***THARANI KUMARESAN*** *|* ***Fort Mill, SC ❏ 803.370.5633 ❏*** [***tharanimscs@gmail.com***](mailto:tharanimscs@gmail.com)

***Senior Data Engineer***

A well-qualified, dynamic, and highly successful Data Engineer with more than 14 years of experience with top-level companies. Skilled in software design, development, and maintenance services using multiple technologies. Adept at building Big Data systems to provide a Unified Analytics Platform (batch/streaming platforms); at conceptualizing and implementing data pipelines.

Able to convert data into informational insights to help organizations make data-driven decisions. Expertise spans banking and finance, communication, and energy & utilization sectors. Comfortable working in vibrant, research-oriented teams with concurrent projects. Bilingual: English and Tamil.

***CORE COMPETENCIES***

Oracle ExaData | SQL | PLSQL | Netezza | SQL Server | Teradata | Hadoop HDFS | YARN | MapReduce | Sqoop | Hive | Pig | Impala | Flume | Spark | Kafka | Storm | Drill | Zookeeper and Oozie | Spark | Scala | NO SQL Databases: HBase, Cassandra, MongoDB | Python | Java | C | C++ | JIRA | BitBucket | GIT | Jenkins | Visual Source Safe | WinCVS | IBM ClearCase | SVN | Amazon AWS-S3 | EC2 | VPC | Informatica | IBM Data Mirror | PLSQL | NZSQL | Autosys | Tidal | Microsoft SQL Studio | Eclipse | PyCharm | Spyder | Cognos BI | Tableau | QlikView

***NOTABLE ACHIEVEMENTS***

* Re-engineered a SAS platform to an Informatica platform with a new database and pipeline solutions to minimize job execution time from 8 hours to 30 minutes (saving 80-90%).
* As a data base performance tuning expert in the Enterprise Data Warehouse area, proactively proposed/implemented a job failure message user story that saved ~70% of debugging time to analyze the failure.
* Worked on the optimization of Spark-based data integration framework performance using best practices, thereby achieving 40% performance improvements, and automated ETL processes making it easier to wrangle data and reduce time by up to 40%.
* Increased the efficiency of the data fetching by ~30% by instituting query optimization and indexing.
* Constructed a new tool for creating, managing, and monitoring metadata for Big Data environments; the tool saved six months of five people’s work within a day; saved $10M annually from this project.
* Built an Operational Deposit Calculation engine that saved $30M in 2020, and saved a five-person resource reduction; operational execution time for calculations was reduced from 14 hours to 30 minutes.
* Received multiple awards including Most Valuable Player from Cognizant Technology Solutions; Most Valuable Player from Infosys Limited; On the Spot from Infosys Limited (multiple times); and the Global Recognition Award from Bank of America N.A. for leading SAS to INFA ETL re-engineering, architecting MLinQ application, architecting operational deposit, and data architecting NZ to Exadata re-engineering.

***EXPERIENCE***

**Senior Data Engineer: Client, Bank of America | Bank of America, N.A. |** Charlotte, NC **11/2016 – Present**

* Lead the Liquidity Risk Reporting Big Data team with 15 onsite/offsite people; interact with business professionals, application developers, and the technical staff in an Agile process and environment.
* Charged with developing technical standards, procedures, and guidelines to govern data models; with documenting success criteria; and with tracking solution effectiveness including system performance, adoption, and other key metrics.
* Author technical documentation, architecture designs, and data flow diagrams.
* Help direct technology initiatives focused on data and analytic solutions for the Marketing and Sales organizations within Global Investment Management.
* Coordinate the Development team and business partners in applying data modeling techniques; utilize Spark API over Cloudera Hadoop YARN to complete data analytics in Hive.
* Develop Scala scripts and UDFFs using data frames/SQL/data sets, and RDD/MapReduce in Spark 1.6 for data aggregation, queries, and writing data back into the OLTP system using Sqoop.
* Key contributor to a POC that compares the processing time of Impala with Apache Hive for batch applications to establish the former in the project.
* Construct a data pipeline with Amazon AWS to extract data from weblogs and then store it in HDFS.
* Work with Data Modelers/Business Analysts to comprehend requirements and the impact of the ETL on the business.
* Contribute to Star Schema Modeling (dimensions/facts) and ETL framework design activities.
* Employ Type I/Type II SCD mappings to update slowly-changing dimension tables; develop UNIX shell scripts and scheduled ETL load applying AutoSys (Job Scheduling Tool); design, develop, test, and maintain Tableau functional reports centered on user requirements; create donut charts; and institute complex features in charts like creating bar charts/tooltips.
* Demonstrate robust dashboard design expertise and an enthusiastic practitioner of effective data visualization; show familiarity with best practices for visualizations and designs.
* Provide production support to Tableau users; write custom SQL to support business requirements; and use Tableau server to publish/share the reports with business users.

**Senior Data Architect: Client, Bank of America | Diversant, LLC |** Charlotte, NC **06/2015 – 11/2016**

* Served as the Architect and Lead Data Engineer for the Enterprise Data Warehouse team that was used to store banking data for Operational Deposit Calculation; oversaw data warehouse plans for product; and partnered with data infrastructure to triage/resolve infra issues; and managed data warehouse plans for business verticals.
* Accountable for informing, influencing, supporting, and executing product decisions and launches.
* Led the delivery of high impact dashboards and data visualizations.
* Helped with Star Schema Modeling (Dimensions & Facts) and ETL framework design activities; assisted with the design and development of ETLs using Informatica Power Center tools: Designer, Workflow Manager, as well as Transformation Developer.
* Utilized Type I/II SCD mappings to update slowly-changing dimension tables.
* Created UNIX shell scripts and scheduled ETL loads using AutoSys (job scheduling tool).

**Big Data Lead: Multiple Clients | Infosys Limited |** Plano, TX **11/2010 – 06/2015**

* Built an enterprise data warehouse for the communication domain, with 12 people; worked with the Data Science and Analytics teams to define/establish project-specific data management approaches; and developed data ETL pipelines for multiple storage solutions including distributed platforms like Hadoop.
* Wrote and accurately maintained documentation on current environment, data locations/sources and models, ETL processes, and code.
* Worked collaboratively with other internal/external Analytics & Engineering teams to direct continuous improvements in metrics quality and processes.
* Served as a technical expert tasked with addressing issues with system/application design, performance, integration, and security; reviewed Hadoop clusters utilizing multiple big data analytic tools such as Pig, Hive, and MapReduce; and contributed to POCs with Apache Spark using Scala to implement Spark in the project, and consumed data from Kafka using Apache Spark.
* Responsible for loading and transforming large sets of structured, semi structured, and unstructured data; helped load data from LINUX file systems to HDFS; and employed Sqoop to import/export data into HDFS and Hive.

**Additional experience**:

**Business Intelligence Developer**, Cognizant Tech Solutions Pvt, LTD, Chennai

***EDUCATION & PROFESSIONAL DEVELOPMENT***

**Master of Science in Computer Science**, University of South Carolina, Columbia, SC

**Master of Engineering in Computer Science and Engineering**, College of Engineering Guindy, Anna University, Chennai, Tamil Nadu, India

**Bachelor of Science in Computer Science and Engineering**, Anna University, Chennai, Tamil Nadu, India

Hadoop Eco System

AWS Cloud Architect

Advance Python Coding

Oracle Certified Associate; Informatica Developer Certification