**Ravi Teja Voleti**

+1 (409) 444-0348

[voleti.raviteja9@gmail.com](mailto:voleti.raviteja9@gmail.com)

**Education** Masters in Computer Science

University Of Central Missouri, Warrensburg, MO

**Summary**

* Extensive experience dealing with tools & technologies such as Talend, Spark, Datastage , ESB, Tableau etc on Big Data & Cloud Environments AWS and Azure.
* Deep Understanding of Data Lake and Data Warehousing , data modeling using Erwin,

HDFS, Spark Data Frames, Spark SQL, Hive, HBase, Postgres, Dremio, Redshift etc

* Developed code for Batch , Real Time , Event Driven data processing using Kafka , ActiveMQ , AWS Lambda etc
* Worked with different components , stages and advanced areas of Talend , Datastage such as Dynamic Schema , Metaservlet APIs , Run Time Column Propagation etc
* Transformed / moved the data between various sources such as files , databases , Amazon S3

ADLS, Azure (Event Hub, Queues), APIs, Salesforce Objects, Platform Events, Salesforce Push topics, Five9 etc

**Technical Skills**

Tools & Technologies – Talend, Spark, Big Data, ETL, Datastage, ESB, Tableau, Erwin,

Kafka, ActiveMQ, HDFS etc

Databases – Hive, Hbase, Postgresql, Redshift, Dremio, Oracle, SQL Server etc

Platforms – Amazon Web Services (AWS), Microsoft Azure

Schedulers – TAC, Datastage Director, Crontab etc

Languages – SQL, Unix Shell, Python etc

**Work Experience**

**Capgemini America, Atlanta, GA** October 2017 – Till Date

Senior Data Engineer

**Client: Virgin Voyages –** Worked on this Project to integrate different source systems and to develop Data Lake on AWS for recommendation engine, customer 360 dashboards and reporting purposes

* Developed ESB routes and Streaming Jobs to integrate ActiveMQ , Kafka with Salesforce
* Created Durable subscribers in ActiveMQ to prevent data loss with streaming jobs
* Developed metadata driven and reusable framework with auditing , alerting processes to save development efforts using Talend, Spark, Big Data etc on jobservers and EMR
* Moved the data from different sources such as files, Oracle, Postgresql , Kafka to AWS S3 (External Hive Tables) and implemented SCDs on the Data Lake for reporting purposes
* Implemented the Partitioning techniques on Hive to Improve the performance

**Client: Royal Caribbean Cruises Ltd –** This is a migration Project from Traditional ETL & Warehouse Technologies to Azure Big Data Cloud for effective performance of applications.

* Worked with Talend Metadata Bridge to convert Informatica jobs to Talend Jobs
* Developed CICD pipelines with Bitbucket , Git , Nexus.
* Developed power shell script to move the files from job servers to ADLS
* Worked with Talend installation and created complex Data Integration & Big Batch Spark jobs using components such as tMap, tJoin, tReplicate, tParallelize, tJava, tJavaFlex, tLogCatcher, cLog, tsqlrow etc
* Developed PySpark Scripts to implement transformations on Parquet files.

**CDI Solutions Inc** Aug 2015 **-** October 2017

Big Data / ETL Developer

**Client: Capital One, Richmond, VA –** This project is to build Portfolio Tracker for Underwriters, Asset Managers etc on AWS to track the auto loans, credit cards, online services etc, undertake trend analysis to provide better customer service and assess current and future profitability

* Developed Real Time and Batch Applications which move the data from Kafka, S3
* Developed Stored Procedures and Functions in SQL Server and Postgresql
* Made use of AWS Lambda to develop Event driven PySpark Code with necessary auditing and alerting which reads the data from SQL Server and writes the data to Postgres
* Developed parallel and server jobs using different processing stages in Datastage like Transformer, Aggregator, Lookup, Join, Merge, Change Apply and Filter ODBC, Shared containers and local containers, reusable mapping parameter sets etc
* Experience using all stages in sequencers and create parallel jobs with RCP and Multi instance

**Client: Elliemae, Pleasanton, CA –** This project is to build mortgage management application that allows banks, credit unions and mortgage lenders to originate and fund mortgages and improve compliance, loan quality and efficiency

* Jobs were developed to Call APIs, windows, python, unix shell scripts, java code.
* Parsed xml, json , csv, delimited , fixed width, .txt files as per the requirement
* Used ETL concepts such as Star and Snowflake Schema to build dimensions and facts and used joblets, trunjobs, context parameters in jobs
* Worked extensively with TAC execution plans, triggers, SMTP, crontab etc
* Developed Tableau Dashboards, Charts, Graphs etc by connecting to oracle database for UI purposes.

**Aparna Group, Hyderabad, India**  August 2013- July 2014

Graduate Engineer Trainee

This project is to provide the leadership by building a data warehouse with the estimates of sales, expenses, revenue forecasts etc to benefit administration, human resource and services for projects of any size.

* Worked on programs for scheduling data loading and transformations
* Developed Master and Child Jobs and specified the sequence on an external file which can modified as per the requirement using Talend
* Worked with different stages like Transformer, Aggregator, Lookup, Join, Merge, Modify, Remove Duplicate, Oracle Stage, Sort, Peek, Row Generator, Column Generator, Sequential File, and Data Set
* Used shared DB connection in Talend to either commit all or rollback data for all jobs
* Worked on different POCs for using open source technologies, developed custom routines etc