

SRINIVAS REDDY ALLURI

+65 82255712 | srinivas96alluri@gmail.com

 [Stackoverflow](#) |  [Github](#)

Professional Summary

- Passionate Software/Data Engineer over 10 years of experience, with proficiency in platform/solution architecture, developing robust data pipelines and software systems.
- Enthusiast of open-source software, contribute back to community whenever possible by writing code to **Apache Kafka** and answering Bigdata (**Apache Spark** and **Hadoop**) questions in Stackoverflow.
- Quick at acquiring of modern technologies and ability to work in a team as well as individual.

Skills Summary

Proficient	Spark, Kafka, HDFS, YARN, Scala, Java, Hive, Alluxio, ANSI SQL
Comfortable	Airflow, Python, Presto, Amazon EMR, Amazon S3, Kubernetes, Gitlab, Git, MongoDB, Docker, JavaScript, PHP, MySQL, Oracle 11g, JAX-RS, JUnit, MRUnit, Antlr 4, Linux, Maven, SBT, IntelliJ
Familiar	Snowflake, Flume, Apache NiFi, Orient DB, Azure, Splunk, Sonar, Nexus, Jenkins, Web Development

Grab, Singapore (Oct 2019 – till date)

Lead Data Engineer

Project: Grab Experimentation Platform.

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

- Leading the team of 3 engineers in building data products/services.
- Written and maintains the core data pipeline sized **~ 80 TB (170 B events) per day** on cloud using **Apache Spark, S3, Airflow, Presto, Terraform, Kubernetes** and **Datadog**.
- Developed modular RBAC solution for experimentation platform user authorisation needs.
- Successfully migrated **cloud data lake (~ 2.5 PB) from AWS to Azure** with minimal interruption to downstream users.
- Improved the data pipelines cost efficiency by 50% with accurate scaling rules and performance tuning of jobs/cluster with better compressions.

Accenture, Singapore (Jun 2018 – Oct 2019)

Associate Bigdata Architect

Project: DBS ADA Platform (Client)

As a Consulting-Engineer/Solutions-Architect for DBS ADA data lake platform,

- Design and develop metadata driven data security framework to provide in house Authorization and ACL solution on datasets/tables with BlueTalon and Protegrity as implementation.
- Solution Architecture the data pipelines in the platform using the tools like **Airflow, Alluxio, Spark, YARN, Hive, Object store and Kafka**.
- Worked in Enterprise Streaming Platform team to setup Kafka/Connect/Schema Registry in data lake then integrate with metadata system and Kerberos authentication with Active directory.
- Written RFC proposals and design docs for data security and streaming systems in platform.

Accenture India (Feb 2017 – Jun 2018)**Senior Data Engineer**

My responsibilities in Accenture, includes but not limited to

- Contributes to the design and implementation of 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 in Bank Danamon on Hadoop, Spark, Hive, HBase, Ozzie and SQL.

Model N India (Jun 2016 – Jan 2017)**Senior Data Engineer****Project: Revenue Management Cloud**

Model N's Pharma RMC suite enables customer to maximize revenue and minimize compliance risk across their product portfolios, customer channels, and geographies.

- Written **Hive Generic UDFs** to implement customized functions in **Spark SQL**.
- Involved in **setting up Hadoop cluster** using Apache components for Dev and QA.
- Rewritten **Oracle SQL** to ANSI SQL compliant **ANTLR v4** grammar utility.
- Tuned jobs performance in data pipelines with the help of YARN web UI.
- Owned **data flow pipeline** for RCP number calculation for 5i drugs in **Amazon EMR** infra.

Verizon India Pvt. Ltd (Feb 2014 – Jun 2016)**Hadoop Spark & Java Developer****Project: Universal Identity Services**

UIS helps customers to provide various authentication methods to verify user identities to systems, applications, and sensitive information so they can work from any device, at any time, with less risk.

- Worked in **scrum team** of 13 to define, design, develop and deploy product features.
- Written **MapReduce** and **Spark jobs** to reveal end user authentication and usage patterns.
- Added Unit test cases with **JUnit** and **MRUnit** for Spark and MapReduce jobs respectively.
- Deployed API monitoring portal using **Splunk engine and REST API** as part of Hackathon.
- Used **Oozie workflow** engine to orchestrate various **data pipelines** in cluster.
- Developed POC with **Spark Streaming** for **Telephone Denial of Service** protection module.
- Optimizing the **Hive** queries using Partitioning and Bucketing for boosting query performance.
- **Mentored juniors** in technically and functionally to accomplish their day-to-day tasks.
- Involved in **Hadoop Eco System** installation and configuration over 40 nodes using CDH5.

Tech4sys Software Solutions (May 2011 - Jan 2014)**Software Engineer**

In tenure with Tech4sys worked with various customers, project leads and team members to defined and develop application features and functionalities. Brief of my roles and responsibilities included,

- Involved in all phases of product development **from an inception to deployable service**.
- Codes in **Java** and **PHP** by keeping an eye towards code abstraction and reuse with OOP.
- **Designed database schema** and provided **end to end architecture** for a jobcoconut.com.
- Programmed in **JAX-RS** and JPA to serve the API request from web application.
- Integrated **Single-Sign-On** and **Payment Gateways** for portals login and purchase modules.
- Boosted database queries performance by playing with indexes in **MySQL**.

- Deployed web applications in **AWS** using cloud products like **S3, EC2, Auto Scaling** and **Elastic Load Balancing**.

Academics

- Masters (Software Engineering) from **National University of Singapore** with certificates in "Architecting Scalable Systems", "Securing Ubiquitous Systems" and "Architecting Smart Systems"

Projects:

- 1. Smart Door Lock** is managed platform to helps users to unlock their using face recognition method along with other authentication factors like passcode.
 - **Contributions:** Platform Architecture, Solutions Architecture for components integration, face recognition module on edge and model training in 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 which enable 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.

Certifications/Independent courses

- AWS Certified Machine Learning Specialty – Amazon Web Services
- Certified Developer on Apache Spark – Databricks and O'Reilly
- Snowflake Certified Associate - Snowflake Academy
- Functional Programming Principles in Scala– Coursera
- M101J: MongoDB for Java Developers - MongoDB University
- Certified Java Developer for J2SE-6.0 - Oracle
- Cyber security and Its Ten Domains - University System of Georgia through Coursera.

Awards/Achievements

- All-time **top 11 StackOverflow answerers** tagged on [Apache Spark](#) and [Hadoop](#) questions.
- INNOVATION award for "**API Monitoring Application**" in **Splunk Hackathon OCT-2015**
- ON-THE- SPOT award for "Gmail bypass Feature Development" JUL-2015
- ON-THE- SPOT award for "Outstanding Performance" NOV-2014
- Participated and organized various technical and cultural events at college level.

"You never fail until you stop trying!" - Albert Einstein