

## Contact

### Phone

+91 9080879079

### Email

vinothkumarav21@gmail.com

### Address

33/11, Sherfudeen Street,  
Choolaimedu,  
Chennai - 600094.

## Education

### B.Tech (IT) - 2010

Saveetha University, Chennai

### 12th Standard - 2006

Holy Cross, Salem

### 10th Standard - 2004

Holy Cross, Salem

## Expertise

- Python, PySpark, SQL, Databricks, AWS, Glue, EMR
- Windows Server OS
- Linux

## Language

- English
- Tamil

# Vinothkumar V

## Data Engineer

## Objectives

To work with my full potential and deliver my best for the development of the company and myself.

## Strengths

- Excellent team player with 4+ years of experience as a AWS Cloud Data Engineer and ETL/SQL Developer.
- Having 10+ years of IT Industry Working Experience.
- Excellent aptitude for understanding New Technologies and bring them into practice.
- Capable of Handling Complex and Time crunching Situations and can create more value for the company.

## Summary

- Having good experience in Data Engineering to create pipelines for various clients to build end to end solutions using distributed computing frameworks and ETL tools.
- Experienced in Python, PySpark, SQL, AWS, Data Warehouse, Shell scripting and BigData components.
- Experienced in AWS Cloud computing platform using the complete pipeline which was built with toolsets such as Databricks, Glue, EMR, Redshift, RDS, Lambda, Dynamo DB, EC2, S3, Airflow, SNS, SQS, also experienced in building pipelines for clients using Glue using Glue Studio and applied custom transformation based on the business needs.
- Worked in creating end to end solution with the transformation applied using SQL and exposed the data in Data Lake for downstream consumption. Used Lambda for automating the pipeline with S3 event trigger and have experienced in RDS Postgres, MySQL for OLTP process implementation.
- Strongly experienced in Python for application development.
- Good knowledge in SQL to achieve the business logics to be implemented based on the mapping requirement.
- Commanding knowledge in scripting (Python) to build and deploy CRON type jobs.
- Good knowledge in creating Lambda functions which have various trigger resources and events triggers with notification to stakeholders using SNS.
- Worked in VCS such as GitHub and bit bucket for maintaining and managing the code versions.
- Good experience in dealing with slowly changing dimensions (SCD) data with the incremental loads to accommodate updated to the data. (Type1, Type2, Type3)
- Good exposure to cost optimization to implement best optimized method with lesser cost
- Good knowledge to work in databases such as Oracle, Postgres.

# Technical Skills

- Python, PySpark, SQL, Databricks, Glue, Athena, EMR, S3, Lambda, SNS, SQS, Cloudwatch, Redshift, RDS, Dynamo DB, Airflow, Github
- Microsoft Windows Servers and Desktop OS.
- Citrix XenApp & XenDesktop

# Tools

- SQL Workbench, Dbeaver, Rundeck, Airflow, Pycharm, Vscod

# Experience

## ○ Hexaware Technologies, Chennai (Jun, 2023 to till date)

### Role: AWS Data Engineer

- Working as Data Engineer to create e2e pipeline for our clients running in Databricks and EMR
- Created Databricks jobs using PySpark.
- Worked in creating pipeline in EMR to submit Spark.
- Used SCD type 2 logic to and stored as medallion arch.
- Created airflow dags for orchestrating the pipeline.
- Creating automation scripts with python and Shell for automation and failure alter mechanism using CloudWatch and SNS.
- Worked on SQL loader to move the flat files from source system to destination system.
- Developed Unix Shell scripts to automate table creation, execute procedures
- Implementation of the application on Production Server.

## ○ Tata Consultancy Services, Chennai (Oct, 2015 to May, 2023)

### Role: AWS Data Engineer (Mar, 2022 to May, 2023)

- Worked as a Data Engineer to extract the data from S3 (json files) which will be extracted from an Oracle database using OGG (Oracle Golden Gate) setup. Then the data will be processed using glue with PySpark and this will push the data to raw zone (S3) as partitioned files with all the initial cleansing done, after that we will push the data to refined zone (S3) with all business logics and transformations are applied from ref zone we will push a data copy to redshift tables using which the dashboard and algorithms will be built by the downstream team.
- Worked with various file formats to extract data and apply data transformation using Glue studio and load in S3.
- Created ETL pipelines with PySpark and submitted in EMR cluster using spark submit and have good understanding in spark architecture to debug and solve issues in jobs.
- Using unload technique data from S3 will be loaded in client specific redshift tables.
- Utilized Athena to apply analytical queries and transformation on data that resides in S3.
- Created lambda functions to enable event trigger that initiates glues job when a object update happened in S3.
- Integrating Amazon cloud watch with EC2 instances for monitoring the log files track files.
- Creating topics in SNS to send notifications to subscribers as per the requirement.
- Created end to end ETL process with the step functions and Glue with the lambda trigger and event bridge.
- Created pipelines with the slowly changing dimension (SCD) and managed the delta updates based on the expectation.
- Used IAM to create new accounts, roles and groups.

# Experience

## Role: Data Engineer (Nov, 2020 to Feb, 2022)

- Created client specific schema in Redshift and multiplexed the data to corresponding redshift tables.
- Worked on putty to create the Unix shell scripts and to run the Unix Shell scripts.
- Worked in AWS cli and Boto3 to play with the AWS resources programmatically.
- Converting the existing Postgres SQL scripts to Athena scripts
- Worked on the creation of the design, development of the code.
- Creation of the complex SQL queries using joins, worked on materialized views.
- Provided the documentation for the release notes to the production team to implement.
- Provided the production support at the time of release to clarify the doubts to production team.
- Created the CRON tab entries in UNIX to run the shell scripts automatically.
- Interaction with the production team to clarify the production issues.
- Created complex OLAP queries with required business logic applied
- Involved in development of Shell scripts to send invoices to customers through mailing.

## Role: SQL/ETL Developer (Sep, 2019 to Oct, 2020)

- Ensure data quality and integrity of the data in RDBMS and in the warehouse and maintain the end to end pipeline.
- Document ETL processes and metadata, Optimize ETL performance.
- Write complex SQL queries for data retrieval, manipulation, and analysis.
- Loaded the data into database tables using SQL\*loader from text and excel files
- Develop and optimize stored procedures and functions for database operations.
- Perform data migration tasks, including data import/export and ETL.
- Worked on Oracle SQL Developer to develop queries and created procedures and packages in Oracle.
- Collaborate with application developers and data analysts to meet database-related requirements.
- Maintain documentation of ETL processes, transformations, and data lineage
- Extract data from diverse sources.
- Transform data for consistency and suitability.
- Load data into storage systems.
- Develop and maintain ETL processes.

## Role: Systems Engineer (Oct, 2015 to Aug, 2019)

- Worked as a Citrix Support Engineer / System Admin. role.
- Implementing, Monitoring and providing support for Citrix Systems and related applications.
- Present Citrix Environment version is 6.5 / 5.0 and supporting around 1K users.
- Having knowledge on Citrix technologies like XenApp and XenDesktop 4.5, 5.0, 5.6, 6.5 and 7.x, Machine Creation Services(MCS), Provisioning Services(PVS) and Web Interface.
- Taking responsibility for escalated technical issues in the environment and providing problem resolution.
- Having good knowledge in accessing to ServiceNow Tool for doing BAU activities.
- Monitoring and updating the environment dashboard regularly in ServiceNow to avoid Incidents, Change and Problem SLA breach.
- Familiar with Windows Server OS/Desktop, Active Directory, Group Policy management, Microsoft Remote Desktop service.
- Ability to isolate the issue to be Citrix related, Microsoft related or 3rd party application related.
- Perform intermediate issue analysis and resolution.
- Identify root cause of issues.
- Respond to Server alerts, system outages, critical monitoring alerts.
- Perform server maintenance and operational procedures.
- Monitoring Citrix profile server session establishment and DFS file synchronization.
- Preparing SOPs, KEDB, RU counts and other Incident analysis reports which needs to be shared to Service Management/Operations team.

# Experience

## ○ Sutherland, Chennai (Nov, 2014 to Oct, 2015)

### Role: Consultant

- Coordinate with site technician and customer to resolve the technical related issues.
- Efficient in providing on-time resolutions to the customer.
- Having knowledge in removing real time viruses, junk applications like spyware/adware and re-installing applications.
- All the basic troubleshooting steps need to follow in by checking host, ping and need to make sure that, the machine is completely functioning without any infections.
- Proficient in managing Windows system related queries.

## ○ Wipro, Chennai (Feb, 2013 - Nov, 2014)

### Role: Consultant

- Need to analyze the request and have to confirm with requestor ( i.e. Project Manager ) before processing, requestor who is completely responsible and have to be in touch with the enterprise level customer to satisfy customers required work to be completed.
- Proficient in managing service related hardware / software / network problem diagnosis resolution.
- Coordinate with client and technicians to configure IP telephones and routers in customer's site.
- Create Administrator account for the customer end users and manage the privileges for an administrator account.

## ○ Reliance, Chennai (Dec, 2011 - Feb, 2013)

### Role: Customer Support

- To support customer over the phone and have to resolve the network speed issues, login errors and technical problems by providing troubleshooting steps.

