

**Sachin Naik**

**Snowflake Data Engineer / Teradata Consultant**

Around 10 years of experience in Teradata data warehouse and Business Intelligence platform, experienced in Bigdata and Cloud technologies.

# **PROFESSIONAL SUMMARY**

* Hands-on experience across all stages of Software Development Life Cycle (SDLC) including business requirement analysis, design and unit testing. Experienced in performance tuning and debugging of existing ETL processes.
* Analyze the current on premises Datawarehouse and ETL framework and design the modern cloud infrastructure to implement the business intelligence solutions to cater current and future business requirements.
* Design data pipelines and storage systems which integrate data across multiple shared applications in on premise systems as well as in cloud. Also design data load pipelines for historical data and current/future data load.
* Implementing effective data integration and optimization mechanisms in on premise data warehouse and on cloud platforms to ensure efficient data processing by following the data security standards in production environments.
* Migrating organization's IT infrastructure, applications, and data from on-premises servers or other hosting environments to cloud-based platforms like Amazon Web Services (AWS), and snowflake Datawarehouse.
* Collaborate with project manager and product owner to understand the business requirements and coordinate with the dependent teams and develop/enhance the applications and complete the testing, migration and production support activities.
* Implement the design, development, testing, optimize and production migration as part of day-to-day activities to complete the deliverables on time.
* Work cross-functionally with teams on data migration, translation, and organization initiatives. Participate in sprint planning meetings and provide estimations on technical implementation. Document every aspect of an application or system as a reference for future maintenance and upgrades.

# **TECHNICAL SKILLS**

* Database: Teradata Datawarehouse, Oracle, Hadoop, HDFS, Hive, Snowflake, AWS S3
* Cloud Platform: Google Cloud Platform (Big Query)
* Data Extraction Tools: Teradata SQL Assistant, Microsoft SQL Developer, HUE
* Streaming Platform: Kafka
* ETL Tools: Multiload, Fast-load, Fast-export, BTEQ, Talend
* Programming Languages: SQL, HQL, Shell Scripting, Python
* Others: Teradata Viewpoint, Putty, Tableau, Control-M

# **PROFESSIONAL EXPERIENCE**

**Anthem (Elevance Health), Virginia Nov’20 – Current**

**ETL Developer/ Teradata Developer**

* Designing and developing the overall data warehouse architecture, including data models, schemas, tables, and relationships.
* Creating and optimizing ETL (Extract, Transform, Load) processes to extract data from source systems, transform it into the desired format, and load it into the data warehouse.
* Design and develop data models. Design database objects, different semantic layers and integrations to ensure scalable and reliable applications in cloud.
* Handling data transformations, data cleansing, aggregations, and enrichment to make data suitable for reporting and analytics.
* Designing and building the data processing pipelines to transform data in Teradata platforms.
* Have strong Hands-on experience in Advanced SQL Data Profiling and Data modelling Knowledge, Implementing Slowly Changing Dimension types (type1, type2 and type3) through Structured Query Language to keep track of historical records.
* Creating and optimizing ETL (Extract, Transform, Load) processes to extract data from source systems, transform it into the desired format, and load it into the data warehouse.
* Created and modified BTEQ, MLOAD, FEXP and FLOAD scripts as business requirements
* Worked on various data loading activities like data loading from local to snowflake and from external stage like AWS s3 to table stage.
* Ensuring data quality by implementing data validation, cleansing, and transformation rules during the ETL process in Teradata and ELT process in AWS and snowflake Datawarehouse
* Implement data partitions, indexes, metadata, temporary and permanent tables for data transfer mechanism.
* Re-design of the process flow, starting with design considerations, technical specification and source to target mapping.
* Transferring, transforming, and validating data during migration. Converting the existing Teradata BTEQ queries to snow SQL, creating DDLs and coming up with strategies for historical data load, change data capture and handling the dependencies across the system.
* Implement data encryption, data masking and customer managed keys for data movement across accounts in cloud.
* Design data sharing and secure views to share data between different accounts and applications.
* Data modelling of the database objects using Erwin Data Modeler
* Performance tuning of complex and long running SQL queries in Teradata by analyzing the explain plan.
* Used Teradata BTEQ to load the day today data load activities.
* Designing and implementing data pipelines to migrate the data from Teradata to Snowflake environment. Worked on creating different objects like Tables, views, stages, integration objects etc. in snowflake environments.
* Prepare Technical specification document (TSD) to outline approach and steps to address a technical problem by designing and building a solution for it.
* Ensuring data integrity and consistency post-migration. Using testing tools to run the test cases to check the consistency of data and monitor the performance.

**Walmart labs, Sunnyvale, California May’ 18 – Aug’2020 Sr. Data Engineer**

* Design and develop data pipeline to pull information from heterogeneous data sources and ingest into data lake to help with product analytics reporting using Hadoop, MPP database, Datawarehouse and related technologies.
* Worked on designing the data flow and building the data pile-line in Big Query/snowflake by fetching the data from various sources such as External APIs, Excel files and RDBMS systems.
* Involved in code migration to Google Cloud Platform, created and modified database objects such as Tables and views in Big Query, converted SQL queries from Teradata platform to Google Big Query, and involved in data validation between Teradata and Big Query
* Converted oracle PLSQL queries to Teradata views and Big Query views.
* Building the data pipeline for implementing the DataMart in Teradata Datawarehouse
* Implementing Slowly Changing Dimension types (type1, type2 and type3) through Structured Query Language to keep track of historical records.
* Worked on python script to fetch the data from the external API and build the data pipeline to load the data to Teradata and Hive
* Streamline the process to create and manage monthly and ad-hoc reports.
* Mapping the columns to personal identifiable information as part of CCPA (California Consumer Privacy Act)
* Data modelling of the database objects using Erwin Data Modeler
* Creating Hive tables, loading data and writing Hive queries which will run internally in MapReduce way.
* Populate the data from heterogeneous sources to HDFS through ETL tools by creating static and dynamic partition tables in hive.
* Writing shell scripts to call the API and streamline the data from API to hive tables.
* Using KAFKA to retrieve the streams of messages from Jason file as well as other sources and populate this data to HDFS.
* Performance tuning of complex and long running SQL queries by analyzing the explain plan.
* Implementing change data capture in Teradata as well as in Hadoop using SQL and HQL respectively
* Cleansing, formatting, and standardizing the data while loading the data, to data warehouse and DataMart.
* Automating the process of Extract, transformation and loading through ETL tool (Talend) by building the workflow and scheduling it through CA atomic platform
* Design and develop reports and dashboards using Business Intelligence tools.

**Johnson & Johnson, Bridgewater, New Jersey Oct’ 16 – Jan’ 18 Teradata Developer**

* Performed comprehensive unit testing on Teradata views as well as Qliksense visualizations.
* Worked on writing expressions using Set analysis in qliksense and unit testing of qliksense visualizations.
* Worked on migrating the code to higher environments through change management process.
* Involved in loading data into Teradata from legacy systems and flat files using Multiload scripts and Fast Load scripts.
* Apply technical knowledge using Teradata Tools (SQL Assistant, viewpoint,) and using Teradata Utilities (BTEQ, Fast Load, Mload,) for development and support of data warehouse.

**PayPal, San Jose, CA Jan’16 – Sep’16 Teradata Developer**

* Worked in Data Profiling based on the business requirements.
* Involved in the complete re-design of the process flow, based on the business logic change.
* Created and modified BTEQ, MLOAD, FEXP and FLOAD scripts as business requirements.
* Developed MLOAD scripts to load data from Hive to Teradata.
* Responsible for creating technical design documents to reflect the changed business logic.

**Apptad Solutions September ’13 – Dec’15**

**Teradata Developer**

* Involved in writing complex BTEQs to load the data from staging to target.
* Error handling and support of existing applications for failed reloads using Teradata MLOAD, BTEQ and Fast Load.
* Provided production support, debugged, and fixed production issues.
* Performance tuning of complex SQL queries.

**EDUCATION:**

Bachelor of Engineering in Computer Science from Dayanand Sagar College of Engineering, Bangalore, India

**BADGE:**