Soumya

**Sr. ETL Informatica Developer**

**Phone: (918)-973-2824| Email: Soumya.etl460@gmail.com**

**Experience Summary:**

* Over 8 years of IT experience in Planning, Analysis, Design, Implementation, Development, Maintenance and Support for production environment in different domains like Insurance, Healthcare, Financial, Retail with a strong conceptual background in Database development and Data warehousing.
* Highly proficient in Development, Implementation, Administration and Support of ETL processes for Large-scale Data warehouses by Using **Informatica Power Center 9.6/9.5/9.1/9.0/8.x/7.x**.
* Experience in working with business analysts to identify study and understand requirements and translated them into ETL code in Requirement Analysis phase
* Experience in creating High Level Design and Detailed Design in the Design phase
* Experienced with industry standard methodologies like **Agile,** Scrum and **Waterfall** methodology within the Software Development Life Cycle (**SDLC**).
* Extensive knowledge on Informatica Client tools like Designer, Workflow Manager, Workflow Monitor, Repository Manager and Server tools - Informatica Server, Repository Server.
* Prepared detail documentation, created and developed complex mappings using various transformations such as Lookup, Joiner, Aggregator, Filter, Sorter, Expression, Router, Union, Rank, Update Strategy, Source Qualifier and Sequence Generator.
* Experienced and comfortable working on **Master Data Management(MDM)** using various databases like **Oracle 12c/11g/10g/9i/8i, MS SQL Server, IBM DB2, XML , Flat Files, Sybase, Teradata, Greenplum** and **Netezza.**
* Very good understanding on Cassandra cluster mechanism that includes replication strategies, snitch, gossip, consistent hashing and consistency levels**.**
* Used various types of inputs and outputs in Pentaho Kettle including Database Tables, MS Access, Text Files, Excel files and CSV files.
* Created and maintained Database Objects like Tables, Views, Materialized views, Indexes, Constraints, Sequence, Table Partitions, Synonyms and Database Link.
* Worked on optimizing the Mappings and implementing the complex business rules by creating re-usable transformations, Mapplets, SQL scripts and triggers and PL/SQL stored procedures.
* Experience in Oracle database **Performance Tuning** of SQL statements and in SQL \* Loader Imports/Exports.
* Experienced in Creating folders, Groups, Roles, Users in Admin console and granting them permissions.
* Strong conceptual skills in Data Warehousing Concepts like Ralph Kimball Methodology, Bill Inmon Methodology, **Star Schema** and **Snowflakes Schema** methodologies used in relational, dimensional modeling and multidimensional modeling using Erwin.
* Implemented Data extracting, Data profiling, Data cleansing, Data integration, Data transformation and loading into Data marts using Informatica.
* Extensive experience on **Pentaho report designer, Pentaho kettle, Pentaho BI server, BIRT report designer.**
* Good knowledge in interacting with **Informatica Data Explorer** (IDE) and **Infromatica Data Quality** (IDQ) for the data management and data cleansing activities.
* Proposed design and developed **Hadoop** migration projects which transformed Complex Informatica and Teradata Code using HDFS, Hive, PIG, SQOOP, OOZIE.
* Designed and developed **UNIX shell scripts** as part of the ETL process to automate the data load processes to target and created Post UNIX scripts to perform the operations like copy, remove and touch files.
* Expertise in using Scheduling tools like **ActiveBatch V9, Informatica Job Scheduler, Control-M, Tidal,** **Autosys** to automate running of Informatica Workflows on a daily/ weekly/ Monthly Batch Jobs.
* Extensively worked on Facts, used Lookup Transformation and Update Strategy Transformation while working with Slowly Changing Dimensions (**SCD**) tables Type-1, Type-2.
* Worked extensively on **JIRA** for issue tracking and project management.
* Successful **code migration** from Development to Testing & Production environment.
* With both On-site and Off-shore model experience have developed skills in system analysis, troubleshooting, debugging, deployment, Team management, prioritizing tasks and customer handling.
* Excellent Documentation Skills, Team problem solving ability, Analytical and Programming skills in High speed, Quality conscious and multitasked environment.
* Outstanding communication and interpersonal skills, ability to learn quickly, high compliance to new technologies and tools.

**Technical Skills:**

|  |  |
| --- | --- |
| ETL | Informatica Power Center 9.6/9.5/9.1/9.0,8.x/7.x, Informatica Power Exchange, Metadata Reporter, Oracle Warehouse Builder, Data Integrator, Master Data Management(MDM), DM Loader, IDQ, IDE, IBM DB2, SSIS 2005, Hadoop |
| Data Modelling | OLTP, OLAP, Star Schema, Snow-Flake Modeling, Erwin 7.1/4 |
| Database | Oracle 12c/11g/10g/9i/8i, DB2, SQL server 2005/2008/2012/2014, MS Access 2000/2005, Sybase, Teradata, Greenplum, Netezza |
| Scheduling Tools | Autosys, Active Batch V9, Tidal, Informatica Job scheduler |
| Pogramming Languages | C, C++, SQL, HTML, UNIX Scripting, Java |
| Reporting Tools | Business Objects (BO), Crystal Report |
| Operating Systems | Windows 2000/XP/7, Linux, MS-DOS |
| Applications | MS-Office, Toad 9.2/8.6, Putty, Winscp, Oracle PL/SQL Developer, SQL Assistant |
| Methodologies | Agile, Scrum and Waterfall |
| Web Technologies | HTML, XML, DHTML, PHP |

**Professional Experience:**

**Northwestern Mutual, Milwaukee-WI July 2017-Till date**

**Sr. ETL Informatica Developer**

**Responsibilities:**

* Involved in complete SDLC phase in collecting and documenting requirements. Prepared technical design/specifications for data Extraction, Transformation and Loading.
* Extensively used Informatica Power Center 9.6.1 as an ETL to extract data from sources like MS SQL Server, Flat files, Oracle, IBM DB2 and load into target database.
* Strong expertise in designing and developing Business Intelligence solutions in staging, populating Operational Data Store (ODS), Enterprise Data Warehouse (EDW), Master Data Management(MDM), Data Marts / Decision Support Systems using Informatica Power Center ETL tool.
* Used MDM tool to support Master Data Management by removing duplicates, standardizing data, and incorporating rules to eliminate incorrect data.
* knowledge on Informatica Power Exchange that brings the data from mainframe structure to the target.
* Worked in a competitive Agile environment to develop standard ETL framework and reusable mappings and Mapplets using various transformations like Expression, Aggregator, Joiner, Router, Lookup and Filter.
* Created Reusable Transformations, Worklets, and made use of the Shared Folder Concept using shortcuts wherever possible to avoid redundancy.
* Implemented standards for naming Conventions, Mapping Documents, Technical Documents, and Migration form.
* Created Pentaho reports and used parameters to pass in the report.
* Experienced in delivering the Technical process flow for different Line of Business. Gathered, analyzed & documented business requirements.
* Extensively worked on developing Informatica Mapplets, Mappings, Sessions, Worklets and Workflows for data loading.
* Experienced Working on different tasks in Workflow Manager like Sessions, Event wait, Decision, E-mail, Command, Assignment, and Scheduling of the workflow.
* Created sessions, batches for incremental load into staging tables and scheduled them to run daily by using Autosys.
* Used Type 1 SCD and Type 2 SCD mappings to update slowly changing Dimension tables
* Used Pushdown Optimization and Partition techniques to push the mapping logic to Source and target to increase session performance.
* Designed and developed UNIX shell scripts as part of the ETL process to automate the data load processes to target.
* Used various Informatica Error handling techniques to debug failed session.
* Used JIRA-Agile project management tool to plan, track and manage progress of the project and also track bugs, tasks and requests.
* Performed Informatica code migration from development to testing and production environment
* Handled Production Support for the monitoring of daily, weekly and Monthly jobs
* Worked with the testers closely in determining both medium and high severity defects that would potentially affect the downstream systems before the release of the project and fixed the defects before moving the jobs into production.

**Environment:** Informatica Power Center 9.6.1, MDM, Informatica PowerExchange, Pentaho, Oracle 12c/11g, IBM DB2, Teradata, Flat files, TOAD, SQL/PLSQL, Autosys, Cherwell, JIRA, UNIX, Windows.

**Hewlett Packard Enterprise, Topeka-KS Jan 2016-June 2017**

**Sr. Informatica Developer/Admin**

**Responsibilities:**

* Experience in Dimensional data modelling techniques, Slow Changing Dimensions (SCD), Software Development Life Cycle (SDLC) and Data warehouse concepts -Star Schema/Snowflake Modeling, FACT & Dimensions tables, Physical & Logical Data Modeling.
* Worked on implementing the new Kansas Modular Medicaid system.
* Assisted in building the ETL source to Target specification documents by understanding the business requirements.
* Involved creating pentaho transformations and jobs from SQL server data base to PostgreSQL
* Experience in several facets of MDM implementations including Data Profiling, metadata acquisition, data migration, validation.
* Worked on data cleansing and standardization using the cleanse functions in Informatica MDM.
* Developed performance tuning on Informatica Mappings/ Mapplets / session/ Workflows/ Worklets for optimum performance.
* Create Batch jobs through Active Batch V9, Tidal tools for auto-run on daily/weekly/ Quarterly basis.
* Check for appropriate stages of the change record and deploy components like Informatica SHELL's and other objects, SQL Scripts from Databases like SQL Server, DB2, Netezza, & Oracle, UNIX shell scripts.
* Design and develop data loading programs to support real time data loading to Netezza Data Marts.
* Developed PL/SQL procedures/packages to load the data into Oracle 11g.
* Worked on IDQ parsing, IDQ Standardization, matching, IDQ web services.
* Involved in massive data profiling using IDQ prior to data staging.
* Designed the automation process of Sessions, Workflows, scheduled the Workflows, created Worklets (command, email, assignment, control, event wait/raise, conditional flows, etc.) and configured them according to business logics & requirements to load data from different Sources to Target.
* Streamlined the Change Management process through standardized documents to meet the stakeholder’s expectations.
* Involved in creation and maintenance of basic Informatica administration such as creating users and privileges, folders, optimizing server settings, deployment groups etc.,
* Designed and developed Big Data analytics platform for processing customer viewing preferences and social media comments using Java, Hadoop, Hive, Pig and provide ETL solution to the requirement using Big Data Hadoop.
* Used Hadoop (HDFS) to provide high-throughput access to application data for processing huge data set.
* Created and maintained the shell scripts and Parameter files in UNIX for the proper execution of Informatica workflows in different environments.
* Worked extensively on JIRA for issue tracking and project management
* Migrated the codes from Dev to Test, Test to Prod environment and wrote up the Team Based Development technical documents to the smooth transfer of the project.
* Prepared ETL technical Mapping documents along with test cases for each Mapping for future developments to maintain SDLC.
* One core responsibility to coordinate with offshore team for Technical implementation of BRD, Design data modeling and Mapping, integrate work procedure and migrating of the code document.

**Environment:** Informatica Power Center 9.5/9.1.0, MDM, SQL Server 2008, DB2, Pentaho, Teradata, Netezza, PL/SQL, IDQ, Active Batch V9, Tidal (Scheduling Jobs), DM Loader, Greenplum, Hadoop, Service Now, JIRA, UNIX Shell Scripting.

**ISBE, Springfield-IL Aug 2014 - Dec 2015**

**Informatica Developer**

**Responsibilities:**

* Worked extensively in full System Development Life Cycle like participating in requirement gathering, business analysis, user meeting.
* Prepared data marts on policy data, policy coverage, claims data, client data and risk codes.
* Developed Naming Standards, Best Practices for ETL development.
* Extensively used Informatica Power Center 9.1 to extract data from various sources, which included flat files, XML Files, Oracle, SQL Server, Sybase, Teradata, MS-Access.
* Worked with Teradata utilities like BTEQ, Fast Export, Fast Load, Multi Load to export and load data to/from different source systems including flat files.
* Worked on Informatica transformations like Source Qualifier, Expression, Filter, Router, Aggregator, Lookup, Update Strategy, Stored Procedure, Sequence generator, Joiner, Union, Normalizer.
* Created Mapplet, Parameter Files wherever necessary to facilitate reusability of business logic.
* Involved in Development and Deployment of Source Code between the Various Repositories (Dev, QA, Prod).
* Extensively used Lookup Transformation and Update Strategy Transformation while working with Slowly Changing Dimensions (SCD) and also performed reading and loading high-volume Type 2 dimensions
* Created Reusable Transformations, Worklets, and made use of the Shared Folder Concept using shortcuts wherever possible to avoid redundancy.
* Developed override SQL statements in Source Qualifier and Lookups to suit extraction of desired data from desired region.
* Used Pentaho Data Integration to cleanse and correct data.
* Used SSIS to take data from a flat file, reformats the data, and then inserts the reformatted data into a fact table.
* Developed test cases for Unit Testing of the Mappings, and was involved in the Integration Testing.
* Scheduling jobs using Autosys to automate the Informatica Sessions.
* Written Unix Shell Scripts to perform several operations on text files and developed for data cleaning and loading process.
* Analyzed and remediated the existing Informatica mapping logic for load processes and system standards.
* Implemented Error Processing for Informatica Mappings and Workflows.
* Used TOAD to develop and debug oracle PL/SQL Functions, Procedures and Packages.
* Monitored and improved query performance by creating views, indexes and hints.

**Environment:** Informatica Power Center 9.1/9.0, SSIS 2005, Data Explorer, Data Quality, PL/SQL, Pentaho-Kettle, Oracle 10g, SQL Server 2008 R2, TOAD for SQL Server, Flat Files, Teradata, Sybase, Cobol,UNIX Shell Scripting, Autosys.

**Syntel, Hyderabad-India Feb 2012 – July 2014**

**ETL Developer**

**Responsibilities:**

* Full lifecycle oversight of code fixes and monthly deployments, including requirements gathering, design, coding, testing, and deployment.
* Analyzing the source data coming from different sources and working with business users and developers to develop the Model.
* Involved in Dimensional modeling and developing Star Schema using Erwin to design Fact and Dimension Tables.
* Extracted, Transformed and Loaded OLTP data into the Staging area and Data Warehouse using Informatica mappings and complex transformations.
* Scheduling jobs using Tidal to automate the Informatica sessions.
* Extracted data from various sources like IMS Data Flat Files and Oracle.
* Extensively Used Environment SQL commands in workflows prior to extracting the data in the ETL tool.
* Wrote complex SQL scripts to avoid Informatica joiners and Look-ups to improve the performance as the volume of the data was heavy.
* Created Sessions, reusable Worklets and Batches in Workflow Manager and Scheduled the batches and sessions at specified frequency.
* Wrote UNIX Shell Scripting for Informatica Pre-Session, Post-Session Scripts.
* Used stored procedures to create a standard Time dimension, drop and create indexes before and after loading data into the targets.
* Removed bottlenecks at source level, transformation level, and target level for the optimum usage of sources, transformations and target loads.
* Captured data error records corrected and loaded into target system.
* Created Mappings, Mapplets and Transformations, which remove any duplicate records in source.
* Implemented efficient and effective performance tuning procedures, Performed benchmarking, and these sessions were used to set a baseline to measure improvements against.
* Tuned Source System and Target System based on performance details, when source and target were optimized, sessions were run again to determine the impact of changes.

**Environment:** Informatica Power Center 8.x, Oracle 9i, MS SQL Server 2005, DB2, SQL, PL/SQL, Unix Shell Scripts, Tidal, Toad.

**HRD Services, Hyderabad-India May 2010-Jan 2012**

**SQL / ETL Developer**

**Responsibilities:**

* Analyzing the data model for source to target mapping.
* Perform optimization of SQL queries.
* Developed Mapplets using corresponding Source and Transformations.
* Created packages using SSIS for data extraction from flat files, Excel files, OLEDB to SQL Server using ETL tool.
* Designed and Developed the Informatica workflows/sessions to extract, transform and load the data from source to Target.
* Developed Informatica Mappings/Sessions to populate the Data Warehouse and Data Mart and also various slowly changing dimensional mappings as per the data mart schemas.
* Used debugger to test the mapping and fixed the bugs.
* Executed sessions, sequential and concurrent batches for proper execution of mappings and worked on workflow manager.
* Designed source qualifier from File, Relational Sources. Assisted in Migrating Repository.
* Prepared the environment for Testing. Prepared Documentation for Production Support.
* Used Autosys, Informatica Job scheduler for scheduling the Jobs. Responsible for defect tracking in ETL, Contact point for defects in ETL team.

**Environment:** Informatica Power Center 7.1, Designer, Repository Manager, Oracle 8i, DB2, PL/SQL, SQL, Business Objects(BO) 6.0, Autosys, TOAD 7.0, UNIX 4.2, Windows XP.