Gyaneshwar

# Experience Summary

**Specialist Technology with 10+ years of professional experience On-Site (United States) as a Data Engineer in critical and large scale DW/BI and AWS cloud data lakes, Big Data, Data-warehouses** and data applications on AWS for large enterprises projects in domains like Insurance, Healthcare, Pharma-Retail, Telecommunication, Banking and currently worked in Insurance Domain Project

* Experience with a complete life cycle of BI/Data warehousing projects – Starting from dimensional data model design (Star Schema), Snowflake using Erwin, database creation (DDL, DML), Design and Development of ETL Architecture and Design and Development of Reports.
* Informatica Power Center specialist for designing and development of ETL, end to end project implementation, providing large volume enterprise solution for data migration, data transformation, and data management.
* Experience in Data Engineering to API data (using Spark, Kafka, streaming, etc)
* Contributor in all stages of (Development, UAT, Production), Database Management, creating complex mapping. Performance & tuning experience on huge multi-terabytes data warehouses.

## Extensive knowledge of BI Technologies and expertise with – Informatica Power Center, IICS, AWS, Redshift, Hive, Greenplum (An MPP Database), Oracle, Unix shell scripting, AWS Cloud, Redshift for Datawarehouse, S3, Hadoop, Map-Reduce, Spark, Athena, Presto, Workbench, Python, Window PowerShell.

* Performing the Administrator work likes to perform cleanup activity on Unix, database and save the huge space.
* Handled databases such as **Oracle, SQL Server, Greenplum, Teradata, Redshift** with large volumes of data and their effective usage within Informatica products with an emphasis on optimal performance.
* Job dependencies, scheduling, and data flow management using **Active Batch tool, Control-M,Tivoli**
* Work on CI/CD process, extensive work on automation through UNIX Shell script and Python Scripting of DW project.
* Hands-on experience in AWS Cloud-based Data Lake formation including Data profiling, Data Ingestion, Data Processing, Transformation and Data Virtualization.
* Experience of **Informatica Intelligence Cloud Service (IICS) ETL tool**, Cloud Data Integration(CDI), Application Integration and Data Quality (DQ).
* Tracking project progress, issues, reporting to stakeholders and deployment coordination. Impact analysis of configuration change in up steam and downstream applications. L1, L2 and L3 level ETL Operation and Support with SLA specifications.
* Project management and effort tracking in E-Jira using Scrum, Kanban, and Hybrid models.

# Professional Experience

* Specialist Technology at Synechron Technologies (On-Site), US (Alpharetta, GA) - 5-Oct-2019 to Till date.
* Technical Lead at Synechron Technologies, Pune India - 26-Mar-2017 to 30-Sep-2019 date.
* Sr. ETL Developer at Wipro Technologies in Credit Suisse, Pune India – 19-Aug-2016 to 24-Mar-2017.
* ETL (DW/BI) Developer at Netlink, Bhopal India– 4th FEB 2013 to 31st March 2016.

## Skills

* ETL Tool – Informatica Power Center 8.x,9.x,10.x and Informatica Integration Cloud Services (IICS) , Talend
* Databases – Redshift, RDS, Greenplum (An MPP database), Oracle, SQL Server, MySQL
* Scripting – Unix Shell, Python, Power shell
* AWS Redshift, Hive (HDFS). S3, Hadoop, Athena, Redshift Spectrum, EC2, RDS, DMS, Lambda, DynamoDB, Kinesis, SQS, Kafka, Aurora, Presto, Workbench, PySpark, DataBricks, Spark Scala.
* Others: GIT , BITBUCKET, TIWOLI (TWS ), Activebatch, Contorl M, Data Modeling – Erwin 7.x, Jenkins, JIRA, Bitbucket, Team-City, Teradata utilities ( BTEQ, Fast-Load, Multi-Load, Fast Export, T-Pump)

# Education

* **B.E. Prof. Ram Meghe institute of Technology and Research ( MAHARASHTRA ),** June 2009, Electronics and Telecommunication

1

# Project Experience

## Client: TransAmerica, Alpharetta GA, USA May-2023 – Till Date Project: Workplace Admin migration in cloud

The Transamerica Corporation is an American holding company for various life insurance companies and investment firms operating primarily in the United States, offering life and supplemental health insurance, investments, and retirement services. As Part of migration project, we are migrating servers and database from On-Prem to AWS cloud.

* Requirements gathering, impact analysis, estimation, and Planning.
* As workplace Informatica Power center admin, Installing the power center 10.5.3 and Power exchange into EC2 server.
* Created the configuration and handling the Unix server, DB2 LUW database installation.
* Created multiple shell scripts for automation (like One go installation, load check, performance check, health check, Repository check and integration check.
* Created the disaster recovery plan setup and provision for informatica server
* Handling P1,P2,P3 related issue, debugging, performance tunning and root cause analysis.

**Environment:** Linux/UNIX Shell Scripting, Informatica PowerCenter 10.5.3, Informatica Power Exchange, IICS (Informatica Intelligent Cloud Services), DB2 LUW, Sql-Server, Oracle

## Client: Morgan Stanley, Alpharetta GA, USA April-2022 – March-2023 Project: Wealth Management (Customer engagement management)

This Project is an integral part of enabling CEM (Customer Engagement Management) solution on Wealth Management (WM) platform using Salesforce Cloud. CEM in turn is building a block for Campaign management and a broader CRM platform. Salesforce (SF) Integration Components are meant to enable the flow of client and Financial Advisor (FA) data from WM Data Lake/Hadoop to Salesforce Financial Services Cloud. This is to provide entitlements and visibility to FAs to send campaigns to clients generated by Next Best Action (NBA) rule engine. Salesforce Marketing Cloud (SFMC) captures all clickstream data about action taken on campaigns sent to clients.

* Interacted with business users for requirement analysis of new development applications and to define business and functional specifications.
* Requirements gathering, impact analysis & estimation for development, Design. coding, testing, implementation, and user acceptance of software applications.
* Designed and developed Informatica ETL process and automated end to end ETL pipeline using Spark data frame and Spark SQL.
* Data is sourced from a variety of sources and loaded into Hadoop (Hive) and loaded into Salesforce.
* Cleansing of the data in client's provided flat files. Load the data in Hive table, creating type1/type2 script to fetch data from HDFS and loading it into hive tables.
* Identify, troubleshoot, and resolve DWH application problems to minimize Applications downtime and optimize the application performance. Proposed design notes & implemented process improvement techniques for the existing ETL routines.
* Created the Hadoop purge Framework independently with the help of Unix scripting.
* Created the CRM purge Framework and report Independently with the help of Informatica and Unix scripting.

**Environment:** Windows, UNIX, Informatica PowerCenter 10.2.0, IICS (Informatica Intelligent Cloud Services), Salesforce CRM, Hadoop (HIVE), Teradata, SQL Server, DB2, Greenplum, TWS Scheduling Tool, Jira, Service Now

## Client: Asurion, Nashville TN, USA March 2017 – March-22 Project: Enterprise Data Platform

Asurion, LLC is a privately held company that provides insurance for smartphones, tablets, consumer electronics, appliances, satellite receivers and jewelry. In 2019, the company operated in 23 countries and had 49 offices with 20,000 employees serving 305 million consumers.

I have worked in Enterprise Data Services team to cater different kinds of data need across organization for BI and Analytics solution.

## Major Role and Responsibilities:

* Designed and developed Data Ingestion framework and pipeline from MS SQL Server, Oracle DB, Flat files to AWS (Amazon Web Services) cloud with the ETL tool Informatica Power center and IICS ( CDI, DQ) for building a Data Lake to support intelligence and analytics capabilities.
* DataMart and Datawarehouse design and implemented in redshift and Hive Amazon S3 with help of technique Dimensional Modeling, Star, Snowflake and Hybrid schema design.
* Data Integration- extraction from SQL Server, Oracle, Flat files, and load with transformation into various data marts for business intelligence solutions. Generating ad-hoc data files and FTP to different destinations.
* Designed, developed, and implemented Custom Scripts using Unix and Python to manipulate data and for other specific uses like Flat File automated framework, Redshift to hive data unload automation utilities.
* Requirement gathering, creating Businesses Requirement and Functional Requirement Documents. Documenting all issues and progress.
* For API Ingestion streaming solution with the help IICS, Amazon Kinesis, SQS, and Kafka ingested the data into data lake, having an experience with REST and SOAP.
* Handling JSON via (IICS) ETL tool with the help of hierarchy parser and hierarchical builder.
* Call the API from CDI (Cloud Data Integration) , having exposure of Swagger file and REST V2 Conn.
* Created S3 Data lakes using Parquet columnar file formats and optimized the process of data scans using Athena & Redshift Spectrum and Experience with Web Services(Rest and SOAP)
* Created best practice modern data lake platform using Lake Formation, PySpark, Redshift, and S3. **Environment**: AWS, ETL Informatica 10.x, IICS, Talend, Database: Redshift, S3 hive, Oracle, SQL Server, RDS, MySQL, DynamoDB, Aurora UNIX: Shell scripting, Python, Databricks, Active Batch., Presto, Workbench.

## Client: Credit Suisse’s Aug-2016 to Mar-2017

**Project: Gateway**

Credit Suisse’s Investment Banking & Capital Markets division offers a broad range of investment banking products and services which include advisory services related to M&A, divestitures, and takeover defense strategies, as well as debt and equity offerings. We also offer derivative transactions related to these activities. Our clients include leading corporations, financial institutions, financial sponsors, ultra-high-net-worth individuals, and sovereign clients.

Working as Sr ETL developer Credit Suisse investment banking domain project, worked on Development, enhancement of already developed mapping, tuning, Automation and support to existing ETL projects.

## Major Role and Responsibilities:

* Development of simple and medium complex mappings.
* Development of Source to target Mapping Design Document
* Testing and Validation of pre-existing mappings.
* Performance improvement in the overall loading process in ETL by using Power Exchange connection with many limitations.
* Performance improvement existing database table and ETL job.
* Improve the file automation framework, without any manual support.

**Environment**: ETL: Informatica 9.x, Database: Oracle, UNIX: Shell scripting

## Client: Rexall Pharmaceutical Feb-2013 to Mar-2016

**Project: Katz EDW project**

Rexall is a member of the Katz Group of Companies; a Canada based privately-held organization. Rexall is a Canada’s most trusted pharmacy brand with more than 425 Rexall and Rexall Pharma Plus stores. Rexall also operates Rexall Specialty Pharmacy, a division focused on providing highly specialized clinical care and complex medication management services in transitional care settings. Transaction data from different applications of stores (POS, Nexxsys, JDA, Kroll) is provided to DWH. This data from different sources are Cleansed, Transformed, and Integrated and aggregated to be used by End users using BI Tool MicroStrategy for Adhoc and canned reporting.

**Description**: Working as an ETL developer worked on Development, production support, Production migration, enhancement of already developed mapping, tuning, testing and support to existing ETL projects.

## Major Role and Responsibilities:

Designed and Developed all ETL layers (Landing, Staging, Data Integration layer and Data ware house layer) with Type II dimension tables which helps users to identify the what (are the changes), who (made the changes), when (changes were made) in that dimension.

* Understand the requirement of ETL loading and define performance-oriented architecture so that every piece of developed ETL can perform well. Automate every possible ETL process so that minimum manual intervention is required.
* Created the entire ETL cycle (LND>STG>IDS>DWH) that use to populate data from Source to Data Mart.
* Creation of Data model with Batch id concept in each layer LND>IDS>DWH that one is best scenario of good practice of data model.
* Identification of bottlenecks and other improvement area in existing projects and perform necessary remediation.
* ETL Tuning, SQL Queries tuning, development of reusable components in Database/ETL, Improvements at report level by optimizing the variables/logical calculation etc.
* Worked on each and every stage of BI from development to till production and based on warranty period.
* Creating Adhoc mapping for Supply Chain and POS.
* SQL Script development and Unit testing for Initial load of Katz Data warehouse. Testing and Validation of pre-existing mappings of Katz project

**Environment**: Informatica 9.x, Database: Greenplum (An MPP database), Data Modeling: Erwin, Linix