

# Vijay V

**Visa:- H1b**

**Current location:- Plainsboro, NJ**

**Ready to relocate anywhere in USA**

**PROFESSIONAL SUMMARY:**

* Overall 9+ **years** of IT experience in Analysis, Design, Developing and Testing and Implementation of business application systems.
* Proficient in designing, developing, validating, and deploying **ETL** processes for data collection, data staging, data movement, data quality, and archiving strategies.
* Expertise in designing **ETL** jobs using various Talend components, creating context variables, joblets, routines, and utilizing them in project-level jobs.
* Strong experience in sourcing data from S3 buckets and loading it into Redshift in an **AWS** environment.
* Skilled in scaling and performance tuning of Talend jobs for optimal efficiency.
* Familiar with **Talend Administration Center** (**TAC**) for job publication and scheduling using Job Conductor & Execution Plans.
* Experience in error handling and debugging Talend jobs to quickly identify and resolve issues.
* Conduct ETL unit and development tests, monitor results, and take corrective actions as needed.
* Prepare detailed system documentation, including requirements, design specifications, test plans, and user manuals/run books.
* Experiences on developing & leading the end-to-end implementation of Big Data projects by using Talend BIGDATA, comprehensive experience as a in Hadoop Ecosystem like **Map Reduce, Hadoop Distributed File System (HDFS), Hive.**
* 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 **Oracle, DB2, SQL server, Teradata, Hive, Hana**, and **flat files, XML** and **Mainframe files and Active MQ.**
* Good Experiences on relational database management systems, experience in integrating data from various data source like **Oracle, MSSQL Server, MySQL** and **Flat files** too.
* Hands on experience on Hadoop technology stack (**HDFS, Map-Reduce, Hive, HBase and Spark**).
* Excellent knowledge in deployment process from **DEV** to **QA, UAT** and **PROD** with both Deployment group and Import/Exports method.
* Excellent working experience in Waterfall, Agile methodologies.
* Familiar with design and implementation of the Data Warehouse life cycle and excellent knowledge on entity-

relationship/multidimensional modeling (star schema, snowflake schema), Slowly Changing Dimensions (**SCD Type 1, Type 2, and Type 3**).

* Debugging ETL jobs errors, ETL Sanity and production Deployment in TAC-Talend **Administrator Console** using **SVN.**
* Experience in Trouble shooting and implementing Performance tuning at various levels such as **Source, Target, Mapping Session** and System in ETL Process.
* Experience in converting the **Store Procedures** logic into ETL requirements.
* Good communication and interpersonal skills, ability to learn quickly, with good analytical reasoning and adaptive to new and challenging technological environment.
* Experience in Big Data technologies like **Hadoop/Map Reduce, HBASE, Hive, Sqoop, Hbase, Dynamodb, Elastic Search and Spark SQL**.
* Experienced in ETL methodology for performing **Data Migration, Data Profiling, Extraction, Transformation and**

**Loading** using Talend and designed data conversions from large variety of source systems including Oracle […] **DB2, Netezza, SQL server, Teradata, Hive, Hana and non-relational sources like flat files, XML** and Mainframe files.

* Experiences on Designed **tables, indexes and constraints** using **TOAD** and loaded data into the database using SQL\*Loader.
* Experience in **AWS** S3, EC2, EMR, Lambda, RDS (MySQL) and Redshift cluster configuration.
* Knowledge on Data Warehousing ETL experience of using Informatica 9.x/8.x/7.x Power Center Client tools - Mapping Designer, Repository manager, Workflow Manager/Monitor and Server tools - Informatica Server, Repository Server manager.
* Experience in **SQL Plus** and **TOAD** as an interface to **databases, to analyze, view and alterdata.**
* **Expertise in Data Warehouse/Data mart, ODS, OLTP and OLAP** implementations teamed with project scope, Analysis, requirements gathering, data modeling, Effort Estimation, ETL Design, development, System testing, Implementation, and production support.
* Experience in Dimensional Modeling using **Star and Snowflake Schema, Identifying Facts and Dimensions**.
* Utilized AWS services (S3, EC2, EMR, Amazon Redshift).
* Experience in implementing Azure data solutions, provisioning storage account, Azure Data Factory, SQL server, SQL Databases, SQL Data warehouse, Azure Data Bricks and Azure Cosmos DB.
* Implementation of data movements from on-premises to cloud in Azure.
* Extensive experience in developing Stored Procedures, Functions, Views and Triggers, Complex SQL queries using SQL Server, **TSQL** and Oracle PL/SQL.
* Extensive experience in writing UNIX shell scripts and automation of the ETL processes using UNIX shell scripting, and also used Netezza Utilities to load and execute SQL scripts using Unix.
* Created **Snowflake** Schemas by normalizing the dimension tables as appropriate, and creating a Sub Dimension named Demographic as a subset to the Customer Dimension.

**TECHNICAL SKILLS:**

|  |  |
| --- | --- |
| **Development Tools** | Talend 7.1, 7.3.1 Talend ESB 6.2, informatica, DataStage, SSIS, Putty, File Zilla, SQL Developer, Toad, Teradata SQL  Assistant, Quality Center, Harvest, SQL Server Management Studio. |
| **Programming Languages** | Core Java, SQL, PL/SQL, C, C++ |
| **Operating Platforms** | MS Windows 2010/2000/XP/NT, Mac OS, Unix and Linux. |
| **Web Technologies** | HTML, XML, CSS, XSD, JavaScript and JSON. |
| **Cloud Technolgies** | Snowflake, AWS Services(S3, EC2, EMR, Amazon Redshift) |

**WORK EXPERIENCE:**

**Client: Ford Direct, Dearborn, MI**

**Apr 2023 – Till date**

**Role: Sr ETL Talend Developer**

**Responsibilities:**

* Designed and implemented **Talend** ETL workflows to support supply chain data integrations, reducing data processing time and enhancing data accuracy for operational reporting.
* Designed and implemented scalable ETL workflows using Talend Big Data components for processing multi-terabyte datasets from diverse sources such as HDFS, Kafka, and relational databases.
* Built and maintained scalable data pipelines in a Big Data environment using Hive, and HDFS.
* Developed ETL processes to integrate data from ERP (SAP), CRM, and third-party logistics systems into a central Snowflake database, enabling real-time data visibility and improved decision-making.
* Collaborated with supply chain analysts to design data mappings and transformations for forecasting, inventory management, and demand planning use cases.
* Worked on Migration of Talend 7.1.1 to Talend 8
* Created Multiple common jobs as a reference project.
* Integrated Azure Technology in Talend jobs to load the file in Azure Blob
* Integrated java routines to check/delete the file in Azure Blob Storage.
* Optimized Talend jobs for higher Performance.
* Implemented Error logging Mechanism using **tLogcatcher** and **tStatcather**.
* The errors and metrics are then sent to ELK logs.
* Implemented Various Filters to filter the data of all the dealers.
* Using CI/CD as pipeline to deploy the jobs in TMC and promotions to the Higher environments.
* Used SFTP shared location to access the data with other teams.
* Implemented Rest API’s Process Parallelly to download around 80 GB of data in a single run.
* Discussed Various scenarios with the team to achieve the perfect job for the process.
* Implement data masking in the job to hind sensitive data.
* Implemented PRE load, Main Logic, post load process for each and every job as per the design process.
* Created and deployed physical objects including custom tables, custom views, stored procedures, and Indexes to SQL Server for Staging and Data-Mart environment.
* Design and Implemented ETL for data load from heterogeneous Sources to SQL Server and Oracle as target databases and for Fact and Slowly Changing Dimensions SCD-Type1 and SCD-Type2.
* Created many complex ETL jobs for data exchange from and to Database Server and various other systems including RDBMS, XML, CSV, and Flat file structures.
* Worked within a cross-functional team to execute data migration initiatives for a major supply chain overhaul, successfully transferring [millions of records] into Oracle and SQL databases.
* Migrated large datasets (e.g., [volume details, if applicable]) from legacy systems to Teradata, ensuring data integrity and minimizing downtime.
* Integrated data from multiple sources, including Oracle and SQL Server, into Teradata environments.
* Working with Talend and HVR for real-time data replication across heterogeneous environments.
* Using the integration of **FiveTran** for automated data synchronization with cloud databases, enhancing reporting capabilities.
* Migrated EHR data to AWS Redshift using **FiveTran** and Talend, ensuring HIPAA compliance.

**Client: Walgreens (TCS)/ Deerfield, IL Nov 2021 – Jan 2023**

**Role: ETL Talend Support/Developer Responsibilities:**

* + Working on Talend big data 7.3.1 Enterprise Version.
  + Design, developed, and deployed of **Talend ETL** processes, adhering to business rules and mapping documents.
  + Worked in a large data warehousing environment, extracting data from various sources and ensuring data quality and accuracy.
  + Designed **ETL** jobs using a variety of Talend components, creating context variables, joblets, and routines for project-level jobs.
  + Successfully sourced data from **S3** buckets and loaded it into Redshift, optimizing data transfer for efficiency.
  + Implemented scaling and performance tuning techniques to enhance the speed and reliability of Talend jobs.
  + Managed **Talend Administration Center (TAC),** publishing and scheduling jobs using Job Conductor & Execution Plans.
  + Utilized Repository Manager for smooth migration of source code between different environments.
  + Installed and set up Talend Server modules and troubleshoot for issues and performance bottlenecks
  + Used TAC to set up users & project authorizations, migrating and scheduling jobs, and monitor job.
  + Monitor Talend jobs which are scheduled in the production environment Closely for any failures and resolve it.
  + Also work on minor enhancements which are recommended by business if it’s required for the production jobs.
  + Worked on different file formats in Talend like XML, CSV, JSON
  + Knowledge of Unix/Linux commands.
  + Designed the data conversion strategy, development of the data mappings, some data profiling and the design of Extraction, Transformation and Load (**ETL**) routines for migrating data from non-relational or source totarget relational.
  + Resolved issues with Nexus Artifact Repository, streamlining code deployment activities and ensuring automation.
  + Testing the product in controlled, real situations before going live and maintaining the systems once they are up and running.
  + Created promotions for the Talend dev jobs after publishing into Talend cloud.
  + Worked on data quality like schema checks using tschemacheck component.
  + As per business requirements we have created test data with pros and corn scenarios.
  + Implemented data quality checks of incoming data before development.
  + Used tRunJob component to run child job from a parent job and to pass parameters from parent to child job.
  + Created context variables and groups to run Talend jobs against different environments.
  + Data ingestion with different data sources and load into Redshift.
  + Loaded the files into s3 bucket and copy into redshift for creating tables.
  + Developed the data ingestion i.e transformation on data source raw data file into aws redshift cluster using talend.
  + Worked with Healthcare Data/Claims 835 and 837 formats for analytical purposes, X12 Enterprise Data Integration (EDI), PHI
  + Worked with X12 EDI standards for healthcare data, HEDIS, and HIPAA.
  + Wrote UNIX shell scripts to work with flat files, to define parameter files and to create pre and post session commands.
  + Prepared comprehensive system documentation, including requirements, design specifications, test plans, and user manuals/run books.
  + Develop stored procedures/views in Snowflake and use in Talend for loading Dimensions and Facts.
  + Develop merge scripts to UPSERT data into Snowflake from an ETL source.

# Client: Indiana Department of Homeland Security (IDHS)/ Indianapolis, IN Apr 2021 – Oct 2021

**Role: ETL Talend Developer Responsibilities:**

* + Develop jobs to ingest client data sets using Talend Big Data components and Amazon Webservices (**AWS**).
  + Working on Talend 7.3.1 Enterprise version.
  + Develop and design of **ETL** (Extract, Transform and Load data) methodology for supporting data transformations and processing in a corporate wide ETL solution using Talend.
  + Developed Jobs in Talend to Load the data into the Amazon S3 buckets in **XML** &**CSV** format.
  + Created **EC2** instances with blob storage in Amazon Web Services (**AWS**).
  + Part of data loading into a data warehouse using big data Hadoop Talend ETL components, AWS S3 buckets and AWS and AWS services for the redshift database.
  + We must design jobs using Bigdata Talend and pick files from AWS S3 Buckets and load it into AWS Redshift.
  + Designed the data conversion strategy, development of the data mappings, some data profiling and the design of Extraction, Transformation and Load (**ETL**) routines for migrating data from non-relational or source totarget relational.
  + Analyze **ETL** requirements, design complex DWH, implement complex ETL jobs according to the given requirement.
  + Create jobs using Talend Bigdata Integration tool for performing client level validations and transformations.
  + Designed a process to download a high volume of data from **AWS** S3 to EMR to save space on the job server.
  + Expand or modify system to serve new purposes or improve workflow.
  + Used Talend Management Console (**TMC**) to Schedule data ingestion and data analyticjobs.
  + Worked on analysis, design, and write complete project documentation.
  + Troubleshoot and solve data integration problems in a timely manner.
  + Ensure data quality checks are in place and routinely adjust processes to uphold quality.
  + Extracted data from flat files/ databases applied business logic to load them in staging databases as well as flat files.
  + Worked on Talend jobs performance tuning.
  + Testing the product in controlled, real situations before going live and maintaining the systems once they are up and running.
  + Developed Unix shell scripts to support **ETL** processes and wrote complex **SQL** queries for data manipulation.
  + Actively participated in an **AWS** environment, collaborating with cross-functional teams to achieve project goals.
  + Demonstrated strong error-handling and debugging skills, quickly identifying and resolving issues in **Talend** jobs.
  + Conducted **ETL** unit and development tests, monitoring results and taking corrective actions as needed.

# Client: Veterans Affair/ Austin, TX Aug 2020 – Feb 2021

**Role: ETL Developer Responsibilities:**

* Worked on **Talend 7.1.1**
* Analyzing the source data to know the quality of data by using **Talend Data Quality**
* Involved in writing SQL Queries and used Joins to access data from **Oracle, and MySQL.**
* 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.
* Developed mappings to load Fact and Dimension tables, SCD Type 1 and SCD Type 2 dimensions and Incremental loading and unit tested the mappings.
* Used **tStatsCatcher, tDie, tLogRow**to create a generic Joblets to store processing stats into a Database table to record job history.
* Created complex mappings with shared objects/Reusable **Transformations/Mapplets** for staging unstructured HL7 files into **Data Vault**.
* Used **Data Vault** for traditional batch loads and incremental build loads.
* 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.

* Used **tRunJob**component to run child job from a parent job and to pass parameters from parent to child job.
* Created Context Variables and Groups to run Talend jobs against different environments.
* Used **tParalleize**component and multi thread execution option to run subjobs in parallel which increases the performance of a job.
* worked extensively in T-SQL for various needs of the transformations while loading the data into Data vault.
* Implemented Third party **Scheduler (Automic Schedular)** to trigger the jobs in TAC server.
* Have used **AWS** components (Amazon Web Services) - Downloading and uploading data files (with ETL) to AWS system using S3 talend components.
* Design and Develop ETL Processes in **AWS** Glue to migrate Campaign data from external sources like **S3**, Parquet/Text Files into **AWS** Redshift.
* Created SSIS package for loading the data coming from various interfaces and also used multiple transformations in SSIS to collect data from sources.

# Client: Clarios, Milwaukee, WI

# Dec 2019 – June 2020

**Role: ETL Talend Consultant**

**Responsibilities:**

* Worked on Talend Data Integration/Big Data Integration (6.1/5.x) / Talend Data Quality.
* 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, tDieetc.

* 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.
* Design and Implemented ETL for data load from heterogeneous Sources to **SQL Server** and Oracle as target databases and for Fact and Slowly Changing Dimensions **SCD**-Type1 and**SCD**-Type2.
* Utilized Big Data components like tHDFSInput, tHDFSOutput, tPigLoad, tPigFilterRow, tPigFilterColumn, tPigStoreResult, tHiveLoad, **tHiveInput**, **tHbaseInput**, **tHbaseOutput**, **tSqoopImport**and**tSqoopExport**.
* Worked Extensively on Talend Admin Console (**TAC**) and Schedule Jobs in Job Conductor.
* Developed Talend jobs to populate the claims data to data warehouse - **starschema**.
* Troubleshot long running jobs and fixing the issues.
* Extensively used **tSAPBapi**component to load and read data from SAP System.
* Created jobs to pass parameters from **child job** to **parent job**.
* Developed **data warehouse model** in **snowflake** for over 100datasets.
* Exported jobs to **Nexus** and **SVN** repository.
* Implemented update strategy on tables and used **tJava**, **tJavarow** components to read data from tables to pull only newly inserted data from source tables.
* Observed statistics of Talend jobs in **AMC** to improve the performance and in what scenarios errors are causing.
* Implemented few **java** functionalities using **tJava** and **tjavaflex**components.
* Part of data loading into data warehouse using big data Hadoop **Talend** ETL components, **AWS S3** Buckets and AWS Services for redshift database.
* We must design jobs using Bigdata Talend and Pick files from **AWS S3** Buckets and Load into AWS Redshift database
* Part of redshift database **AWS** maintenance, we must vacuum and analyze our **AWS** redshift tables.
* Designed and developed **SSIS** packages to move data from various sources into destination flat files and databases.
* Worked on **SSIS** package, DTS Import/Export for transferring data from database(oracle and text format data) to SQL server.

# Client: Simon Property Group/ Indianapolis, IN Apr 2017 – Nov 2019

**Role: ETL Talend Developer Responsibilities:**

* Developed ETL (data transformations and data movement) using **Talend integration, SQLserver**.
* Supported production release during migration and high per care period.
* Created Java Routines, Reusable transformations, Joblets using **Talend** as an **ETL** Tool
* Developed mappings to load Fact and Dimension tables, **SCDType 1** and **SCD Type 2** dimensions and Incremental loading.
* Developed **CDC (Change Data Capture)** jobs to process inserts, updates and deletes in talend using routines.
* Refined and performed information relocations with **SQL** Server and complex Excel macross
* Made **SQL** put away systems and capacities for information purifying, organizing, and the consolidation of business rules.
* Performed issue examination and relocation issue goals with every day customer calls If customer support all through movement.
* Migrating from Cast Iron (**Data stage**) Jobs to**Talend**.
* Arranged documentation for the help assets on new usefulness and material for examination of client issues Striking Accomplishments.
* Designed and implemented product **data vault** to extract, transform and load the product data with history of changes on the products attributes.
* Made new SQL put away methodology to improve precision and speed of existing Forms.
* Extensively used database connections, file component, tMap, tJoin, tReplicate, tParallelize, tConvertType,

# tflowtoIterate, aggregate, tSortRow, tFlowMeter, tLogCatcher, tRowGenerator, tNormalize, tDenormalize, tSetGlobalVar, tHashInput, tHashOutput, tJava, tJavarow, tAggregateRow, tWarn, tLogCatcher.

* Involved in design, development of **Talend** mappings.
* For dividing rows used **tSplitrow**component and for removing duplicates used**tUniquerow.**
* Experienced with structured query nomenclature. Examples include **MSSQL**, DB2, SQL, etc.

**Client: Usine Technologies / Hyderabad, INDIA**

**Apr 2014 – June 2015**

**Role: SQL Developer**

**Responsibilities:**

* Created and managed schema objects such as Tables, Views, Indexes, and referential integrity depending on user requirements.
* Actively involved in the complete software development life cycle for a design of database for new Financial Accounting System.
* Successfully implemented the physical design of the new designed database into **MSSQL** Server 2008/2005.
* Used **MS SQL Server 2008/2005** to design, implement and manage data warehouses OLAP cubes and reporting solutions to improve asset management, incident management, data center services, system events support and billing.
* Utilized T-SQL daily in creating customs view for data and business analysis.
* Utilized Dynamic T-SQL within functions, stored procedures, views, and tables.
* Used the **SQL Server Profiler** tool to monitor the performance of SQL Server particularly to analyze the performance of the stored procedures.
* Stored Procedures and Functions were optimized to handle major business crucial calculations.
* Implementation of data collection and transformation between different heterogeneous sources such as flat file, Excel and SQL Server 2008/2005 using **SSIS**.
* Migrated all DTS packages to **SQL Server Integration Services (SSIS**) and modified the package according to the advanced feature of SQL Server Integration Services.
* Defined Check constraints, rules, indexes and views based on the business requirements.
* Extensively used SQL Reporting Services and Report Builder Model to generate custom reports.
* Designed and deployed Reports with Drop Down menu option and Linked reports.
* Created subscriptions to provide a daily basis report and managed and troubleshoot report server related issues.

**Environment:** MS SQL Server 2005/2008, SSIS, SSRS, Microsoft.NET, MS Office Access2003, Excel, Erwin.