

Chirag Vachhani

990-980-8397

vcd.vachhani@gmail.com

Professional Summary:

- **10.5 years** of professional experience in **Data Engineering** field of Automotive B2C/B2B, Travel/Hospitality B2B, ERP Service, E-Commerce solution, Legal Law firm, Market Research Solution and Real Estate.
- Proficient in analyzing business processes/requirements, data modeling and translating them into data warehouse solutions.
- Implemented data analytics warehouse solutions using **Cloud platform/SAAS/ETL tool** like **Domo, Apache Spark, Kafka, AWS Glue/Data Pipeline, Microsoft SSIS, Pentaho** with various SQL/NoSQL data import sources like **PostgreSQL, SQL Server, MySQL, ELK stack, Apache Solr, MongoDB, Redis, CSV, Excel, Rest API**.
- Developed ETL custom solutions using **Python** libraries like **PySpark, Pandas, Numpy** to overcome ETL limitations and provide **In Memory** data computation solutions. Prepared data visualizations Admin UI using **Django framework** for managing configurations, visualizing error log/output log/data log and managing history of all operations.
- Designed a highly complex and logical **algorithmic solution** for the **ECommerce Recommendation Platform** by using Alternating least Square(**ALS**), Singular Value Decomposition(**SVD**), **CUR** matrix approximation, **Latent factor** modeling and various data mining algorithms with in memory data processing(**Apache Spark**) based on collaborative filter technique.
- Experienced in designing/developing/optimizing complex data mapping solutions of **structured SQL/Nosql** data source.
- Proficient in **Data Cleaning/Mining, Data Integration, Data Migration, DBA** technique to empower the right data warehouse solutions.
- Have got good exposure working on **Agile** based projects and followed **Agile Scrum** methodology/code versioning tools such as **SVN, Jira, Git, Bitbucket and Basecamp**.
- Excellent communication and interpersonal skills, good analytical reasoning and high adaptability of new tools and technologies.
- Spiritually believe in a team working and building professional relationships.

Objective:

Aiming to utilize my highly research/analytical/algorithmic problem solving and decision making skills in the field of computer science. I would be very enthusiastic to take on new challenges, ambitious about learning new technologies, observing / implementing new things everyday which puts me to draw my life with high satisfaction.

Educational Qualification:

Course	Institution / University	Year of Passing	Grades
B.Tech (Computer Engineering)	U V Patel College Of Engineering, Ganpat University, Kherva (Mehsana), Gujarat	2012	First class

Technical Skills:

Languages: Python, PHP, SQL/PL SQL, Javascript
Databases: PostgreSQL, Mysql, SQL Server, MongoDB, Redis, Elastic Cache.
Web Servers: NGINX, uWSGI
Operating Systems: MAC OS, Microsoft Windows, Unix/Linux (Ubuntu, Fedora)
IDE: Eclipse, Netbeans, Pycharm, Sublime, Visual Studio
Integration Tools: SVN, Git, Bitbucket
Frameworks: Apache Spark, Django(Python)
Cloud Services: Amazon AWS (EC2, RDS, S3/Glacier, Glue, Data Pipeline, Batch, Lambda, Step Function, SQS/SNS, SES, EventBridge, API Gateway, Cloudwatch)

Professional Experience:

Simform Solutions, Ahmedabad, Gujarat

Tech Lead

Jun' 20 – Nov' 22

(2.5 Years)

key Experience:

- Worked as a **Tech lead** in a data engineering team to provide solutions for automotive industries projects.
- Transformed routine ETL jobs to serverless fault tolerance using **AWS Analytics, Lambda, SQS, SNS** services.
- Prepared **data model** strategy to prepare customer repo for automated Email Marketing platform named Mautic.
- Prepared data warehouse data import strategy for B2C automotive project using **Domo cloud service** as an ETL and reporting tool.
- Prepared customer desiredness model based on passing and filtering certain defined rules on **Google Analytics** row data. Prepared matched inventory priority JSON based on customer's desiredness using **Pyspark, Pandas** Python script.
- Worked on complex data transforming jobs with Microsoft SSIS, Stored Procedures and Materialized View of PostgreSQL and SQL Server.
- Prepared demo for optimizing data warehouse solution to work with **Databricks** as a custom cloud service provider which built on **Apache Spark** framework.

Travel Designer Group, Ahmedabad, Gujarat

Sr. Database Developer

Dec' 17 – May' 20

(2.5 Years)

key Experience:

- Designed data model, data dictionary and ER diagram of large scale B2B platform which provides services to travel industry domain, which is running over more than 12 years worldwide in more than 38 countries with an **OLTP DB** size of more than **150GB**.
- Database schema design, Database migration, data verification, data transforming and loading.
- Optimized system by implementing highly performing algorithm to give benefit in terms of speed, memory consumption and highly reliable to recover any loss of data.
- Database analysis and monitoring by using **Percona Monitoring and Management toolkit, Grafana, SQL DM(Monyog), MyTop, InnoTop** kind of software toolkit.
- Tuning Mysql global parameter by monitoring and analyzing the output of **mysqltuner script**.
- Optimized performance by changing database schema, alter table engine, alter data type, apply correct composite B-tree, fulltext index as required, optimized stored procedure, trigger and implement event as when required to remove unusual data, refactoring query structure using explaining the execution plan, changing the join order by using STRAIGHT join.
- Decrease query per time second (QPS) by analyzing slow query log, using Percona **PT-Query digest** toolkit.
- Analyzing Index duplication and usage of that index by using Percona **PT-index-duplicate** and **PT-index-usage** toolkit.
- Found and resolved deadlock situation by analyzing Innodb engine status, by query profiling, by using error log of previous regular interval, finding pattern of problems.
- Extracting, transforming and loading massive data by using **Pentaho ETL** tool, sending daily BI reports to management.
- Optimized date-range by altering schema design and implementing Merge and Split compression algorithms.
- **Database Migration** of different DB applications into single DB since serving the same purpose.
- Prepared statistical analysis report by analyzing data volume and data pattern which helpful to save business loss by preplanning.
- Massive data of more than 15GB json file loading into **MongoDB Nosql** db. Convert JSON based embedded document set of **MongoDB to MySql** relational structure DB with **Pentaho ETL** tool.
- Configured and Implemented **Elastic Logstash/Kibana(ELK stack)** to make data transform-loading(Mysql to Redis, Mysql to elasticsearch)/data analysis(Configure Kibana dashboard) on heavy volume.
- Developed data Extract/Transform/Loading scripting component in **Python** to overcome ETL limitations, **Django web framework** for ready Admin UI to manage configurations, to visualize error log/output log/data log and manage history of all operations.
- Massive in memory data computations with the help of **Pandas dataframe** and **Numpy** library of python.

key Experience:

- Worked with business analysts for requirement gathering, analyzed requirements and coordinated with multiple projects.
- Designed project workflow, Data flow, ERD and Activity diagram, Prepared project Architecture, Finalized coding-convention by standard regulation pattern.
- Designed database. (RDBMS, NoSql)
- **Optimized performance** of Mysql database of more than **100 GB** by analyzing the session logs, understanding various bottlenecks, tuning the session and global variable and using techniques like Innodb locking, partition, replication (master-slave architecture).
- Designed highly logical **Mysql recursive** structured query, only using Mysql session variable. (**B-tree** searching all child nodes of given node, B-tree searching only leaf node, B-tree finding all element of given node, pagerank algorithm, review rating algorithm, (remove Special characters, Non-ASCII, hidden characters leads to spaces, accents etc))
- **Optimized Mysql query** output by using B-tree Index, spatial Index for geo enable application, full-text Index in boolean mode, stored procedure with exception handling and dynamic query using prepared/execute statement, trigger, function, auto schedule event, innodb exclusive and shared locking, query cache etc.
- Designed a highly complex and logical **algorithmic solution** by extracting-transforming-loading **massive datasets** from various sources into **Apache Spark**. (Recommender system using Alternating least Square(**ALS**), Singular Value Decomposition(**SVD**), **CUR** matrix approximation, **Latent factor** model algorithms with memory + model based on collaborative filter technique)
- Configured and implemented Amazon **AWS cloud service**. (AWS S3 bucket policy to authorize only specific domain, AWS Glacier for uploading and restoring file, EC2, RDS, CDN, Auto Scaling, Load Balancer)
- Configured and Implemented **searching platform** by using **Apache Solr** to make searching experience easier and better with massive dataset.
- **OCR content searching** technique from PDF/Image file using **python pytesseract** package.
- Implemented online **radio streaming** service by using **SHOUTcast** streaming server and winamp plugin.

key Experience:

- Handled a dedicated live project with 6 team members as a tech lead to provide technical solutions and engaged with clients to give daily updates and progress reports.
- On changing the business requirement, update and maintain a Database designed of RDBMS by using various concepts of database.
- Designed technical documents to understand high level algorithmic solutions.
- Configured and launched **AWS EC2, RDS** service.
- Designed dynamic Filter-adding algorithm on searching platform.
- **Optimized Mysql query** to enhance the performance of the system.
- Provided an **SEO solution** to get the highest rank of a website on the first page of Google search engine. Optimized quality of web content to score more than 90% on Google web quality tool.
- Designed and developed Mobile Android App by using **Google App sheet** tool. Data synchronized with ERP by using **Google spreadsheet, Google drive client API**.
- Designed and developed a complex **BIS report** for ERP which has helped a lot to take an immediate proper decision by comparing necessary data at one place.
- Designed and developed various **graphical, statistical and mapping view BIS reports** by visualizing properly, which vertical of business require and missing what amount of effort in terms of time and money.
- Had a good opportunity to work with different jQuery Plugins.
- Integrated third party **REST API** by using CURL method, like Facebook, Twitter, Google Calendar, Pinterest, SMS Gateway, Paypal, CCavenue payment gateway solution.

Declaration:

I hereby declare that the above mentioned information is true to the best of my knowledge.

Chirag Vachhani