**Summary**

* 9 years of experience in all phases of SDLC including Requirements, application architecture, designing, development, testing and post implementation support in Data warehousing technologies using IBM InfoSphere DataStage 8.5/8.1 ( Information Server, Information Analyzer, Metadata Workbench, Quality Stage), Ascential DataStage 7.X/6.X/5.2 using components like DataStage Designer, DataStage Manager, DataStage Director, DataStage Administrator.
* 7+ years of experience in end-to-end IT solution offerings in all phases of the development lifecycle involving prototyping, Requirements Gathering, Analysis, Data Profiling, Design, Development, Implementation, Acceptance Testing, coherent with Data Cleansing, Data Conversion, Performance Tuning.
* Strong experience with Star and Snowflake schema, Data conversions, Dimensional Data Modeling, Fact and Dimensional tables, Slowly Changing Dimensions, CDC (CDC Change Data Capture).
* Excellent experience in working with Parallel jobs (pipelining & partitioning) to distribute the load to different processors. Experience in Sales force Datastage service pack.
* 8+ years of Oracle 11g/10g/9i/8i/8.0/7.0,DB2, SQL, PL/SQL, Microsoft SQL Server 2000/7. Experience in Integration of various data sources like Teradata, DB2 UDB, SQL Server, Oracle 10g and Netezza database.
* Good experience in dealing with the Oracle Database applications by writing queries and stored procedures.
* Experience in providing Data Warehouse data modeling, design and architecture, Creating Indexes, Triggers for data manipulation and data consistency. Involved in Performance Fine Tuning of ETL programs.
* Extensively worked on Huge Databases in Production environments, experience in Integration of various data sources like Oracle, Sequential File and Flat Files into a staging area. Exposer in Cognos reporting tool. Involved in the schedules the DataStage jobs and upgrades the environment. Involved in initial and incremental runs, Validation, monitor documents, monitor scripts.
* Good experience in Transforming Business specific rules into functional Specs.
* Extensive experience in design and development of Decision Support Systems (DSS), experience in FTP & SFTP Protocols.
* Developed Test Cases for Unit Testing and System Integration testing. Seamlessly migrated the code from Development to Testing, UAT and Production. Experience in scheduling the jobs names and script names by using scheduling toll- Control-M job. Experience in Autosys scheduling tool.
* Knowledge of Big Data and Hadoop.
* Experienced in scheduling sequence, parallel and server jobs using DataStage Director, UNIX scripts and scheduling tools.
* Good experience in scheduling Jobs using AutoSys and Corntab
* Good communication, interpersonal, presentation, analytical and problem solving skills, self-motivated, willing to take initiative and able to learn quickly and apply new tools and technologies in the projects.
* Working experience in interacting with business analysts and developers to analyze the user requirements, functional specifications and system specifications.
* Completed IBM Datastage 8.5 Certification in 2012. 

**Education**

* Bachelor of Technology in Computer Science and Engineering (B. Tech)

**Software Skills**

|  |  |
| --- | --- |
| DataStage Tools | Datastage 9.1/8.7/8.5/ 8.1/ 8.0.1/ 7.5.3, IBM Information Analyzer 8.5, Quality Stage |
| Warehousing and Data modeling | Dimensional Data Modeling, Data Modeling, Star Schema Modeling, Snow-Flake Modeling, FACT and Dimensions Tables, Physical and Logical Data Modeling, ERwin 3.5.2/3.x. |
| RDBMS | Oracle 11g/10g/9i/8i/8.0/7.0, DB2, MS SQL Server 2008/2000, SQL, PL/SQL, ANSI SQL, Transact SQL. |
| Environment | Suse Linux, RHEL, UNIX, HP-UX, IBM AIX 4.2/4.3, OS/390 |

**Professional Experience:**

**Nielsen, Schaumburg, IL May’2014 – Till Date**

**Senior BI/Datastage Developer**

1) Panel Athena

Project Description:

This project deals with gathering the households data from Panels installed across US. These panels are of Hispanic and Mega. The usage of consumer products will be tracked based on the data sent through panels. The consumer will scan the receipts once they performed a trip at various outlets. The purchased products from each outlet are collected from various shops. The data will be sorted and projected in order to analyze based on the elementary periods with a cumulative period. It involves complex calculations to manipulate the data before storing into dataware house and distribute to several reporting databases.

**Responsibilities:**

* Created Low Design Document which illustrates detailed flow for all the interfaces.
* Implemented Framework which is used for audit and error logging purposes.
* Understanding the requirements and preparing Low Level Design documents and there by developing jobs.
* Developed Shared containers and Server Routines
* Written complex queries to facilitate the supply of data to other teams.
* Responsible for using different types of Stages such as Sequential File, Sort Aggregator, Transformer and ODBC to develop different jobs.
* Designed jobs using different parallel job stages such as Join, Merge, Lookup, Remove Duplicates, Copy, Filter, Funnel, Dataset, Lookup File Set, Change Data Capture, Modify, and Aggregator
* Used DataStage Director for controlling and sequencing jobs.
* DataStage Enterprise Edition for parallel processing as part of ETL conversion process integrating multiple source systems.
* Prepared Unit Test Cases and executing them.
* Developing Datastage Parallel and Sequence Jobs.
* Developed UNIX shell Scripts used to validate the files.
* Trained team members in learning Datastage.
* Understand the existing PL/SQL procedures to extract data loaded by ODS and DIM sourcing.
* Responsible for documenting user requirements and translated requirements into system solutions.
* Leading team size of 5members.

**Environment:** IBM WebSphere Datastage 8.1v, UNIX shell script, Netezza, PL/SQL, Aginity Workbench, Erwin, Office, Visio.

**General Motors, Troy, MI Sept’2011 – May’2014**

**Technical Lead**

1. **ADP Canada Payroll Project**

**Project Description:**

General Motors (GM) ADP payroll project is an outsourcing project wherein all the payroll processing will be migrated from Legacy application to 3rd party vendor – ADP. This migration was planned to be implemented in multiple phases.

In the first phase, the entire US region payroll was successfully migrated to ADP during the period – October 2012 to January 2013.

In the second phase, entire Canada region’s payroll processing is being migrated to ADP. This migration, being similar to the US Payroll migration, follows the same overall architecture as implemented for US region with necessary customization.

**Responsibilities:**

* Created Architecture Design Document which illustrates the flow of project for all the interfaces.
* Used Information Analyzer tool for data profiling and to understand data structure, format and relationships
* Analyzed the source data from various sources using Column analysis, Primary Key, Foreign Key, Cross-Domain Analysis methods through Information Analyzer.
* Analyzed the columns, keys and relationships using Analyzer tool
* Applied normalization and business rules to publish analysis reports
* Implemented Framework which is used for audit and error logging purposes.
* Created Project, users and assigned roles to users.
* Understanding the requirements and preparing Low Level Design documents and there by developing jobs.
* Developed Shared containers and Server Routines which are used across the project in most of the interfaces.
* Involved in Each Phase of project – End to End module.
* Written complex queries to facilitate the supply of data to other teams.
* Responsible for using different types of Stages such as FTP, Hashed File, Sequential File, Sort Aggregator, Transformer and ODBC to develop different jobs.
* Designed jobs using different parallel job stages such as Join, Merge, Lookup, Remove Duplicates, Copy, Filter, Funnel, Dataset, Lookup File Set, Change Data Capture, Modify, and Aggregator
* Used DataStage Director for controlling and sequencing jobs.
* DataStage Enterprise Edition for parallel processing as part of ETL conversion process integrating multiple source systems.
* Prepared Unit Test Cases and executing them.
* Provide support to Testing Team and fix defects raised in different phases of testing.
* Developing Datastage Parallel and Sequence Jobs.
* Developed UNIX shell Scripts used to validate the files.
* Trained team members in learning Datastage.
* Understand the existing PL/SQL procedures to extract data loaded by People Soft.
* Provide Post Implementation Support.
* DataStage Enterprise Edition for parallel processing as part of ETL conversion process integrating multiple source systems.
* Responsible for documenting user requirements and translated requirements into system solutions.
* Created Error Files and Log Tables containing data with discrepancies to analyze and re-process the data.
* Used Investigate, Standardize and Matching quality stages to cleanse the source data
* Experience in design and implementation of Star, Snowflake schemas and multi-dimensional modeling
* Tuned DataStage transformations and jobs to enhance their performance.
* Lead team size of 8members.

**Environment:** IBM InfoSphere Datastage 8.5v, Information Analyzer 8.5, Quality Stage, UNIX shell Script, ORACLE 10g, PL/SQL, SQL Developer Client for Oracle 10g, Toad, Erwin, Office, Visio.

1. **ADP US Payroll Program**

**Project Description:** GM ADP Payroll project will outsource the payroll processing data to a third party provider - ADP which will process the Employee Payrolls and sends back the data that will be used by various Downstream Payroll related Applications.

Currently 2 legacy applications are doing the payroll processing - HPS (Hourly Payroll System) and SPS (Salaried payroll system) which involves lot of manual activities and need very long time to process the complete Payroll data. With increase in number of employees, this process is expected to become more cumbersome.

So General Motors decided to decommission the existing HPS/SPS and carry out processing through a 3rd party automated application called ADP

After analyzing the requirements of the ADP applications as well all the Payroll related source systems earlier sending data to HPS and SPS, processes data as required by ADP and sends it to ADP.

After ADP generates the payroll, it sends post-payroll data back which will then be maintained by TCS. This data then will be sent to various Payroll related downstream applications.

The Payroll Project has been implemented first of US Region followed by Canada.

**Responsibilities:**

* Gathering requirements and Preparing Specifications for ETL Job modules.
* Interacting with the clients to understand the business and mapping rules of all interfaces.
* Analyzed the source data from various sources using Column analysis, Primary Key, Foreign Key, Cross-Domain Analysis methods through Information Analyzer.
* Analyzed the columns, keys and relationships using Analyzer tool
* Designing and developing Master controlling sequencer jobs using the DataStage Job Sequencer
* Build the code based on the requirements and develop middleware layer to provide data to ADP Global View. It is proposed to route data from all the source applications to GV through GIF. And similarly route all data from GV to downstream applications through GIF. New interfaces would be developed & deployed to facilitate this data interchange.
* Develop Data Flow diagrams and Source-to-Target mapping.
* Extracting, cleansing, transforming, integrating and loading data into data warehouse using DataStage Designer.
* Prepare Low Level Design (LLD) document based on the mapping document and Data Model.
* One time migration of initial data to ADP.
* Develop Data Flow diagrams and Source-to-Target mapping.
* Creation of CTRL tables for frame work development and STG tables while loading data before target.
* Came-up with an efficient recovery, rollback and restart mechanism at the DataStage job level as well as database level.
* Developing DataStage Parallel and Sequence Jobs based on Low Level Design Document.
* Involved in the testing of the various jobs developed and maintaining the test log.
* Migrating jobs from one environment to other.
* Involved in the implementation of UNIX shell scripts to invoke the DataStage jobs and automate file manipulation.
* Created Shared Containers for Re-using the Business functionality.
* Creation of Unit testing strategy according to the test plans.
* Used Investigate, Standardize and Matching quality stages to cleanse the source data
* Creation of Integration testing strategy according to the test plans.
* Involved in ETL Testing Process and Bug Fixing.
* Interacting with FTP Tools to upload/download the files into/from UNIX.
* Managing the versioning of Documents and DataStage Code using iCollab.
* Provide Post Warranty Support.

**Environment:** IBM InfosphereDataStage 8.5, Information Analyzer 8.5, Oracle 10g and UNIX Shell Script and Control M

**Pfizer, UK Nov’10 – Aug’11**

**Senior Datastage Developer**

***DIM Project***

**Project Description:** DIM (Drug inventory management) project deals with the various drugs specified by unique drug codes across globe. In order to integrate and centralize the sales for every region, we have developed datamarts and ODS to load the existing as we a new drug information, store information and the sales details.

**Responsibilities:**

* Visited Business and Operations to gather requirements and document the same.
* Participates in Analysis of the Project and prepared technical specifications document.
* Created and Populated the Data ware house and Data Marts for the BI Developers to create there reports.
* Participated in Group discussions, leading group of members - part of work.
* Was responsible for creating Data model using Erwin, Toad, Visio and other data modeling tools.
* Worked on DB2 extensively.
* Involved in designing, developing and testing of the ETL (Extract, Transformation and Load) strategy to populate the data from various source systems.
* Responsible for creation of parallel and Sequence Jobs.
* Used Datastage Jobs to extract and customize data from different sources.
* Monitored Test Cases in HP Quality Center, Supplied Comments and close the defects once they are fixed.
* Handled Performance Tuning of Jobs to ensure faster Data Loads.
* Designed sequence jobs using the activities such as Job Activity, Nested Condition, Notification Activity, Sequencer Activity, Terminator Activity and Execute Command.
* Performed the Integration and System testing on the ETL jobs.
* Provided support in different phases of testing.
* Migrating – Data stage 8.1v to Data stage 8.5v.
* Migrating code from one environment to other.
* Used the Data Stage Designer to develop processes for extracting, cleansing, transforming, integrating, and loading data into data warehouse.
* Worked with Data Stage Manager to import/export metadata from database, jobs and routines between Data Stage projects.
* Used multiple stages like SAP ABAP R/3 Extract, Sequential File, Transformer, Aggregator, Join, Lookup, Sort and Filter during ETL Data Quality development.
* Used Data Stage Director to schedule, monitor, cleanup resources and run jobs.
* Designed jobs using different parallel job stages such as Join, Merge, Lookup, Remove Duplicates, Copy, Filter, Funnel, Dataset, Lookup File Set, Change Data Capture, Modify, and Aggregator.
* Extensively used SQL tuning techniques to improve the performance in Data Stage Jobs.
* Performed through data profiling by using the Investigate stage of Quality Stage.
* Created many PL/SQL stored procedures and packages while extracting the data from sources into staging area.
* Developed user defined Routines and Transforms by using Data Stage Basic language.
* Used Data Stage Director for controlling and sequencing jobs.
* Used Data Stage Shared Containers, Local Containers for DS jobs.
* Design, build, and manage complex data integration process.
* Used SQL\*Loader, and Shell Scripts to import data into data staging area.
* Involved in writing triggers, stored procedures in PL/SQL to extract data from various source systems into Data Staging Area.
* Extensively used ETL to load data from Oracle, SQL Server, and DB2, Flat files to SQL Server.
* Analyze newly converted data to establish a baseline measurement for data quality in data warehouse.

**Environment:** Ascential Datastage 7.5.3, IBM WebSphere Datastage 8.1v, Db2 9, Shell Scripting, Oracle 9i, HP Quality Centre, Control M, Teradata, Ms-Access, XML.

**Arizona Public Services, AZ, US Jan’10 – Nov’10**

**Datastage Developer**

***Power Distribution and Consumption Analysis***

**Project Description:** APS (Arizona Public Services) is a global leader in the area of power distribution and saving through Residential Services and Business Services in various aspects. This is a power distribution and analysis solution which can be utilized by energy saving parties, energy consumers to make effective energy utilization in day to day life. The source for this ETL process is CIS, EDIS in the form of flat files and database tables.

**Responsibilities:**

* Interacting with onsite to understand the requirements , Mapping Document and Data model
* Create Low level design document.
* Develop parallel and sequence jobs, unit testing them.
* Build EDW (Enterprise Data Warehouse) and Datamarts for subject areas like Sales, Stock, and Purchasing to have effective reporting to power users.
* Data from SAP ECC sends data in XML format through middleware, SAP BI data is read from SAP pre-built views. The data from these source systems are cleansed, integrated and loaded into EDW layer using Datastage. There are 12 source systems / interfaces in EDW.
* Review, Defect Analysis and Defect logging
* Maintain the test cases and test results in Quality Center Tool, Update the defect status as on when required.
* Deploy the developed code to QA, UAT, Production environments and validate the code.
* Worked with Parallel Extender using Parallel processing techniques to improve job performance while working with bulk data sources.
* Used many stages like Teradata Enterprise stage, Dataset , Lookup Stage, Merge, Modify, Transformer, Funnel, Filter, Aggregator , Sort stage etc. to design the jobs.
* Generated the Surrogate keys for composite attributes while loading data into data warehouse.
* Involved in writing the data loading stored procedures, functions using PL/SQL while extracting the data from various sources systems into staging area.
* Responsible for importing metadata from repository, transforming them and loading as per requests.
* Created the job sequencers to run the designed jobs in sequence.
* Involved in every phase of module. Participated in group discussions and requirement gathering.

**Environment:** DataStage 8.0.1v, Oracle 9i, HP Quality Centre, UNIX Shell Scripting, Oracle Data modeler, Visio.

**ELBIS, Irvine, California Jun’08 –Dec’09**

**Datastage Developer and Tester**

***ELBIS Project***

**Project Description:** Edwards Life Sciences (NYSE: EW) is a global leader in products and technologies to treat advanced cardiovascular disease, the global leader in acute hemodynamic monitoring and the number-one heart valve company in the world. Headquartered in Irvine, California, Edwards has more than 5,600 employees’ worldwide, selling medical technologies in nearly 100 countries with 2007 sales of $1.091 billion. Edwards Life Sciences leverages its research, design, development and marketing expertise to produce products that address specific cardiovascular opportunities including heart valve disease, vascular disease and critical care technologies.

The purpose of ELBIS is to provide a single common repository of corporate information in support of tactical and strategic decision-making processes across the organization. The initial goals are to collect daily sales and inventory data from global customer facing invoice and warehouse management systems and make this information available to meet executives’ and knowledge workers’ need for user-friendly access to data on a global basis.

A process is in place to extract data from global systems and send the data to Irvine where the data is quality controlled, cleaned, transformed and loaded into the ELBIS database.

**Responsibilities:**

* Understanding the requirements of the ETL process.
* Involved in requirement collection.
* Involved in system Design, testing and implementation.
* Designed the ETL processes using DataStage Designer tool.
* Based on the Business Requirements developed Data Stage Jobs for transformation and cleansing of Data.
* Used the Data Stage Designer to develop process for extracting, cleansing, transforming, integrating and loading the data into Oracle Data warehouse.
* Used Transformer, Remove Duplicates, Copy, Funnel, Lookup, and Change Capture Stages in designing jobs.
* Involved in unit testing.
* Created the job sequencers to run the designed jobs in sequence.
* Involved in every phase of module. Participated in group discussions and requirement gathering.

**Environment:** DataStage 7.5.3 px, Oracle 9i, UNIX Shell Scripting, SQL, PL/SQL.

**StateStreet, CA, US Jan’06 –May’08**

**Datastage Developer and Tester**

***IMO Standard Offering***

**Project Description:** The transaction for the trades done by portfolios for different clients will be populated in Enterprise Data Warehouse (EDW), which is running on top of Sybase Data Base. Data will be extracted from the source tables/views and brought into the ETL server. This project Provide the State Street Business Operations with a standard set of configurable IMO Interfaces that can be offered to clients.

This project will allow the business (via a GUI) to configure the Standard Offering on the Control Database to meet their client’s requirements (with no additional coding) and allow other applications with different requirements leverage the same extract logic to improve data consistency and reduce the overall custom coding.

All business rules, data validation rules and code translations will be applied during ETL process within ETL server. The Flat files are generated for the confirmed trades. After the Extraction is complete the flat files will be transmitted out to the clients FTP server using FTP feature

**Responsibilities:**

* Developing DataStage jobs to load data into the Warehouse
* Analysing the existing system and preparation of Design Documents.
* Preparation of ETL mapping specifications.
* Prepared AutoSys jobs to schedule DataStage jobs
* Performing Unit testing, reviewing peer’s code and peer testing.

**Environment:** DataStage 7.1, Putty, Winscp, Autosys, DBArtisan and sun solaris 6.0.

***AXA IMS***

**Project Description:** The transaction for the trades done by portfolios will be populated in Enterprise Data Warehouse (EDW), which is running on top of Sybase Data Base. Data will be extracted from the source tables/views and brought into the ETL server. All business rules, data validation rules and code translations will be applied during ETL process within ETL server. The Flat files are generated for the confirmed trades. After the Extraction is complete the flat files will be transmitted out to the clients FTP server using Data Stage’s FTP feature.

**Responsibilities:**

* Developing DataStage jobs to load data into the Warehouse
* Analysing the existing system and preparation of Design Documents.
* Preparation of ETL mapping specifications.
* Prepared AutoSys jobs to schedule DataStage jobs
* Performing Unit testing, reviewing peer’s code and peer testing.

**Environment:** DataStage 7.1, Putty, Oracle 7.0, WinSCP, Autosys, DBArtisan and sun solaris 6.0.

***RKS-EDW Adapter***

**Project Description:** This project being undertaken for the biggest custodian bank in U.S.A. Security and fund related data are received as RKS (Record keeping system) Generic files. The “RKS Generics” consist of a set of interfaces developed by CTI Boston to support the requirements of clients who wish to access details of the data held on RKS for their funds.

The new ETL processing system using DataStage is replacing the existing Perl system framework. Initial phase of the project was to analyse the existing system and preparation of design documents. DataStage is used as ETL tool to implement business transformations to the input data and generate the output interface files.

ETL generated Output interface files are validated against the existing Perl system output files to ensure data quality. Then these files are loaded into EDW using DSTi HiInvest loader.

**Responsibilities:**

* Developing DataStage jobs to load data into the Warehouse
* Analysing the existing system and preparation of Design Documents.
* Preparation of ETL mapping specifications.
* Prepared AutoSys jobs to schedule DataStage jobs
* Performing Unit testing, reviewing peer’s code and peer testing.

**Environment:** DataStage 7.1, Oracle 7.0, DSTi HiInvest loader, DBArtisan, Solaris 6.0, Oracle, Shell Scripting and AutoSys