**Resume**

**Rajaputra Pavan Kumar Singh**

**E-mail:**[pavankumarsingh.raj@gmail.com](mailto:rpavank1982@gmail.com)  **Mobile: +1-5132521087**

**EXPERIENCE SUMMARY**

* 16+ years of experience in Data Engineering & Analytics as Senior ETL Developer / Data Engineer.
* Experience in design and development of Data Analytics solutions using Informatica IDMC (IICS), Snowflake, Teradata and Informatica PowerCenter.
* Experience working with cloud-based database solutions including Snowflake.
* Designed, Developed, and Implemented ETL processes using IDMC(IICS) integration.
* Extensively used cloud transformations – Aggregator, Expression, Filter, Joiner, Routers, Union, Unconnected / Connected Lookups, Update Strategy
* Experience in working on Agile driven projects.
* Experience with CI/CD using Jenkins.
* Performed Loads into Snowflake instance using Snowflake connector in IICS to support data analytics.
* Bulk loading from the external stage (AWS), internal stage to snowflake cloud using the COPY command.
* Experience with applying Performance Optimization techniques for Data Analytics workloads.
* Experience in building Data Pipelines
* Experience in Migration from Teradata to Snowflake
* Experience in Migration from PowerCenter to IDMC (IICS)
* Loading data into snowflake tables from the internal stage using snowsql.
* Used COPY, LIST, PUT and GET commands for validating the internal stage files.
* Used import and Export from the internal stage (snowflake) from the external stage (AWS).
* Writing complex snowsql scripts in snowflake cloud data warehouse for business analysis and reporting.
* Used FLATTEN table function to produce a lateral view of VARIANT, OBJECT, and ARRAY column.
* Performed data quality issue analysis using Snow SQL by building analytical warehouses on Snowflake.
* Building Datawarehouse applications using Snowflake (AWS).
* Good Experience in core AWS services (S3, EC2, ELB, VPC, Auto scaling, ECS etc.) and deployment services.
* Experienced in performing Incremental Loads and Data cleaning in Informatica. Expert in handling errors, logging.
* Worked on Push Down Optimization, Partitioning, Versioning at Informatica.
* Worked with XML Sources and Targets at Informatica
* Worked on Performance Tuning techniques while loading into Snowflake using IICS.
* Analysis for cause identification and tuning measures like change PIs, additional SI, collect stats, apply MVC, rewrite queries, identify ETL process bottlenecks, design changes etc.
* In-depth knowledge and understanding of dimensional modelling (Star schema and Snowflake schema, SCD types- 1, 2, 3), and data modelling (star schema) at logical and physical level.
* Worked on Shell scripts at UNIX environment.
* Experience in handling Production Releases.
* Excellent problem-solving skills with strong technical background and good inter-personal skills.

**EDUCATION**

* Master of Business Administration in Marketing & Systems from Osmania University, 2005
* Bachelor of Science in Computers from Sri Venkateswara University, 2003

**CERTIFICATIONS**

* SnowPro Core
* Teradata 12 Certified Professional

**WORK EXPERIENCE**

1. **BetMGM, NJ, USA**  **From** 12/2019 To Till Date

**Role**: Senior Software Engineer

**Responsibilities**:

1. Analyzed existing ETL Datawarehouse process and created designed specification based on new target cloud Datawarehouse.
2. Redesigned the existing Informatica Power Center logic to IICS with target as Snowflake.
3. Experience integrating data using IICS for reporting needs.
4. Connected to different sources like files in google cloud and Relational databases.
5. Used SNOW PIPE for continuous data ingestion from the AWS bucket.
6. Developed complex Informatica Cloud Task flows(parallel) with multiple mapping tasks and workflows.
7. Cloned Production data for code modifications and testing.
8. Redesigned existing Teradata procedures to Snowflake procedures.
9. Created clone objects to maintain zero-copy cloning.
10. Performed troubleshooting analysis and resolution of critical issues.
11. Designed and developed scalable and efficient data pipelines, ETL processes, and data integration solutions.
12. Involved in loading the raw data and processing it through Landing, Sanitized, and Curated/Presentation layers.
13. Developed and integrated various data pipelines and reduced data processing time by 40%, allowing for quick analysis and decision-making.
14. Written complex joins to process the ingested data in data lakes efficiently and reduce the end-to-end processing time by 20%.
15. Investigated various financial report discrepancies raised by the stake holders due to incorrect business logic/in-correct data/missing key data and resolved most of the critical issues in less than a day and met the deadlines.

**Technologies used**: IICS (IDMC), Power Center, Snowflake, Maven, GITHub, Teradata, Docker.

1. **KOHL’S Inc, WI, USA**  **From** 06/2018 To 12/2019

**Role**: Senior Technical Lead

**Responsibilities**:

Implemented Type 1 and Type 2 Dimensions ETL logic using Airflow and Snowflake.

Interacted with Client, BA team, ST team for the different project related activities.

Implemented Snowflake data warehouse for a client, resulting in a 30% increase in query performance.

Migrated on-premises data to Snowflake, reducing query time by 50%.

Built and optimized ETL processes to load data into Snowflake, reducing load time by 40%.

Designed and developed a real-time data pipeline using Snowpipe to load data from Kafka with 99.99% reliability.

Leveraged Sqoop to import data from RDBMS into HDFS

Involved in code migration to QA for testing and till production.

Interacted with the Business users to resolve the user incident tickets.

1. Generated reports to compare the amounts based on various data fields using window and ranking functions.
2. Developed various controls to ensure record count and amounts match between the Landing, Sanitized, and Curated/Presentation layers and trigger an email when there is a discrepancy or a failure.
3. Built the scripts which will archive stored data and DB tables data to meet the compliance and regulatory requirements.
4. Designed airflow DAGs using python scripts.
5. Distributing a large dataset efficiently across multiple nodes in a Spark cluster for parallel processing.
6. Optimized Spark jobs for faster data processing.
7. Worked on the data frame transformations and actions in Spark.
8. Developed ETL framework using Python (including daily runs, error handling, and logging) to clean data.
9. Developed data flows to interact with various APIs and integrate the data.

Co-ordinate with ML engineers and Data Scientists and provide quality data and resolve any discrepancies.

**Technologies used**: Informatica, Cloud Composer (Airflow), Snowflake, HDFS, Hive.

1. **DBS, USA**  **From** 11/2015 To 06/2018

**Role**: Senior Software Developer

**Responsibilities**:

Understanding the project architecture, ETL framework, standards and processes.

Developed the Informatica sessions, tasks and workflows to extract data from Oracle

database and flat files sources and load into warehouse tables.

Extensively worked in tuning the sessions to optimize the performance of Informatica objects.

Used PMCMD commands to start workflows from Unix server.

Interacting with the Business users to resolve the user incident tickets.

Investigating and solving the data quality and data mismatch issues between

Upstream-EDW-Downstream systems by providing the root cause analysis.

Analyzed the Business requirements to be converted into the ETL loads from Source to Stage layer and then to Semantic Layer.

Interaction with Client, BA team, ST team for the different project related activities.

Extensively worked with   Informatica objects using Push down Optimization.

Developed the mappings using xml Targets, Xml sources.

Extensively worked on External loaders (Fast Load, Multi Load, Fast Export and Tpump)

Extensively worked on the data extracting with database and flat files using informatica

Create workflows using event wait and event raise objects to provide the dependencies.

Prepare test cases for unit testing.

Involved in requirement analysis, design, and performance analysis for the project.

Designing Standards, naming Conventions for ETL Code development.

**Technologies used**: Informatica, Teradata.

1. **VOLVO, USA**  **From** 04/2012 To 10/2015

**Role**: Technical Consultant

**Responsibilities**:

Understanding requirements, design of solution and preparation of Design documents.

Involved in preparing the knowledge sharing and handover documents.

Involved in production support activities like monitoring the daily jobs.

Proactively involved in creating scripts for automating a few processes.

Analyze logs, scripts, BTEQs, high volume tables, database performance and identify the performance tuning opportunities.

Analysis for cause identification and tuning measures like change PIs, additional SI, collect stats, apply MVC, rewrite queries, identify ETL process bottlenecks, design changes etc.

Modify and unit test existing code/ database objects to apply tuning measures.

Set up parallel environments for existing/baseline and modified code/database objects.

Execution of both processes and compare the results for accuracy and Total IO, CPU Time for improvement.

Created TPT scripts for loading high volume data into tables.

**Technologies used**: Informatica, Teradata.

1. **JPMC, USA**  **From** 10/2009 To 04/2012

**Role**: ETL Consultant

**Responsibilities**:

Coordinating with Onsite Team and client for Requirements gathering and analysis.

Creating and Reviewing of Macro Design and Micro Design Documents based on the JPMC ICDW solution outline document.

Participating in peer review of Informatica objects

Coding & testing the Informatica Objects as per client’s BI standards.

Developed Informatica mappings, mapplets using various transformations like Expression, Router, Update strategy etc.

Involved in performance tuning of the mappings using pushdown Optimization.

Involved in Unit Testing, Integration Testing and System Testing.

**Technologies used**: Informatica, Teradata.

1. **Credit Suisse, USA**  **From** 01/2008 To 10/2009

**Role**: Sr Software Engineer

**Responsibilities**:

Interacting with the client for enhancements and changes

1. Involved in writing High level design (HLD) and low-level design (LLD)
2. Developed Informatica mappings Type-I & Type-II SCDs using various transformations.
3. Developed mapping for Fact loading and implemented error handling strategy.
4. Prepared the required Unit test cases for assigned tasks to self-review the test cases.

Involved in Control–M Desktop Setup, Configuration and Monitoring the Control M Jobs.

**Technologies used**: Informatica, Teradata, Control-M.