



Siby Abin Thomas *Senior Bigdata Consultant*

✉ sibyabin@gmail.com

☎ +91 984 76 454 76

📍 Kottayam, India

in <https://www.linkedin.com/in/sibyabin/>

SUMMARY

- Passionate data engineer with expertise in designing and developing big data processing solutions on the cloud.
- 12 Years of Total IT experience with 10+ Years of experience in ETL development and DWH modernization.
- Developed scalable data engineering pipelines utilizing delta lake technology.
- Certified expert in Apache Airflow . Implemented efficient and manageable airflow pipelines
- Key skills include Spark, AWS Cloud, Apache Airflow, Databricks and Python
- Proficient in AWS technology stack. Expert in designing re-usable data processing solutions
- Certified Pentaho Data Integration ETL developer.
- Developed highly tuned and customizable ETL solutions in Pentaho - MySQL technology stack.
- Worked for many international clients in onsite and remote mode and mentored resources across the globe.

SKILLS

AWS Cloud

● ● ● ● ●

S3 , Athena , Glue , EMR , Redshift , RDS , EC2 , DMS ,
Codecommit , Code Deploy , Lambda , Cloudwatch

Delta Lake

● ● ● ● ●

Databricks Delta , DeltaIO

Apache Airflow

● ● ● ● ●

Databases

● ● ● ● ●

MySQL , Redshift

ETL Tools

● ● ● ● ●

Pentaho ETL

Other Scheduling & Automation

● ● ● ● ●

Control M , Jenkins

Bigdata

● ● ● ● ●

Apache Spark , PIG , Hive , MongoDB

Python

● ● ● ● ●

Version Control

● ● ● ● ●

Github , Bit-bucket , AWS codecommit

Scripting

● ● ● ● ●

Bash Scripting , Groovy

Agile Tools

● ● ● ● ●

Jira , confluence

IDEs and Development Tools

● ● ● ● ●

Visual Studio Code , Databricks notebook , Jupiter
notebook

CERTIFICATIONS

- AWS Certified Cloud Practitioner (<https://cutt.ly/snkZglo>)
- Astronomer Certification for Apache Airflow Fundamentals (<https://cutt.ly/Lv7Hu64>)
- Pentaho Data Integration Consultant
- Mongo DB for Python Developers (M101P) - From Mongo DB university

AWARDS & RECOGNITIONS

- **Most Valuable Player:** Awards for Excellence at Infosys in Q4 2019
- **Most Valuable Player:** Awards for Excellence at Infosys in Q4 2018
- **STAR Awards - 2018:** Awarded by client AIMIA (<https://www.aimia.com/>) considering the Individual contribution to its global data platform design & architecture.
- **Most Valuable Player:** Awards for Excellence at Infosys in FY' 14
- **Global Customer Focus Award 2011:** Awarded by Satmetrix for successful implementations of their critical customers ING & Standard Life in a short span and making them valuable customers.

EDUCATION

- Master of Computer Application from RVS CAS - Coimbatore, India with 83.5% in April 2009
- Bachelor of Science from M.G University, Kerala, India with 85.4 % in March 2006

PROFESSIONAL EXPERIENCE

- | | |
|--|---------------------|
| Senior Consultant, Infosys Ltd | 2013 Mar – present |
| <ul style="list-style-type: none">▪ Data Lake implementation in AWS using Spark as a data processing tool.▪ Re-architected BI solutions using open source technologies & cloud and helped organizations to reduce the operational/licensing costs.▪ Developed analytics solutions on top of big data using Pentaho, Apache Pig, Hive, and Mongo DB.▪ Open source COE Spoc and mentored resources on Data and Analytics skill sets.▪ Pentaho ETL consultant for various clients | |
| Software Engineer, Satmetrix | 2010 Sep – 2013 Mar |
| <ul style="list-style-type: none">▪ Data migrations and data transformation using the Pentaho Data Integration tool.▪ Development of supporting plug-ins, solutions to integrate various products.▪ Implementation of Xperience product for Global customers.▪ Development of various tools for the professional services team to fasten the implementation activities.▪ Execution of Customer Experience Management program for various international clients. | |
| Software Engineer, Ryancom Infotech | 2009 Apr – 2010 Sep |
| <ul style="list-style-type: none">▪ Performed end to end delivery of the product "Athithi POS", from requirement analysis, development, and successful implementations of the product for 50+ clients in a short span.▪ Developed software products for touch screen machines and programmed s/w interfaces for electronic devices such as scanners, barcode readers, weighing machines, etc. | |

PROJECTS

- | | |
|--|--------------------|
| Data lake implementation in AWS for a Top 5 US Retailer | 2019 Apr – present |
| <p>Role : Senior Bigdata Consultant - Trivandrum , India</p> <p>Technology: Python, Spark, Apache Airflow, AWS data stack (S3, Athena, Glue, Redshift), Boto3</p> <ul style="list-style-type: none">• Proposed ideas and developed utilities to bring down the cloud operational cost.• Tuned spark jobs to improve the performance and thereby reduced costs.• Developed utilities to scale up/down the EMR clusters using AWS API's• Development of Pyspark jobs for data processing and transformation.• Designed Dynamic Airflow DAGs for orchestration of jobs | |

Data lake implementation - Aimia Global Data Platform | Loyalty Analytics

2016 Jun – 2019 Mar

Client : AIMIA (<https://aimialoyalty.com/>)

Role: Senior ETL developer (AWS) - London, United Kingdom

Technology: Pentaho, AWS Cloud (S3, DMS, Redshift), Python, Bash scripting

- Developed highly reusable ETLs with metadata injection framework which avoided the need for specific technical knowledge for developers and helped clients to save 40% on the development budget.
- Performed POC's and evaluated the latest tools and technologies to deliver a scalable and cost-effective product.
- Mentored teams on open source ETL tools.

DW Modernization for a US Retailer

2013 Mar – 2015 May

Client: Sears Holdings Corporation (<https://www.sears.com/>)

Role: Business Intelligence Lead - Chicago, USA

Technology: Pentaho, HDFS, Apache PIG, Apache Hive, MongoDB, Python, CTools Dashboards

- Involved in Re-architecture of BI solutions using open source technologies and helped the organization to reduce the operational/licensing costs.
- Gathered requirements, designed and developed highly customizable and maintainable analytics solutions using open technology stack(Pentaho/MongoDB/Python).
- Delivered Pentaho reports and dashboards to key stakeholders which helped them in making decisions to drive profitability for the business.
- Managed clients Pentaho server infrastructure and performed patch installations and code deployments

Various Projects at Satmetrix

2010 Sep – 2013 Mar

Xperience Product Implementation

Clients: Standard Life Germany, ING Banking, and Insurance (Worked with more than 10 systems across the globe), Colfax, Urban Science, Neustar, eBay, Philips, Experian, and Sodexo.

Role:

- Responsible for the company's KEY customer implementations - ING & Experian.
- Designed transformations and jobs using Pentaho based on customer requirements.
- Implemented a Customer Experience Management program for various international clients.
- Involved in modeling and development of internal tools for automating complex workflows in the Professional Services Department.

Data Migration projects

Clients: Cricket Communications, Experian, Sony Middle East, and ING accounts

Role:

- Helped to analyze client systems, gathered requirements, and implemented the migration framework and solution to seamlessly move data across the source system to the target system (Xperience).
- Worked with different teams across the globe and client's technical teams to discuss plans for migrations, risks, etc.

REFERENCES

Anand Raghavan, *Senior Technology Architect*, GenPact
anand.ragha@gmail.com

Mr. Ganesh Babu Vasantharajan, *CEO*, AgilizTech Software Services Pvt. Ltd.
vganeshbabu@outlook.com