**Vidyasagar Reddy Jonnala**

**Data Architect**

**vsrjonnala@outlook.com**

**+1-945-227-8365**

**Overview**

* Throughout my extensive career spanning over 14 years, I have consistently demonstrated excellence in roles including Data Architect, Data Engineer, and Data Analyst. My primary focus has been on designing and building robust ETL integration/migration frameworks, implementing pattern-based integrations, and delivering efficient data warehousing, data management, and data visualization solutions. Notably, I have successfully led projects across a wide range of industries including retail, e-services, F&B, manufacturing, and finance.

**Skills & Capability**

* Designing conceptual, logical and physical models for the target system is a critical component of the job.
* Expertise in designing ETL Data Migration (Cloud &On-Prem) & Integration frameworks and Proficient in building complex transformation rules & Data Cleansing rules.
* Extensive experience in design and maintain of Enterprise Data Warehouse / data marts.
* Solid understanding of Data Warehousing, Data marts, Dimensional Modeling, Star, Snowflake schema, OLTP and OLAP concepts.
* Drawing on Experience in all aspects of analytics/data warehousing solutions (Data Modeling, Data mapping, ETL Development, metadata management, data migration and Analytical solutions) I have been keying in delivering innovative database /data warehousing solutions to the clients.
* Proficient in building complex transformation rules, Data Quality (DQ), Master Data Management, Data Profiling, Data Governance, Data Lineage, Cleansing, and Data Conversion techniques.
* Experience in developing of Dashboards and Reporting solutions for the senior leadership team and the business in achieving better service quality, workforce management and business process optimizations.
* Strong Knowledge in AWS, GCP data components.
* Hands on Experience in various ETL tools like DataStage, Informatica, Pentaho, Talend, SSIS, DBT and Cloud based data integration tools.
* Hands on experience in Relational databases, Teradata, Snowflake, Vertica, No SQL database, HDFS and Cloud-based data storage.
* Experience in Data migration projects, especially data migration to CRM applications like Salesforce
* Hands on experience in various BI tools like Tableau, QlikView, PowerBI, Pentaho BI and Business Objects.
* Hands on experience on various Reporting tools like SSRS, Crystal Reports, Pentaho Report Designer, BIRT Reports and TIBCO.
* Hands on Experience in QLIK, DataBricks, Snowflake, Python, Oracle ERP,SAP ERP, Git, Pyspark, Spark SQL, Unix, SaaS and Shell scripting.
* Strong Knowledge on Net Suite, Matillion and Fivetran
* Excellent technical, analytical, and problem-solving skills. Strict attention to detail and ability to work independently and within a team environment.

**Academic and Professional Qualifications**

* Master of Computer Applications (M.C.A) from Anna university, India - 2010
* Bachelor's in computer science (B.sc Computer Science) from Acharya Nagarjuna University, India. - 2007

**Certifications**

* Azure Data Engineer Associate
* Hands on Essentials Snowflake
* Hands on Essentials DBT

**Experience Summary:**

**Data Capital INC /Wellbore Integrity Solutions-Houston-Onsite** **Mar 2024 - Till Date**

**Data Architect**

* Data Architecture Development: Develop and maintain data architecture for supply chain, manufacturing, and finance management using Snowflake, Oracle, NetSuite ERP, CRM, Databricks, and AWS Glue, ensuring data governance and adherence to industry standards.
* Data Modeling: Design and implement data models for OLTP and OLAP solutions, leveraging Snowflake for warehousing, Snowpipe for real-time ingestion, and NetSuite ERP to optimize processes in supply chain, manufacturing, and finance.
* ETL & Data Integration: Utilize Snowflake’s data sharing,AWS,Lambda, AWS Glue, Databricks, and dbt for ETL processes, integrating data from multiple sources into a unified platform tailored for supply chain, manufacturing, and finance efficiency.
* Migration Leadership: Lead the migration of SQL Server, Oracle, and NetSuite ERP to Snowflake and Databricks, managing database access control and optimizing data processing for supply chain, manufacturing, and finance systems.
* Real-time Data Processing: Implement real-time data processing solutions using Snowflake’s integration capabilities, Snowpipe, and Databricks, ensuring timely insights into supply chain, manufacturing, and financial data.
* Scalable Architecture: Design scalable architectures with performance optimization for growing data volumes in supply chain, manufacturing, and finance, utilizing Snowflake, Databricks, Oracle, and NetSuite ERP.
* Data Governance & Security: Ensure compliance with data security and privacy regulations across Snowflake, Databricks, and NetSuite, implementing robust governance policies for supply chain, manufacturing, and finance data protection.
* Advanced Analytics Solutions: Collaborate with cross-functional teams to deliver analytics solutions using Power BI, Databricks, and machine learning models, optimizing operations in supply chain, manufacturing, and finance while providing key insights to stakeholders.

**Technologies:** Snowflake, AWS Glue,DBT, SQL Server, Oracle ERP,Net Suite ERP, GEMS, FMS, S3 and Power BI

**Data Capital INC/ Stanley Black & Decker -India-US Onsite Apr 2023 - Feb 2024**

**Responsibility: Data Architect**

* Lead Data Architecture: Develop and implement the overall architecture for managing supply chain, manufacturing, and inventory data within Snowflake and Databricks, focusing on data modeling, data sharing, and governance for optimal data storage and processing.
* Data Integration: Integrate data from sources such as production lines, inventory systems, and logistics platforms into a unified platform using Snowflake, Databricks, dbt, and Snowpipe for data ingestion and transformation.
* SQL Database Migration: Spearhead the migration of SQL databases to Snowflake for supply chain, manufacturing, and inventory data, managing access control and governance during the transition from legacy systems.
* ETL & Data Pipelines: Utilize dbt, ADF,AWS Glue, Databricks, and PySpark for ETL processes to ensure data quality and build scalable, maintainable pipelines for manufacturing, supply chain, and inventory data.
* Data Sharing & Governance: Implement governance policies and best practices for data sharing within Snowflake and Databricks, ensuring data security and privacy compliance for supply chain, manufacturing, and inventory operations.
* To implement and optimize data models for Salesforce and Microsoft Dynamics while integrating data from various systems, you need to follow best practices focused on performance, scalability, reliability, and data integrity.
* Performance Optimization: Design scalable architecture to handle growing data volumes in supply chain, manufacturing, and inventory, utilizing performance optimization techniques with Snowflake, Databricks, and PySpark.
* Analytics Solutions: Develop analytics solutions for production efficiency, supply chain performance, and inventory management using Snowflake and Databricks for warehousing and transformation.
* Collaboration & Reporting: Collaborate with cross-functional teams to develop Power BI dashboards using Snowflake and Databricks, ensuring data governance and monitoring across supply chain, manufacturing, and inventory workflows.

**Technologies:** Azure Data Factory,AWS GLUE Azure Data Lake Gen2, Blob storage, Databricks, Snowflake, DBT, S3, Pyspark, Salesforce, Oracle, SQL Server, Azure SQL, ERWIN, Power BI and Tableau

**Reliable Business Technologies Sdn Bhd/Teras Malaysia · On-site Sep 2022 - Mar 2023**

**Data Architect/Data Engineer**

**Project : PLUS, Unified Data Platform**

* Collaborate with various business teams and stakeholders to design and implement BI solutions on Azure using services like Azure Data Lake, Azure Data Factory, Data Lake Analytics, Stream Analytics, Cosmos DB, Blob Storage, and Databricks, ensuring alignment with business goals.
* Design, develop, and maintain data integration pipelines to collect data from sources such as toll booths, parking facilities, and public transport systems, utilizing Databricks for advanced data processing and machine learning tasks.
* Implement real-time data processing solutions to handle continuous data streams, including toll booth transactions and parking spot availability, ensuring data is stored and processed efficiently with Azure and Snowflake.
* Utilize Databricks and Azure Data Factory (ADF) for advanced data processing and machine learning, ensuring data is efficiently transformed and analyzed for valuable business insights. Develop ETL processes using ADF and Databricks to convert and cleanse raw data, preparing it for structured analysis and reporting.
* Use Databricks to create scalable data pipelines that can handle large volumes of data, ensuring high performance and reliability for both batch and real-time processing.
* Design and create data models using Erwin that represent the structure and relationships of data entities, ensuring compliance with data security and privacy regulations within Azure, Databricks, and Snowflake environments.
* Collaborate with data analysts and visualization experts to develop Power BI dashboards and reports, providing insights into transaction data and public transport usage, using Azure services, Databricks, and dbt transformations.
* Monitor and optimize the performance of data pipelines and databases to ensure high throughput and low latency for both real-time and batch processing, leveraging Databricks and Azure Monitor.
* Implement security measures and compliance standards within Azure, Databricks, and Snowflake to ensure the confidentiality, integrity, and availability of sensitive data, particularly for financial transactions.
* Respond to data-related issues, troubleshoot problems, and ensure the high availability and reliability of mission-critical data solutions, particularly within the toll collection and public transport systems.

**Technologies:** ADF, Databricks, DBT, Azure Data Lake Gen2, Blob storage, Snowflake, Cosmos DB, Azure SQL, MySQL, Oracle, Pyspark, Power BI and Erwin

**Reliable Business Technologies Sdn Bhd /Telecom Malaysia- Malaysia · On-site** **Jan 2021 - Aug 2022**

**Data Architect/Data Engineer**

* Design and implement data pipelines to ingest, process, and transform CDR data from various sources into Azure data storage solutions (Azure Data Lake, Blob Storage, Azure SQL, Azure SQL Data Warehouse).
* Design and implemented data pipelines to ingest data from various sources, including network equipment, customer databases, and external sources, into Azure data storage solutions.
* Used Azure services like Azure Data Factory and Azure Event Hubs to facilitate data ingestion.
* Developed Pentaho ETL (Extract, Transform, Load) processes to cleanse, enrich, and transform CDR data into a usable format for analytics and reporting.
* Choose and manage appropriate data storage solutions on Azure, considering factors like cost, performance, and scalability. This may include Azure SQL Database, Azure Cosmos DB and Azure Synapse Analytics.
* Implemented data security and access controls to ensure that CDR data is protected and compliant with data privacy regulations.
* Monitor and optimize the performance of data pipelines and databases to ensure efficient data processing and reporting.
* Design and maintain data models and data warehouses using Azure Synapse Analytics.
* Created data models and schemas that support efficient querying and reporting for telecom CDR data.
* Implement data quality checks and validation processes to identify and rectify data anomalies and discrepancies.
* Implement automation and orchestration techniques to schedule, monitor, and manage data pipelines and workflows.
* Collaborated with data analysts and data scientists to create Power BI dashboards and reports that provide insights into CDR data.
* Collaborate with cross-functional teams, including data scientists, analysts, and business stakeholders, to understand their data requirements and deliver solutions accordingly.

**Technologies:** Azure Data Factory, Pentaho, Azure Synapse Analytics, Data Lake Gen2, Cosmos DB, Azure SQL, MySQL, Oracle, Pyspark, Power BI, SQL Server, Oracle

**Reliable Business Technologies Sdn Bhd /JKR -- Malaysia · On-site** **Sep 2019 - Dec 2020**

**Data Architect**

* Collaborated with end users to identify analytical KPIs and documented wire-framed dashboards based on their requirements.
* Utilized Erwin for designing data models and ensuring data integrity and relationships within the data warehouse architecture.
* Utilized Pentaho Data Integration and Talend to migrate data from legacy systems (Postgres) to the target database (Oracle) and designed a data mart for data visualization.
* Developed data models and centralized storage to facilitate efficient data retrieval and analysis, ensuring optimized performance in querying and reporting.
* Led the migration of historical and unstructured data, such as design documents, notes, construction videos, and pictures, into HDFS (Private Cloud) for seamless integration into the analytical system.
* Used Pentaho Data Integration and Talend for migrating data from legacy systems (Postgres) to target databases (Oracle) and designing data marts.
* Experienced in working with unstructured data migration, including design documents, notes, construction videos, and pictures into HDFS.
* Designed and implemented a new data warehouse to fulfill business requirements and support analytical needs.
* Used Pentaho BI for developing dashboards and visualizations, incorporating analytical KPIs and wire-framed dashboards based on end-user requirements. Proficient in integrating Google Maps for consolidated views based on geographical locations.
* Used Pentaho Report Designer for real-time reports, providing stakeholders with valuable insights and facilitating informed decision-making.
* Worked in both hybrid cloud and on-premises environments, ensuring effective data management and security.

**Technologies:** Pentaho Data Integration, HDFS(Cloud), Data Lake, Oracle, PostgreSQL, MySQL, Pentaho BI and Pentaho Report Designer.

**Reliable Business Technologies Sdn Bhd / MMDIS, Malaysia · On-site** **Oct 2018 - Aug 2019**

**Data Engineer for JLM**

* Played a key role in implementing JLM, an ERP system focused on registering ships in Malaysian Ports, encompassing modules such as Ship Registration, Ship Clearance, Target Positions, Document Management, and Human Resource.
* Actively participated in client meetings to gather requirements, ensuring a thorough understanding of their needs and expectations.
* Analyzed requirements and formulated Key Performance Indicators (KPIs) for data migration and visualization, aligning them with business objectives.
* Prepared mapping documents for data migration, ensuring accurate and seamless transfer of data from legacy systems (MySQL and MSSQL) to the target database (MySQL) using Pentaho.
* Collaborated with cross-functional teams to gather data requirements and designed data models to meet business needs.
* Ensured data quality and integrity through data validation techniques and data cleansing procedures.
* Developed migration scripts and designed the DataMart once data migration was successfully completed, enabling efficient data storage and retrieval.
* Utilized Pentaho Business Intelligence to create interactive dashboards and generate reports, providing stakeholders with valuable insights and facilitating informed decision-making.

**Technologies:** Pentaho Data Integration, Pentaho business intelligence, MySQL, SQL Server

**MYEG Services- Kuala Lumpur, Malaysia Dec 2016 - Sep 2018**

**Data Engineer - MYEG Data Analysis**

* Collaborated closely with cross-functional teams, including IT, business analysts, and management, to understand data requirements and ensure the accuracy and relevance of the implemented solution.
* Implemented a star-schema data warehousing approach to support comprehensive data analysis across all MYEG services, customers, and transactions.
* Developed a robust data integration solution using Pentaho to efficiently extract, transform, and load data from multiple production databases into the data mart.
* Conducted thorough testing and performance optimization to ensure the efficient and reliable functioning of the data warehousing system.
* Utilized identifier keys within the data warehousing system to optimize data storage and retrieval, minimizing storage space requirements.
* Leveraged QlikView and Crystal reports to create interactive and visually appealing dashboards, reports enabling stakeholders to easily explore and analyze the data.

**Technologies:** Pentaho Data Integration, SAP HANA, PostgreSQL, QlikView, and Crystal Reports

NEXGEN APS SDN BHD/**Coles Myer** **Sep 2015- Nov 2016**

**Migration Consultant**

* Use SQL Server Integration Services (SSIS) to design and develop data integration processes that extract, transform, and load (ETL) sales data into a SQL Server database.
* Maintain and optimize the SQL Server database to ensure data integrity, performance, and availability.
* Implement an Enterprise Data Warehouse to centralize and store sales data for decision support purposes. Ensure that the EDW supports data quality and consistency.
* Set up processes to provide data access and updates, enabling stakeholders to make decisions based on the most current information.
* Implement data quality checks and validation procedures to minimize errors and ensure data accuracy and reliability.
* Created visually engaging data visualizations that help users quickly understand trends, patterns, and insights from sales data.
* Utilize tools such as QlikView and Crystal Reports to design and create interactive and visually appealing dashboards and reports for sales and business stakeholders.
* Enable stakeholders to explore and analyze sales data through user-friendly dashboards, allowing them to make data-driven decisions.

**Technologies:** SSIS, MySQL, Oracle, SQL Server, SSRS and QlikView.

**Migration Consultant and Data Analytics Dec 2010 - May 2015**

**Systel · Full-time (with various clients)**

**Capgemini /Schneider Electrics**  **Apr 2014 - May 2015**

**Migration Consultant**

* Design, develop, and maintain ETL (Extract, Transform, Load) processes using Informatica PowerCenter or other related tools.
* Create and optimize data integration workflows and mappings.
* Collaborate with business analysts and data analysts to understand data requirements.
* Debug and troubleshoot ETL processes to ensure data accuracy and reliability. Perform performance tuning and optimization of ETL processes.
* Have imported data from applications such as NetSuite to build data warehouses to serve end reporting needs used for tracking subscription sales, invoices, subscription expiries, renewals and open/closed opportunities.
* Conducted testing of data load results and error log status in RDBMS Oracle using SQL queries, ensuring data accuracy and integrity.
* Utilized Informatics power center to extract data from source transactional databases, perform transformations, and load it into the target system.
* Played a role in performance tuning of Informatics jobs, identifying and resolving performance issues to optimize job execution.
* Conducted Unit, Systems Integration testing, and User Acceptance Testing (UAT) to ensure the successful implementation of migration projects.

**Technologies:** Informatica Power Center, Oracle, PL/SQL, SAP ECC

**Capgemini/Target**  **Jan 2012- Apr 2014**

**Migration Consultant**

* Analyzed existing jobs and successfully migrated them to the AIX server, adhering to development standards and best practices.
* Developed Server and Parallel jobs utilizing Oracle, ODBC, FTP, Pivot, Hashed file, Aggregator, STP, Filter, Funnel, Copy, Change Capture, Merge, Lookup, Join, Sort, and other stages within IBM Infosphere DataStage.
* Employed Quality Stage stages such as Standardize, Match, Investigate, and Survive to ensure data validation in accordance with business requirements.
* Conducted thorough unit testing and optimized performance by fine-tuning data warehouse tables, utilizing IBM Infosphere DataStage Director for job monitoring and troubleshooting.
* Created the data warehouse repository using IBM Infosphere DataStage Manager, importing source and target database schemas.
* Implemented shared containers, enabling their reuse in multiple jobs to promote efficiency and consistency.
* Utilized IBM Infosphere DataStage Job Sequencer extensively for scheduling jobs and ensuring proper job sequencing.
* Conducted unit testing and facilitated deployments to ensure successful project implementations.
* Collaborated with cross-functional teams to identify and resolve script-related issues, ensuring smooth operation of scheduled jobs and monitoring processes.
* Assisted in issue tracking, monitoring, and reporting throughout the migration process.

**Technologies:** DataStage, Oracle, PL/SQL, DB2, SAP ECC, SAP BO,Cognos, Linux, Shell Scripting

**Capgemini/Target**  **Dec 2010 - Dec 2011**

**Migration Consultant**

* Developed Server and Parallel jobs utilizing Oracle, ODBC, FTP, Pivot, Hashed file, Aggregator, STP, Filter, Funnel, Copy, Change Capture, Merge, Lookup, Join, Sort, and other stages within IBM Infosphere DataStage.
* Employed Quality Stage stages such as Standardize, Match, Investigate, and Survive to ensure data validation in accordance with business requirements.
* Conducted thorough unit testing and optimized performance by fine-tuning data warehouse
* tables, utilizing IBM Infosphere DataStage Director for job monitoring and troubleshooting.
* Created the data warehouse repository using IBM Infosphere DataStage Manager, importing source and target database schemas.
* Implemented shared containers, enabling their reuse in multiple jobs to promote efficiency and consistency.
* Utilized IBM Infosphere DataStage Job Sequencer extensively for scheduling jobs and ensuring proper job sequencing.
* Conducted unit testing and facilitated deployments to ensure successful project implementations.
* Assisted in issue tracking, monitoring, and reporting throughout the migration process.
* Implemented and maintained shell scripts for job scheduling and automation, ensuring timely execution of critical tasks and reducing manual intervention.

**Technologies:** DataStage, Oracle, DB2, Linux, PL/SQL, Shell Scripting