practices like code samples and documents to application teams for the utilization of tools in the platform.

Skills: Team Leadership · Data Engineering · Software Development · Stakeholder Management · Apache Spark · Communication · Database Design · Mentoring · Agile Methodologies · Apache Kafka · Apache Airflow · Scala · Information Security · Data Structures · Multithreading · Spring Boot · Web Development · Java · Git · Alluxio · REST APIs · Problem Solving · Kerberos · Big Data Analytics · Data Governance

## > Senior Bigdata Engineer - Spark/Hadoop Developer

### Accenture

Feb 2017 - Jun 2018 (1 year 5 months)

My responsibilities in Kogentix(Acquired by Accenture AI) include but are not limited to

- Contributed to designing and implementing data pipelines for in-house products and utilities using Oozie, Spark, HDFS, YARN, HBase, Kafka, Hive, and Sqoop.
- Worked for Neilson on a migration project from premise to Azure cloud environment and explored tools like Docker and Azure Batch Services, Azure Blob, and HDInsights.
- Given onsite training in Jakarta for new/experienced hires of Banking Client on Hadoop. Spark, Hive, HBase, Spark, Ozzie, and SQL.
- Trained 30+ people of various levels like Interns, Mid and Senior level people in Kogentix on Hadoop and Spark.

Skills: Team Leadership · Data Engineering · Software Development · Apache Spark · Communication · Mentoring · Agile Methodologies · Scala · Data Warehousing · Data Analysis · Spring Boot · Hadoop · Web Development · Java · Python · Git · Problem Solving · Hive · Kerberos

## Model N Senior Member of Technical Staff - I | Hadoop/Spark Developer

### Model N

Jun 2016 - 2017 (1 year 7 months)

Model N's Pharma RMC suite enables customers to maximize revenue and minimize compliance risk across their product portfolios, customer channels, and geographies. It delivers measurable ROI through the end-to-end applications designed to manage the gross-to-net lifecycle.

- Written Hive Generic UDFs to implement customized functions in Spark SQL.
- Involved in setting up Hadoop cluster using Apache components for Dev and QA.
- Rewrite Oracle SQL to ANSI SQL compliant ANTLR v4 grammar utility.
- Tunes jobs performance by using YARN Application Master and Spark History Server.
- Written Spark applications for Government Pricing calculation and submitted over EMR-YARN.
- Transforms the data in S3 from CSV to Parquet using Spark SQL external tables.
- Attends Scrum call with a team of 10 members.

Skills: Data Engineering · Software Development · Stakeholder Management · Apache Spark · Amazon Web Services (AWS) · Agile Methodologies · Scala · Data Warehousing · Data Analysis · Hadoop · Java · Git · Problem Solving · Hive



## Software Engineer - III | Hadoop/Spark Developer

### Verizon

Feb 2014 - Jun 2016 (2 years 5 months)

Srinivas worked in a team of 5 to define and develop data/ML pipelines for Verizon's security products, Universal Identity Services(UIS) and Managed Security Services(MSS).

- Written MapReduce and Spark jobs to reveal UIS user authentication/usage patterns and MSS's telephone denial of service protection module.
- Deployed an API monitoring portal using Splunk engine and REST API as part of the Hackathon.
- Boosted QR code authentication performance by 60% using binary techniques.
- Responsible for managing data coming from different sources and involved HDFS maintenance and loading structured and unstructured data with Sqoop and Flume, respectively.
- Mentored juniors in designing and developing REST APIs (JAX-RS) and UI layer.

Skills: Team Leadership · Data Engineering · Software Development · Stakeholder Management · Apache Spark · Communication · Database Design · Mentoring · Agile Methodologies · Information Security · Data Structures · Data Analysis · Hadoop · Web Development · MongoDB · Java · MySQL · Python · Git · REST APIs · Problem Solving · Hive · Big Data Analytics



## **Software Engineer - FullStack Developer with Java/PHP**

### **Tech4sys Software Solutions**

May 2011 - Jan 2014 (2 years 9 months)

Roles and Responsibilities:

- Collaborated with the client, project lead, and team members to define and develop application features and functionalities.
- Designed database and provided the architecture for a job portal (JobCoconut).
- Integrated payment gateways for Maventus and Planlab applications.
- Worked on Facebook graph API and Twitter API to post multimedia content through Maventus.
- Added Social Networking (Facebook, Google, Twitter, and LinkedIn) authentication and registration for e-commerce and job portal applications.
- Involved in all phases of software development, from application design to deployment.
- Programmed in Java and PHP with an object-oriented programming mindset, keeping an eye on code abstraction and reuse.

Skills: Software Development · Database Design · Amazon Web Services (AWS) · Mentoring · Data Structures · Web Development · Java · MySQL · Git · REST APIs · Problem Solving · PHP · HTML5 · JavaScript

## **Education**



### **National University of Singapore**

#### **Masters, Software Engineering**

Certificates in 'Architecting Scalable Systems', 'Securing Ubiquitous Systems', and 'Architecting Smart Systems'

Projects:

1. Smart Door Lock is a managed platform to help users unlock their face recognition and other authentication factors like a passcode.
  - Contributions: Platform Architecture, Solutions Architecture for components integration, face recognition module on edge, and model training in the cloud.
  - Tools/Technologies: Keras, TensorFlow, Anaconda, Docker, Eclipse Kapua, Eclipse Kura, Raspberry Pi, Pi Camera, AWS, NodeJS, and ReactJS.

2. Reach52 Data Analytics Platform is a cost-effective org level datalake system that enables Data Analysts and Business Users to get reports from various data sources in one place.

- Contributions: End to End Platform Architecture, Data pipelines design, and development, Data Quality check implementation and Platform setup.
- Technologies: Apache Spark, Spring Boot, Apache Airflow, Amazon S3, Presto, MySQL, ReactJS and NodeJS.


## Licenses & Certifications

 **M101J: MongoDB for Java Developers** - MongoDB

 **Oracle Certified Professional, Java SE 6 Programmer** - Oracle

 **ITIL® Foundation Certificate in IT Service Management** - EXIN  
5140126.20311315

 **Certified Developer for Apache Spark - Databricks & O'Reilly** - O'Reilly Media  
1.x-0920

 **AWS Certified Machine Learning – Specialty** - Amazon Web Services (AWS)

## Skills

Team Leadership • Data Engineering • Software Development • Stakeholder Management • Apache Spark • Communication • Amazon Web Services (AWS) • Mentoring • SQL • Representational State Transfer (REST)