**Renuka D**

**Mobile: (732) 930 – 1801**

**Email:** [**renukadevi0588@gmail.com**](mailto:renukadevi0588@gmail.com)

**Professional Summary:**

* 8+ years of experience in Information Technology as an Sr ETL IICS Developer with a strong background in ETL Data warehousing experience using Informatica Power Center 10.4/10.2/ 9.5x/9.x/8.x, Snowflake, SnowSQL, Informatica Data Quality (IDQ) 10.4/10.2/9.6, Intelligent Cloud Services (IICS), Informatica Meta Data Manager, Azure ML, Power BI, Tableau, Oracle PL/SQL, Bash, SQL Server Management Studio (SSMS), Unix.
* 3+ years of experience using Talend Data Integration/Big Data Integration (6.1/5.x) / Talend Data Quality.

# Good experience in Informatica Installation, Migration, and Upgrade Process.

* Skilled in cloud platforms including Azure and AWS, with experience in DevOps practices and Git repositories.

# Good experience in Informatica Installation, Migration, and Upgrade Process.

# Extensive experience in ETL methodology for performing Data Profiling, Data Migration, Extraction, Transformation and Loading using Talend and designed data conversions from wide variety of source systems including Netezza, Oracle, DB2, SQL server, Teradata, Hive, Hana and non - relational sources like flat files, XML and Mainframe files.

# Expertise in creating mappings in TALEND using tMap, tJoin, tReplicate, tParallelize, tConvertType,, tflowtoIterate, tAggregate, tSortRow, tFlowMeter, tLogCatcher, tRowGenerator, tNormalize, tDenormalize, tSetGlobalVar, tHashInput, tHashOutput, tJava, tJavarow, tAggregateRow, tWarn, tLogCatcher, tMysqlScd, tFilter, tGlobalmap, tDie etc.

# Experience in using Informatica Power Center Transformations such as Source Analyzer, Transformation Developer, Mapplet Designer, Mapping Designer, Workflow Manager, Workflow Monitor, and Repository Manager.

# Experience with Snowflake cloud data warehouse and AWS S3 bucket for integrating data from multiple source system which include loading nested JSON formatted data into snowflake table.

# Written Spark applications using Pyspark for real-time data analysis by connecting to the multiple data warehouse like Hive and HBase.

# Information Technology including analysis Design, Development, Implementation of Data Warehousing in ETL Worked extensively in developing ETL for supporting Data Extraction, SSIS, Transformation and Loading using Oracle Data Integrator (ODI) and Data stage.

# Involved in converting Hive/SQL queries into Spark transformations using Spark RDDs, PySpark and Scala.

# Analyzed the SQL scripts and designed the solution to implement using PySpark.

# Created Talend ETL jobs to receive attachment files from pop e-mail using tPop, tFileList and tFileInputMail and then loaded data from attachments into database and achieved the files.

# Hands - on experience in Azure Analytics Services - Azure Data Lake Store (ADLS), Azure Data Lake Analytics (ADLA), Azure SQL DW, Azure Data Factory (ADF), Azure Data Bricks (ADB) etc.

* Created IICS connections using various cloud connectors in IICS administrator.
* Installed and configured Windows Secure Agent register with IICS org.

# Experience in building BI DW’s/performing Data Integration from source systems to EDW, ODS, BI Reporting solution systems using Oracle Data Integrator v.11g/10g (v 11.1.1.7/11.1.1.6 v/11.1.1.3/v.10.1.3.5).

# Expertise in Oracle Data Integrator and ETL loads.

# Experience of having worked with ELT Tool ODI against different databases Oracle (9i, 10g), DB2, MS SQL Server (2005, 2008) and MS Access.

# Well versed with Talend Big Data, Hadoop, Hive and used Talend Big data components like tHDFSInput, tHDFSOutput, tPigLoad, tPigFilterRow, tPigFilterColumn, tPigStoreResult, tHiveLoad, tHiveInput, tHbaseInput, tHbaseOutput, tSqoopImport and tSqoopExport

# Good experience with Snowflake Cloud Data warehouse, AWS S3.

# Good experience with Snowflake utility SnowSQL

# Worked extensively with complex interfaces using ODI.

# Wrote scripts in Python for Extracting Data from JSON and XML files.

# Developed the back-end web services for the worker using Python Flask REST APIs

* Experience in the development of mappings in IDQ to load the cleansed data into the target table using various IDQ transformations.
* Actively involved in migrating the data warehouse to Snowflake and re-platforming the ETL to Informatica Intelligent Cloud Services (IICS).
* Designed, Developed and Implemented ETL processes using IICS Data integration.
* Extensively used performance tuning techniques while loading data into Azure Synapse using IICS.
* Successfully migrated multiple terabytes of data to Snowflake, achieving improvements in data processing times by over 40%.
* Developed and refined ETL frameworks that increased data accuracy and reduced processing time, supporting critical business decisions

# In addition has knowledge of ODI 12 C and other tools like OBIEE, Informatica.

# Extensive development experience in Extraction, Transformation and Loading of data from multiple sources such as Oracle, MS SQL Server,DB2, Legacy systems, Flat Files and XML Files into ODS (Operational Data Store) and EDW (Enterprise Data Warehouse) systems.

* Experience with Snowflake cloud data warehouse for integrating data from multiple source systems which include loading nested JSON formatted data into Snowflake table.
* Experience in data profiling and analyzing the scorecards to design the data model.
* Proﬁcient knowledge and hands-on experience in building Data Warehouses, Data Marts, Data Integration, Operational Data Stores and ETL processes.
* Good Exposure to Teradata DBA utilities Teradata Manager, Workload Manager, Index Wizard, Stats Wizard, and Visual Explain.
* Proficient in building and optimizing data solutions using Snowflake and Informatica Power Center, with extensive experience in designing, developing, and managing robust ETL pipelines. Skilled in data warehousing and database tuning to enhance data retrieval and processing speeds.
* Experienced in Agile methodologies, utilizing Jira for project management, and proficient with GitLab CI/CD pipelines for efficient and error-free code deployment.
* Designed and developed Power BI graphical and visualization solutions with business requirement documents and plans for creating interactive dashboards.
* Utilized Power Query in Power BI to Pivot and Un-pivot the data model for data cleansing and data massaging.
* Prepared the complete data mapping for all the migrated jobs using SSIS.

# Good knowledge of Python and Pyspark, and Dimensional Data Modeling, ER Modeling, Star Schema/Snowflake Schema, FACT and Dimensions Tables, and Physical and Logical Data Modeling.

# Strong experience in analyzing large amounts of data sets writing PySpark scripts and Hive queries.

* Created ODI Packages, Jobs of various complexities and automated process data flow.
* Used ODI Operator for debugging and viewing the execution details of interfaces and packages.
* Experienced in Installing, Managing, and configuring Informatica MDM core components such as Hub Server, Hub Store, Hub Cleanse, Hub Console, Cleanse Adapters, and Hub Resource Kit.
* Database experience using Oracle 19c/12c/11g/10g/9i, Teradata, MS SQL Server 2008/2005/2000 and MS Access.
* Experience in UNIX Operating System and Shell scripting.
* Experience in implementing Azure data solutions, provisioning storage accounts, Azure data factory, and Azure data Bricks.
* Highly proﬁcient in using SQL for developing complex Stored Procedures, Triggers, Tables, Views, User deﬁned Functions, User processes, Relational Database Models and Data Integrity, SQL joins, Functions like Rank, Row Numbers, Dense\_ Rank, etc., indexing and Query Writing
* Extensive experience using database tools such as SQL \*Plus, SQL \*Developer, Autosys, and TOAD.
* Effective working relationships with the client team to understand support requirements, and effectively manage client expectations.
* Strong understanding of the principles of Data Warehousing concepts using Fact tables, Dimension tables and Star / Snowflake Schema modeling.
* Performed loads into Snowflake instance using Snowflake connector in IICS for a separate project to support data analytics and insight use case for Sales team
* Knowledge in Reporting Services (SSRS), Integration Services (SSIS), and Analysis Services (SSAS).
* Worked with teams and helped them in code build for chat board formation within the internal servers using Python and Java.
* Worked on creating a framework using Java API for implementing reusable components.

Technical Skills :

|  |  |
| --- | --- |
| Data Warehousing/ETL Tools | (Source Analyzer, Data Warehousing Designer, Mapping Designer, Mapplet, Transformation, Sessions, SSIS, Informatica MDM, kafka, Workﬂow Manager-Workﬂow, Task, IICS, Commands, Work let, IDQ, Transactional Control, Constraint-Based Loading, SCD I, II, Data Flux, Data mart, OLAP, ROLAP, MOLAP, OLTP. Informatica Intelligent Cloud Services (IICS) Snowflake, Redshift, Teradata |
| ETL Tools | ODI 11g/10g, Data stage 8.1, Informatica 8.6.1 |
| Cloud | Snowflake, Azure, SnowSQL, Amazon AWS services such as Redshift, RDS, S3, EC2 etc. |
| Data Modeling | Physical Modeling, Logical Modeling, Relational Modeling, Dimensional Modeling (Star Schema, Snow-Flake, Fact, Dimensions), Entities, Attributes, Cardinality, ER Diagrams. |
| Databases | Oracle 19c,12c,11g/10g/9i/8i, Teradata, MS SQL Server 2008/2005/2000, MS Access, Denodo 6 and 7and DB2 |
| Programing Languages | SQL, PLSQL C, C++, Data Structures, T-SQL, Unix Shell Script, Visual Basic, Java and Python/Pyspark |
| Web Technologies | XML, HTML, JavaScript |
| Tools | Toad, SQL\* Developer, Autosys, Erwin |

**Domain Knowledge:**

* Healthcare, Insurance and Telecommunications, Banking/Financials.

**Professional Experience :**

**Client: Santander Bank – Boston, MA OCT’2022 to Till Date**

**Role: Sr ETL IICS Informatica Developer**

**Responsibilities:**

* Modified existing developed new complex Informatica Power Center Mappings to extract and pull the data according to the guidelines provided by the business users and populate the data into Target Systems.
* Responsible for pulling the data from XML files, Flatfiles Fixed width Delimited, and COBOL files by using complex transformations like Normalizer, XMLSourceQualifieretc. Created Talend jobs to copy the files from one server to another and utilized Talend FTP components.
* Created and managed Source to Target mapping documents for all Facts and Dimension tables
* Used ETL methodologies and best practices to create Talend ETL jobs. Followed and enhanced programming and naming standards.
* Provide guidance to development team working on PySpark as ETL platform.
* Makes sure that quality standards are defined and met.
* Optimize the Pyspark jobs to run on Kubernetes Cluster for faster data processing.
* Provide workload estimates to client
* Developed framework for Behaviour Driven Development (BDD).
* Created tables, views, secure views, user defined functions in Snowflake Cloud Data Warehouse.
* Extracted and loaded CSV files, json files data from AWS S3 to Snowflake Cloud Data Warehouse.
* Migrated Oracle database tables data into Snowflake Cloud Data Warehouse.
* Designed and created optimal pipeline architecture on Azure platform.
* Writing Python scripts to parse XML documents as well as JSON based REST Web services and load the data in database.
* Implemented several DAX functions for various fact calculations for efficient data visualization in Power BI utilized Power BI gateway.
* 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.
* Writing ORM’s for generating the complex SQL queries and building reusable code and libraries in Python for future use.
* Created pipelines in Azure using ADF to get the data from different source systems and transform thedata by using many activities.
* Used Talend most used components (tMap, tDie, tConvertType, tFlowMeter, tLogCatcher, tRowGenerator, tSetGlobalVar, tHashInput & tHashOutput and many more).
* Working closely with software developers and debug software and system problems.
* Experience with Snowflake cloud data warehouse and AWS S3 bucket for integrating data from multiple source system which include loading nested JSON formatted data into snowflake table. Professional knowledge of AWS Redshift.
* Develop framework for converting existing PowerCenter mappings and to PySpark(Python and Spark) Jobs.
* Create Pyspark frame to bring data from DB2 to Amazon S3.
* Profiling Python code for optimization and memory management and implementing multithreading functionality
* Interact with BusinessAnalyst to assist him in understanding the Source and Target System.
* Responsible for pulling the data from XML files, Flatfiles Fixed width Delimited, and COBOL files by using complex transformations like Normalizer, XMLSourceQualifieretc.
* Migrating the data warehouse to Snowflake and re-platforming the ETL to Informatica Intelligent Cloud Services (IICS).
* Utilized Power BI (Power View) to create various analytical dashboards that depicts critical KPIs such as legal case matter, billing hours and case proceedings along with slicers and dicers enabling end-user to make filters.
* Hands-on experience with Snowflake utilities, SnowSQL, SnowPipe, Big Data model techniques using Python / Java.
* Integrated Informatica Intelligent Cloud Services (IICS) with Kafka for real-time data ingestion and processing, enhancing agility and responsiveness.
* Migrating data for reporting and reference data availability using tools and technologies including Tidal, Oracle, VB, Teradata, UNIX, and SSIS.
* Good experience with Snowflake Cloud Data warehouse, AWS S3 and Snowflake utility SnowSQL.
* Configured Kafka connectors within IICS to streamline data movement between Kafka topics and target systems, enabling efficient streaming and analytics.
* Used Azure Devops & Jenkins pipelines to build and deploy different resources(Code and Infrastructure)in Azure.
* Development, Implementation of Data Warehousing in ETL Worked extensively in developing ETL for supporting Data Extraction, Transformation and Loading using Oracle Data Integrator (ODI) and Data stage.
* Experience in building BI DW’s/performing Data Integration from source systems to EDW, ODS, BI Reporting solution systems using Oracle Data Integrator v.11g/10g (v 11.1.1.7/11.1.1.6 v/11.1.1.3/v.10.1.3.5).
* Expertise in Oracle Data Integrator and ETL loads.
* In-depth knowledge of Snowflake Database, Schema and Table structures.
* Define virtual warehouse sizing for Snowflake for different type of workloads.
* Experience of having worked with ELT Tool ODI against different databases Oracle (9i, 10g), DB2, MS SQL Server (2005, 2008) and MS Access.
* Experience managing Azure Data Lakes (ADLS) and Data Lake Analytics and an understanding of howto integrate with other Azure Services.
* Designed and developed a new solution to process the NRT data by using Azure stream analytics,Azure Event Hub and Service Bus Queue.
* Hortonworks Hadoop cluster to run multiple ETL jobs developed in python, Pig and spark in orderly manner.
* Hands on experience on dealing with ODI Knowledge Modules like, LKM, IKM and JKM, CKM.
* Experience in debugging Interfaces. Identified bugs in existing Interfaces.
* Migration of on premise data (Oracle/ Teradata) to Azure Data Lake Store(ADLS) using Azure DataFactory(ADF V1/V2).
* Good knowledge of Data Marts, Data warehousing, Operational Data Store (ODS), OLAP, Data Modeling like Dimensional Data Modeling, Star Schema Modeling, Snow-Flake Modeling, FACT and Dimensions Tables using Analysis Services.
* Developed Spark code using scala and PySpark -SQL for batch processing of data..Utilized in-memory processing capability of Apache Spark to process data using Spark SQL, PySpark Streaming using PySpark and Scala scripts
* Designed and implemented high-performance data processing pipelines in Snowflake, optimizing ETL processes and significantly reducing data load times.
* Utilized Snowflake's advanced features for managing large-scale data layers, ensuring efficient data storage and retrieval.
* Wrote scripts in Python for Extracting Data from JSON and XML files.
* Developed the back-end web services for the worker using Python Flask REST APIs
* Managed cloud-based data warehousing projects, leveraging AWS and Snowflake to handle complex data sets and enable scalable analytics solutions.
* Performed advanced SQL query tuning and schema refinement to enhance system performance and support complex data analysis requirements.
* Expertise’s in developing Slowly Changing Dimensions mappings using Change Data capture (CDC) Type1, Type2 and Type3 as a part of Performance Tuning.
* Experience in analyzing, requirement gathering, documenting and editing Business/User Requirements.
* Created several scripts in UNIX for data transformations to load the base tables by applying the business rules based on the given data requirements.
* Strong in analysis, design, coding and a good Team player in full life cycle / SDLC support from Initial design phase to production support.
* Experience working with IICS transformations like Expression, joiner, union, lookup, sorter, filter, normalizer, and various concepts like macro fields to templatize column logic, smart match fields, renaming bulk fields, and more.
* Data movement Control System was created between salesforce and Oracle tables are correctly getting loaded as a data checkpoint through IICS.
* Generated ad-hoc reports in Excel Power Pivot and sheared them using Power BI to the decision makers for strategic planning.
* Worked on Power Center Designer tools like SourceAnalyzer, TargetDesigner, MappingDesigner, MappletDesigner, and TransformationDeveloper.
* Created Linked service to land the data from different sources to Azure Data Factory.
* Worked on SQL tools like TOAD and SQLDeveloperto to run SQL queries and validate the data.
* Scheduled Informatica Jobs through the Autosysscheduling tool.
* Assisted the QA team in fixing and finding solutions for production issues.
* Prepared all documents necessary for knowledge transfer such as ETL strategy, ETL development standards, ETL processes, etc.

**Environment:** IICS, Informatica Power Center 9.1, PowerExchange 9.1, Oracle 11g, Teradata, Erwin, UNIX, PL/SQL, Autosys, MS-SQL Server 2008, Python, Talend Data Integration 6.1/5.5.1, Talend Enterprise Big Data Edition 5.5.1, Toad, MS-Visio,Snowflake, Snowflake, Redshift, SQL server, SSIS, SSRS, SSAS, Power BI, AWS, AZURE, TALEND, JENKINS and SQL UNIX scripting.

**Client: Spectrum – Stamford, CT AUG‘2020 to SEP’2022**

**Role: ETL Informatica IICS Developer**

**Responsibilities**

* Performed the roles of ETL Informatica developer and Data Quality (IDQ) developer on a data warehouse initiative and was responsible for requirements gathering, preparing mapping documents, architecting end-to-end ETL flow, building complex.
* 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.
* Worked on reading and writing multiple data formats like JSON,ORC,Parquet on HDFS using PySpark.
* Creating Reports in Looker based on Snowflake Connections.
* ETL procedures, developing strategy to move existing data feeds into the Data Warehouse (DW), perform data cleansing activities using various IDQ transformations.
* Developed jobs in Talend Enterprise edition from stage to source, intermediate, conversion and target.
* Implemented Azure Data Factory (ADF) extensively for ingesting data from different source systems like relational and unstructured data to meet business functional requirements
* Experience with Snowflake Multi - Cluster Warehouses.
* Collaborated with data architects, BI architects, and data modeling teams during data modeling sessions.
* Developed jobs, components and Joblets in Talend. Designed ETL Jobs/Packages using Talend Integration Suite (TIS).
* Extensive experience in building high-level documents depicting various sources, transformations, and targets.
* Extensively used Informatica transformations- Source qualifier, expression, joiner, filter, router, update strategy, union, sorter, aggregator and normalizer transformations to extract, transform, and load the data from different sources into DB2, Oracle, Teradata, Netezza, and SQL Server targets.
* Optimization of Hive queries using best practices and right parameters and using technologies like Hadoop, YARN, Python, PySpark.
* Understanding of SnowFlake cloud technology.
* Extensively used Informatica Data Explorer (IDE) & Informatica Data Quality (IDQ) profiling capabilities to profile various sources, generate scorecards, create and validate rules, and provide data for business analysts to create the rules.
* Created complex mappings in Talend using tHash, tDenormalize, tMap, tUniqueRow. tPivotToColumnsDelimited as well as custom component such as tUnpivotRow.
* Created numerous pipelines in Azure using Azure Data Factory v2 to get the data from disparate sourcesystems by using different Azure Activities like Move &Transform, Copy, filter, for each, Databricks etc.
* Maintain and provide support for optimal pipelines, data flows and complex data transformations and manipulations using ADF and PySpark with Databricks.
* Created Talend Mappings to populate the data into dimensions and fact tables. Frequently used Talend Administrative Console (TAC).
* Extensively used ETL Informatica to integrate data feed from different 3rd party source systems - Salesforce and Touch Point.
* Created internal and external stage and transformed data during load.
* Used Polybase to load tables in Azure synapse.
* Developed spark applications in python(PySpark) on distributed environment to load huge number of CSV files with different schema in to Hive ORC tables.
* Developed data warehouse model in snowflake for over 100 datasets using whereScape.
* Orchestrated the setup of Kafka connectors within Informatica Power Center workflows, enabling seamless data movement between Kafka topics and diverse target systems.
* Scheduled, automated business processes and workflows using Azure Logic Apps.
* Implemented Azure, self-hosted integration runtime in ADF.
* Leveraged Kafka streams within Informatica environments to support real-time data processing and analytics, meeting evolving business needs for timely insights
* Unit tested the data between Redshift and Snowflake.
* Used Informatica Data Quality transformations to parse the “Financial Advisor” and “Financial Institution” information from Salesforce and Touchpoint systems and perform various activities such as standardization, labeling, parsing, address validation, address suggestion, matching, and consolidation to identify redundant and duplicate information and achieve the MASTER record.
* Experience in building Azure stream Analytics ingestion spec for data ingestion which helps users to get sub second results in Real Time.
* Created data sharing between two snowflake accounts.
* Extensively used Standardizer, Labeler, Parser, Address Validator, Match, Merge, Consolidation transformations.
* Extensively worked on performance tuning of Informatica and IDQ mappings.
* Created Informatica workflows and IDQ mappings for - Batch and Real Time.
* Converted and published Informatica workflows as Web Services using Web Service Consumer transformation as source and target.
* Involved in designing, developing, and deploying reports in MS SQL Server environment using SSRS-2008 and SSIS in Business Intelligence Development Studio (BIDS).
* Analysed the sql scripts and designed it by using PySpark SQL for faster performance.
* Redesigned the Views in snowflake to increase the performance.
* Used ETL (SSIS) to develop jobs for extracting, cleaning, transforming, and loading data into the data warehouse.
* Worked on the ETL process using Informatica BDM.
* Involved in the Migration of Databases from SQL Server 2005 to SQL Server 2008.
* Prepared the complete data mapping for all the migrated jobs using SSIS.
* Designed SSIS Packages to transfer data from flat files to SQL Server using Business Intelligence Development Studio.
* Extensively used SSIS transformations such as Lookup, Derived column, Data conversion, Aggregate, Conditional split, SQL task, Script task Send Mail task, etc
* Created reusable components, reusable transformations, and applets to be shared among the project team.
* Used ILM&TDM to mask sensitive data in Dev, and QA environments.
* Used XML & MQ series as the source and target.
* Used-in-built reference data such as token sets, reference tables, and regular expressions to build new reference data objects for various parse/cleanse/purging needs.
* Extensive experience in integration of Informatica Data Quality (IDQ) with Informatic Power Center.
* Worked closely with the MDM team to identify the data requirements for their landing tables and designed the IDQ process accordingly.
* Created Informatica mappings keeping in mind Informatica MDM requirements.
* Extensively used XML, and XSD/schema files as source files, parsed incoming SOAP messages using XML parser transformation, and created XML files using XML generator transformation.
* Worked extensively with Oracle external loader- SQL loader - to move the data from flat files into Oracle tables.
* Worked extensively with Teradata utilities- Fast load, Multi load, Tpump, and Teradata Parallel Transporter (TPT) to load huge amounts of data from flat files into the Teradata database.
* Created BTEQ scripts to invoke various load utilities, transform the data, and query against the Teradata database.
* Proficient in performance analysis, monitoring, and SQL query tuning using EXPLAIN PLAN, Collect Statistics, Hints, and SQL Trace both in Teradata as well as Oracle.

**Environment:** Informatica Power Center 10.4,10.2,9.6, Oracle 19c,12c,11g, 10g, T-SQL, IDQ, Snowflake, Redshift Informatica MDM 10.1/10.2, Informatica MDM Data Director 10.1/10.2, Azure ML, MS SQL Server 2008, UNIX (Sun Solaris5.8/AIX), Data Marts, Erwin Data Modeler 4.1, Agile Methodology, Teradata 13, FTP, MS-Excel, SQL Integration Services (SSIS).

**Client: Conduent – Austin, TX FEB 2018 to JULY’2020**

**Role: ETL Developer**

**Responsibilities:**

* Gathered user Requirements and designed source-to-target data load specifications based on business rules.
* Evaluate Snowflake Design considerations for any change in the application.
* Experience in building ETL(Azure Data Bricks) data pipelines leveraging PySpark, Spark SQL.
* Used Informatica Power Centre 9.6 for extraction, loading, and transformation (ETL) of data in the data mart
* Participated in the review meetings with the functional team to sign the Technical Design document.
* Experience in building the Orchestration on Azure Data Factory for scheduling purposes.
* Developed complex Talend ETL jobs to migrate the data from flat files to database. Implemented custom error handling in Talend jobs and also worked on different methods of logging.
* Developed Python code to gather the data from HBase(Cornerstone) and designs the solution to implement using PySpark.
* Created ETL/Talend jobs both design and code to process data to target databases
* Involved in the Design, Analysis, Implementation, Testing, and support of ETL processes
* Worked with Informatica Data Quality 9.6 (IDQ) toolkit, Analysis, data cleansing, data matching, data conversion, exception handling, and reporting and monitoring capabilities of IDQ 9.6.
* Designed, Developed, and Supported Extraction, Transformation, and Load Process (ETL) for data migration with Informatica Power Center.
* Solid experience in implementing complex business rules by creating re-usable transformations and robust mappings using Talend transformations like tConvertType, tSortRow, tReplace, tAggregateRow, tUnite etc.
* Developed Talend jobs to populate the claims data to data warehouse - star schema.
* Build the Logical and Physical data model for snowflake as per the changes required.
* Developed various mappings using Mapping Designer and worked with Aggregator, Lookup, Filter, Router, Joiner, Source Qualifier, Expression, Stored Procedure, Sorter and Sequence Generator transformations.
* Created complex mappings that involved Slowly Changing Dimensions, implementation of Business Logic, and capturing the deleted records in the source systems.
* Create PySpark scripts to load data from source files to RDDs, create data frames from RDD and perform transformations and aggregations and collect the output of the process.
* Integrated java code inside Talend studio by using components like tJavaRow, tJava, tJavaFlex and Routines.
* Experienced in using debug mode of talend to debug a job to fix errors. Created complex mappings using tHashOutput, tHashInput, tNormalize, tDenormalize, tMap, tUniqueRow. tPivotToColumnsDelimited, etc
* Experience working with Azure Logic APP Integration tool.
* Define roles, privileges required to access different database objects.
* Worked extensively with the connected lookup Transformations using dynamic cache.
* Created Context Variables and Groups to run Talend jobs against different environments.
* Worked with complex mappings having an average of 15 transformations.
* Expertise on working with databases like Azure SQL DB, Azure SQL DW.
* Coded PL/SQL stored procedures and successfully used them in the mappings.
* Coded Unix Scripts to capture data from different relational systems to flat files to use as a source file for the ETL process and schedule the automatic execution of workflows.
* Implemented FTP operations using Talend Studio to transfer files in between network folders as well as to FTP server using components like tFileCopy, TFileAcrchive, tFileDelete, tCreateTemporaryFile, tFTPDelete, tFTPCopy, tFTPRename, tFTPut, tFTPGet etc.
* Define virtual warehouse sizing for Snowflake for different type of workloads.
* Scheduled the Jobs by using Informatica scheduler & Job Trac.
* Worked on Oracle Databases, RedShift and Snowflakes, Defining virtual warehouse sizing for Snowflake for different type of workloads.
* Created and scheduled Sessions and jobs based on demand, run on time, and run only once
* Monitored Workflows and Sessions using Workflow Monitor.
* Performed Unit testing, Integration testing, and System testing of Informatica mappings.
* Involved in enhancements and maintenance activities of the data warehouse including tuning, and modifying stored procedures for code enhancements.
* Experienced in Building a Talend job outside of a Talend studio as well as on TAC server.
* Leveraged Kafka streams within Informatica environments to support real-time data analytics and monitoring, ensuring responsiveness and agility in ETL processes.
* Develop stored procedures/views in Snowflake and use in Talend for loading Dimensions and Facts.
* Collaborated with Kafka infrastructure teams to optimize Kafka cluster performance and troubleshoot any issues related to data streaming within the ETL workflows.
* Used ETL methodologies and best practices to create Talend ETL jobs.
* Responsible for determining the bottlenecks and fixing the bottlenecks with performance tuning at various levels like mapping level, session level, and database level.
* Introduced and created many project-related documents for future use/reference.
* Designed and developed ETL Mappings to extract data from Flat files and Oracle to load the data into the target database.
* Developing several complex mappings in Informatica a variety of Power center transformations, Mapping Parameters,
* Mapping Variables, Mapplets & and Parameter files in Mapping Designer using Informatica Power Center.
* Created complex mappings to load the data mart and monitored them. The mappings involved extensive use of Aggregator, Filter, Router, Expression, Joiner, Union, Normalizer, and Sequence generator transformations.
* Ran the workflows on a daily and weekly basis using a workflow monitor.

**Environment:** Informatica 9.6, PL/SQL, Talend 5.5/5.0, Informatica Data Quality IDQ 9.6, Snowflake, Redshift, Informatica, 9.5, Oracle 9i, UNIX, SQL, PL/SQL, Informatica Scheduler, SQL\*loader, SQL Developer, Framework Manager, Transformer, Teradata.

**Client: ITC Infotech – India AUG’2015 to DEC’2017**

**Role: ETL Engineer**

**Responsibilities**

* Involved in designing the mapping document according to HIPAA 5010 and ICD10 standards.
* Created views on all required claims tables and got required data into the staging area.
* Created 837 outbound tables in the database and populated all the data in it from the staging area using Informatica Power Center.
* Converted 837 claims table data into EDI format using data transformation.
* Developed different mappings for healthcare institutional and professional data.
* Converted EDI X12 format file claims to XML files using a parser in B2B data transformation.
* Filtered the XML claims files by using filter conditions on the D9 segment and converted back the filtered XML claim files to EDI format using a serializer in B2B data transformation.
* Used B2B Data Exchange for end-to-end data visibility through event monitoring and to provide a universal data transformation supporting numerous formats, documents, and files.
* Populated the Acknowledged information about the claims in the database tables.
* Developed Technical Specifications of the ETL process flow.
* Extensively used ETL to load data from Flat files, XML files SQL server, and Oracle sources to SQL server target.
* Created mappings using different transformations like Expression, Unstructured Data transformations, Stored Procedure, Filter, Joiner, Lookup, and Update Strategy.
* Created Triggers to update the data in the 837 outbound tables.
* Used Tidal for scheduling.
* Involved in Integration, system, and performance testing levels.

**Environment:** Informatica Power Center 9.1.0, SQL Server 2008, Oracle 11g, Toad 8.0, B2B Data Transformation, FACT, T-SQL, Windows7.