|  |
| --- |
| MD.Imran  Phone: 9989949194  Email: mohammadimran688@gmail.com |

Sr. Data Engineer with 8+ years of experience in building a complete Software Development Life Cycle (Business Intelligence/Data Warehousing/ Databricks pipelines, Azure Cloud Services) including requirement gathering, analysis, design, development, testing, performance tuning, implementation, production support, Hadoop-data Migration and Cloud migration.

|  |
| --- |
| Skills |

Core Java

JMS Kafka

**Cloud-Azure**

Java 8 Spring

Spring Boot

JWT

**Python**

Spring Data

**PySpark/DBFS**

JWE Encryption

**Carrier Profile:**

|  |  |
| --- | --- |
| Sr. Tech Engineer  Opus Technologies  Mar 2023 – Present  Proficiently managing data ingestion processes within Databricks, maintaining Delta tables, and seamlessly developing new workflows in alignment with the established framework, all while harnessing the power of Azure services.   * Azure Analytics / Databricks * Azure SQL server / Postgres SQL | Sr. Data Engineer  Digile  Feb 2022 – Mar2023  Designed and implemented a Devops pipelines and  Employed GitHub actions within the Azure  environment to automate and streamline various  processes.   * Azure Analytics * Business Intelligence (power BI) * Data Modeling * Cloud Migration |
| Sr. Data Engineer  UHG (Optum)  Apr 2016 – Jan2022  Developed and Implemented Data Modeling for a Crucial Project with path breaking Master Data Management.   * Business Analytics * Business Intelligence * Data Modeling * Cloud Migration | Software Engineer  Magna Infotech  Feb 2015 - Apr2016  Hands on experience in handling and processing ETL application with extensive knowledge of testing and validating data warehouse.   * Business Intelligence * ETL/Report testing * Processes Integration * ETL-Reporting Integration |

**Professional Profile:**

* Having 8+ years of Professional IT experience in covering all the phases of Software Development Life Cycle (on DW Project) including requirement gathering, analysis, design, development, testing, performance tuning, implementation, and production support.
* Proficient in designing, developing, and maintaining data warehouse systems.
* Working on Azure Cloud Migration projects
* Having experience in Agile Framework process.
* Working with ETL leads to formulate ETL approach in data warehouse systems.
* Skills Core Java, Spring Spring Boot, Spring Data, JPA, JMS Kafka, JWE Encryption.
* Having experience in data ingestion, storage, querying, processing, and analysis of data in Hadoop Ecosystem using Map Reduce, HDFS, Hive, Sqoop, HBase, Oozie and PySpark.
* Very good understanding on developing ETL pipeline for customers using Python, Spark, Sqoop, Hive and various BIGDATA and Data Lake.
* Managed data coming from different sources and involved in HDFS.
* Having hands on experience in HDFS Data Flow to process the data in batch mode.
* Hands-on experience in writing Sqoop jobs, to implement the import and export data from HDFS to RDBMS and vice versa.
* Working knowledge in different file formats like Parquet, CSV files processing in Hive and spark.
* Hands-on experience in writing hive queries for data analysis to meet the business requirements with the improve performance by implementing the Partitioning, Bucketing concept.
* Hands-on experience in creating Datasets and Data Frames in spark SQL.
* Experienced in creating Hive tables and file ingestion into hive
* Having knowledge in Hadoop Distributed File System.
* Experienced in integration of various data sources with Multiple Relational Databases RDBMS systems Oracle, Db2 and SQL Server
* Involved in Production Support and Operations teams to resolve production issues in a timely and efficient manner.
* Proficient in unit testing, system integration testing.
* Experienced in leading and mentoring team members on both functional and technical aspects.
* Flexible, enthusiastic, and project-oriented team player with solid communication and leadership skills to develop creative solutions for challenging client needs.
* Hands-on experience in creating Datasets and Data Frames using pyspark and handling the ingestion framework and code changes as required.
* Having knowledge on creation of new environment setup for databricks and make the changes as required for the new environment.
* I am well-equipped to operate as a prompt engineer, ensuring the timely and efficient execution of data processing tasks.
* Able to work independently and collaborate proactively & cross functionally within a team.

**Work Experience:**

* Working as Sr. Tech Engineer in Opus Technologies from Mar’2023 to Till Date.
* Working as Sr. Data Engineer in Digile from Feb’2022 to Mar’2023.
* Worked as Sr. Data Engineer in UHG (Optum) from Apr’2016 to Jan’2022.
* Worked as Software Engineer in Magna Infotech from Feb’2015 to Apr’ 2016.

**Qualification:**

* MBA from JNTUK University 2012.

**Technical Skills:**

|  |  |
| --- | --- |
| **Concepts** | ETL, Data warehousing, Business Intelligence, Relational Database Management, Data Integration, HDFS, Hive, PySpark, Python and Databricks, Azure Services |
| **Domain Experience** | Payments Domain, Manufacturing, Health Insurance |
| **Database** | Db2, SQL server, Hive, Netezza, Postgres SQL |
| **Reporting tool** | Cognos, Mstr, Power BI |
| **ETL tool** | Informatica. |
| **Data Modeling** | Dimensional Modeling for Data warehouse |
| **Scheduling tool** | Airflow, Oozie |
| **Database tools** | SQL Server, SQL\*plus., Hive, MySQL, Postgres SQL |
| **Other IT skills** | SQL, PYSPARK, Python, HDFS, Unix |

**Projects Summary:**

|  |  |
| --- | --- |
| **Project** | ICBA |
| **Customer** | **Opus Technologies** |
| **Period** | March ,2022 to till date |
| **Role** | Team Member   * Process Cobol files twice a week using the PySpark and SQL Server ingestion framework and load them into Delta tables for downstream consumption. Modify the code as needed to meet requirements. * Establish the necessary tables in PostgreSQL for the API team, leveraging Delta tables. * Include information on performance tuning and optimization efforts to enhance data processing efficiency. * Analyze and implement client requests as they arise. * Manage key vault services for enhanced security. * Set up a new Databricks environment for UAT testing. * Deliver data to the sales force team for loading data objects in accordance with their specific requirements. * Emphasize the importance of maintaining detailed documentation for the entire process, including code, configurations, and procedures.   Industry Practice: Payments Domain  Service Practice: IT Infrastructure Services  Special Software: Azure Services, Databricks, Pyspark, Sql, python, Azure SQL server, Postgres SQL, GitHub  Project Location: Hyderabad, India |
| **Description** | The **Independent Community Bankers of America** is dedicated exclusively to representing the interests of the community banking industry and its membership through effective advocacy, best-in-class education, and high-quality products and services.  They provide support services for all these small community banks and help them sustain and grow in this ever so competitive world. **ICBA Bancard** delivers flexible, innovative payments and digital commerce solutions for community banks.  Build a Modern Data Platform that would process high volume and low latency financial data. Integrate with other SaaS solutions like IVR and Salesforce using MuleSoft. Ingested legacy data files from 5 Financial Institutions and process on Databricks Spark Engine.  As a Data Engineer I was responsible for developing end-to-end data pipelines from ingestion transformation and consumption for all orchestration MuleSoft was leveraged for moving data from sftp servers to Azure BLOB Storage and from the data platform to Salesforce. |
| **Skills** | Databricks, Azure SQL server, Postgres SQL, Pyspark, Python, SQL   * Creating Data model structure. * Efficiently oversee the ingestion process pipelines to ensure timely and successful data loads. * Creating scripts to load the data in target location. * Creating delta tables on top of raw location. * Providing Data solutions to existing problems. * Creating and standardizing complete data process. |

|  |  |
| --- | --- |
| **Project** | Azure Devops |
| **Customer** | **Daimler** |
| **Period** | February ,2022 to March ,2023 |
| **Role** | Team Member   * Enhance the required changes from use case teams. * Extracting/Fetching data from different marketplaces as per the customer needs and provide the data in required Azure blob storage and create the tables. * Loading the Data into azure by using Databricks/Data-factory using Pyspark and SQL * Creating Reconciliation framework. * Working on different service desk tickets and resolve them on time.   Industry Practice: Mercedes Benz Vehicles  Service Practice: IT Infrastructure Services  Special Software: Azure, Databricks, Data factory, GitHub, Devops  Project Location: Hyderabad, India |

|  |  |
| --- | --- |
| **Project** | Azure Cloud Migration |
| **Customer** | **OPTUM Rx** |
| **Period** | 1st January,2021 to 25th January 2022 |
| **Role** | Team Member   * Migrate the data from On-prem to Azure cloud * Extracting/Fetching data from On-prem sources (Oracle/Netezza/DB2/SQL server) by using Sqoop/Integration-Runtime/ETL tool/Data-factory * Loading the Data into azure by using Databricks/Data-factory/spark-Scala * Creating Reconciliation framework.   Industry Practice: Health Insurance  Service Practice: IT Infrastructure Services  Special Software: Azure, Databricks, Data factory, Sqoop  Project Location: Hyderabad, India |

|  |  |
| --- | --- |
| **Project** | UHG [INET](https://arajira.optum.com/confluence/display/RXTEAM/Exploratory+Analysis+-+INET) FORMULARY |
| **Customer** | **OPTUM Rx** |
| **Period** | 10thJune ,2019 to 10th January 2022 |
| **Role** | Team Lead (Sr data engineer)   * Performed Team Lead activities to handle the Bdpass related activities. * Identified performance bottlenecks in the jobs and implemented the changes. * Created a new SQL script with Standard code design. * Having experience in agile (Jira) framework process. * Performed Root Cause Analysis, code fix/data fix for Data issues raised by the business in the Bdpass. * Analyze the production issues regularly and bug fixing.   Industry Practice: Health Insurance  Service Practice: IT Infrastructure Services  Operating System: Windows Server 2003 Enterprise, Linux  Special Software: Sqoop, SQL server, Linux, HDFS  Project Location: Hyderabad, India |
| **Description** | INET is a backend relational database that houses (among other things) formulary data to be used by downstream systems. It records changes in clients' formularies like tier changes and introduction/removal of step therapy, prior authorization, quantity limit etc. at an NDC level. The INET relational database cannot be accessed directly, but the data is replicated and updated frequently in BDPaaS.  UHG has common enterprise data warehouse and enterprise business Analytics owns multiple sources in Bdpass for Prod Ops Analytics, Manufacturing Analytics, HR Analytics, Finance Analytics, Marketing Analytics, and Procure to Pay Analytics, Sales Analytics, and IT Analytics. |
| **Skills** | Sql server, shell scripting, Sqoop   * Creating Data model structure. * Creating Sqoop scripts to load the data in target location. * Creating hive tables on top of raw location. * Providing Data solutions to existing problems. * Creating and standardizing complete data process. |

|  |  |
| --- | --- |
| **Project** | **UHG Rxclaim DB2 tables migration to Bdpass** |
| **Customer** | UHG |
| **Period** | 9thJuly ,2017 to 1st January 2020 |
| **Role** | Team Member   * Performed Team Lead activities to handle the Off-Shore team. * Identified performance bottlenecks in the jobs and implemented the changes. * Creating RxclaimDb2 tables in Bdpass environment without using any manual intervention using Sqoop. * Having experience in Grid migration projects and Db2 to Big Data migration projects. * Having experience in agile framework processes. * Performed Root Cause Analysis, code fix/data fix for Data issues raised by the business in the HDFS. * Analyze the production issues regularly and bug fixing. * Co-ordination with various source teams during business impacting changes and deployments.   Industry Practice: Health Insurance  Service Practice: IT Infrastructure Services  Operating System: Windows Server 2003 Enterprise, Linux  Special Software: Sqoop, Hive, HDFS, Db2, Oozie, HBase, Airflow.  Project Location: Hyderabad, India |
| **Description** | UnitedHealth Group Inc. is an American managed health care company based in Minnetonka, Minnesota. It is 6th in the United States on the Fortune 500.UnitedHealth Group offers products and services through two operating businesses, UnitedHealthcare and Optum, both subsidiaries of UnitedHealth Group. Optum is a new business brand of UnitedHealth Group and operating business and Business Process Outsourcing (BPO) services. UnitedHealth Group serves approximately 70 million individuals throughout the United States.  UHG has common enterprise data warehouse and enterprise business Analytics owns multiple Hadoop solutions for Prod Ops Analytics, Manufacturing Analytics, HR Analytics, Finance Analytics, Marketing Analytics, and Procure to Pay Analytics, Sales Analytics, and IT Analytics. |
| **Skills** | Sqoop, Hive, HDFS, Db2, Oozie, HBase, Airflow   * Creating Airflow scheduler and migrating file for production on daily basis. * Optimizing Bdpass jobs and its runtime for any performance issues. * Developing the new Sqoop jobs. * Optimizing jobs from server to parallel. * Providing enhancement support. (Data issues, performance issues and new deployment) |

|  |  |
| --- | --- |
| **Project** | **UHG Rxmax DB2 tables migration to Bdpass** |
| **Customer** | **UHG** |
| **Period** | 4th Dec,2019 to 6thJune,2020 |
| **Role** | Team Member   * Identified performance bottlenecks in the jobs and implemented the changes. * Migrating RXmax Db2 tables in Bdpass environment without using any manual intervention using sqoop. * Performed Root Cause Analysis, code fix/data fix for Data issues raised by the business in the Data Warehouse. * Analyze the production issues regularly and bug fixing. * Resolving the production issues within the time period. * Co-ordination with various third-party vendors during business impacting changes and deployments. * Assign the issue tickets and Tasks to the team members * Attending status meetings with client   Industry Practice: Storage  Service Practice: IT Infrastructure Services  Operating System: Windows Server 2003 Enterprise, Linux  Special Software: Sqoop, Hive, HDFS, Db2, Oozie, HBase, Airflow.  Project Location: Hyderabad, India |
| **Description** | RXmax—storage systems and software that help customers around the world store, manage, protect, and retain one of their most precious corporate assets: their data. Rebate data is recognized throughout the industry for continually pushing the limits of today’s technology.  Over all 1100 tables migrated in Bdpass environment and has a common Hadoop enterprise business Analytics owns multiple business teams and BI reporting solutions for Prod Ops Analytics, Manufacturing Analytics, HR Analytics, Finance Analytics, Marketing Analytics, and Procure to Pay Analytics, Sales Analytics, and IT Analytics. |
| **Skills** | Sqoop, Hive, HDFS, Db2, Oozie, HBase, Airflow:   * Migration of the DB2tables from one environment to the HDFS raw layer. * IBM CDC is enabled for Incremental process. * Creating hive tables on top of raw layer for business use. * Maintain sub volumes to store raw data. * Developing the new Airflow jobs. * Optimizing jobs from server to parallel. * Providing enhancement support. (Data issues, performance issues and new deployment) |

|  |  |
| --- | --- |
| **Project** | **UHG IDW NETEZZA tables migration to Bdpass** |
| **Customer** | **UHG** |
| **Period** | 28th Sep,2018 to 4thJan 2021 |
| **Role** | Team Member   * Pulling historical data from Netezza and push it to Data Lake platform * Abintio ETL team will take care of Incremental process by pushing the load IDS files. * Creating hive tables on top of raw layer for business use. * Creating dedup views as per the Netezza source logic. * Maintain sub volumes to store raw data. * Creating reports to monitor the daily load status. * Recon dashboard is created to monitor the daily load failures. * Scheduling jobs in Airflow and Oozie. * Worked as shadow resource in development of Business Objects and Bdpass   Industry Practice: Storage  Service Practice: IT Infrastructure Services  Operating System: Windows Server 2003 Enterprise, Linux  Special Software: Sqoop, Hive, HDFS, Netezza, Oozie, HBase, Airflow.  Project Location: Hyderabad, India |
| **Description** | IDW — storage systems and software that help customers around the world store, manage, protect, and retain one of their most precious corporate assets: their data. Rxclaim data is recognized throughout the industry for continually pushing the limits of today’s technology.    Over all 100 tables migrated in Bdpass environment and has a common Hadoop enterprise business Analytics owns multiple business teams and BI reporting solutions for Prod Ops Analytics, Manufacturing Analytics, HR Analytics, Finance Analytics, Marketing Analytics, and Procure to Pay Analytics, Sales Analytics, and IT Analytics. |
| **Skills** | Sqoop, Hive, HDFS, Db2, Oozie, HBase, Airflow:   * Migration of the DB2 tables from one environment to the HDFS raw layer. * IBM CDC is enabled for Incremental process. * Creating hive tables on top of raw layer for business use. * Maintain sub volumes to store raw data. * Developing the new Airflow jobs. * Optimizing jobs from server to parallel. * Providing enhancement support. (Data issues, performance issues and new deployment) |
| **Achievements** | Winner of Idea Max challenge, appreciations from Client and Higher Management. |

Personal Details:

|  |  |
| --- | --- |
| **Language** | English, Telugu, Hindi, and Urdu |
| **Hobbies** | Playing cricket, Praying namaz. |
| **Date of Birth** | 28thAug,1988 |
| **Sex** | MALE |
| **Nationality** | INDIAN |
| **Marital Status:** | Married |
| **Preferred Location:** | Hyderabad |
| **Contact Number:** | +91 9989949194 |