**Sree R. Pinnamaneni**

**Professional Summary:**

A self-driven and organized data enthusiast with around **20 years** of IT experience, including **15 years** working with **Enterprise Data Warehouse** systems in distributed computing model. Experienced in solutioning & implementing in various areas of enterprise data lifecycle & mgmt. Have been working with MPP systems like Teradata, Netezza to modern distributed computing systems in Big Data and Cloud arena like Hadoop/Hive, S3, Snowflake etc. Experienced in leveraging modern storage formats/methods & building/managing data pipelines in distributed processing frameworks like Snowflake warehouse, Teradata, Spark, Hadoop. Worked in Platform assessment, workload analysis, Data Mgmt. Strategy, Data Collection & Engineering, On-prem to Cloud Migration, Performance Tuning, Data Governance etc. Ability to compare/contrast different toolset and understand the architectural tradeoffs. Team player with excellent work ethics, analytical skills coupled with good business communication/interpersonal skills. Committed to learn new skills along the way & willing to coach the junior team members.

* Experienced in Data Mgmt. & Engg. on MPP platforms like **Snowflake,** **Redshift, Databricks,** **Teradata, Netezza, S3, Hadoop**
* Built/Managed pipelines in distributed compute frameworks **Snowflake, Databricks, Teradata, Spark**
* Experienced with Data Engg. toolset SQL, Snowflake, Databricks, Glue, EMR, PySpark, DBT, Informatica
* Experienced in migrating data workloads from legacy on-prem systems to **Cloud Platform Snowflake**
* Very good understanding of dynamics of modern storage formats and indexing, partitioning/clustering etc.
* Experienced in migrating ETL workloads to Hadoop/Hive Platform
* Worked with Data Masking/Tokenization along with Row Access Policies in conjunction with RBAC security
* Designed & Built cloud native application on AWS using it’s serverless services in event-based framework
* Involved in platform & workload assessment and developed Data Mgmt. Strategy & Roadmaps
* Involved in Solutioning POCs/Proto-Typing when integrating new technologies
* Leveraged Rest API for exchanging the data with 3rd party vendors
* Experienced in hydrating various physical & reference data modelsIBM’s IIW, 3NF, Dimensional/Star etc.
* In Collaboration with stakeholders, built Analytics to realize value out of the enterprise data assets

**Technical Skills**:

|  |  |
| --- | --- |
| **Data Platforms** | **S**nowflake, Databricks, Redshift, AWS S3, Teradata, Netezza, Hadoop/Hive, Oracle, SQL Server |
| **Cloud Services** | Snowflake, Databricks, Redshift, AWS S3, AWS Glue, EMR, EC2, AWS Lambda, Dynamo DB, Event Bridge, SQS/SNS, Step Functions, CloudWatch, ELK stack, Terraform etc. |
| **Data Engg. Toolset** | SQL, Snowflake Scripting, DBT, Python, PySpark, AWS Lambda, AWS Glue, AWS EMR, API Gateway, AWS Event Bridge, SQS/SNS, AWS Step Functions, Rest API, Informatica, DataStage, BTEQ, Teradata Utilities, Hive, PL/SQL, Syncsort DMX etc. |
| **OS & Languages** | Linux, SQL, Python, Type Script, Java Script, PL/SQL etc. |
| **Other Tools** | GitLab, VS Code, Jupyter Notebook, Snowsight, SQL Assistant, TOAD, Teradata Utilities, Airflow, Control M, Autosys, Maestro**,** Subversion etc. |

**Education & Certifications:**

Bachelor of Technology in Engineering, Jawaharlal Nehru Technological University, India.

* Snowflake SnowPro Advanced: Architect Certification
* AWS Solutions Architect Associate
* DCAM (Data Mgmt. Capabilities Assessment Model) Certified by EDM Council.
* Cloudera Certified Hadoop & Spark Developer
* Teradata Certified Master

**Work History:**

**1) Apexon, Dallas, TX Nov ‘19 – Current**

**Roles(s) Played:** Solutions Architect

**At Apexon,** engaged multiple clients and delivered Platform Assessments, Data Migration/Modernization, Strategic Roadmaps and execute the roadmaps for our Customers in phased approach. Helped customers with platform & application modernization and realize true value of their data assets and align modern tools /products to their goals in a strategic manner.

**Responsibilities & Accomplishments:**

* Involved in Assessment of current on-prem Data platforms and data assets along with the tech-stack in use
* Analyzed the sources, ingestion patterns, data processing patterns, consumption patterns etc.
* Identify gaps & opportunities in the current platform to be remediated & position for future readiness
* Involved in developing comprehensive road map for platform migration (Data & The Work Loads)
* Designed medallion architecture in building secured Lakehouse on Databricks
* Converted legacy Parque files to Delta Tables
* Designed custom framework to migrate TBs of on-prem data to Snowflake with data reconciliation
* Built a new data collection framework to collect diverse feeds to AWS S3 for curation and aggregation
* Involved in building Data Quality framework along with controls, scorecards, and audit components
* Involved in rewriting on-prem workloads using cloud native toolset Snowflake, Databricks, Redshift etc.
* Involved in analyzing/fine tuning complex long running or sub-optimal performing SQLs in distributed computing
* Leveraged Snowflake Data Sharing & zero-copy clone features
* Built & Maintained data pipelines using DBT, AWS Glue, EMR, PySpark and other cloud native services
* Worked with AWS serverless services for event driven application like Lambda, Glue, Event Bridge, SQS/SNS, Step Functions, Dynamo DB, EC2, ECS, Fargate etc.

**Technology Stack:** Snowflake, Databricks, Redshift, Netezza, SQL, DBT, Hadoop, Python, PySpark, AWS Glue, EMR, S3, AWS Lambda, Step Functions, DynamoDB, Event Bridge, SQS/SNS, CloudWatch, ELK stack, Rest API, Airflow, Control-M, Terraform, GitLab etc.

**Bank Of America, Plano, TX Nov ‘15 – Nov ‘19**

**Roles(s) Played:** Tech Lead, Lead Data Engineer

**At Bank Of America,** played different rolesin various projects and initiatives around its massive parallel processing DW systems on Teradata & Hadoop to support Retail & Consumer Banking division at core and beyond. Involved in its data modernization journey by working in migration projects with intent to offload some of the workloads from Teradata and migrate them to Hadoop platform.

**Responsibilities & Accomplishments:**

* Lead team of developers onsite & offshore by providing solutions and technical guidance to carryout various projects and initiates around data processing on Teradata & Hadoop platforms
* Lead migration team in offloading ETL workloads from Teradata and migrate them to Hadoop platform
* Leveraged Syncsort DMX tool to develop MapReduce jobs for the Hadoop workloads
* Involved in re-design of data backup and archival practices between the data platforms on Teradata & Hadoop
* Involved in compute & storage savings initiatives on Teradata platform by analyzing heavy workloads
* Developed and documented Architectural Decision Documents as needed
* Involved in design & implementation of Data Quality Framework to gain insights and enriching the data quality
* Involved in developing the DR strategy for the Teradata platform between the data centers
* Involved in planning and carrying out annual DR tests for the systems responsible on both Teradata & Hadoop
* Involved and supported audit and ad-hoc regulatory type requests

**Technology Stack:** Teradata, Oracle, Cloudera Hadoop, Syncsort DMX, Hive, Sqoop, Informatica, Python, PySpark, SQL, BTEQ, TPT, Mainframe, Linux, Autosys etc.

**Prior clients worked with:**

Frontier Communications, Richardson, TX Mar ‘14 – Nov ‘15

Nationwide Insurance, Columbus OH Aug ‘02 – Mar ‘14