



Bhagwan Chaubey



E-Mail: bhagwan.chaubey@gmail.com

Mob: +91 9142105441 / +91 7596871311

Profile Summary:

- Having around 14+ years with strong background in Data warehousing using Enterprise ETL Tools Talend Data Integration, Talend Big Data, Informatica, Databricks and Snaplogic
- Extensively used ETL Methodology for performing data extraction, transformation and loading using Talend and designed data conversions from wide variety of source systems including flat files, Oracle, SQL Server, Snowflake
- Good Experience in Development of **Talend, Snowflake.**
- Worked on snowflake functionality Stream, Snowpipe, Timetravel, Masking, Clone, Row access Policy, Merge, Internal and External Stage, Error Handling.
- Hands-on experience in development, testing, bug fixing and maintenance phases
- Extensive knowledge in writing SQL statements to meet the business requirements
- Capable of analyzing and understanding the business requirements
- Worked in creating documents like Technical Specification, Design and Migration documents.
- Significant experience with Data Extraction, Transformation and Loading (ETL) from disparate data sources such as multiple relational databases and worked on integrating data from flat files (CSV, PSV, txt files) and XML, JSON files into a common reporting and analytical data model.
- Having experience on TAC (Talend Administration Centre)
- Understand the client business requirement
- With excellent user communication, quick analyzing, planning, strong analytical and problem solving and presentation skills
- Self-motivated, team player, highly dedicated, adaptable to the demands of the projects and zeal to learn New Technologies and Tools
- Experience in scheduling and maintaining ETL jobs
- Worked with Spark ecosystems such as **Spark Core, Pyspark, Spark SQL, Spark Streaming**
- Experience with Messaging System **Kafka**
- Knowledge of **Snaplogic**
- Experience in **PL/SQL** and **T-SQL**.
- Having experience on **Azure Data Factory(ADF)**.
- Good knowledge on latest technology like Azure Synapse Analytics, Azure Streaming Analytics and in Big Data Platform like Azure Databricks
- Migrating Data from **Talend** to **Azure Databricks**
- Hands-on experience in **Talend ESB** using **Route**
- Experience in **DBT**.

Professional Experience:

- Working as a Senior Software Engineer in Intraedge from May - 2024 to 2 Jan 2025
- Worked as a Technical Architect in Virtusa from Aug - 2023 to May 2024
- Working as a Technical Architect in Perficient from May - 2022 to April 2023.
- Worked as a Sr. Data Engineer in Nityo, Kolkata from Nov - 2020 to Feb 2022
- Worked as a Sr. Associate in Cognizant, Kolkata from October - 2014 to June 2020
- Worked as a Sr. Software Engineer at Accenture, Bangalore from July – 2011 to Sep – 2014.
- Worked as an Associate at Datamatics Software Services Limited, Bangalore from Dec – 2010 to June – 2011

Technical Skills:

Databases	: Snowflake, Oracle, MS SQL Server
Software Tools	: Putty, FileZilla, VS Code and S3 Browser
Software/Languages	: PL/SQL, T-SQL, Scala, Python
Operating System	: Windows, Linux
Cloud Technology	: Amazon Web Services (S3, EC2, Lambda, API Gateway), Azure (Storage Account, ADLS, Database), Snowflake
Data Warehouse DB	: Amazon Redshift, Teradata
ETL Technologies	: Talend Cloud Big Data platform (7.3), Talend DI/Big data 6.3/6.5, Talend ESB Informatica, Snaplogic, Snowflake
Big Data	: Azure Databricks (Spark and Scala), Azure Synapse Analytics
Agile Tools	: Confluence, Jira

Certification: : Talend Data Integration Certified Developer-2023
Databricks Certified Associate Data Engineer-2023

Project # 9

Project Title	: Ebay
Environment	: Snowflake, Talend, Oracle, AWS
Role	: Sr. Data Engineer

Description:

This project involves designing and implementing an ETL pipeline that extracts retail-related data from eBay's platforms (such as listings, orders, inventory, and customer data), transforms the data to meet business intelligence requirements, and loads it into a data warehouse or data lake for analysis and reporting

Roles and Responsibilities:

- Accountable for all the project and program management and delivery at offshore and onshore for multiple projects in Agile format, using JIRA board.
- Involved in customer project management for retail domain.
- Involved in various Agile Ceremony like Sprint Planning, Retrospective Meeting, Daily Scrum etc.
- Preparation of various metrics like Sprint burndown chart, Release burndown chart.
- Participated in Presales activity and involved in preparation of RFP for various proposals.
- Sharing the project knowledge through Confluence tools.
- Involved in planning activities for various projects in the account.
- Involved in resource allocation in various projects.
- Managing and approving the team timesheet and issues related to billing, invoices, budget and P&L.
- Involved in estimation of efforts for various projects.
- Mentored the team member technically on Talend, Snowflake, Azure Data Factory, Data Visualization and Data Techniques.
- Meeting with various stakeholders of the project to discuss and close the outstanding issues and to discuss the project status.

Project # 8

Project Title : Citi
Environment : Informatica, Talend, Oracle
Role : Lead Data Engineer

Description:

Citi is one of the world's leading international banks, with a global network of financial services in nearly 160 countries. Its financial products and services to institutions, including corporate and commercial banking, investment banking, and treasury and trade services.

Roles and Responsibilities:

- Accountable for all the project and program management and delivery at offshore and onshore for multiple projects in Agile format, using JIRA board and Confluence.
- Involved in customer project management for banking domain.
- Involved in various Agile Ceremony like Sprint Planning, Retrospective Meeting, Daily Scrum etc.
- Preparation of various metrics like Sprint burndown chart, Release burndown chart.
- Sharing the project knowledge through Confluence tools.
- Involved in planning activities for various projects in the account.
- Involved in resource allocation in various projects.
- Managing and approving the team timesheet and issues related to billing, invoices, budget and P&L.
- Involved in estimation of efforts for various projects.
- Meeting with various stakeholders of the project to discuss and close the outstanding issues and to discuss the project status.
- Provide technical solution to our team member.
- Worked on query optimization in Oracle and Code optimization in Talend.
- Involved in Performance tuning in SQL and PL/SQL, Talend.

Project # 7

Project Title : EDH Platinum
Environment : Databricks, PySpark
Role : Sr. Data Engineer

Description:

The project involves on Azure Cloud Platform, and data ingestion of data using Spark and migration of Hadoop (Cloudera) to Azure Synapse Analytics in Wealth Management Domain. The data is used to provide insights, which can be used by organization

Roles and Responsibilities:

- Accountable for all the project and program management and delivery at offshore and onshore for multiple projects in Agile format, using JIRA board.
- Involved in customer project management for banking domain.
- Involved in various Agile Ceremony like Sprint Planning, Retrospective Meeting, Daily Scrum etc.
- Preparation of various metrics like Sprint burndown chart, Release burndown chart.
- Participated in Presales activity and involved in preparation of RFP for various proposals.
- Sharing the project knowledge through Confluence tools.

- Involved in planning activities for various projects in the account.
- Involved in resource allocation in various projects.
- Managing and approving the team timesheet and issues related to billing, invoices, budget and P&L.
- Involved in development of script using Scala and Python in Azure Databricks, Azure Synapse.
- Involved in creation of Data pipelines on Azure and implementing CI-CD(DevOps) on Azure platform.
- Involved in estimation of efforts for various projects.
- Mentored the team member technically on Azure Databricks, Azure Synapse, Azure Data Factory, Spark, Scala, Data Visualization and Data Techniques.
- Meeting with various stakeholders of the project to discuss and close the outstanding issues and to discuss the project status.

Project # 6

Project Title : IVCI
 Environment : Talend, AWS, Snowflake
 Role : Sr. Talend Data Engineer

Description:

Enterprise Data Hub (EDH) solution is an Enterprise-Wide Data Store Solution for data integration using Talend to extract data from various source system into different zones in Azure Data Lake Storage and finally to the Azure Synapse. There are 2 zones (containers) in the Azure Data Lake Storage, namely Raw and Enrich. The Raw zone stores data in its raw form. While Enrich (Enrich Raw) is an enriched raw data that has been undergo 'Enterprise Standardization & Quality Checking'. This will also help to enable data quality reporting based on the profiling of data attributes relevant to the EDH provide key input into metrics and reports

Roles and Responsibilities:

- Interacted with Business Analysis team in gathering the requirements and created Mapping documents.
- Involved in Testing the ETL process that meets the business requirement.
- Involved in Production deployment support and KT.
- Ability to meet deadlines and handle multiple tasks.
- Implemented CI/CD pipeline in current project with following tools
 - Azure DevOps
 - Maven
 - Nexus Artifacts Repository
- Actively monitor, research and analyze ways in which the services in AWS can be improved.
- Involved in Testing the ETL process that meets the business requirement.
- Involved in Production deployment support and KT.
- Experience in importing Azure Synapse Relational Database Systems and vice-versa.
- Extensively used Talend Administration Centre (TAC) for publishing jobs.
- Good hands-on TAC scheduling jobs scripts, monitoring and writing triggers to schedule the jobs
- Involved in all phases of agile board to create sprints and perform retro specification meeting from requirement gathering, design, development, testing, Production, user training and support for production environment.

Project # 5

Project Title : Family Dollar
Environment : Talend, Oracle
Role : Sr. Talend Developer

Description:

Family Dollar is an American variety store chain. With over 8,000 locations in all states except Alaska, Hawaii, Oregon and Washington, it is the second largest retailer of its type in the United States. Family Dollar offers a compelling mix of merchandise for the whole family. Ranging from an expanded assortment of refrigerated and frozen foods and health and beauty items to home décor and seasonal items, Family Dollar offers the lowest possible price, the name brand and quality private-brand merchandise customers need and use every day.

Roles and Responsibilities:

- Interacted with Business Analysis team in gathering the requirements and created Functional documents.
- Involved in migrating Data stage jobs to Talend
- Involved in migrating Talend on prem to Talend cloud (AWS).
- Involved in configuring AWS services EC2, S3, Redshift as per business needs.
- Involved in building lambda functions with python and deployed.
- Actively monitor, research and analyze ways in which the services in AWS can be improved.
- Involved in Testing the ETL process that meets the business requirement.
- Involved in Production deployment support and KT.
- Experience in importing and exporting data using Sqoop from HDFS to Relational Database Systems and vice-versa.
- Creating Hive tables and working on them using Hive QL.
- Extensively used Talend Management Console (TMC) for publishing jobs.
- Extensively used UC4 for scheduling the jobs.
- Ability to meet deadlines and handle multiple tasks.
- Involved in all phases of SDLC from requirement gathering, design, development, testing, Production, user training and support for production environment.

Project # 4

Project Title : Nationwide
Environment : Talend Big Data platform, AWS, Hive, HDFS
Role : Talend Developer

Description:

Nationwide Mutual Insurance Company and affiliated companies is a group of large U.S. insurance and financial services companies based in Columbus, OH. The company also operates regional headquarters in Scottsdale, AZ; Des Moines, IA; San Antonio, TX; Gainesville, FL; Raleigh, NC; Sacramento, CA, and Westerville, OH

Nationwide Financial Services (NFS), a component of the group, was partially floated on the New York Stock Exchange prior to being repurchased by Nationwide Mutual

Roles and Responsibilities:

- Worked with Business Analysis team in gathering the requirements and created Functional and ETL specification documents.

- Analyzed the Business Requirement Documents (BRD) and laid out the steps for the data extraction, business logic implementation & loading into targets.
- Need to collect the requirement and do the analysis and give the estimation ETA to the business users.
- Need to analyze the STM (source to target mapping) from data analyst team and do the validation to start the mapping in Talend.
- Involved into end-to-end Development, Implementation & production support process from dev to prod environment.
- Involved in Testing the ETL process that meets the business requirement.
- Developed Talend jobs to populate the claims data to data warehouse - star schema.
- Worked with different sources such as Relational DB and flat files.
- Extracting, Transforming and Loading the data from Source to Staging and Staging to Target according to the Business requirements.
- Ingestion of huge files from various sources into S3 bucket and processing the data
- Involved in all phases of SDLC from requirement gathering, design, development, testing, Production, user training and support for production environment.
- Developed the mappings using needed Transformations in Talend tool according to technical specifications.
- Ability to meet deadlines and handle multiple tasks.
- Extensively used Talend Administrative Console (TAC) tool for Scheduling and monitoring.
- Performance tuning in PL/SQL

Project # 3

Project Title : ArjoHuntleigh
 Environment : Informatica, Oracle
 Role : Informatica Developer

Description:

ArjoHuntleigh is a leading global provider of medical technology products and solutions for patients with reduced mobility and related conditions due to for example age, illness, overweight or in post-surgery situations. This is migration project from ERP system to M3. ETL tool Informatica is used in between for data translation. Data from ERP system is extracted into flat file (DAT). Informatica are used to load the data in staging area, use mapping rule and load the data in work table, validate and load the data in target and reject table.

Roles & Responsibilities:

- Interacting with Business analysis in understanding the Database requirements whenever debugging is required without changing the code.
- Create and maintain Informatica jobs to transform and load data into DWH tables.
- Create and maintain Informatica jobs to extract data from Oracle Tables and load into Target fact tables used for reporting.
- Extensively worked with Lookup Transformation, Update Strategy Transformation, Expression Transformation, Transaction Control and Router, Source Qualifier and Mapplate
- Creating Mapplets, Parameter & Variables in mappings.
- Extensively worked with delimited files.
- Create and maintain Mapping and Documentation.
- Implemented Change Data Capture (CDC) techniques like Slowly Changing Dimensions.
- Preparing Unit Test Case for mappings.
- Create and maintain Mapping and Documentation.

- Involved in Performance Tuning.
- Done the Unit testing and UAT for the developed jobs for different iterations
- Involved in Production deployment support and KT.
- Worked on unit testing and performance tuning of the Jobs.
- Writing Store procedure, cursor in PL/SQL

Project # 2

Project Title : Medical Health Enterprise
 Environment : Informatica, Oracle
 Role : ETL Developer

Description:

This project is for health services named as Alaska Medicaid Health Enterprise for US government. Which is providing web enabled solution for administrating all the medical assistant program. It also has self-service features so users can access the system through the user-friendly ways. The company requires the different level of analysis regarding the health insurance in different states. These aspects will play the major role for taken the better decision for Alaska Medicaid Health Enterprise to find out the analytical study that in which ratio the people owning the health insurance

Roles & Responsibilities:

- Create and maintain Informatica jobs to transform and load data into DWH tables.
- Create and maintain Informatica jobs to extract data from Oracle Tables and load into Target fact tables used for reporting.
- Extensively worked with Lookup Transformation, Update Strategy Transformation, Expression Transformation, Transaction Control and Router, Source Qualifier and Mapplate`
- Creating Mapplets, Parameter & Variables in mappings.
- Extensively worked with delimited files.
- Create and maintain Mapping and Documentation.
- Implemented Change Data Capture (CDC) techniques like Slowly Changing Dimensions.
- Preparing Unit Test Case for mappings.
- Create and maintain Mapping and Documentation.
- Involved in Performance Tuning.
- Done the Unit testing and UAT for the developed jobs for different iterations
- Creating Mapplets, Parameter & Variables in mappings.

Project # 1

Project Title : EDW
 Environment : Informatica, Oracle
 Role : ETL Developer

Description:

This Project is meant for the analysis of data for the business requirement of client. This project is involved in the Creation of Data Warehouse for my client who is one of the lead financial companies across the country. There are different types of data marts such as Customer, Loans, Bank, Insurance and real estate itself etc. Company's Management to decide the type of Potential Customers who are beneficial to the company and those who have an amount of risk associated with them.

Roles & Responsibilities:

- Create and maintain Informatica jobs to transform and load data into DWH tables.

- Create and maintain Informatica jobs to extract data from Oracle Tables and load into Target fact tables used for reporting.
- Extensively worked with Lookup Transformation, Update Strategy Transformation, Expression Transformation, Transaction Control and Router, Source Qualifier and Mapplet
- Creating Mapplets, Parameter & Variables in mappings.
- Extensively worked with delimited files.
- Implemented Change Data Capture (CDC) techniques like Slowly Changing Dimensions.
- Preparing Unit Test Case for mappings.
- Create and maintain Mapping and Documentation.
- Involved in Performance Tuning.
- Done the Unit testing and UAT for the developed jobs for different iterations
- Writing Store procedure, cursor, Function in PL/SQL

Academic Qualification

- Master of Computer Application.