Vinod Singh

Snowflake Developer | Senior Data Engineer

vinodsinghinbox22@gmail.com

(302)-345-2729

Wilmington, DE, 19801

Work Authorization: USA Citizen

Summary:

- **10+ Yrs**. of experience in SQL Database/Data Analysis, Data Engineering, Data Modeling, Database Design, Database Programming, Implementation, Development, Snowflake and Testing of Data Warehousing Systems & ETL Applications using IICS and Informatica Power Center tools.
- Expertise in the complete software development life cycle process, including Analysis, Design, Development, Testing, and Implementation.
- Developed understanding about major components of AWS like EC2, AWS Glue, Airflow, DMS, KMS, Secrete Manager, IAM Role, S3 bucket in using snowflake and their ecosystem.
- Data Architecture Design: Design and implement the data architecture, including data models, databases, and data pipelines, ensuring efficient data storage, retrieval, and processing.
- Data Integration: Integrate data from various sources, such as transactional systems and external APIs into target Database, ensuring data consistency and integrity.
- Data Transformation and ETL: Develop and maintain Extract, Transform, Load (ETL) processes to cleanse, transform, and load data into the appropriate formats and structures required by the platform.
- Data Security and Privacy: Implement data security measures, such as encryption and access controls, to ensure the confidentiality and integrity of sensitive customer data. Performance Optimization: Optimize data retrieval and processing performance, implementing indexing strategies, query optimizations, and caching mechanisms to enhance the platform's overall performance.

- Data Quality Assurance: perform data quality checks, validation, and monitoring to ensure the accuracy, completeness, and consistency of data used within the digital bank and digital credit card platform.

- Extensive experience in Extraction, Transformation, and Loading (ETL) data from various data sources into Data Warehouse and Data Marts using IICS, Informatica PowerCenter tools (Repository Manager, Designer, Workflow Manager and Workflow Monitor)

- Bulk loading from the external stage (AWS S3), internal stage to snowflake cloud using the COPY command.

- Loading data into snowflake tables from the external stage using snow pipe.

- Used COPY, LIST, PUT and GET commands for validating the internal stage files.

- Used import and Export from the internal stage (snowflake) from the external stage (AWS S3).

- writing complex Snow SQL scripts in snowflake cloud data warehouse to business analysis and reporting.

- Worked on Tasks, Streams, clustering keys, virtual warehouse creation, time travel, failsafe, zero copy cloning and micro partitions.

- Work with structured/semi-structured data ingestion and processing on AWS using S3, Python.

- Design and implement secure data pipelines into a Snowflake data warehouse from different source systems and cloud data sources.

- Creation of best practices and standards for data pipelining and integration with Snowflake data warehouses

- Design, develop, and implement advanced data pipelines that bring together data from disparate sources, making it available to data scientists, analysts, and other users using a variety of programming languages (Unix, Python, etc.).

- Worked in Agile and Waterfall mode of Software Development methodology.

- Expertise in Data Warehouse/Data mart, ODS, OLTP and OLAP implementations, Analysis, requirements gathering, data modeling, ETL Design, development, System testing, Implementation, and production support.

- I have expertise in writing complex SQL quires, Oracle, PL/SQL programming, Stored Procedures, Package, Functions, Triggers, Views, Materialized Views, cursors and tuning the SQL query.

- Expertise working on Informatica transformations such as Source Qualifier, Joiner, Filter, Router, Expression, Lookup, Aggregator, Sorter, Normalizer, Update Strategy, Sequence Generator and Stored Procedure transformations.
- Strong SQL skills including queries, triggers, stored procedures, cursors, and functions, particularly with Oracle databases.
- Experience in the concepts of Data Warehousing, Data Lakes, Data Marts, ER Modeling, Dimensional Modeling, Fact, and Dimensional Tables, Star and Snowflake schema, Normalization/De Normalization.
- Experienced in tools like JIRA, Postman, GIT, and integration testing.
- Demonstrated leadership experience, including leading teams and managing project workflows.
- Capable of analyzing epics and translating them into user stories.
- Business knowledge spans Insurance, Telecom, Logistics and Healthcare domains.
- Motivated team player with excellent communication, interpersonal, analytical, and problem-solving skills.
- Experience in troubleshoot and debug issues related to DBT models and data pipelines.

Technical Summary:

Languages	Python, SQL, PL/SQL, Snow SQL
Databases/DW	SQL Server, Oracle, Snowflake
AWS Services	Amazon S3, AWS EC2, AWS Secrete Manager, AWS Glue, AWS Airflow, DMS
IDE's	PyCharm, Visual Studio
Other Tools	Apache Airflow, Snowflake, MuleSoft, GitHub, confluence, stash, Jira, control-m, AutoSys
ETL Tools	PL/SQL, SQL Server, Informatica 8.x, 9.x, 10.x, IICS Data Integration
Cloud	Amazon Web Services (AWS)

Professional Experience:

JP Morgan Chase, Wilmington, DE   10/2018 – 07/2024

Snowflake Developer | Senior Data Engineer

➢ Worked as a Sr. Data Engineer in part of team, using different ingestion methods loaded the data into raw zone in **Snowflake cloud**.

➢ Experience with **Snowflake cloud** data warehouse and **AWS S3** bucket for integrating data from multiple sources into **Snowflake table**.

➢ Designed and implemented **Snowflake stages** to efficiently load data from various sources into **Snowflake tables**.

➢ Created and managed different types of tables in Snowflake, such as **transient, temporary, and persistent tables.**

➢ Optimized **Snowflake warehouses** by selecting appropriate sizes and configurations to achieve optimal performance and cost efficiency.

➢ Developed complex **SnowSQL queries** to extract, transform, and load data from various sources into Snowflake.

➢ Implemented **partitioning techniques** in Snowflake to improve query performance and **data retrieval.**

➢ Configured and managed **multi-cluster warehouses** in Snowflake to handle high-concurrency workloads effectively.

➢ Defined roles and access privileges in Snowflake to ensure proper **data security** and **governance**.

➢ Implemented **Snowflake caching mechanisms** to improve query performance and reduce data transfer costs.

➢ Utilized **Snowpipe** for real-time data ingestion into Snowflake, ensuring continuous data availability and automated data loading processes.

- Leveraged Snowflake's **time travel** features to track and restore historical data for auditing and analysis purposes.
- Developed ETL workflows using **AWS Glue** to extract, transform, and load data from various sources into Snowflake.
- Implemented and maintained **DBT projects**, including version control and collaboration workflows.
- Worked on troubleshoot and debug issues related to **DBT models** and data pipelines.
- Used **DBT** to test the data (schema tests, referential integrity tests, custom tests) and ensures data quality.
- Documented transformation logic and maintained documentation for the **DBT models**.
- For version control of DBT projects used **git hub**.
- Created **DMS** End points to and **serverless replication** to load the data from source to S3 bucket location.
- Created **IAM** roles for Secrete Manager ID's, and for S3 buckets.
- Crated **Glue jobs** for each canonical to load the data from S3 bucket to snowflake Raw zone.
- Created **airflow instance**, created **airflow DAGs** to schedule the AWS glue jobs.
- Utilized **AWS CloudWatch** for monitoring, **logs** and managing resources, setting up alarms, and collecting metrics.
- Implemented complex SQL quires, Oracle, **PL/SQL programming**, Stored Procedures, Package, Functions, Triggers, Views, Materialized Views, cursors and tuning the SQL query.
- In-depth understanding of **Data Warehouse**, ETL concept and modeling structure principles
- Responsible for maintaining quality reference data in source by performing operations such as cleaning, transformation and ensuring integrity in a relational environment by working closely with the stakeholders and solution architect.
- Performed end-to-end **Architecture** and implementation assessment of various AWS services like S3, Glue, EC2, Airflow.
- Involved in migration of **Oracle objects** including queries, views, materialized views, tables structures, schemas, stored procedures and functions into **Snowflake cloud**.
- Involved in migration of **Tibco Data Virtualization** integration tool code into **Snowflake cloud**, it contains five different layers and each layer contains oracle objects including

views, materialized views, tables structures, stored procedures and functions into snowflake cloud, we have migrated all the objects in each layer and done the end-to-end testing.

➢ Involved in migration of **informatica mappings** code into snowflake cloud by converting them into **SQL code** and implemented.

➢ Bulk loading from **external stage** (AWS S3), internal stage to snowflake cloud using the **COPY** command.

➢ Loading data into **snowflake tables** from the internal stage using **snowsql**.

➢ Used **COPY, LIST, PUT, GET** commands for validating the internal stage files.

➢ Developed **snowflake procedure** for Incremental values, branching & looping.

➢ Created clone objects to maintain **zero-copy cloning**.

➢ **Data Validations** have been done through information schema.

➢ Involved in Migrating Objects from on prem to **Snowflake** and building up **snowpipe**, involved in doing data sharing, building up Table structures, schema and snowflake clone and involved in catapulting data from **oracle** to **snowflake** to consume.

➢ **Investigated** data sources to identify new data elements needed for **data integration**.

➢ **Collaborated** with cross-functional teams globally, ensuring efficient offshore-onsite coordination and project delivery.

➢ Created **effective test data**, developed, and performed thorough Unit test cases to ensure successful execution of the data loading.

➢ Developed **Python frameworks** to load data from AWS S3 to Snowflake, Oracle DB to AWS S3 and Rest API to AWS S3.

➢ Developed **test cases** to ensure data quality and integrity.

➢ Created **documentation** for the models and transformations to maintain clarity and ease of use for other team members.

➢ **Optimized** SQL queries to ensure efficient data processing.

Environment: Snowflake, Python, S3, Airflow, AWS Glue, DMS, JIRA, GIT, IICS, DBT, Confluence, Informatica PowerCenter, IICS, Oracle, Snow SQL, Snow Pipe, Agile, OLTP, OLAP.

State of CT – Department of Social Services, Hartford, CT   05/2015-10/2018

Senior Data Engineer

- ➢ Analyzing the source system based on business requirements and designed the **informatica mappings**.
- ➢ Developed **Mappings**, Mapping Configuration Tasks, Data Synchronization Tasks, Data Replication Tasks, PowerCenter Tasks, File Listener Tasks, Tasks Flows etc. using **IICS Cloud Data Integration Service**.
- ➢ Implemented commission calculation logic in informatica mappings based on the customer roles.
- ➢ Experience in almost all transformations in IICS like Source, Target, Aggregator, joiner, Normalizer, Sequence, **Lookup**, Hierarchy Builder, Hierarchy Parser, and Transactional Control
- ➢ Determine source, target, or mappings bottle necks for slow mappings.
- ➢ Experience in Creating different types of **Task flows** like Parallel tasks with Decision, Sequential Tasks with Decision and Linear Task flow.
- ➢ Performed Unit testing and has created test scripts for the interfaces developed.
- ➢ Created different types of Components like Saved Query, Hierarchical Schema and Fixed width file format, Shared Sequence and File Listener.
- ➢ Experience in **Error handling Techniques** in Task flows using Throw Step.
- ➢ Experience in Creating IICS complex mappings using Dynamic, Un-connected lookup, Incremental load, SCD Type-1 and Type-2 Techniques.
- ➢ Prepared Design documents as per the Business requirements for enhancements
- ➢ Involved in ETL process in development and production environments.
- ➢ Used tasks such as email, command task, file Listener task, decision, Assignment to build Workflows and Worklets.
- ➢ Implemented efficient performance tuning procedures to eliminate the **bottlenecks**.
- ➢ Worked on Dynamic Target Files generation.
- ➢ Implemented data Integration process from **Rest APIs** related to Workday and OKTA into snowflake.

- ➢ Worked on **Indirect File** method to load the flat files with the same structure.
- ➢ Bulk loading from external stage (**AWS S3**), internal stage to snowflake cloud using the COPY command.
- ➢ Loading SQL server data into snowflake tables using DMS.
- ➢ Used COPY, LIST, PUT, GET commands for validating the internal stage files.
- ➢ Used Snow Pipe for continuous data ingestion from the S3 bucket.
- ➢ Developed snowflake procedure for Incremental values, branching & looping.
- ➢ Created clone objects to maintain **zero-copy cloning**.
- ➢ Involved in implementation of Rest API data ingestion into Snowflake from OKTA and Workday.
- ➢ Migrated the Oracle code related to internal and external customers into **Snowflake cloud**.
- ➢ Fixed the migration issues and to make sure code is working fine we have **validated** the data in both Oracle and Snowflake cloud.

Environment: JIRA, GIT, IICS, Confluence, Control-M, Informatica PowerCenter, Oracle, Snowflake, SnowPipe, AWS, Agile, OLTP, OLAP.

Chevron, Pittsburgh, PA   6/2013-3/2015
ETL Informatica Developer

- ➢ **Collaborated** with cross-functional development team members to analyze potential system
  solutions based on evolving client requirements.
- ➢ **Identified** and **documented** detailed business rules and use cases based on requirements analysis.
- ➢ Developed code fixes and enhancements for inclusion in future code releases and patches.
- ➢ Collaborated with developers and performance engineers to enhance supportability and identify

performance bottlenecks.

➢ Created detailed **ETL mapping** documentation, Unit Test Plans, and System Integration test plans
to ensure comprehensive testing and easy maintenance.

➢ Designed comprehensive test cases, encompassing data from all relevant sources and publications, to validate the functionality of the new portal.

➢ Managed **unit testing** and collaborated closely with the testing team for Business Intelligence
(BI) testing and UAT.

➢ Involved in discussions on requirements with Onsite Functional and Technical leads. Worked in different kinds of source systems like SAP and JDE to analyze the data model.

➢ Prepared Requirement analysis documents and Design documents.

➢ Effectively handled critical business scenarios and developed mappings.

➢ Involved in analysis and **bug fixing** during SIT and UAT.

➢ Involved in **complex data loads**.

➢ Supporting for SPVC and BDP R&M activities.

➢ Used Informatica Power Center for (**ETL**) extraction, transformation and loading data from heterogeneous source systems into target database.

➢ Includes in designing the mappings between sources and targets by using different **Transformations**.

➢ Developed Informatica **workflows/worklets/sessions** associated with the mappings using Workflow Manager

➢ Prepared the Test Cases and Performed Unit Testing for Developed **Mappings**.

➢ Involved in Documentation of Mappings As per Standards.

➢ Created Informatica **schedules** to execute the Workflows.

➢ Monitoring the jobs using workflow monitor.

➢ Applied continuous integration best practices in general (SCM, build automation, unit testing, dependency management)

➢ Contributed to planning and continuous improvement activities.

➢ Supported the PO, Analyst and scrum master.

➢ Has extensive experience with oracle database, **SQL** and **PLSQL**.

- ➢ Knowledge and experience with widely successful Agile techniques: user stories, acceptance test driven development, continuous integration and continuous testing, pair working.
- ➢ Created and enforced the policies for effective data management.
- ➢ Formulated management techniques for quality data collection to ensure adequacy, accuracy, and legitimacy of data.

Environment: Informatica PowerCenter, Control-m, Oracle.

AIG, New York, NY   **7/2012-5/2013**

ETL Developer

- ➢ Analyzed client requirements and prepared **technical specification** documents to accurately capture project needs.
- ➢ Collaborated with the Data Modeling team during the design phase to ensure a seamless
  integration of the developed solution.
- ➢ Developed and executed **SQL queries** to retrieve and compare source and target tables, effectively identifying any missing records.
- ➢ Established and maintained source and target connections, configuring the application and
  documenting all necessary processes.
- ➢ Updated **Informatica** code based on User Acceptance Testing (UAT) feedback, facilitating,
  successful completion of UAT.
- ➢ Conducted thorough data analysis and developed robust mappings, including the implementation of **Slowly Changing Dimensions (SCDs)** and various transformations like joiner,
  aggregate, expression, filter, lookup, and union.
- ➢ Created ETLs, **mappings, sessions, and workflows** for final installation and release to the client,
  ensuring smooth implementation.

- ➤ Efficiently **monitored**, **scheduled**, and restarted ETL jobs for optimal performance and timely,

  execution.
- ➤ Performed seamless **migration** of Informatica objects and scripts across repositories without any disruptions.

Environment: Informatica PowerCenter, SQL, Oracle, Control-M.

Bank of America, Charlotte, NC   4/2010-6/2012

Data Analyst

Lead Data Analyst for the Bank of America (BoA) Global Corporate, Commercial, and Investment Banking Technology ("GCCIBT") organization. Worked on departments Enterprise Regulatory Data Modeling Practices for federal reporting. Established a granular reporting procedure in which bank reported every single loan by providing 95 credit, credit risk and financial data details to the authorities. Overcame the flaws of redundant and parallel processing steps for the same reporting data. Removed inconsistencies in the reporting, complexity at the banks and high costs. Worked the various levels of reports granularity to make them comparable and combinable within the supervisory authorities. Harmonized data management between the bank and supervisory authorities and to request granular data.

Involved in an effort to decommission 2 data marts and expand the data model of an existing Oracle data warehouse used as their DMM (Data Mining and Metrics) repository. Implemented a top-down/bottom-up analysis approach to assess the as-is data warehouse environment (7 legacy source systems and over 600 reports) and extend the to-be data models. Designed conceptual and logical target OLAP data models using Power Designer15.

Environment: Oracle, Toad, Power Designer15

Almarai-KSA   3/2004-2/2007

Data Analyst

Data Analyst support for the leading food manufacturing Gulf company. Analyzed data sent by 3rd party systems. Interacted with more senior level internal and external stakeholders. Facilitated review sessions with management, business users and other team members. Analyzed business reports that involved sales and shipping data.

Environment: SQL Server, MS Access, Erwin, DB2

Xerox, Gurgaon, India   4/2000-2/2004

Senior Business Analyst/Data Analyst

Hands on experience with data analysis, data profiling and data governance identifying data sets, source data, source metadata, data definitions and data formats.

Prepared reports by collecting, merging, analyzing, and summarizing information. Documented the complete process flow to describe program development, logic, testing, and implementation, application integration, coding. Wrote the test cases and technical requirements and got them electronically signed off. Performed gap analysis to check the compatibility of the existing system infrastructure with the new business requirement. Created pivot tables and charts using worksheet data and external resources, modified pivot tables, sorted items and group data, and refreshed and formatted pivot tables.

EDUCATION:

Masters in Bioengineering, University of Hawaii at Manoa USA, 2009

Training & Certifications in Snowflake:

1. **Snowpro Advanced** Architect Certified.
2. **Snowpro Core** Certified.