

VISA STATUS: VISA - H1

Please send **DIRECT CLIENT Requirements ONLY**

LOOKING FOR CONTRACT / CORP-TO-CORP POSITIONS ONLY

STRICTLY NOT INTERESTED IN FULL-TIME POSITIONS

SHAR D

571-299-3446

shard287@hotmail.com

SUMMARY:

- **Overall, 12 Years of total IT experience** and Technical proficiency in the Data Warehousing involving Application Design, Data Modeling, Development, Testing and Documentation
- **2 years of experience as Azure** Data Engineer in Azure Cloud, Azure Data Factory, Azure Data Lake Storage, Azure Synapse Analytics, Azure Analytical services, Azure Cosmos, NO SQL DB, Azure HDInsight, Big Data Technologies (Hadoop, Hive, Impala, Oozie etc).
- **9 years of experience as ETL Lead in** Informatica Power center design, implementation, and maintenance mainly in large scale Data Warehouse environments, experience in Oracle RDBMS and Unix shell scripting.
- Experience in designing Azure Cloud Architecture and Implementation plans for hosting complex application workloads on MS Azure.
- Providing Azure technical expertise including strategic design and architectural mentorship, assessments, POCs, etc., in support of the overall sales lifecycle or consulting engagement process.
- Excellent knowledge on integrating Azure Data Factory V2/V1 with variety of data sources and processing
- the data using the pipelines, pipeline parameters, activities, activity parameters, manually/window based/event - based job scheduling.
- Have Experience in designing and developing Azure stream analytics jobs to process real time data using Azure Event Hubs.
- Experience with different cloud-based storage systems like S3, Azure Blob Storage, Azure Data Lake Storage Gen 1 & Gen2.
- Experience in implementing data pipelines using Azure Data Factory
- Worked on Data Warehouse design, implementation, and support (SQL Server, Azure SQL DB, Azure SQL Data warehouse, Teradata).
- Experience in implementing in ETL and ELT solutions using large data sets
- Expertise in querying and testing RDBMS such as Teradata, Oracle and SQL Server using SQL for data integrity.
- Extensive Experience in Oracle server for developing ETL and ELT architectures
- Proficient in Data Modeling Techniques using Star Schema, Snowflake Schema, Fact and Dimension tables, RDBMS, Physical and Logical data modeling for Data Warehouse and Data Mart.
- Technical expertise in ETL methodologies, Informatica 9.x/8.6/8.5/8.1 - Power Center, Power Mart, Client tools - Mapping Designer, Workflow Manager/Monitor and Server tools.

- Excellent communication and interpersonal skills, Proactive, Dedicated and Enjoy learning new Technologies and Tools.
- Strong commitment towards quality, experience in ensuring compliance to coding standards and review process.

CERTIFICATION:

- Microsoft Certified Azure Data Engineer Associate in Jan-2021

TECHNICAL SKILLS:

Cloud Providers

Azure\

**DevOps Tools **

Jenkins, Jfrog, Postman\

Tools

HDFS, Hadoop, Sqoop, JIRA, Maven, Github, \ Bitbucket, Hue, VS code, SQL Developer, \ Azure storage explorer, Azure data studio, \ Zeppelin, Ambari, oozie\

Databases

Azure Synapse, Hive, Oracle, Teradata and \ MS SQL Server\

ETL Tools

Informatica Power Center 8/9/10, SSIS\

Shell Scripting

Unix shell Scripting, Python\

Healthcare Domain

HIPAA, Claims and Members Process, EOBC, \ EPIC\

PROFESSIONAL EXPERIENCE:

Client: Confidential

Oct 2021 to Till Date

ETG (Episode Treatment Grouper), Social Determinants of Health, Thrive Local Data Engineer

The ETG (Episode Treatment Grouper) is a software application owned by Optum/Symmetry that categorizes episodes of medical care into predefined categories, using the ICD categories as a baseline.

An episode of care can be:

- Acute (e.g., bone fracture),
- Chronic (e.g., cancer, diabetes), or
- Acute and chronic (e.g., heart attack)

In addition to tracking utilization variables, the ETG also tracks the financial data associated with each category, used to compare/contrast costs among KP's various U.S. regions to observe trends and identify anomalies. Specifically, management scrutinizes cost-saving trends and anomalies to determine whether medical staff can leverage optimized treatment modalities from other regions.

Responsibilities:

- Worked on creating hive tables. Have knowledge and hands on experience in writing Hive SQL queries
- Created Internal and External tables in Hive. Worked on performance tuning on Hive database on partitioning.
- Familiar with sqoop commands. Used sqoop queries for import and export of data from hive database to oracle database.
- Converted the Hive sql queries to Impala sql queries. Compatible to run them in impala for faster execution
- Created batch log process to load the data from databases to Hive on daily basis and automated the scripts through oozie scheduler
- Hands-on experience in Unix shell scripting and Python scripting.
- Used bitbucket for version control on the source code. This helped in the code migration process.
- Created Jenkins pipelines on the existing java code application to build and deploy the code to the artifactory.
- Used github for saving source code and versioning. Have hands on experience on github commands
- Used JIRA for creating stories and development tasks for the code delivery process.
- Worked on a Devops project called FHIR. The idea is to receive the docker images from a third party vendor into our storage zone and process and produce the json files using Jenkins pipelines and load the data to the Azure SQL database.
- Used a postman application for storing the sensitive information in config. Have knowledge on Jfrog artifactory.
- Worked on On-premises for Azure Cloud migration project. Created the data pipelines for loading the data from the Flat files to hive database
- Created Data pipelines to load data from Hive database and ADLS to Azure synapse data warehouse.
- Designing and Developing Azure Data Factory (ADF) extensively for ingesting data from different source systems like relational and Non-relational to meet business functional requirements.
- Creating pipelines, data flows and complex data transformations and manipulations using ADF
- Ingested huge volume and variety of data from disparate source systems into Azure Data Lake Gen2 using Azure Data Factory V2.

Technologies:

Hadoop, Hive, Impala, Azure Data Factory, ADLS, Hue, Zeppelin, Unix, Python, Jenkins, GitHub, Stash, sqoop and Jira

Client: Confidential

Apr-2017 to Oct-2021

CDW (Claims Data Warehouse)

Data Engineer

The Claims Data Warehouse(CDW) program provides a central, consistent industry standard view of finalized claims data for all regions to support reporting, analytical and operational functions. The scope includes integrating all internal and external claims data where adjudicated claims are in final

status into the CDW and provide end users ability to generate reports from one central repository with common view of claims data across all regions.

Responsibilities:

- Worked as lead developer from onshore to process the claims data with the business logic provided.
- Used Oracle Data loader tool to process the data from around 500 flat files
- Design and approach method is prepared to load the data from Raw zone to Stage layer and then to reporting Layer.
- Analyzed the Design and prepared the Source to Mapping document.
- Used Huge unix scripts to process and load the data continuously to the database without failing.
- Performance enhancement process included loading the files to the database on a schedule of first come first load.
- Worked on Informatica ETL Tool to load the data from staging layer to the reporting layer.
- Worked and implemented the complex oracle joins and queries by using Analytical functions.
- Experience on the stored procedures in oracle
- Implemented Performance tuning and caching in oracle database
- Good knowledge in writing **HIVE COMPATIBILITY** sql queries(Oracle vs Hive)
- Knowledge in preparing Test data in non-prod environments. Worked as ETL code reviewer in the team and responsible for code deliveries to PROD within SLA.
- Experience in tools like servicenow, Tivoli scheduler and crontab

Technologies:

Informatica Power Center, Hive, Unix, Oracle, Ambari, python, Teradata, Erwin

Client: Kaiser Permanente - Hyderabad, Telangana, INDIA

Feb-2013 to Mar-2017

MDW – EBOC (Members Data Warehouse – Employee benefits part C)

Data Engineer

The Members Data Warehouse(MDW) program provides a central, consistent industry standard view of finalized Members data for all regions to support reporting, analytical and operational functions. The scope includes integrating all internal and external Members data where adjudicated claims are in final status into the MDW and provide end users ability to generate reports from one central repository with common view of claims data across all regions.

Responsibilities:

- Worked as lead developer from onshore to process the claims data with the business logic provided.
- Used Oracle Data loader tool to process the data from around 500 flat files
- Design and approach method is prepared to load the data from Raw zone to Stage layer and then to reporting Layer
- Analyzed the Design and prepared the Source to Mapping document
- Used Huge unix scripts to process and load the data continuously to the database without failing.
- Performance enhancement process included loading the files to the database on a schedule of first come first load.

- Worked on Informatica ETL Tool to load the data from staging layer to the reporting layer.
- Worked and implemented the complex oracle joins and queries by using Analytical functions.
- Experience on the stored procedures in oracle
- Implemented Performance tuning and caching in oracle database
- Knowledge in preparing Test data in non-prod environments. Worked as ETL code reviewer in the team and responsible for code deliveries to PROD within SLA.
- Experience in tools like servicenow, Tivoli scheduler and crontab

Technologies:

Informatica Power Center, Unix, Oracle, Teradata, Erwin

IMServ - Pune, INDIA

Oct-2011 to Jan-2013

ETL Developer

IMServ is an industry leader in providing carbon and energy management solutions that help organizations in all sectors to Save Energy, Reduce Costs, Control Carbon and Be Greener.

IMServ foresees data services to empower the customers to reduce emissions, cut bills and in turn help IMServ to benefit financially. To address this market, IMServ wants to provide an end to end energy management solution which the customers can use to analyze their carbon footprint, consumption and utility costs.

IMServ's energy management solution helps its partners and customers achieve new levels of energy management and carbon control. The energy management solution, which is a web-based reporting platform called EDV (Energy Data Vision) allows customers to analyze the data collected from meters and other instruments. It currently reports channel and meter data, which is obtained from its own database and a number of external systems either directly, or via STAR's data load and exchange components. The output of this system is consumed as services by internal users, external users and systems.

Responsibilities:

- Worked in Creating Informatica mappings using Flat files as Sources
- Developed Critical Informatica Mappings to populate data from tables to single Data warehouse.
- Imported Several relational and non-relational sources and targets into the informatica power center.
- Worked with flat file sources such as Delimited flat files (Comma delimited, Pipe delimited) and Fixed width flat files.
- Good Knowledge and Experience in creating Informatica objects such as Mappings, Workflows, Maplets and Worklets.
- Created Deployment groups and used them in migrating code from lower environments to higher environments.
- Worked in code migration by using the Labelling concept in informatica
- Good knowledge in creating the Informatica mappings that involves Slowly changing dimensions concepts – SCD Types
- Did Performance tuning on the existing code for long running sessions in informatica
- Worked in Development of Informatica code from the scratch till the production go-live.
- Worked in writing simple to complex Unix scripts in the project for handling several scenarios.
- Experience in writing Unix scripts for file processing – Archival process and Analyzing data.

- Exposure to Oracle database and DB2 databases, several database objects have been created in databases.
- Worked on preparing complex SQL queries for code validations
- Exposure to several flat file scenarios, date-time conversion logics, usage of flat files as lookup transformations.
- Worked on JAVA transformation for splitting the records scenario
- Knowledge on Informatica transformations – Normalizer, Transaction control, Rank, HTML and JAVA transformations.
- Worked on both static and dynamic lookup transformations
- Knowledge in preparing Test data in non-prod environments. Worked as ETL code reviewer in the team and responsible for code deliveries to PROD within SLA.
- Created the Informatica webforms using Tivoli scheduler properly without affecting the run flow.
- Prepared test-case documents SIT and UAT documents for each piece of code
- Used Platinum code reviewer tool to validate the informatica code by importing the XML code into the tool by setting up the static and dynamic rules in the tool.
- Created work orders, Incidents and change requests for getting software accesses, in case of any failures in the process, privileges issues, code migrations etc.
- Created one to one, one to many mappings by using Microsoft Visio templates in the informatica power center.

Technologies:

Informatica Power Center, Informatica Power Exchange, Unix, Oracle, DB2, Teradata, MS SQL server, TWS, Crystal Reports and BMC remedy

EDUCATION:

- **Bachelor of Technology in Civil Engineering**