**Parikshit**

**Sr. Data Architect/Scientist**

**Summary:**

* Over **10+ years** of Experience **in Data Architecture, Design, Development and Testing** of business application systems, **Data** **Analysis** and developing **Conceptual, logical models** and **physical** database design for Online Transactional processing (**OLTP**) and Online Analytical Processing (**OLAP**) systems.
* Experienced working with data modeling tools like **Erwin**, **Power Designer** and **ER Studio**.
* Experienced in designing **star schema**, **Snowflake schema** for **Data** **Warehouse**, and **ODS architecture**.
* Experienced in **Data Modeling** &**Data Analysis** experience using **Dimensional Data Modeling** and **Relational Data Modeling, Star Schema/Snowflake Modeling, FACT & Dimensions tables, Physical & Logical Data Modeling**.
* Experienced **in big data analysis** and developing data models using **Hive, PIG, and Map reduce, SQL** with strong data architecting skills designing data-centric solutions.
* Experienced in **Data Profiling, Analysis** by following and applying appropriate database standards and processes, in definition and design of enterprise business data hierarchies.
* Hands on experience with big data tools like **Hadoop, Spark, Hive, Pig, Impala, Pyspark, SparkSql.**
* Very good knowledge and experience on **AWS, Redshift , S3 and EMR.**
* Proficient with **Data** **Analysis**, **mapping source** and target systems for **data** **migration** efforts and resolving issues relating to **data** **migration**.
* Excellent development experience **SQL, Procedural Language(PL)** of databases like **Oracle, Teradata , Netezza and DB2**
* Very good knowledge and working experience on big data tools like **Hadoop, Azure Data Lake, AWS Redshift**.
* Experienced in **Data Scrubbing/Cleansing, Data Quality, Data Mapping, Data Profiling, Data Validation** in **ETL**
* Experienced in creating and documenting **Metadata** for **OLTP** and **OLAP** when designing a systems.
* Expertise in managing entire data science project life cycle and actively involved in all the phases of project life cycle including **data acquisition, data cleaning, data engineering, features scaling, features engineering, statistical modeling (decision trees, regression models, neural networks, SVM, clustering), dimensionality reduction using Principal Component Analysis and Factor Analysis, testing and validation using ROC plot, K-fold cross validation and data visualization.**
* Excellent Knowledge of **Ralph Kimball** and **BillInmon's** approaches to **Data Warehousing**.
* Expertise in synthesizing **Machine learning, Predictive Analytics and Big data technologies** into integrated solutions.
* Extensive experienced in working with structured data using **Hive QL**, join operations, writing custom UDF's and experienced in optimizing **Hive Queries.**
* Extensive experience in development of **T**-**SQL**, **DTS**, **OLAP**, **PL**/**SQL**, Stored Procedures, Triggers, Functions, Packages, performance tuning and optimization for business logic implementation.
* Experience in using various packages in Rand **python** like **ggplot2, caret, dplyr, Rweka, gmodels, RCurl, tm, C50, twitteR, NLP, Reshape2, rjson, plyr, pandas, numpy, seaborn, scipy, matplotlib, scikit-learn, Beautiful Soup, Rpy2.**
* Experienced using query tools like **SQL Developer, PLSQL Developer, and Teradata SQL Assistant.**
* Excellent in performing data transfer activities between **SAS** and various databases and data file formats like **XLS,CSV,DBF,MDB etc.**
* Excellent knowledge of **Machine Learning, Mathematical Modeling and Operations Research**. Comfortable with **R, Python, SAS and Weka, MATLAB**, **Relational databases**. Deep understanding & exposure of **Big Data Eco-system.**
* Expertise in designing complex **Mappings** and have expertise in performance tuning and **slowly-changing Dimension Tables** and **Fact tables**
* Extensively worked with **Teradata** utilities **BTEQ, Fast export, and Multi** Load to export and load data to/from different source systems including flat files.
* Hands on experience in implementing LDA, Naive Bayes and skilled in **Random Forests, Decision Trees, Linear and Logistic Regression, SVM, Clustering, neural networks, Principle Component Analysis.**
* Expertise in extracting, transforming and loading data between homogeneous and heterogeneous systems like **SQL Server, Oracle, DB2, MS Access, Excel, Flat File and etc. using SSIS packages.**
* Experienced in Performance tuning on **Oracle** databases by leveraging explain plans, and tuning **SQL queries**.
* Proficient in **System Analysis**, **ER**/**Dimensional** **Data** **Modeling**, **Database design** and implementing **RDBMS** specific features.
* Experience in **UNIX shell scripting, Perl scripting** and automation of **ETL Processes.**
* Extensively used **ETL** to load data using **Power Center / Power Exchange** from source systems like **Flat Files** and **Excel Files** into staging tables and load the data into the target database **Oracle**. Analyzed the existing systems and made a Feasibility Study.
* Strong experience and knowledge in **Data Visualization** with **Tableau** creating: **Line and scatter plots, Bar Charts, Histograms, Pie chart, Dot charts, Box plots, Time series, Error Bars, Multiple Charts types, Multiple Axes, subplots etc.**
* Excellent understanding and working experience of industry standard methodologies like **System Development Life Cycle (SDLC), as per Rational Unified Process (RUP), AGILE Methodologies.**
* Experience in source systems analysis and data extraction from various sources like **Flat files, Oracle 12c/11g/10g/9i IBM DB2 UDB, XML files.**
* Proficiency in SQL across a number of dialects (we commonly write **MySQL, PostgreSQL, Redshift, SQL Server, and Oracle)**
* Experienced in developing **Entity-Relationship diagrams** and modeling **Transactional Databases** and **Data** **Warehouse** using tools like **ERWIN**, **ER/Studio** and Power Designer and experienced with modeling using **ERWIN** in both **forward** and **reverse** engineering cases.

**Technical Skills:**

* **Data Analytics Tools/Programming:** Python (numpy, scipy, pandas,Gensim, Keras), R ( Caret, Weka, ggplot), MATLAB, Microsoft SQL Server, Oracle PLSQL**,** Python**.**
* **Analysis and Modeling Tools**: Erwin, Sybase Power Designer, Oracle Designer, Erwin, Rational Rose, ER/Studio, TOAD, MS Visio, SAS
* **Data Visualization:** Tableau, Visualization packages, Microsoft Excel**.**
* **Big Data Tools:** Hadoop, MapReduce, SQOOP, Pig, Hive, NOSQL, Cassandra, MongoDB, Spark, Scala.
* **ETL Tools**: Informatica Power Center, Data Stage 7.5, Ab Initio, Talend
* **OLAP Tools**: MS SQL Analysis Manager, DB2 OLAP, Cognos Powerplay
* **Languages**: SQL, PL/SQL, T-SQL, XML, HTML, UNIX Shell Scripting, C, C++, AWK, JavaScript.
* **Databases**: Oracle12c/11g/10g/9i/8i/8.0/7.x,Teradata14.0,DB2 UDB 8.1, MS SQL Server 2008/2005, Netezaa 4.0 and Sybase ASE 12.5.3/15,Informix 9, AWS RDS.
* **Operating Systems**: Windows 2007/8, UNIX (Sun-Solaris, HP-UX), Windows NT/XP/Vista, MSDOS
* **Project Execution Methodologies**: Ralph Kimball and Bill Inmon data warehousing methodology, Rational Unified Process (RUP), Rapid Application Development (RAD), Joint Application Development (JAD)
* **Tools & Software**: TOAD, MS Office, BTEQ, Teradata SQL Assistant
* **Methodologies**: Ralph Kimball, COBOL
* **Reporting Tools:** Business ObjectsXIR2/6.5/5.0/5.1, Cognos Impromptu 7.0/6.0/5.0,Informatica Analytics Delivery Platform, Micro Strategy, SSRS, Tableau
* **Tools**: MS-Office suite (Word, Excel, MS Project and Outlook), VSS
* **Programming Languages**: SQL, T-SQL, Base SAS and SAS/SQL, HTML, XML

**IRI Chicago, IL Mar 17 to till date**

**Sr. Data Architect/Scientist**

**Responsibilities:**

* Design, Develop and implement Comprehensive **Data Warehouse Solution** to extract, clean, transfer, load and manage quality/accuracy of data from various sources to **EDW Enterprise Data Warehouse.**
* Architect framework for data warehouse solutions to bring data from source system to **EDW** and provide data mart solutions for Order/Sales operation, Salesforce activity, Inventory tracking, in depth data mining and analysis for market projection etc.
* Utilized **Apache** **Spark with Python** to develop and execute Big Data Analytics and Machine learning applications, executed machine Learning use cases under **Spark ML and Mllib.**
* Handled importing data from various data sources, performed transformations using **Hive, Map Reduce, and** loaded data into **HDFS.**
* Utilized **Spark, Scala, Hadoop, HBase, Cassandra, MongoDB, Kafka, Spark Streaming, MLLib, Python**, a broad variety of machine learning methods including classifications, regressions, dimensionally reduction etc. and Utilized the engine to increase user lifetime by 45% and triple user conversations for target categories.
* Designed and developed **NLP models** for sentiment analysis.
* Designed and provisioned the platform architecture to execute Hadoop and machine learning use cases under Cloud infrastruture, **AWS, EMR, S3.**
* Developed and configured on **Informatica MDM** hub supports the Master Data Management (**MDM**), Business Intelligence (**BI**) and **Data Warehousing** platforms to meet business needs.
* Transforming staging area data into a STAR schema (hosted on Amazon **Redshift**) which was then used for developing embedded Tableau dashboards
* Worked on machine learning on large size data using **Spark and MapReduce.**
* Let the implementation of new statistical algorithms and operators on **Hadoop** and **SQL platforms** and utilized optimizations techniques, **linear regressions, K-means clustering, Native Bayes and other approaches.**
* Developed **Spark/Scala, Python** for regular expression (**regex**) project in the **Hadoop/Hive** environment with Linux/Windows for big data resources.
* Proficiency in **SQL** across a number of dialects (we commonly write **MySQL, PostgreSQL, Redshift, Teradata, and Oracle**)
* Responsible for full data loads from production to **AWS** **Redshift** staging environment and Worked on migrating of **EDW** to **AWS** using **EMR** and various other technologies.
* Deep understand of **deep learning algorithms** and workflows, in particular working with large scale visual data and **agile** integration of **deep learning, data collection and failure analysis.**
* Worked on **TeradataSQL queries, Teradata Indexes, Utilities such as Mload, Tpump, Fast load and FastExport.**
* Application of various machine learning algorithms and statistical modeling like **decision trees, regression models, neural networks, SVM, clustering to identify Volume using scikit-learn package in python, Matlab.**
* Worked on data **pre-processing and cleaning** the data to perform feature engineering and performed data imputation techniques for the missing values in the dataset using **Python.**
* Developing new **algorithms** and prototyping them for **deep learning training** and inference and visualization and deeper understanding of **deep learning networks.**
* Created **Data Quality** Scripts using **SQL** and **Hive** to validate successful data load and quality of the data. Created various types of data visualizations using **Python and Tableau.**
* Worked with **BTEQ** to submit **SQL** statements, import and export data, and generate reports in **Teradata**.
* Build and maintain scalable data pipelines using the **Hadoop** ecosystem and other open source components like **Hive** and **HBase**.
* Build advanced **Deep Learning (DL)** based systems and drive their adoption into next-generation of MultiOmyx image analytics applications and Worked with **TensorFlow, Keras, CNTK, Caffe** or other **Deep Learning frameworks.**
* Created **Hive** architecture used for real time monitoring and **HBase** used for reporting and worked for map reduce and query optimization for **Hadoop hive and HBase architecture.**
* Involved in **Teradata** utilities (**BTEQ, Fast Load, Fast Export, Multiload, and Tpump**) in both **Windows and Mainframe platforms.**
* Built analytical data pipelines to port data in and out of **Hadoop/HDFS** from structured and unstructured sources and designed and implemented system architecture for **Amazon EC2** based cloud-hosted solution for client.

**Environment:** Erwin9.6.4, Oracle 12c, Python, Pyspark, Spark, Spark MLLib, Tableau, ODS, PL/SQL, OLAP, OLTP, AWS, Hadoop, MapReduce, HDFS, Python, MDM, Teradata 15, Hadoop, Spark, Cassandra, SAP, MS Excel, Flat files, Tableau, Informatica, SSIS, SSRS, AWS EC2, AWS EMR, Elastic Search.

**SEI Investments Malvern PA July 16 – Feb 17**

**Sr. Data Architect /Data Scientist**

**Responsibilities:**

* Provide **data architecture** support to enterprise data management efforts, such as the development of the enterprise data model and master and reference data, as well as support to projects, such as the development of **physical data models, data warehouses and data marts.**
* Create new data designs and make sure they fall within the realm of the overall **Enterprise BI Architecture** and Building relationships and trust with key stakeholders to support program delivery and adoption of **enterprise architecture.**
* Implemented **HQL Scripts** in creating Hive tables, loading, analyzing, merging, binning, backfilling, cleansing using **hive.**
* Used **R, SQL** to create **Statistical algorithms** involving **Multivariate Regression, Linear Regression, Logistic Regression, PCA, Random forest models, Decision trees, Support Vector Machine** for estimating the risks.
* Used **R** for Exploratory Data Analysis, A/B testing, Anova test and Hypothesis test to compare and identify the effectiveness of Creative Campaigns.
* Providing technical leadership, mentoring throughout the project life-cycle, developing vision, strategy, **architecture** and overall design for assigned domain and for solutions and responsible for the development of target data architecture, design principles, quality control, and data standards for the organization
* Developed and maintains **data models and data dictionaries**, data maps and other artifacts across the organization, including the **conceptual and physical models**, as well as metadata repository
* Performed **extensive Data Validation, Data Verification against Data Warehouse** and performed debugging of the **SQL-Statements** and **stored procedures** for **business scenarios.**
* Used **Spark Data frames, Spark-SQL, Spark MLLib** extensively and developing and designing **POC's** using **Scala, Spark SQL and MLlib libraries.**
* Working on a **Map RHadoop** platform to implement **Bigdata solutions** using **Hive, Map reduce**, **shell scripting and Pig.**
* Worked with cloud based technology like **Redshift, S3, AWS, EC2 Machine**, etc. and extracting the data from the **Oracle** financials and the **Redshift** **database**.
* Used **Teradata OLAP** functions like **RANK, ROW\_NUMBER, QUALIFY, CSUM and SAMPLE.**
* Involved in designing and developing **Data Models** and **Data Marts** that support the **Business Intelligence Data Warehouse.**
* Extensively used **Aginity Netezza** work bench to perform various **DML, DDL**etc operations on **Netezza database**.
* Developed multiple **MapReduce** jobs in java for **Data Cleaning** and **pre-processing analyzing** data in **PIG**.
* Worked on predictive and what-if analysis using **R from HDFS** and successfully loaded files to **HDFS** from **Teradata**, and loaded from **HDFS to HIVE.**
* Designed the schema, configured and deployed **AWS Redshift** for optimal storage and fast retrieval of data.
* **NLTK, Stanford NLP, RAKE** to preprocess the data, entity extraction and keyword extraction.
* Transforming staging area data into a **STAR schema (hosted on Amazon RedShift)** which was then used for developing embedded **Tableau dashboards**
* Developed **SQL scripts** for loading data from staging area to **Target tables** and worked on **SQL** and **SAS script mapping**.
* Performed transformations of data using **Spark and Hive** according to business requirements for generating various analytical datasets.
* Performed **Multinomial Logistic Regression, Random forest, Decision Tree, SVM** to classify package is going to deliver on time for the new route and Performed **data analysis** by using **Hive** to retrieve the data from **Hadoop cluster, Sql** to retrieve data from **Oracle database.**
* Proposed the **EDW data design** to centralize the data scattered across multiple datasets and Worked on migrating of **EDW** to **AWS** using **EMR** and various other technologies.
* Worked on the development of **Data Warehouse, Business Intelligence architecture** that involves data **integration** and the conversion of data from multiple sources and platforms.
* Let the implementation of new statistical algorithms and operators on **Hadoop** and **SQL platforms** and utilized optimizations techniques, **linear regressions, K-means clustering, Native Bayes and other approaches.**
* Created **mapreduce** running over **HDFS** for data mining and analysis using **R** and Loading & Storage data to **Pig Script and R for MapReduce** operations.
* Worked on **Teradata SQL queries**, **Teradata** Indexes, Utilities such as **Mload, Tpump, Fast load and Fast Export.**
* Used **Metadata** tool for importing metadata from repository, new job categories and creating new data elements and proficiency in **SQL** across a number of dialects (we commonly write **MySQL, PostgreSQL, Redshift, SQL Server, and Oracle)**

**Environment**: Oracle 12c, SQL Plus, Erwin 9.6, MS Visio, SAS, Source Offsite (SOS), Hive, PIG, Windows XP, AWS, QC Explorer, Share point workspace, Teradata, Oracle, Agile, PostgreSQL, Data Stage, MDM ,Netezza, IBM Infosphere, SQL, PL/SQL, IBM DB2, SSIS, Power BI, AWS Redshift, Business Objects XI3.5,COBOL,SSRS, QuickData, Hadoop, MongoDB, HBase, Hive, Cassandra, JavaScript.

**TSYS, Atlanta GA Jan15 – June16**

**Sr. Data Architect/Data Modeler**

**Responsibilities:**

* **Architect** and design, solutions for complex business requirements, including data processing, analytics and ETL and reporting processes to improve performance of data loads and processes.
* Develop a high performance, scalable data architecture solution that incorporates a matrix of technology to relate **architectural decision** to business needs.
* Participated in the design, development, and support of the corporate operation data store and enterprise data warehouse database environment.
* Conducting strategy and **architecture** sessions and deliver artifacts such as **MDM strategy (Current state, Interim State and Target state) and MDM Architecture (Conceptual, Logical and Physical) at detail level.**
* Owned and managed all changes to the **data models, Created data models, solution designs and data architecture** documentation for complex information systems.
* Design and development of dimensional data model on **Redshift** to provide advanced selection analytics platform and developed Simple to complex **Map Reduce Jobs using Hive and Pig**.
* Analyze change requests for mapping of multiple source systems for understanding of **Enterprise** wide **information architecture** to devise **Technical Solutions.**
* Worked on **AWS Redshift** and **RDS** for implementing models and data on **RDS** and **Redshift.**
* Worked with SME's and other stakeholders to determine the requirements to identify **Entities** and **Attributes** to build **Conceptual**, **Logical** and **Physical** **Data** **Models**.
* Provided data sourcing methodology, resource management and performance monitoring for data acquisition.
* Designed and implemented Near Real Time **ETL** and Analytics using **Redshift database**
* Supported and followed information governance and data standardization procedures established by the organization. Documents reports library as well as external data imports and exports.
* Prepared **Tableau reports** and **dashboards** with calculated fields, parameters, sets, groups or bins and publish on the server.
* Developed mappings to load Fact and Dimension tables, **SCD Type 1 and SCD Type 2** dimensions and Incremental loading and unit tested the mappings.
* Extensively used **Netezza** utilities like **NZLOAD and NZSQL** and loaded data directly from **Oracle** to **Netezza** without any intermediate files.
* Performed analysis of data sources and processes to ensure data integrity, completeness and accuracy.
* Created a **logical** design and **physical** design in **Erwin.**
* Enforced referential integrity in the **OLTP** **data** model for consistent relationship between tables and efficient database design.
* Implemented **Hive** Generic UDF's to incorporate business logic into **Hive Queries** and Creating **Hive** tables and working on them using **Hive QL.**
* Developed **DataMapping**, **DataGovernance**, and transformation and cleansing rules for the **Master Data Management** Architecture involving **OLTP**, **ODS** and generated ad-hoc reports using **OBIEE**.
* Responsible for migrating the data and data models from **SQL server** environment to **Oracle** environment.
* Analysis and designing the ETL architecture, creating templates, training, consulting, development, deployment, maintenance and support.
* Created **SSIS** Packages which loads the data from the **CMS** to the EMS library database and Involved in **data modeling** and providing technical solutions related to **Teradata** to the team.
* Build a real time event analytic systems using dynamic **Amazon redshift schema**.
* Designed the **physical model** for implementing the model into **Oracle** 11g **physical** data base and Developed **SQL Queries** to get complex data from different tables in Hemisphere using joins, database links.
* Wrote **SQL queries, PL/SQL** procedures/packages, triggers and cursors to extract and process data from various source tables of database.
* Created **Hive** Tables, loaded transactional data from **Teradata using Sqoop** and created and worked Sqoop jobs with incremental load to populate **Hive** External tables.
* Developed **LINUX Shell scripts** by using **NZSQL/NZLOAD** utilities to load data from flat files **to Netezza database**.
* Worked with cloud based technology like **Redshift, S3, AWS, EC2 Machine**, etc. and extracting the data from the **Oracle** financials and the **Redshift database.**
* Used **Erwin** to create **logical** and **physical** **data** models for enterprise wide **OLAP** system and Involved in **mapping** the data elements from the User Interface to the Database and help identify the gaps.
* Designing and customizing data models for **Data warehouse** supporting data from multiple sources on real time. Requirements elicitation and **Data analysis**. Implementation of **ETL Best Practices**.
* Generated comprehensive **analytical reports** by running **SQLqueries** against current databases to conduct **data analysis**.
* Created data models for **AWS Redshift** and **Hive** from dimensional data models.
* Developed **complex SQL scripts** for **Teradata** database for creating BI layer on **DW** for **Tableau reporting.**
* Extensively used **ETL** methodology for supporting data extraction, transformations and loading processing, in a complex **EDW** using **Informatica.**
* Created Active Batch jobs to load data from distribution servers to **PostgreSQL** DB using \*.bat files and worked on CDC schema to keep track of all transactions.

**Environment**: Erwin 9.5, MS Visio, Oracle 11g, Oracle Designer, MDM , Power BI, SAS, SSIS, Tableau, Tivoli Job Scheduler, SQL Server 2012, DATAFLUX 6.1, JavaScript, AWS Redshift, PL/SQL, SQL/PL SQl, SSRS, PostgreSQL, Data Stage, SQL Navigator Crystal Reports 9, Hive, Netezza, Teradata, T-SQL, Informatica.

[**Molina Healthcare**](https://www.indeed.com/cmp/Molina-Healthcare) **Long Beach CA Mar 13 – Dec 14**

**Sr. Data Architect/Data Analyst/Data Modeler**

**Responsibilities:**

* Design and develop **data warehouse architecture**, **data modeling/conversion solutions**, and **ETL mapping solutions** within structured data warehouse environments
* Reconcile data and ensure data integrity and consistency across various organizational operating platforms for business impact.
* Define best practices for **data loading** and **extraction** and ensure **architectural alignment** of the designs and development.
* Used **Erwin** for effective model management of sharing, dividing and reusing model information and design for productivity improvement.
* Involved in preparing **Logical Data Models/Physical Data** **Models**.
* Worked extensively in both **Forward** Engineering as well as **Reverse** Engineering using **data modeling** tools.
* Involved in the creation, maintenance of **Data** **Warehouse** and repositories containing **Metadata**.
* Involved using ETL tool **Informatica** to populate the database, data transformation from the old database to the new database using **Oracle and SQL Server.**
* Identifying inconsistencies or issues from incoming **HL7** messages, documenting the inconsistencies, and working with clients to resolve the **data inconsistencies**
* Resolved the data type inconsistencies between the source systems and the target system using the **Mapping Documents** and **analyzing** the **database** using **SQL queries**.
* Extensively used both **Star Schema** and **Snow flake** schema methodologies in building and designing the **logical** data model in both **Type1** and **Type2Dimensional Models**.
* Worked with **DBA** group to create **Best**-**Fit Physical** **Data** **Model** from the Logical **Data** **Model** using **Forward Engineering**.
* Worked with Data Steward Team for designing, documenting and configuring **Informatica DataDirector** for supporting management of MDM data.
* Conducting **HL7** integration testing with clients systems that is testing of business scenarios to ensure that information is able to flow correctly between applications.
* Extensively worked with MySQL and **Redshift** performance tuning and reduced the ETL job load time by 31% and DW space usage by 50%
* Developed **Data Migration** and **Cleansing** rules for the **Integration Architecture (OLTP, ODS, DW)**.
* Used **Teradata SQL** Assistant, **Teradata** Administrator, **PMON** and data load/export utilities like **BTEQ, Fast Load, Multi Load, Fast Export, Tpump** on **UNIX/Windows** environments and running the batch process for **Teradata**.
* Created Dashboards on **Tableau** from different sources using data blending from **Oracle, SQL Server, MS Access and CSV at single instance.**
* Used the **Agile Scrum methodology** to build the different phases of Software development life cycle.
* Documented **logical, physical, relational** and **dimensional** **data** **models**. Designed the **data** **marts** in dimensional data modeling using star and snowflake schemas.
* Created dimensional model based on **star schemas** and designed them using **ERwin**.
* Performed match/merge and ran match rules to check the effectiveness of **MDM** process on data.
* Carrying out **HL7** interface unit testing aiming to confirm that **HL7** messages sent or received from each application conform to the **HL7** interface specification.
* **Used tools such as SAS/Access and SAS/SQL to create and extract oracle tables.**
* **Data modeling and design** of data warehouse and data marts in **star schema methodology** with confirmed and granular dimensions and **FACT tables**.
* Developed **SQL Queries** to fetch complex data from different tables in remote databases using joins, database links and Bulk collects.
* Enabled the **SSIS** package configuration to make the flexibility to pass the connection strings to connection managers and values to package variables explicitly based on environments.
* Responsible for Implementation of **HL7** to build **Orders, Results, ADT, DFT** interfaces for client hospitals
* Connected to **Amazon RedShift** through **Tableau** to extract live data for real time analysis.
* Developed **SQL Queries** to fetch complex data from different tables in remote databases using joins, database links and Bulk collects.
* Developed Slowly Changing Dimensions Mapping for Type **1 SCD and Type 2 SCD** and Used **OBIEE** to create reports.
* Worked on **data modeling** and produced **data mapping** and data definition specification documentation.

**Environment:** Erwin, Oracle, SQL server 2008, Power BI, MS Excel, Netezza, Agile, MS Visio, Rational Rose, Requisite Pro, SAS, SSIS, SSRS, Windows 7, PL/SQL, , SQl Server, MDM, Teradata, MS Office, MS Access, SQL, SSIS, MS Visio, Tableau, Informatica, Amazon Redshift.

**Capital Bank, FL Jan 12 – Feb 13**

**Sr. Data Modeler/Data Analyst**

**Responsibilities:**

* Designed **logical** and **physical data models** for multiple **OLTP** and Analytic applications.
* Involved in analysis of business requirements and keeping track of data available from various data sources, transform and load the data into Target Tables using **Informatica Power Center.**
* Extensively used the **Erwin design tool** &**Erwin model manager** to create and maintain the **Data Mart.**
* Extensively used **Star Schema** methodologies in building and designing the **logical** **data** **model** into Dimensional Models
* Created stored procedures using **PL**/**SQL** and tuned the databases and backend process.
* Involved with **Data Analysis** primarily **Identifying Data Sets, Source Data, Source Meta Data, Data Definitions and Data Formats**
* Performance tuning of the database, which includes indexes, and optimizing **SQL** statements, monitoring the server.
* Developed **Informatica** mappings, sessions, workflows and have written Pl SQL codes for effective and optimized data flow coding.
* Wrote **SQL Queries**, Dynamic-queries, sub-queries and complex joins for generating Complex Stored Procedures, Triggers, User-defined Functions, Views and Cursors.
* Created new **HL7** interface based on the requirement using **XML, XSLT technology.**
* Experienced in creating **UNIX scripts** for file transfer and file manipulation and utilized **SDLC** and **Agile methodologies** such as **SCRUM**.
* **DataStage** jobs were scheduled, monitored, performance of individual stages was analyzed and multiple instances of a job were run using **DataStage Director.**
* Led successful integration of **HL7** Lab Interfaces and used expertise of **SQL** to integrate **HL7** Interfaces and carried out detailed and various test cases on newly built **HL7** interface.
* Wrote simple and advanced **SQL queries** and **scripts** to create standard and **ad hoc reports** for senior managers.
* Involved in collaborating with **ETL/Informatica** teams to source data, perform data analysis to identify gaps
* Used Expert level understanding of different databases in combinations for **Data extraction** and **loading**, joining data extracted from different databases and loading to a specific database.
* Designed and Developed **PL/SQL** procedures, functions and packages to create Summary tables.

**Environment**: SQL Server, UML, Business Objects 5, Teradata, Windows XP, SSIS, SSRS, Embarcadero, ER studio, Erwin, DB2, Informatica, HL7, Oracle, Query Management Facility (QMF), SSRS, Data Stage, Clear Case forms, SAS, Agile, Unix and Shell Scripting.

**GrowelSoftech, India Aug 07 – Dec 11**

**Data Analyst/Data Modeler**

**Responsibilities:**

* Developed **Data Mapping**, **Data Governance** and transformation and cleansing rules for the **Master Data Management** Architecture involving **OLTP**, **ODS**.
* Created new **conceptual, logical and physical data** models using **ERWin** and reviewed these models with application team and modeling team.
* Performed numerous data pulling requests using **SQL** for analysis and created databases for **OLAP Metadata** catalog tables using forward engineering of models in **Erwin**.
* Enforced referential integrity in the **OLTP data model** for consistent relationship between tables and efficient database design.
* Proficient in importing/exporting large amounts of data from files to **Teradata** and vice versa.
* Developed **Data Mapping, Data Governance,** and **Transformation** and cleansing rules for the Master Data Management Architecture involving **OLTP, ODS.**
* Identified and tracked the **slowly changing dimensions, heterogeneous sources** and determined the hierarchies in dimensions.
* Utilized **ODBC** for connectivity to **Teradata** & **MSExcel** for automating reports and graphical representation of data to the Business and Operational Analysts.
* Extracted data from existing data source, Developing and executing departmental reports for performance and response purposes by using oracle **SQL**, **MS Excel**.
* Extracted data from existing data source and performed ad-hoc queries and used **BETQ** to run and **Teradata SQL scripts** to create physical data model.

**Environment**: UNIX scripting, Oracle SQL Developer, SSRS, SSIS, Teradata, Windows XP, SAS data sets.