**Lakshman R M**

Cloud Senior Data Architect

Bangalore, India • (91) 9676128646 • lakshanjobhit@gmail.com • https://www.linkedin.com/in/ml-rao-3a376b1b2/

**A seasoned 9 years Senior Big Data Architect with 18+ years of ITS experience in designing and implementing scalable data solutions. Demonstrates a strong command of big data technologies and cloud platforms.**

**EXPERIENCE**

**Minfy Technologies (Cloud Technology Partners), Bangalore** Feb 2022 -Sept 2024

**>> Senior Cloud Data Architect**

As a senior cloud data architect: Key responsibility is toalign client’s data strategies with business plans through detailed technical roadmaps, for a world-class technology services firm that incorporates industry insights and experience to deliver solutions for Cloud Technology Platform Customers.

* Define data architecture frameworks by translating business requirements into databases, data warehouses, and streams.
* Establish technical solutions for data delivery, integration, and consumption to augment transactional, operational, and analytical system performance.
* Execute clustering and fail-over strategies to assure availability of critical software.
* Develop, consolidate, and maintain data backup, redundancy, and storage plans to guarantee enterprise-wide data integrity, accuracy, and accessibility.
* Enhance integration packages for inbound data import and outbound data deliveries.
* Create application programming interfaces to retrieve data to drive for real-time analytics.
* Participate in the planning, implementation, and exercises concerning disaster recovery.

**Customers Served:** Cyble, Zolve, BMS, Dalmia Cements, Zelar (GCP Partner), Impelsys, Indus-Towers & Nanonets & Data Analytics Project on Bloomberg/Reuters Market Data Providers.

Thorough knowledge of end-to-end projects, implementing the Medallion Architecture using Microsoft Fabric tools like Dataflows, Power BI, and Azure Synapse. This architecture organizes data into Bronze, Silver, and Gold layers for enhanced scalability, clarity, and performance (Customer: Impelsys, Bangalore customer)

Latest PoC:

* AWS Project: Build an AI-Powered Chatbot with Amazon Lex, Bedrock, S3 and RAG

Learning & Development:

* Creating an Amazon Bedrock Agent
* Setting Up the Agent
* Creating a Knowledge Base
* Testing the Agent
* Adding Action Groups
* Final Testing and Troubleshooting
* Creating an Alias

# Completed mastering ‘**ArchitectureBytes Course on**/ **Amazon Bedrock Agents Tutorial - Architecture and Orchestration’ & Generative AI Fundamentals, Science Behind LLMs, Advanced Data Analysis using AI & Excel, Creation of Conversational AI Bots, Research-Write&Design a full presentation using AI, Created a personalized email workflow,**

**Latest PoC & Go-Live:**

**(Capital Market Trading Domain Project)**

**Capital Market Bonds, Trading market dataset was used to produce**

Key Metrics to Visualize**:**

* **Portfolio-level yield and duration distribution.**
* **Risk-adjusted returns (Sharpe Ratio).**
* **VaR(Value at Risk) by portfolio and asset class.**

**Data Ingestion:**

**Tools**:

* + Use **Databricks Autoloader** for streaming data ingestion from cloud storage (AWS S3, Azure Blob Storage).
  + Process **CSV, JSON, or Parquet** data files.

**Implementation Steps:**

* + **Configure Databricks Autoloader for continuous file ingestion.**
  + **Use Spark Structured Streaming to process real-time data streams.**
  + **Apply schema inference to parse data dynamically.**

**Data Transformation:**

* **Key Activities**:
  + Cleanse and enrich the data.
  + Compute derived metrics (e.g., modified duration, convexity, risk measures like Value at Risk - VaR).
  + Aggregate data by portfolio, asset class, or maturity buckets.

**Implementation Steps:**

* **Used Spark’s DataFrame API or SQL API for transformations.**
* **Created Delta Lake tables for enriched and aggregated data.**

**(Telecom Towers Operations Management Domain Project)**

* Azure Databricks PoC Project for Realtime/Streaming Analytics Product Development Expert.
* Data Transformation: Proficient in SQL, Python, and Spark for data transformation and preparation for analytics.
* Ex: for an ‘Infra development firm’: Extensively worked on Azure Databricks for data ETL Data Pipelines/from multiple file formats, also Configured and implemented the ADF triggers and scheduled the pipelines & created Azure stream analytics jobs to replicate the real-time data to load to Azure SQL Data /Azure DW.
* The tools used in the PoC project are: 1. Azure Data Factory 2. Azure Data Lake Storage Gen2 3. Azure Databricks 4. Azure Synapse Analytics 5. Azure Key vault 6. Azure Active Directory (AAD) and 7. Microsoft Power BI

Adapted the **ToC Analytics** use case to leverage **Google Cloud Platform (GCP)** as the cloud platform while maintaining **Databricks** for data processing and integrating other GCP services. This Project Architectural approach & Components were:

**GCP Components Used:**

1. **Google Cloud Storage (GCS)**: For raw and processed data storage.
2. **Databricks on GCP**: For Spark-based data processing.
3. **Delta Lake**: For storing processed data.
4. **BigQuery**: For SQL analytics.
5. **Cloud Composer**: For orchestrating the ETL pipeline.
6. **Looker Studio**: For data visualization.

Technical Approach:

1. **Reading nested JSON** data in Databricks.
2. **Flattening** nested structures for analysis.
3. Storing data in **Delta Lake** for reliability and performance.
4. Automating the workflow with **Airflow**.

Led migration project of big data workflows to Google Cloud Platform (GCP), resulting in a 30% increase in data processing efficiency

Implemented real-time data processing frameworks to process and analyse terabytes of data sets, increasing data accuracy by 45% using Cloud Technology Services (AWS).

Designed and optimized SQL queries and ETL operations, cutting down operation times by 35%

Championed data governance and security protocols on cloud platforms, mitigating security risks by 75%

Trained and mentored a team of 5 junior data engineers, improving team productivity by 20%

**SRS Business Solutions/CISCO Partners, Bangalore**  Jan 2020 – Jan 2022

**>> Cloud Data Architect**

Spearheaded the adoption of innovative Nextgen forecasting solutions software platform built on Spark/Scala/Cassandra DB/SnowflakeDB/Salesforce.com CRM, improving query performance by 60%

Leveraged GCP Dataflow to build high-throughput, fault-tolerant data pipelines, increasing managing data pipeline reliability by 45%

Automated data extraction and integration processes using GCP Data Fusion, reducing man-hours spent by 20%

Snowflake modelling – roles, schemas, and databases

Expert in Snowpipe (a continuous data ingestion service provided by Snowflake)

Set Up AWS DMS for CDC from **Oracle to Snowflake (Process CDC Data in Snowflake Using Streams and Tasks)**

**Innominds - Bangalore** Feb 2018 – Sept 2019

**>> Sr Data Architect**

<https://www.linkedin.com/company/vectorflow/>

**Project:** Security Operations Controlling Application frameworks development/Reports platform using latest cutting-edge tools – ex: Splunk

**Other Customers Served:** Hertz Rental/Aris Global Healthcare, NetSPI Security Solutions, JemStep Financials

**Responsibilities:**

TCoE team member to execute various CRM modules flawlessly/with zero- defects.

Led and Developed Bigdata/Data Science/EDW BI Solutions Architecture for Client Product Software (Security).

Implemented a data warehousing solution for NetSPI dashboard reporting/KPI Matrix Reporting.

Created and deployed Matrix, Grid, Chart, and Gauge Reports Utilizing ETL PostgreSQL Scripts. Created Parameters in PowerBI Desktop Reporting Manager.

Analyzing facts and dimensions for data analysis and design within PostgreSQL. Created KPIs for reporting of scorecards and/or performance management.

Designed and deployed ETL solutions using scripts. Imported data from OLTP, flat files, or other data sources and created data flow solutions to cleanse and validate imported data inserted and/or updated newly cleansed and validated data into OLTP or OLAP databases.

Created stored procedures, functions, common table expressions, and views from T-SQL; queries OLTP data with T-SQL and OLAP (fact and dimensions) DW expert

Completed usefully a pilot project using Data Science Anaconda/Jupyter/Power BI Desktop tool to ‘Key indication of Vulnerabilities & Findings, Issues as per Customer needs, trends.

Expert in Spark with Python & acted as a team member in Pilot/POC projects of PySpark for analyzing diverse datasets.

Report generation using Tableau, MicroStrategy Reporting Tool.

Led the development of Provisional Layer Application under the DWH Application frameworks (built using Talend, Power BI, Oracle 12c Database, and ExtJS Platform) for their Bioinformatics Analytics Division. The parallel processing of SQLs using Spring Java Batch Platform was developed along with an AngularJS UI interface.

PTG India (TechMinfy AWS Partners) Jan 2017 to Feb 2018

>> Sr Manager (AWS Cloud Platform)

Product development on Next Generation Solution: Using Sqoop connected to Oracle 12c DB, Hadoop/HDFS -> Hive Big Data Framework Platform, Tableau for Report Generation (POC), using AWS – Centos Linux OS. (POC completed)

Happiest Minds Technologies Aug 2014 – Jan 2017

>> **Associate Architect**

IT Services & CRM Applications using AMDOCS CRM development, upgrade, Oracle, Business Objects BI, Maintenance & Support

**Sep 2009 – July 2014:**

Worked as a self-consultant to establish Sales & Distribution Channel Partner Integrations between SAP S&D modules & Patanjali FMCG Stores (Across A.P. State & KA States).

I underwent the PGWPM 2013-2015 MBA Course from Great Lakes Institute of Management/CUOA Foundation University, Alta Villa Vincentia, Italy.

Hitachi Data Systems Aug 2000 – Aug 2009

>> Sr. Programmer & Analyst

IT Services & CRM Applications using AMDOCS CRM development, upgrade, Oracle, Business Objects BI, Maintenance & Support

**EDUCATION**

PGWPM from Great Lakes Institute of Management

Bachelor of Technology in Engineering

**SKILLS**

Programming Languages: Python, PySpark, Scala, Visual Basic – Clear Basic, SQL

Big Data Tools: AWS, Google Cloud Platform, Hadoop, Spark, Hive, Kafka, Flume, Pig

Databases: MySQL, PostgreSQL, Google Cloud Firestore, Google Cloud Spanner, NoSQL, Oracle

Data Visualization: Looker Enterprise BI, Business Objects BI, PowerBI, Amazon Quick Sight

Certifications:

Google Certified Professional Solutions Architect

AWS Solutions Architect & Data Analytics Speciality