**Sai Puppala**

**Mobile: 612-666-0788**

**Email: puppala0789@gmail.com**

**SUMMARY**

* Over 7+ years of progressive hands-on experience in analysis, ETL processes, design and development of enterprise level data warehouse architectures, designing, coding, testing, and integrating ETL.
* Experience in Dimensional data modelling techniques, Slow Changing Dimensions (SCD), Software Development Life Cycle (SDLC) (Requirement Analysis, Design, Development & Testing), and Data warehouse concepts - Star Schema/Snowflake Modelling, FACT & Dimensions tables, Physical & Logical Data Modelling.
* Experienced in integration of various data sources like Oracle 11g, 10g/9i/8i, MS SQL Server, XML files, Teradata, Netezza, Sybase, DB2, Flat files, into staging area and different target databases.
* Expertise in the ETL Tool Informatica and have extensive experience in Power Center Client tools including Designer, Repository Manager, Workflow Manager/ Workflow Monitor.
* Extensively worked with complex mappings using various transformations like Filter, Joiner, Router, Source Qualifier, Expression, Union, Unconnected / Connected Lookup, Aggregator, Stored Procedure, XML Parser, Normalize, Sequence Generator, Update Strategy, Reusable Transformations, User Defined Functions, etc.
* Extensively worked on Relational Databases Systems like Oracle11g/10g/9i/8i, MS SQL Server, Teradata and Source files like flat files, XML files.
* Worked with Informatica Data Quality 9.6.1 (IDQ) toolkit, Analysis, data cleansing, data matching, data conversion, exception handling, and reporting and monitoring capabilities using IDQ 9.6.1.
* Experienced in mapping techniques for Type 1, Type 2, and Type 3 Slowly Changing Dimensions
* Strong experience in SQL, PL/SQL, Tables, Database Links, Materialized Views, Synonyms, Sequences, Stored Procedures, Functions, Packages, Triggers, Joins, Unions, Cursors, Collections, and Indexes in Oracle.
* Experience on Netezza Performance Server (NPS) and Netezza Tools, Experience involved in Performance Tuning, Error Handling, Production Support and debugging mappings.
* Sound knowledge of Linux/UNIX, Shell scripting, experience in command line utilities like pmcmd to execute workflows in non-windows environments.
* Designed complex Mappings and have expertise in performance tuning and slowly-changing Dimension Tables and Fact tables.
* Expertise in Informatica Data Quality (IDQ) tool installation and Standardization, Address Validation and Deduplication.
* Good knowledge of Hadoop, HBase, Pig, Hive, Hadoop, HDFS, Big Data.
* Have hands on experience in tools like Address Doctor which is used for Address validations
* Good Experience in working with Teradata utilities like (BTEQ, FASTLOAD, FASTEXPORT, MULTILOAD).
* Extensive experience using database tool such as SQL \*Plus, SQL \*Developer, Autosys and TOAD.
* Proficient in Oracle Tools and Utilities such as TOAD and SQL\*Loader.
* Identified and fixed bottlenecks and tuned the complex Informatica mappings for better Performance.
* Excellent analytical, problem solving, technical, project management, training, and presentation skills.

**TECHNICAL SKILLS**

|  |  |
| --- | --- |
| ETL Tools | Informatica Power Center 9.x/8.x/7.x (Source Analyzer, Repository Manager, Transformation Developer, Mapplet Designer, Mapping Designer, Workflow Manager, Workflow Monitor, Warehouse Designer and Informatica Server), Informatica Data Quality (IDQ). |
| Databases | Oracle12C/11g/10g/9i/8i/8.0 Teradata14.1, DB2 UDB 8.1, MS SQL Server 2008/2005, Netezza 7.1 |
| Languages | SQL, PL/SQL, XML, UNIX Shell Scripting, Python Scripting |
| Operating Systems | Windows 95/98/NT/2000/XP, UNIX, Linux, NCR MP-RAS UNIX |
| Data Modeling | Dimensional Data Modelling, Star Schema Modelling, Snow Flake Modelling, FACT, Dimensions, Relational Modelling, Physical, Logical Data Modelling, and ER Diagrams. |
| Tools/Utilities | PLSQL Developer, TOAD, Hadoop , Hive, Pig, SQOOP, SQL Developer, Erwin, ER Studio, Microsoft Visio, Talend, DataStage, Mainframes |

**PROFESSIONAL EXPERIENCE**

**Client: RX Crossroads, Louisville, KY Feb 2017 - Till Date**

**Role: Sr. Informatica/ETL Developer**

**Responsibilities:**

* Involved in gathering and analysis business requirement and writing requirement specification documents and identified data sources and targets.
* Communicated with business customers to discuss the issues and requirements.
* Designed, documented and configured the Informatica MDM Hub to support loading, cleansing of data.
* Involved in implementing the Land Process of loading the customer/product Data Set into Informatica MDM from various source systems.
* Worked on data cleansing and standardization using the cleanse functions in Informatica MDM.
* Developed ETL programs using Informatica to implement the business requirements.
* Developed several IDQ complex mappings in Informatica a variety of Power Center, transformations, Mapping Parameters, Mapping Variables, Mapplets & Parameter files in Mapping Designer using Informatica Power Center.
* Profiled the data using Informatica Data Explorer (IDE) and performed Proof of Concept for Informatica Data Quality (IDQ)
* Used Informatica Power Center to load data from different data sources like xml, flat files and Oracle, Teradata, Salesforce.
* Migrated servers from 1and1 to AWS Elastic Cloud (EC2), databases (RDS).
* Developed the SQL scripts for the business rules using Unix Shell and NZSQL for Netezza.
* Imported the IDQ address standardized mappings into Informatica Designer as a Mapplets.
* Utilized of Informatica IDQ to complete initial data profiling and matching/removing duplicate data.
* Used relational SQL wherever possible to minimize the data transfer over the network.
* Extracted data from Oracle and SQL Server then used Teradata for data warehousing.
* Stored data from SQL Server database into Hadoop clusters which are set up in AWS EMR.
* Identified and validated the Critical Data Elements in IDQ.
* Built several reusable components on IDQ using Standardizers and Reference tables which can be applied directly to standardize and enrich Address information.
* Effectively worked in Informatica version based environment and used deployment groups to migrate the objects.
* Used debugger in identifying bugs in existing mappings by analyzing data flow, evaluating transformations.
* Worked on CDC (Change Data Capture) to implement SCD (Slowly Changing Dimensions) Type 1 and Type 2
* Extensively used various active and passive transformations like Filter Transformation, Router Transformation, Expression Transformation, Source Qualifier Transformation, Joiner Transformation, and Look up Transformation, Update Strategy Transformation, Sequence Generator Transformation, and Aggregator Transformation.
* Designed workflows with many sessions with decision, assignment task, event wait, and event raise tasks, used Informatica scheduler to schedule jobs.
* Used the Teradata fast load utilities to load data into tables
* Used SQL tools like TOAD to run SQL queries and validate the data.
* Converted all the jobs scheduled in Maestro to Autosys scheduler as the per requirements
* Worked on maintaining the master data using Informatica MDM
* Wrote UNIX Shell Scripts for Informatica Pre-Session, Post-Session and Autosys scripts for scheduling the jobs (work flows)
* Performed tuning of queries, targets, sources, mappings, and sessions.
* Used Linux scripts and necessary Test Plans to ensure the successful execution of the data loading process.
* Worked on AWS Data Pipeline to configure data loads from S3 to into Redshift. Have used AWS components (Amazon Web Services) - Downloading and uploading data files(with ETL) to AWS system using S3 components.
* Worked with the Quality Assurance team to build the test cases to perform unit, Integration, functional and performance Testing.

**Environment:** Informatica Power Center 10.1, UNIX, SQL, IDQ, IDE, CDC, MDM, Java, Linux, Perl, AWS, WINSCP, Shell, PL/SQL, Netezza 7.1, Teradata, Microsoft SQL Server 2008, and Microsoft Visual studio.

**Client: Intermountain HealthCare – Salt Lake City, UT Jan 2016 to Feb 2017**

**Role: Sr. Informatica/ETL Developer**

**Responsibilities:**

* Worked in Agile development methodology environment and Interacted with the users, Business Analysts for collecting, understanding the business requirements.
* Worked on building the ETL architecture and Source to Target mapping to load data into Data warehouse.
* Involved in the installation and configuration of Informatica Power Center 10.1 and evaluated Partition concepts in Power Center 10.1
* Developed various mappings using Mapping Designer and worked with Aggregator, Lookup, Filter, Router, Joiner, Source Qualifier, Expression, Stored Procedure, Sorter and Sequence Generator transformations.
* Created stored procedures, views, user defined functions and common table expressions.
* Generated underlying data for the reports through SSIS exported cleaned data from Excel Spreadsheets, Text file, MS Access and CSV files to data warehouse.
* Managed the Metadata associated with the ETL processes used to populate the Data Warehouse.
* Involved in IDS Services like building Business logics, analyzing the structure and data quality, creating a single view of data etc.
* Worked on Informatica cloud for creating source and target objects, developed source to target mappings.
* Involved in importing the existing Power center workflows as Informatica Cloud Service tasks by utilizing Informatica Cloud Integration.
* Involved in Data integration, monitoring, auditing using Informatica Cloud Designer.
* Worked on Data Synchronization and Data Replication in Informatica cloud.
* Worked with Data Architect Tools like Embarcadero ER Studio to maintain the MDM Data models and also and also managed database objects like triggers, stored procedures, sequences and indexes containing PL/SQL in the database and monitored overall performance management.
* Written PL/SQL scripts, created stored procedures and functions and debugged them.
* Created Mapplets, reusable transformations and used them in different mappings. Used Type 2 SCD mappings to update slowly Changing Dimension Tables.
* Involved in Production Support by performing Normal Loads, Bulk Loads, Initial Loads, Incremental Loads, Daily loads and Monthly loads and Developed reports based on issues related to the data warehouse.
* Used different Informatica Data Quality transformations in the Developer and Configured match properties match paths, fuzzy match key, fuzzy and exact match columns.
* Experience in optimization and query performance tuning in Netezza using various techniques involving distribution analysis, data-type analysis, zone maps etc.
* Created profiles, rules, scorecards for data profiling and quality using IDQ.
* Used Informatica Data Quality for addresses and names clean-ups and developed error handling & data quality checks to pull out the right data
* Used IDQ to cleanse and accuracy check the project data, check for duplicate or redundant records.
* Used debugger to test the mapping and fix the bugs and identified the bottlenecks in all levels to tune the performance and Resolved the production support tickets using remedy.
* Developed monitoring scripts in UNIX and moved Data Files to another server by using SCP on UNIX platform.
* Extensively used Teradata Utilities like Fast-Load, Multi-Load, BTEQ & Fast-Export.
* Created Teradata External loader connections such as M Load, Upsert and Update, Fast load while loading data into the target tables in Teradata Database.
* Involved creating the tables in Teradata and setting up the various environments like DEV, SIT, UAT and PROD.

**Environment:** Informatica Power Center 10.1, Oracle12C, Informatica Cloud, IDS 9.6.1, IDQ 9.6.1 Teradata 14.0, SQL Server 2014, Teradata Data Mover, Autosys Scheduler Tool, Netezza, Data modeling (Embarcadero ER Studio), UNIX, Toad, PL/SQL, SSIS, Power Connect, DB2, Business Objects XI3.5.

**Client: Vanguard, Malvern, PA Jan 2015 to Dec 2015**

**Role: ETL Developer**

**Responsibilities:**

* Worked on Power Center client tools to perform ETL to extract, transform and load data from Oracle, SQL Server, and flat files to Netezza.
* Designed and Developed the ETL (Informatica Power Center 9.x) solutions using Power Center Designer, Workflow Manager, Repository Manager, Workflow Monitor, Syncsort (DMExpress), DB2 9.x, Oracle 10g, Teradata Fast load/Multiload Utilities, PL/SQL Procedures, Unix Shell Scripts, Hyperion Interactive Reporting Studio 11.x and ER Studio Data Architect Tools.
* Created various transformations such as aggregator, Update Strategy, Look Up, Joiner, Filter and Router Transformations. Created and Monitored Workflows using Workflow Manager and Workflow Monitor.
* Performance Tuning of Sessions and Mappings.
* Involved in Unit Testing & Integration testing of the mappings.
* Created Relational Connections, Migrated the Mappings from Dev to Test and Test to Production.
* Involved in writing shell scripts by connecting to Database to perform various database operations.
* Developed, implemented and maintained Universes using Business Objects Designer.
* Worked with NZLoad to load flat file data into Netezza tables.
* Created and Configured Workflows, Worklets, and Sessions to transport the data to target warehouse Netezza tables using Informatica Workflow Manager
* Used Informatica Developer IDQ tool to provide data quality and Metadata management solutions.
* Created test plans, test data for extraction and transformation processes and resolved data issues following the data standards.
* Experienced in using IDQ tool for profiling, applying rules and develop mappings to move data from source to target systems.
* Used Stored Procedures as source for Business Objects querying, analysis and reporting.
* Created complex reports and published on Web using Web I. Created Ad-hoc reports using Business Objects functionalities like Queries, Slice and Dice, Drill Down, @Functions, Cross Tab, Master Detail and Formulas etc.
* Experience in Web publishing and scheduling documents using Broadcast Agent.
* Extensively involved in writing triggers, cursors, procedures, functions, loops in Oracle 9i/8i for various purposes and optimizing the SQL to improve performance.

**Environment:** Informatica PowerCenter 9.6.1/9. 5.1, Informatica developer IDQ 9.1/9.5.1, 7.1.4, Oracle 9i, SQL, PL/SQL, Teradata, MS SQL server, XML, SQL\* Loader, Shell Scripts, Netezza SQL, TOAD, Netezza SQL Remedy Incident Management, Data modeling on (Erwin & Embarcadero ER Studio), Windows XP, Linux, Informatica Scheduler, Business Objects XIR2 (desktop and WEBI).

**Client: Tech Mahindra, Mumbai –India May-2012 to Jun 2014**

**Role: ETL Developer**

**Responsibilities:**

* Conducted source System Analysis and developed ETL design document to meet business requirements.
* Developed Informatica Mappings and Workflows to extract data from PeopleSoft, Oracle, CSV files to load into Teradata staging area using Fast Load/Tpump utilities.
* Developed ETLs to load data from source to 3NF, stage to 3NF and Stage area to Work, work to 3NF using Informatica Push Down optimization technique to utilize Database processing power.
* Designed and developed custom Data Quality audits to identify and report the data mismatch between source and target systems and alert Operations Team.
* Tuned Teradata Sql queries and resolved performance issues due to Data Skew and Spool space issues.
* Developed Flat files from Teradata using fast export, BTEQ to disseminate to downstream dependent systems.
* Supported System Integration and User acceptance tests to obtain sign off.
* Post go live Production Support and Knowledge Transfer to Production Support team

**Environment:** TeradataV2R6/12, Oracle9i/n10g, InformaticaPowerCenter8/8.6, LINUX, Business Objects, UNIX, TeradataV2R6/V2R5, Teradata SQL Assistant, ARCHMAIN, MLOAD, BTEQ, BOXIR3.1, Teradata Manager, MainframesDB2, Erwin Designer, UNIX, Windows2000, Control M, Clear Case, Shell scripts

**Client: HSBC, Hyderabad, India Jun 2011 - April 2012**

**Role: Informatica/Teradata/ETL Developer**

**Responsibilities:**

* Used Informatica Power Center Designer to analyze the source data to Extract & Transform from various source systems (oracle 10g, DB2, SQL server and flat files) by incorporating business rules using different objects and functions that the tool supports.
* Created Mappings using the transformations like Source Qualifier, Aggregator, Expression, Lookup, Filter, Router, Joiner, etc.
* Used Informatica Power Center to create mappings and mapplets to transform the data according to the business rules.
* Implemented slowly changing dimensions (SCD) for some of the Tables as per user requirement.
* Documented Informatica mappings in Excel spread sheet.
* Tuned the Informatica mappings for optimal load performance.
* Created and Configured Workflows and Sessions to transport the data to target warehouse Oracle tables using Informatica Workflow Manager.
* Used BTEQ, FEXP, FLOAD, MLOAD Teradata utilities to export and load data to/from Flat files.
* Migrated data with help of Teradata Fast Export, Insert/ Select, flat files from one system to another.
* Involved in test data set up and data quality testing and Unit testing and Integration Testing.
* Used BTEQ and SQL Assistant (Query man) front-end tools to issue SQL commands matching the business requirements to Teradata RDBMS.
* Performance tuning, monitoring and index selection while using PMON, Teradata Dashboard, Statistics wizard and Index wizard and Teradata Visual Explain to see the flow of SQL queries in the form of Icons to make the join plans more effective and fast.
* Provided suggestions for best join plans while visualizing SQL queries with Visual Explain and Explain while recommending best Join Indexes such as Single table or multi table join indexes.
* Supported jobs running in production for failure resolution with tracking failure reasons and provided best resolution in timely manner.
* Interacted with different teams to find failure of jobs running in production systems and provided solution, restarted the jobs and confirmed if jobs completed in the specified time window.
* Constantly interacted with business users to discuss requirements.

**Environment:** Informatica Power Center Designer 8.6/8.1, Erwin, Informatica Repository Manager, Teradata, shell scripting, SQL Assistant, Windows Vista, Autosys.