**Professional Summary:**

* High-energy **Data Engineer** with 9+ years of experience in Analysis, Design, Development, Implementation and Support of various client server and decision support system environments with focus on **Data Warehousing** and **Business Intelligence**
* Strong **Ab Initio** experience with **Data mapping**, **Transformation** and **Loading** from Source to Target Databases and in developing strategies for Extraction, Transformation and Loading (ETL) mechanism using Ab Initio.
* **AWS Certified Developer – Associate.**
* Strong knowledge of Data Warehousing concepts and Dimensional modeling like Star Schema and Snowflake Schema. Experience with **Incremental** & **Full** load to Target databases
* Good experience in translating business requirements into transformations in AWS EMR.
* Experience in replacing ETL transformation to Cloud ETL transformation.
* Proficient in AWS services like VPC, EC2, S3, ELB, Autoscaling Groups (ASG), EBS, RDS, SNS/SQS, IAM, CloudFormation, Route 53, CloudWatch, CloudFront, CloudTrail.
* Loaded data from AWS S3 to Snowflake and used NoSQL for querying the data.
* Good experience working with very large databases and **Performance tuning**.
* Experience in providing **production support** to various ETL jobs and developing various **UNIX shell scripts** to run ETL and Data base jobs.
* Good Experience working with various Heterogeneous Source Systems like **Oracle, MS SQL Server, Flat files** and **Legacy Systems.**
* Worked with **Transformations** such as **Aggregate, Sorter, Filter, Expression, Join, Normalize and Scan Components.**
* Knowledge about developing prototype for parsing high volume XML files using Hadoop and storing output in HDFS and HIVE for Ab Initio.
* Experience with Ab Initio Co-Operating System, application tuning, and debugging strategies.
* Expert in writing UNIX Shell Scripts including Korn Shell Scripts, Borne Shell Scripts.
* Very good experience with SQL Server & Oracle Databases.
* Good understanding of Apache Hadoop ecosystem components like HDFS, Hive, Sqoop.
* Experienced in handling different file formats like Text file, Avro data files, Sequence files, Xml and Json files.
* Good knowledge with Sqoop to export analyzed data from HDFS environment into RDBMS for report generation and visualization purpose.
* Worked in Agile/Scrum environment and agile tools like Rally and Jira board
* Experienced in all phases of Software Development Life Cycle (**SDLC**).
* Experience in **Data Modeling**, **Data Extraction**, **Data Migration**, **Data Integration**, **Data Testing** and **Data Warehousing**.
* Experience working in **Financing and Banking Domain** **and Retail Domain**.
* Strong hands on experience using Teradata utilities (SQL, B-TEQ, Fast Load, MultiLoad, FastExport, Tpump, Visual Explain, Query man), Teradata parallel support and Unix Shell scripting.
* Proficient in coding of optimized Teradata batch processing scripts for data transformation, aggregation and load using BTEQ.
* Experience in developing Entity-Relationship diagrams, Logical and Physical Data Models, and the use of Data modeling tools.
* Excellent analytical & problem solving and business interaction skills.
* Excellent Written and verbal communication skills.
* Worked in Agile methodology.

**Technical Skills:**

|  |  |
| --- | --- |
| **ETL Tools** | Ab Initio, Informatica PowerCenter |
| **AWS Tools** | S3, Athena and AWS EMR, NoSQL DBs(RedShift) etc. |
| **Data Modeling** | Erwin 7.0/5.1 |
| **Databases** | Oracle 12c/11g/10g, MS SQL Server 2005/2000, OLTP, Teradata |
| **Programming** | SQL, shell scripting, Python |
| **Environment** | UNIX, Windows. |

**Education**:

Master’s in Computer Science, Northwestern Polytechnic University, Fremont, CA, USA.

Bachelor of Engineering in Computer Science, JNTU, India.

**Professional Experience:**

**Client: Kroger OCT 2021 – Till Date**

**Role: Data Engineer.**

**Responsibilities**:

* Preparation of Functional Design documents and development tasks for **ETL**.
* Translating business rules and functional requirements into **ETL** procedures
* Development of mappings, workflows, session and tasks using **Ab Initio**
* Examining the session logs if the workflows fail and debugging mappings using Debugger.
* Participated in testing Data matching between source to target Databases.
* Worked on transformations like Lookup, Join, Filter, Reformat and Scan etc.
* Involved in Data loading to staging tables and from staging tables to Target tables i.e. Data loading from source to target Database.
* Testing the codes and deploying it to the QA environment for UAT.
* Perform system and requirement analysis, coding, unit testing of load into EDW.
* Code promotion and working on Permanent fixes of code as a part of production implementation.
* Worked as part of enhancements team.
* Unit testing the reports, debugging and fixing the errors.
* Supported SIT/UAT in preparing test plans and working with them if any defects are raised.
* End to End testing of Reports for data and functional behavior.
* Providing support and fixing issues related to Data on Reports and data mappings
* End to End testing of Reports for data and functional behavior.
* Configuring **BMC** **Control M** scheduler to execute the ETL workflows and monitoring daily ETL data load.
* Involved in graph migration from DEV to QA, QA to Prod.
* Involved in mock deployment activity prior to the main release activity where the code is migrated to new environment.
* Experience in working in production environment during major outages.
* Worked in the production Team as a POC (Production On Call).
* Monitored via BMC Control-M over 2000+ daily jobs, weekly, quarterly, monthly and request jobs in production Enterprise Data Warehouse (EDW), making sure that jobs are running smoothly (no long-running or delayed jobs) and meeting deadline as per Service Level Agreement (SLA);
* Immediate action on abends and performed extraction of error/log details, and primary analysis of job abends and fixing job abends.
* Coordinated with/escalated to Level 3 developers, DBAs, data analysts, schedulers, or other on-call staff to diagnose and rectify if any serious issues;
* Tracked and updated emergency change management tickets when job failures necessitate change in ETL codes or deleting duplicate entries to resolve, following up for approvals, and closing where appropriate

**Client:** **Discover Financial Services, Riverwoods, IL Sep 2020 – Sep 2021**

**Role: Data Engineer.**

**Responsibilities**:

* Worked on analysis and design part of business requirements and business rules based on given preliminary design documentation and working closely with tech leads and data analysts in understanding the current system.
* Developed Abinitio graphs for the extraction of data matching various business rules from different files. Daily jobs were developed, thoroughly tested on the test data, and then were run on production data. The output of these graphs would mainly be mainframe files.
* Automated some existing applications.
* Worked on Creating the Detailed Design Addendum once the Applications are developed.
* Created the deployment documents and run books for generic jobs.
* Designed SCD type 2-dimension data flows for data warehouse and designed Fact data flows to Data warehouse.
* Worked on extracting data from the archer tool and W-desk data to transform it as per the business requirements and loading into data warehouse.
* Support and Maintenance of more than five ETL applications in production environment.
* Well versed in recovering applications for any technical failure or due to data issue.
* Involved in performance tuning of Ab Initio graphs.
* Provided on-call support.
* Coordinated with source, downstream and business teams with respect to data issues.
* Acting as Subject Matter Expert (SME) for 5 critical applications.
* Worked on generic AWS load scripts to move data from AWS S3 to Teradata, Teradata to SNOWFLAKE as part of data migration.
* Recreating Teradata SQL to Snowflake SQLS.
* Have experience working in onsite-offshore model.

Environment: Ab Initio (GDE 3.3.2 Co>Op Sys 3.3.2.1, Control Center), EME, Oracle 12c, Linux, Terdata, SnowFlake, AWS S3, EC2, EMR.

**Client: JPMorgan Chase & Co, Columbus , OH Oct 2019 – Aug 2020.**

**Role: Sr. Ab Inito Develpoer / Data Engineer**

**POC on AWS**: I have done the POC to integrate the AWS into Chase environment.

* Set up and built AWS infrastructure with various services available by writing cloud formation templates
* Implemented Amazon RDS multi-AZ for automatic failover and high availability at the database tier.
* Implemented CloudTrail to capture the events related to API calls made to AWSinfrastructure.
* Implemented Amazon RDS multi-AZ for automatic failover and high availability at the database tier.
* With the help of IAM created roles, users and groups and attached policies to provide minimum access to the resources.
* Created topics in SNS to send notifications to subscribers as per the requirement.
* Collaboration with DeVOps team and created CI/CD tool Jenkins pipeline.
* Created EBS volumes for storing application files for use with EC2 instances whenever they are mounted to them.
* Built S3 buckets and managed policies for S3 buckets and used S3 bucket and Glacier for storage and backup on AWS.
* Management and Administration of AWS Services CLI, EC2 , S3 , Route 53 , IAM , and Trusted Advisor services.
* Initiating alarms in **CloudWatch** service for monitoring the server's performance, **CPU** Utilization, disk usage etc. to take recommended actions for better performance.
* Configured AWS Multi Factor Authentication in IAM to implement 2 step authentication of user's access using PING ID and AWS Virtual MFA.
* Designed, built, and coordinate an automated build & release CI/CD process using Gitlab, Jenkins and Puppet on hybrid IT infrastructure.

**Responsibilities:**

* Involved in analysis of end user requirements and business rules based on given preliminary design documentation and working closely with tech leads and data analysts in understanding the current system.
* Used Ab Initio GDE to generate complex graphs for transformation and loading of data into Staging and Target Database areas for Chase Applications.
* Used Conduct>IT like Plan, Program, Graph, Conditinal Task’s and used Signal Event, Wait for Event and wait for file.
* Created a generic Ab initio job for automation testing like compare source and target tables, duplicate check in target.
* Worked on Extreme Dialogue to build the account statements for chase account holders.
* Enhanced the Existing Application with new changes as per the Business requirements.
* Worked on Creating the Design Addendum once the Applications are developed.
* orked on the Production Enhancement fixes and High priority Immediate fixes for issues in Production.
* Created the detail design documents and run books for generic jobs.
* Developed Ab Initio jobs for loading and extracting data from various data stores like DB2, Oracle.
* Designed SCD type 2-dimension data flows for data warehouse.
* Performed Unit Testing and responsible for the promotions to QA environment.
* Used PDL and Shell interpretations to resolve the graph level parameters.
* Worked with Production Service and Automation teams to debug issues in production and for migrations from Pre-prod to Production.
* Involved in Load testing in different environments and System Integration Testing (SIT) and User Acceptance Testing (UAT).
* Used Ab Initio Control Center as a job scheduler to automate the daily, weekly and monthly jobs.
* Analyzed schedules in Control-M to check for dependencies, modified to add dependencies on new jobs using xmls.

**Environment:** Ab Initio (GDE 3.3.2 Co>Op Sys 3.3.2.1, Control Center), EME, Oracle 12c, Linux, PL/SQL, Dialogue, AWS S3, EC2, EMR.

**Client: Capital One Financial Corporation, Richmond, VA May 2018 – Sep 2019**

**Role: Data Engineer.**

**Responsibilities:**

* Developed and Implemented extraction, transformation and loading the data from the legacy systems using Ab Initio to replacing ETL transformation to Cloud ETL transformation (ECE modernized system)
* Knowledge on Amazon Web Services, to load the target S3 buckets and Snowflake tables.
* Involved in creating the Snowflake loader script, which is used to load data to snowflake tables from S3 buckets.
* Created detail data flows with source to target mappings and convert data requirements into low level design templates.
* Responsible for data cleansing from source systems using Ab Initio Components such as join, dedup sorted, Denormalize, Normalize, Reformat, Filter by expression, Rollup
* Worked with departition Components like Concatenate, Gather, Interleave and Merge to departition and repartition data from Multifils accordingly.
* Worked with Partition Components like Partition by key, Partition by Expression, Partition by Round Robin to partition the data from serial file.
* Worked on Generic graphs for data validation and data transformation.
* Involved in System and Integration testing of the project.
* Tuning of Ab Initio graphs for better performance.
* Used phases/checkpoints to avoid deadlocks to improve efficiency.
* Gathering the knowledge of existing operational sources for future enhancements and performance optimization of graphs.
* Used UNIX environment variables in all the Ab Initio graphs, which comprises of specified locations of files to be used as source and target.
* Developed shell scripts to automate file manipulation and data loading.
* Replicate operational table into staging tables, transform and load data into warehouse tables using Ab Initio GDE.

**Environment**: Abinitio Co >Op, GDE, Oracle 12c, SQL, PL/SQL, Shell Scripting, Control- M, Arow, AWS S3, Snowflake.

**Client:** **Discover Financial Services, Riverwoods, IL Nov 2016 – April 2018**

**Role: Sr. AbIntio developer.**

**Responsibilities:**

* Understanding the Business requirements based on Functional specification to design the ETL methodology in technical specifications.
* Involved in requirement analysis, ETL design and development for extracting data from the heterogeneous source systems like Oracle, Teradata, flat files, and loading into Staging and Data Warehouse Star Schema.
* Involved in massive data cleansing prior to data staging.
* Designed and developed logic for handling Slowly Changing Dimension Table Type 1/ Type 2.
* Extensively used Ab initio client tools Graphical Development Environment, Co-ops, and Enterprise Metadata Environment (EME).
* Developed graphs for the ETL processes using Join, Rollup, Scan, Normalize, Denormalize and Reformat transform components as well as Partition and Departition components extensively.
* Involved in extracting the data from the Flat Files and Relational databases into staging area.
* Worked with UNIX commands in command prompt, created and executed different types of shell scripts.
* Worked on improving the performance of Ab Initio graphs by using Various Ab Initio performance techniques like using lookup s (instead of joins), In-Memory Joins and rollups to speed up various Ab Initio Graphs.
* Implemented slowly changing dimension methodology for accessing the full history of accounts.
* Used debugger in identifying bugs in existing graphs by analyzing data flow, evaluating transformations.
* Optimizing performance tuning at source, target, and graph and plan level.
* Modified the Ab Initio graphs to utilize data parallelism and thereby improve the overall performance to fine-tune the execution times.
* Implemented Lookup's, lookup local, In-Memory Joins and rollups to speed up various Ab Initio Graphs.
* Worked with SQL Joins, Materialized Views, Indexes and written PL/SQL procedures and functions for processing business logic in database
* Used continuous flow points like subscribe, publish, continuous role up, continuous scan etc.
* Moved the data in the enterprise like a queue system using continuous flows.

**Environment:** Ab Initio (GDE 3.1.5 Co>Op Sys 3.1.5), EME, Oracle 11g, Linux, Teradata V16, Meastro, PL/SQL, Share point.

**Client: JPMorgan Chase & Co, Columbus, OH Feb 2016 – Oct 2016**

**Role: Ab Initio developer.**

**Responsibilities:**

* Analyzed business needs and documented functional and technical specifications based upon user requirements with extensive interactions with business users
* Extensively worked with the Ab Initio Enterprise Meta Environment EME to obtain the initial setup variables and maintaining version control during the development effort.
* Worked with data modeler to create Logical and Physical data model
* Extracted data from various sources like databases Teradata, Oracle, delimited flat files and Salesforce.com
* Designed and developed parameterized generic graphs
* Designed Ab Initio graphs that would harness Teradata capabilities ELT and balance resources between database and Ab Initio
* Analyzed SQL explain plans to determine optimization techniques join strategy, indexes, collect stats etc.
* Worked on how to bring the data from salesforce.com and load into .CSV or Flat files.
* Developed prototype for parsing high volume XML files using Cloudera Big Data and storing output in HDFS and HIVE for Ab Initio.
* Responsible for the detailed design and documentation.
* Used Informatica designer for designing mappings and mapplets to extract data from various sources like **Oracle, flat files, and XML files.**
* Extensively used Source Qualifier Transformation to filter data at Source level rather than at Transformation level. Created different transformations such as Source Qualifier, Joiner, Expression, Aggregator, Rank, Lookups, Filters, Stored Procedures, Update Strategy and Sequence Generator.
* Designed complex mappings with a slowly changing dimensions (type2: version) to keep track of current and historical data.
* Experience in the usage of complex sql queries, cursors in pl/sql blocks, and bulk load processing.
* Used Debugger to test the data flow and **fix the mappings**.
* Partitioned Sessions for concurrent loading of data into the target tables.
* Tuned the workflows and mappings.
* Involved in identifying the bottle necks and tuning to improve the Performance.
* Created workflows using Workflow manager for different tasks like **sending email notifications**, **timer that triggers** when an event occurs, and sessions to run a mapping.
* Executed Workflows and Sessions using Workflow Monitor.
* Used mapplets, parameters and variables to facilitate the reusability of the code.
* Coordinated efforts with different teams and stakeholders
* Developed, tested and reviewed complex Ab Initio graphs, sub-graphs, DML, Psets, XFR, deployed scripts, DBC files for connectivity, create Package and exports.

**Environment:** Ab Initio, Informatica PowerCenter, Teradata, UNIX, Ctrl-M scheduler, Oracle 10g, Salesforce.com, SAS Marketing Automation.

**Client: Harte Hanks, Burlington, MA June 2015 – Jan 2016**

**Role: Ab Initio Developer**

**Responsibilities:**

* Developed and Implemented extraction, transformation and loading the data from the legacy systems using Ab Initio
* Metadata mapping from legacy source system to target database fields and involved in creating Ab Initio DMLs and written complex XFRs for implementing the business logic transformation.
* Implemented various levels of parameter definition like project parameters and graph parameters instead of start and end scripts.
* Extensively Used Transform Components Aggregator, Match sorted ,Join, Denormalize sorted, Reformat, Rollup and Scan Components
* Involved in Unix Korn shell wrapper scripts to accept parameters and scheduled the processes using Crontab, Job Scheduler, Database Load Interface, and Denormalization using Ab Initio.
* Worked with EME / sandbox for version control and did impact analysis for various Ab Initio projects across the organization.
* Developed Ab Initio scripts for data conditioning, transformation, validation and loading.
* Gathering the knowledge of existing operational sources for future enhancements and performance optimization of graphs.
* Assisted in developing various reports using PL/SQL Stored Procedures.
* Involved in writing wrapper scripts to run the graphs, to load into Data Warehouse and to verify the counts while loading.
* Used Parallelism techniques to partition and process large data simultaneously
* Created Auto Sys job stream to schedule jobs by creating box jobs and jil templates.
* Developed shell scripts to automate file manipulation and data loading.
* Replicate operational table into staging tables, transform and load data into warehouse tables using Ab Initio GDE.
* Coordinated with software programmers and users for utilization of application resources.
* Conducted investigation and analysis of technical problems relating to databases and software applications.
* Implemented procedures for testing of new releases and related installation of systems.
* Assisted in installation, assembly and patching of Exadata and Oracle applications.
* Suggested appropriate processes for tuning and upgrading of databases and relevant components.
* Provided technical assistance for management and execution of database backup and recovery services.
* Supported technical team members in administration and control access services on database servers.
* Executed effective processes for integration of database support services with overall system applications.

Environment: Ab Initio GDE 3.1.6,Co-op , Oracle PL/SQL and Exadata, Autosys ,Toad 7.6.0.11, HP UX 11.x, IBM PCOM 4.2, TSO/ISPF, Korn Shell Scripting.

**Client : IMS Health, Hyderabad, India March 2012 – Dec 2013**

**Role: ETL Informatica Developer**

**Responsibilities:**

* Understanding of the business requirements and enhancing the existing data warehouse architecture design for a better performance.
* Analyzed business requirements and worked closely with the various application teams and business teams to develop ETL procedures that are consistent across all application and systems.
* Used Informatica designer for designing mappings and mapplets to extract data from various sources like Oracle, flat files and XML files
* Extensively used Source Qualifier Transformation to filter data at Source level rather than at Transformation level. Created different transformations such as Source Qualifier, Joiner, Expression, Aggregator, Rank, Lookups, Filters, Stored Procedures, Update Strategy and Sequence Generator.
* Implemented SCD methodology including Type 1 and Type 2 changes to keep track of historicaldata.
* Created PL/SQL Stored procedures and implemented them through the Stored Procedure transformation.
* Extensively worked on loading dimensions and fact tables in a star schema.
* Wrote complex SQL queries (Joins, sub queries, aggregations etc.) for data analysis, validate data that was populated by ETL applications.
* Implemented the Incremental loading of Dimension and Fact tables.
* Involved in identifying the bottle necks and tuning to improve the Performance.
* Created workflows using Workflow manager for different tasks like sending email notifications, timer that triggers when an event occurs, and sessions to run a mapping.
* Executed Workflows and Sessions using Workflow Monitor.
* Involved in Unit testing, System testing to check whether the data loads into target are accurate, which was extracted from different source systems according to the user requirements.
* Migrated repository objects, services and scripts from development environment to production environment.

**Environment:** Informatica PowerCenter 8.6.1, oracle 11g, SQL, PL/SQL, Windows.