A blue sign with white text

Description automatically generated with medium confidence

|  |
| --- |

‍SRUJANA

Dallas Fort Worth, TX | [srujanakaranam02@gmail.com](mailto:srujanakaranam02@gmail.com) |+1 ‪(682) 900-1758‬ |<https://www.linkedin.com/in/srujana1>

**SUMMARY:**

* Around 10 years of extensive IT Experience especially in Data Warehousing and Business Intelligence applications in financial, banking, and Healthcare Technology Solution industries. I was Snowflake certified.
* Have good Knowledge of ETL and hands-on experience in ETL.
* Development of Jobs in the Matillion ETL tool.
* ETL pipelines in and out of the data warehouse using a combination of Python, Snowflake, and Matillion Writing SQL queries against Snowflake.
* Experience in building Snow pipes.
* Define virtual warehouse sizing for Snowflake for different types of workloads.
* Used ETL methodologies and best practices to create Talend ETL jobs. Followed and enhanced programming and naming standards.
* Develop highly scalable, fault tolerant, maintainable ETL data pipelines to handle the vast amount of data.
* Interface with business customers, gathering requirements and delivering complete Data Engineering solutions.
* Experience in using Snowflake Clone and Time Travel.
* Progressive experience in the field of Big Data Technologies, Software Programming, and Developing, which also includes Design, Integration, and Maintenance.
* Hands-on experience with Snowflake utilities, SnowSQL, SnowPipe, and Big Data model techniques using Python / Java.
* Experienced in creating Hive, Pig, and custom map-reduce programs for analyzing data.
* Experience in validating and analyzing Hadoop log files.
* Experience in loading multiple larger datasets into HDFS and processing the datasets by using Hive and Pig.
* Familiar with data architecture including data ingestion pipeline design, Hadoop information architecture, data modeling and data mining, machine learning, and advanced data processing.
* Created Implicit, local, and global Context variables in the job. Worked on Talend Administration Console (TAC) for scheduling jobs and adding users.
* Worked on various Talend components such as tMap, tFilterRow, tAggregateRow, tFileExist, tFileCopy, tFileList, tDie, etc.
* Experience in validating tables with Partitions, bucketing, and Loading data into HIVE tables.
* Created Talend ETL jobs to receive attachment files from pop e-mail using tPop, tFileList, and tFileInputMail and then loaded data from attachments into the database and achieved the files.
* Organizing and managed the technical DATA team of data architects, data modelers, and Engineers, and providing leadership, guidance, and mentoring.
* Deploying, managing, and operating scalable, highly available, and fault-tolerant systems on AWS.
* Worked on snowflake connector for developing python applications.
* Hands-on experience on Hadoop technology stack (HDFS, Map - Reduce, Hive, HBase, Pig, Cassandra, Flume, Kafka, and Spark).
* Well versed with Talend Big Data, Hadoop, and Hive and used Talend Big data components like tHDFSInput, tHDFSOutput, tPigLoad, tPigFilterRow, tPigFilterColumn, tPigStoreResult, tHiveLoad, tHiveInput, tHbaseInput, tHbaseOutput, tSqoopImport, and tSqoopImport.
* Strong experience in migrating other databases to Snowflake.
* Participate in design meetings for the creation of the Data Model and provide guidance on best data architecture practices.
* Analyzing the source data to know the quality of data by using Talend Data Quality.
* Broad design, development, and testing experience with Talend Integration Suite and knowledge in Performance Tuning of mappings.
* Developed jobs in Talend Enterprise edition from stage to source, intermediate, conversion, and target.
* Participates in the development, improvement, and maintenance of snowflake database applications.
* Design and code required Database structures and components.

**TECHNICAL SKILLS:**

| **Programming Languages** | Python, Bash |
| --- | --- |
| **Query Languages** | SQL, PL/SQL, T-SQL |
| **ETL Tools** | Matillion, Talend |
| **Databases** | Snowflake, Hive, Oracle 10g/9i/8i, DB28.x, SQL Server 2012/2014 |
| **Reporting Tools** | Tableau |
| **Version control** | Git, SVN, |
| **Cloud Platform** | AWS, Azure |
| **Cloud Technologies** | EC2, S3, EMR, Redshift Spectrum, RDS SQL Server, Lambda, Step-functions, DynamoDB, SQS, SNS. |
| **Scheduling Tools** | Control-M, Crontab, Autosys, TAC |

**PROFESSIONAL WORK EXPERIENCE:**

**BCBS, Irving, TX Jan 2022 to Till Date**

**Senior Snowflake Developer**

**Responsibilities:**

* Created data sharing between two snowflake accounts (Prod—Dev).
* Migrate the database 500 + Tables and views from Redshift to Snowflake
* Redesigned the Views in snowflake to increase the performance.
* Unit tested the data between Redshift and Snowflake.
* Creating Reports in Tableau based on Snowflake Connections.
* Performed daily admin tasks of user registration, security configuration, and usage monitoring.
* Validated the Map Reduce, Pig, and Hive Scripts by pulling the data from the Hadoop and validating it with the data in the files and reports.
* Utilized SQOOP, Kafka, Flume, and Hadoop File System API for implementing data ingestion pipelines from heterogeneous data Sources.
* Experience with Data Analytics, Data Reporting, Ad-hoc Reporting, Graphs, Scales, pivot tables, and OLAP reporting.
* Developed spark applications in python(PySpark) on the distributed environment to load a huge number of CSV files with different schema into Hive ORC tables.
* Worked on reading and writing multiple data formats like JSON, ORC, and Parquet on HDFS using PySpark.
* Involved in End-to-End migration of 800+ Objects with 4TB Size from Sql server to Snowflake.
* Data moved from Sql Server 🡪 Azure 🡪snowflake internal stage🡪Snowflake with copy options.
* Created roles and access level privileges and took care of Snowflake Admin Activity end to end.
* Converted 230 views queries from Sql server🡪snowflake compatibility.
* Publishing customized interactive reports and dashboards, and report scheduling using the Tableau server.
* Administered users, and user groups, and scheduled instances for reports in Tableau.
* Involved in the installation of Tableau desktop 8.1, and Tableau server Application software
* Worked on snowflake connector for developing python applications.
* Installing python connectors and python connector API.
* Knowledge of Azure Site Recovery and Azure Backup Installed and Configured the Azure Backup agent and virtual machine backup, Enabled Azure Virtual machine backup from the Vault, and configured the Azure Site Recovery (ASR).
* Implemented a CI/CD pipeline using Azure DevOps (VSTS, TFS) in both cloud and on-premises with GIT, MS Build, Docker, and Maven along with Jenkins plugins.
* Experience in writing Infrastructure as a code (IaC) in Terraform, Azure resource management, and AWS Cloud formation. Created reusable Terraform modules in both Azure and AWS cloud environments.
* Experience writing Hive Queries for analyzing data in Hive warehouse using Hive Query Language (HQL).
* Validating the data from SQL Server to Snowflake to make sure it has Apple to Apple match.
* Consulting on Snowflake Data Platform Solution Architecture, Design, Development, and deployment focused to bring the data-driven culture across the enterprises
* Used different AWS Data Migration Services and Schema Conversion Tool along with Matillion ETL tool.
* Driving is replacing every other data platform technology using Snowflake with the lowest TCO with no compromise on performance, quality, and scalability.
* Created a Snow pipe for continuous data load.
* Very good knowledge of RDBMS topics, ability to write complex SQL, PL/SQL.
* Experience working with various distributions of Hadoop like Cloudera, Horton Works, and MapR.

**Environment:** Snowflake, Redshift, SQL server, BI Architect, AWS, AZURE, TALEND, JENKINS, and SQL

**WalMart, Charlotte, NC Nov 2020 to Dec 2021**

**Snowflake Developer**

**Responsibilities:**

* Created data pipeline for several events of ingestion, aggregation, and load consumer response data from AWS S3 bucket into Hive external tables.
* Worked on Oracle Databases, Redshift, and Snowflake.
* Core Technologies: End-to-end data pipeline with Snowflake, Matillion, Tableau, Python
* Creating lambda SQS queues to create event-based triggers for Matillion
* Create conceptual, logical, and physical models for OLTP, Data Warehouse Data Vault and Data Mart Star/Snowflake schema implementations.
* Used Alteryx for Data Preparation and then Tableau for Visualization and Reporting.
* Processed data in Alteryx to create TDE for tableau reporting.
* Working experience with Kimball Methodology and Data Vault Modeling.
* Define virtual warehouse sizing for Snowflake for a different type of workloads.
* Experience in Tableau Server administration tasks including Tableau server optimization and performance tuning.
* Major challenges of the system were to integrate many systems and access them which are spread across South America; creating a process to involve third-party vendors and suppliers; creating authorization for various department users with different roles.
* CSM certified, Worked in Scaled Agile (SAFE) environment as System/DBT QA, hands-on experience with Rally, JIRA.
* Involved in the design and development of GLS application developed in C/C++ on HP UNIX.
* Validated the data load process for Hadoop using the HiveSQL queries.
* Evaluate Snowflake Design considerations for any change in the application
* Build the Logical and Physical data model for snowflake as per the changes required
* Define roles, and privileges required to access different database objects.
* Define virtual warehouse sizing for Snowflake for a different type of workloads.
* Design and code required Database structures and components
* Build the Logical and Physical data model for snowflake as per the changes required.
* Worked on scalable distributed data system using the Hadoop ecosystem in AWS EMR, and MapR distribution.
* Educate developers on how to commit their work and how can they make use of the CI/CD pipelines that are in place
* Setup full CI/CD pipelines so that each commit a developer makes will go through the standard process of the software lifecycle and gets tested well enough before it can make it to production.

**Environment:** Snowflake, SQL server, AWS and SQL

**JPMC, Chicago, IL Jan 2019 to Oct 2020**

**Senior Data Engineer**

**Responsibilities:**

* Performance monitoring and Optimizing Indexes tasks by using Performance Monitor, SQL Profiler, Database Tuning Advisor, and Index tuning wizard.
* Acted as point of contact to resolve locking/blocking and performance issues.
* Wrote scripts and indexing strategy for migration to Confidential Redshift from SQL Server and MySQL databases
* Worked on AWS Data Pipeline to configure data loads from S3 to Redshift
* Used JSON schema to define table and column mapping from S3 data to Redshift
* Helped individual teams to set up their repositories in bit bucket and maintain their code and help them set up jobs which can make use of CI/CD environment.
* Wrote indexing and data distribution strategies optimized for sub-second query response.
* Primarily worked on a project to develop an internal ETL product to handle complex and large volume healthcare claims data. Designed ETL framework and developed several packages.
* Performing Data source investigation, developed source-to-destination mappings and data cleansing while loading the data into staging/ODS regions
* Involved in various Transformation and data cleansing activities using various Control flow and data flow tasks in SSIS packages during data migration
* Applied various data transformations like Lookup, Aggregate, Sort, Multicasting, Conditional Split, Derived column, etc.
* Developed Mappings, Sessions, and Workflows to extract, validate, and transform data according to the business rules using Informatica.
* Designed the data models to be used in data-intensive AWS Lambda applications which are aimed to do complex analysis creating analytical reports for end-to-end traceability, lineage, and definition of Key Business elements from Aurora.
* Worked on scalable distributed data system using the Hadoop ecosystem in AWS EMR and MapR (MapR data platform).
* Developed Mappings, Sessions, and Workflows to extract, validate, and transform data according to the business rules using Informatica.
* Worked with Various HDFS file formats like Avro, Sequence File, and various compression formats like snappy, and Gzip.
* Worked on data ingestion from Oracle to hive.
* Involved in fixing various issues related to data quality, data availability, and data stability.

**Environment**:SQL Server 2016, SSRS, SSIS, T-SQL, Snowflake, ELT Maestro, Oracle 11g, SQL \* Loader, Visual Studio 2015, XML, Excel

**Ratna Global Tech, India Oct 2016 to Dec 2018**

**Data Engineer**

**Responsibilities:**

* Designed and developed a database to house data from various data sources for all reporting needs.
* Created SSIS packages to extract, transfer and load data into the data warehouse on a daily basis.
* Created SSIS Packages for Incremental loading and slowly changing dimensions.
* Created SSIS Packages to import and export data from Excel Spreadsheets, People Soft, Text files, CSV files, and various direct sources.
* Executed the DTSX Packages by using DTExec Command.
* Created logging, custom logging, and error handling packages.
* Identified the Data sources and defined them to build the Data Source Views.
* Created and scheduled SQL jobs to run SSIS packages daily.
* Experienced in extracting data from various data sources like Oracle and SQL servers, raw files, flat files, Excel, and OLE DB sources from Client servers.
* Worked on different packages with various types of transformation tools like Lookup, fuzzy Lookup, derived column, join, merge join, conditional split and slowly changing dimension.
* Expertise in troubleshooting ETL packages by implementing SSIS functionalities, like breakpoints, checkpoints, and event handlers.
* Utilized configuration tables to update multiple SSIS packages simultaneously.
* Tune ETL procedures and STAR schemas to optimize load and query Performance.
* Prepare all ETL processes according to business requirements and incorporate all business requirements into all design specifications.
* Experienced in using SSIS catalog, File System SQL Server deployment process to deploy ETL packages.
* Utilized configuration tables to update multiple SSIS packages simultaneously.
* Implemented Custom Logging, Auditing, and custom Email notifications using SSIS.

**Environment**: MS SQL Server 2012/2008, SQL Server Integration Services (SSIS), SQL Server Analysis Services (SSAS), SQL Server Migration Assistant (SSMA), Oracle 11G, SQL Profiler, T-SQL, Power Bi Desktop, Power Bi pro, Visual Studio 2010, HTML5, Java Script.

**Infosun Technologies, India June 2014 to Sep 2016**

**ETL Developer**

**Responsibilities:**

* Involved in the requirement gathering, project plan, and effort estimations leading to successful product delivery.
* Create and maintain the database for Server Inventory, Performance Inventory
* Working with SQL, and T-SQL, VBA
* Involved in creating tables, stored procedures, indexes
* Creating / Running jobs with packages.
* Design, Develop, and Deploy Packages with WMI Queries.
* Importing Data from various sources like Excel, SQL Server, and Front base
* Creating Linked Servers to other databases like Front base, and importing data
* Ensuring Data Consistency, Analyzing the Data
* Deploy various reports on SQL Server 2005 Reporting Server
* Migrated hundreds of Physical Machines to Virtual Machines
* Conduct System Testing and functionality after virtualization
* Extensively involved in new systems development with Oracle 6i.
* USED SQLCODE returns the current error code from the error stack SQLERRM returns the error message from the current error code.

**Environment**: MS SQL Server 2012/2008R2, SQL server integration services (SSIS) 2012, SQL server reporting services (SSRS) 2012, and TSQL, MS Office, Notepad++

**Education Details:**

* Bachelors in Electronics and Communication Engineering, India May 2014