**Lakshman**

Email Id: lakshmanofficial41@gmail.com

Phone No: +1 913-(730)-7367

**PROFESSIONAL SUMMARY:**

* **9 years** of IT experience with a focus in designing, developing, and implementing **Data Warehouses** with proven experience in **Cloud migration.**
* Implemented solutions using **snowflake’s** snow pipes, stages, tasks, data sharing, cloning and time travel.
* Used containers in **Azure** to create stages for copying the data from stage to **snowflake**.
* Used **Azure cosmos DB** containers and collections to load and analyze the data.
* Experience with **Snowflake Virtual Warehouses**.
* Experience in building **Snow pipe**.
* In-depth knowledge of Data Sharing in Snowflake.
* In-depth knowledge of Snowflake Database, Schema and Table structures.
* Experience in using Snowflake Clone and Time Travel.
* Expertise in using **DBT** (Data Build Tool) to efficiently transform data and create scalable models in Snowflake.
* Created models in **DBT** moving the data incrementally from Stage to Final Consumption Marts.
* Worked on **Secure views** and Data Shares that are created for client specific data and other down streams.
* Extensive experience with **Amazon S3 bucket creation, AWS IAM** access configuration
* Having experience on **AWS (EMR, RDS, S3)**
* Having Experience in **AWS, Redshift.**
* Created different DBT models, macros and persist the data in different data marts in Snowflake.
* Good hands-on **CI/CD** process for deploying new and modified existing components to higher environments.
* Quite comfortable with **Teradata** Tools and utilities like **BTEQ, Fast Load, Multiload and Fast export.**
* Involved in performance tuning for the complex **SQL** by following various handling processes like collecting stats, by creating temporary tables, Indexes.
* Expertise in **Dynamic SQL, Collections and Exception handling**.
* Experience in **Oracle packages, Records**and**PL/SQL Tables**.
* Experience in **Oracle** SQL, high-performance **PL/SQL** code development, SQL query tuning, Data migration and Data replications with **Oracle 11g,12c and 19c**.
* Created Views as per the Business recommendations with all the filters to help downstream applications to capture data and generate reports.
* Thorough understanding of **Python** concepts like **Lists, Strings, Indexing, Slicing, Tuples** and **Dictionaries**
* Have Experience in Onsite Client Interaction, having good understanding of onsite and offshore model.
* Experience in leading and managing teams, Handled multiple roles – Tech Lead and Developer.
* Good Analytical, Strong interpersonal and excellent communication and managerial skills.
* Extensive experience with analysis, design, development, customizations, and implementation of software applications.
* Having very good working knowledge on versioning tools like **TFS, CVS, Clear case, and Clear Quest.**
* Good Experience on Scheduling tools like **Control-M and Tivoli Work Schedule (TWS).**
* Having number of appreciations to my credit by revealing customer delight and satisfaction, for the appropriate communication with users and timely deliverables.

**SKILLS:**

|  |  |
| --- | --- |
| **Cloud Computing** | Snowflake, Azure containers, Azure Cosmos DB, Azure Events, **AWS s3, IAM, AWS EC2, AWS SNS, AWS RDS** |
| **ETL Tools** | IBM WebSphere, DataStage 8.0,8.5, **DBT**, **Matillion** |
| **Databases** | Teradata, Oracle |
| **Languages/Technologies** | SQL, Python, PL/SQL, Snow SQL, Unix, Linux |
| **Version Control** | GIT, Team Foundation Server (TFS), Rational Clear Case, Sub Version (SVN), CVS, Clear Quest |
| **Operating Systems** | Linux, Windows  |
| **Scheduling** | Control-M, Airflow |

**PROFESSIONAL EXPERIENCE**

**Conoco Phillips, TX Sep 2022 – Present**

**Sr Snowflake Developer**

**Responsibilities:**

* Involved in the backlog grooming calls with BA for understanding the requirements and giving inputs on the technical feasibility as and when required.
* Data Integration team works on migration of FDR data from legacy databases to Snowflake database. The migration process involves generating feed files to Azure blob storage. Data from Azure blob is loaded into stage tables using snow pipe and further moved to ODS and Row Versioning Tables
* Created DDLs of Stage and ODS table DDLs and Repeatable scripts.
* Loading data from Azure Blob storage to Snowflake Staging and then to ODS layer Using repeatable scripts.
* Have used COPY statements to ingest data from stage to target tables.
* Implemented snow pipe for real-time data ingestion.
* Involved in creating standardization views to consumption layer based on mapping sheet.
* Created different DBT models, macros and persist the data in different data marts in Snowflake.
* Worked on Secure views and Data Shares that are created for client specific data and other down streams.
* Good hands-on CI/CD process for deploying new and modified existing components to higher environments.
* Good hands-on writing and testing the code as part of requirement.
* Establish and ensure adoption of best practices and development standards.
* Involved in Planning meetings and daily stand-up meetings.
* Communicate with peers and supervisors routinely, document work, meetings, and decisions.

**Environment:**Snowflake, DBT, Azure blob, Azure Cosmos Db, CI/CD, GIT, Kubernetes, Jira

**Cigna Healthcare, Remote June 2020 – Aug 2022**

**Snowflake Developer**

**Responsibilities:**

* Involved in Requirement Analysis & preparation of Functional Design documents.
* Table DDLs creation in snowflake development database.
* Have used COPY statements to ingest data from stage to Tables.
* Implemented snow pipe for real-time data ingestion.
* Cloned Production data for code modifications and testing.
* Work with multiple data sources.
* Created data sharing out of snowflake with testing team.
* Worked with streams for change data capture and implemented SCD.
* Implemented solutions using snowflake’s data sharing, cloning and time travel.
* Design, develop, and maintain ETL workflows in Matillion to extract data from various sources, perform necessary transformations, and load it into Snowflake.
* Create transformations within Matillion to clean, enrich, aggregate, and reshape data according to business requirements before loading it into Snowflake.
* Developed Orchestration job to orchestrate the flow in Matillion
* Used different transformation components like Join, Filter, Detect Change, Calculator, Sort, Distinct etc.
* Used different target components like Table Update, Table Output, Rewrite Table etc.
* Used different types of Variables like Job Variable, Grid Variable, Environmental Variable, Built-In/Automatic Variables while developing jobs to make job more configurable.
* Used S3 Load components to Load file from S3 to Snowflake.
* Have loaded data from external Database, Oracle to Snowflake using Database Query component.
* Performed Unit testing on each jobs developed.
* Involved in unit testing and Integration Testing.
* Establish and ensure adoption of best practices and development standards.
* Communicate with peers and supervisors routinely, document work, meetings, and decisions.

**Environment:** Snowflake, Matillion, Python, SQL, **AWS** services, CI/CD, GIT, Kubernetes, Jira.

**IQVIA , PA May 2019 - March 2020**

**DataEngineer**

* Implemented **Apache Airflow** for authoring, scheduling and monitoring **Data** **Pipelines.**
* Designed several **DAGs** (**Directed Acyclic Graph**) for automating **ETL** **pipelines.**
* Performed data extraction, transformation, loading, and integration in data warehouse, operational data stores and master data management.
* Strong understanding of AWS components such as EC2 and S3
* Performed **Data Migration** to GCP
* Responsible for data services and data movement infrastructures
* Experienced in **ETL** concepts, building ETL solutions and **Data** **modeling**
* Worked on architecting the **ETL transformation** layers and **writing spark jobs** to do the processing.
* Aggregated daily sales team updates to send report to executives and to organize jobs running on **Spark** **clusters**
* Loaded application **analytics** **data** into **data** **warehouse** in regular intervals of time
* Designed & build infrastructure for the **Google Cloud environment** from scratch
* Experienced in fact dimensional modeling (**Star** **schema**, **Snowflake** schema), transactional modeling and SCD (Slowly changing dimension)
* Leveraged cloud and **GPU** computing technologies for automated machine learning and analytics **pipelines**, such as **AWS**, **GCP**
* Worked on **confluence** and **Jira**
* Designed and implemented configurable **data delivery pipeline** for scheduled updates to customer facing data stores built with Python
* Proficient in **Machine Learning** techniques (Decision Trees, Linear/Logistic Regressors) and Statistical Modeling
* Compiled data from various sources to perform complex analysis for actionable results
* Measured Efficiency of **Hadoop/Hive** environment ensuring SLA is met
* Optimized the **TensorFlow** Model for efficiency
* Analyzed the system for new enhancements/functionalities and perform Impact analysis of the application for implementing ETL changes
* Implemented a Continuous Delivery **pipeline** with **Docker**, and **Git** **Hub** and AWS
* Built performant, scalable **ETL** processes to load, cleanse and validate data
* Participated in the full software development lifecycle with requirements, solution design, development, QA implementation, and product support using **Scrum** and other **Agile** methodologies
* Collaborated with team members and stakeholders in design and development of data environment
* Preparing associated documentation for specifications, requirements, and testing

**Environment:** AWS, GCP, Big query, Gcs Bucket, G-Cloud Function, Apache Beam, Cloud Dataflow, Cloud Shell, Gustily, By Command Line Utilities, Dataproc Cloud SQL, MySQL, Postgres, SQL Server, Python, Scala, Spark, Hive, Spark-SQL

**Verizon, Irving TX Oct 2017 – Apr 2019**

 **Data Engineer**

**Responsibilities:**

* Involved in Requirement Gathering, Business Analysis, and Translated Business requirements into Technical Design in Hadoop and Big Data.
* Great Hands-on experience working on different Hadoop ecosystem components like Hive, Sqoop, Spark, Kafka.
* Designed and executed data pipelines for collecting, processing, and storing telecom network performance data and telematics data from connected vehicles.
* In-depth understanding of Spark Architecture including Spark Core, Spark SQL, Data Frames, Spark Streaming, Spark.
* Data Ingestion to one or more Azure Services - (Azure Data Lake, Azure Storage, Azure SQL, Azure DW) and processing the data in In Azure Databricks.
* Developed JSON Scripts for deploying the Pipeline in Azure Data Factory (ADF) that process the data using the SQL Activity.
* Extensive experience in data integration using Informatica PowerCenter, a leading enterprise data integration and ETL tool.
* Utilize Java's capabilities to handle complex ETL scenarios and custom data manipulations.
* Imported and exported data into HDFS from the database and vice versa using Sqoop.
* Extract Transform and Load data from Sources Systems to Azure Data Storage services using a combination of Azure Data Factory, T-SQL, Spark SQL and U-SQL Azure Data Lake Analytics.
* Developed Spark Python modules for machine learning & predictive analytics in Hadoop on Azure.
* Expertise in designing and implementing data transformations, including data cleansing, filtering, aggregation, lookup, and join operations, to ensure data quality and consistency in the integration process.
* Created Partitions and Buckets in Hive for both Managed and External tables in Hive for optimizing performance.
* Utilized real-time data streaming technologies (e.g., Apache Kafka) to ingest and process live telecom and telematics data for immediate analysis.
* Develop automated ETL processes using Java to ensure consistent, reliable, and scheduled data movements and transformations.
* Used Talend to run the ETL processes instead of Hive queries.
* Successfully moved data from Hadoop to Cassandra using Bulk output format class.
* Extracted data from Teradata database and loaded it into Data warehouse using spark-JDBC.
* Implemented predictive maintenance models for telecom infrastructure based on telematics data, reducing downtime by 20%.
* Experienced in code repositories.
* Good Understanding of NoSQL database, ETL process and hands-on experience in writing applications on NoSQL databases like HBase, Cassandra, and MongoDB.

**Environment**: Hadoop YARN, Spark Core, Spark Streaming, Spark SQL, Python, java, Kafka, Hive, Sqoop, Azure cloud: ADF, Data Lake Gen2, Databricks, Impala, Cassandra, snowflake, Tableau, Talend, Oozie, Jenkins, Cloudera, Oracle 12c, Linux.

**Cygnet Infotech, Hyderabad, India Dec 2014 – Sep 2017**

**ETL Developer**

**Responsibilities:**

* Understanding the business requirements from data analysts. Based on the requirements defining the data required from source team for better usage in BIP.
* Designing and Understanding structure of source data behavior and their relationship for their mappings to target data.
* Work with DAs for Functional Spec Sign off.
* Involved in Impact Analysis of the Migration.
* Extracting the data from different kind of source files (Mainframe and Flat Files) and loaded to OIA (operational image area) tables and created the User views for their extraction.
* Developing the Mappings, Sessions and Workflows based on the user requirements.
* Involved in creating the Unix Scripts and jobs to handle Informatica workflows and Teradata utilities like Bteq, Fast Export and TPT scripts.
* Create and Use project’s best practices and coding standards
* Coordinate with End Users, Client and Source system team to meet the timely deliverables
* Extensively involved in different kind of testing environments and supporting to user verification.
* Maintaining the version of latest artifacts through Clear Case and extensively involved in release management activities.
* Responsible for delivering each system phase wise which includes coding, testing and production implementation.
* Interaction with warehouse analysts on domain related issues. This involved understanding of user requirements, customizations, enhancements and implementation.
* Manage/Monitor the production run and Change requests.

**Environment:** Informatica 9.x, Informatica Power Exchange, Teradata 13, UNIX AIX, TWS, Quality Centre, RATIONAL CLEARCASE, RATIONAL CLEARQUEST, RATIONAL TEAM CENTRE.

**Educational Details:** Bachelors, Computer Science, JNTUK-2015