**Akeesh Email: akeesh32@gmail.com**

**Phone:6035210997**

**Profile Overview:**

* A Dynamic IT professional with 6+years of work experience in Data Integration for Data Warehouses including IBM Info Sphere/WebSphere/AscentialData stage (9.x / 10. X /11.X) (8.7/8.5/8.0.1/7.5.1/7.5.2) Server Edition, Enterprise Edition (Designer, Director, Administrator)
* Worked in various business domains including Retail, Healthcare, Investing, Logistics, Telecom industries
* Experience in Analysis, Design, Development and Implementation of Relational Database (OLTP) and Data Warehousing Systems (OLAP)
* Extensive experience on the Data Modeling phases conceptual, logical, physical
* Experienced in working with Casandra
* Strong experience in Data warehouse development life cycle and Design of Data marts with Star Schemas, Snowflake Schemas and Integrated Schemas
* Used the Data Stage Designer to develop processes for extracting, transforming, and loading data into data warehouse databases
* Experience in client/server business systems in design/analysis/testing, GUI prototyping, Data Warehousing, and application development
* Experience in Pipeline and Partitioning Concept.
* Converting Large XML files to Multiple XML files as required by downstream application.
* Significant experience with Data Extraction, Transformation and Loading (ETL) from disparate data sources such as multiple relational databases and also worked on integrating data from flat files, CSV files and XML files into a common reporting and analytical data model.
* Good working knowledge on Protegrity encrypted/Tokenized data for handling PII, PCI and sensitive info
* Strong experience developing SSIS ETL process for a data warehouse using Microsoft SQL server.
* Extensively worked with Parallel Extender for parallel processing to improve job performance while working with bulk data sources
* Integrated the email feature in Talend which would send a report to the targeted user when a data integration job fails.
* Worked extensively with different database stages of Data Stage like Oracle Enterprise, Dynamic RDBMS, Stored Procedure, ODBC
* Worked extensively with different file stages of Data Stage like Sequential, Dataset, File set, Lookup File set
* Experience in different Hadoop distributions like Cloudera (Cloudera distribution CDH3, 4 and 5), Horton Works Distributions (HDP)
* Worked extensively with different processing stages of Data Stage like Filter, Sort, Aggregator, Copy, Transformer, Funnel, Join, Look Up, Merge, Pivot, Modify, Switch, Change Capture, Surrogate Key Generator, Remove Duplicates
* Worked extensively with different development/debug stages of Data Stage like Head, Tail, Row Generator, Column Generator, Peek
* Worked with Local Containers, Shared Containers and Job Sequences
* Knowledge in SQL & PL/SQL and expertise in writing Stored Procedures, functions and Packages
* Used python web framework Django apis to access databases.
* Excellent in creating SSIS Packages for integrating data using OLE DB connection from heterogeneous sources (Excel, CSV, flat file, Text Format Data) by using multiple transformations provided by SSIS such as Data Conversion, Conditional Split, Bulk Insert, Merge Join, Lookup, Fuzzy Lookup, Slowing Changing Dimension, CDC Incremental Load, etc.
* Extensive experience in dealing with High Volume Data, Performance Tuning, Maintaining, Multiple Job Compile, Surrogate Key Generator Stage, Job Reports, Message Handler Options
* Developed efficient mappings for data extraction/transformation/loading (ETL) from different sources to a target data warehouse, Informatica powerMart 7x/6x, Power Connect, Power Exchange
* Good Knowledge on Hadoop Cluster installation, architecture and monitoring the cluster.
* Extensively Worked on Job Sequences to control the flow of job execution using various activities like Job Activity, Email Notification, Sequencer, Routine activity and Exec Command
* Wrote Python code using ansible Python API to automate cloud deployment process.
* Excellent problem-solving and trouble-shooting capabilities. Quick learner, highly motivated, result oriented and an enthusiastic team player. Good interpersonal skills, experience in handling communication and interactions between different teams.

**Tech Skill Set:**

|  |  |
| --- | --- |
| **ETL Packages** | IBM Info Sphere/Web Sphere Data Stage and Quality Stage 10.x,8.7, 8.5, 8.0.1, Ascential Data Stage 7.5.1/2 (Designer, Director, Manager, Administrator) |
| **Databases** | Teradata 13.11 client, Casandra, Oracle 10g/11g, MS SQL Server 2005/2008, DB2 UDB 8.1/7.2 |
| **Database Tools** | SQL\* Plus, SQL Loader, Toad, DB2. |
| **Database Modeling** | Anchor Modeling, Star-Schema Modeling, Snowflakes Modeling, Integrated Schema, E-R Modeling, FACT and Dimension Tables,ERWIN 4/7.1/8 |
| **Operating Systems** | Windows NT / 2000 / XP Pro / Vista, 7, Windows Server 2000 / 2003 / 2008, UNIX – Solaris/AIX/UX, Linux – Mint, Ubuntu |
| **Languages** | SQL, PL/SQL, UNIX Shell scripting, C,C++,java, XML, HTML, Javascript |
| **Reporting Tools** | OBIEE, Micro strategy, Business Objects, IBM Cognos, ClearCase |

**Education:**

**B. tech: JNTU – INDIA**

**Masters: Rivier university - USA**

**Professional Experience:**

**Blue Cross Blue Shield Association July 2018 – Present**

**DataStage Developer**

As the health care industry continues to rapidly transform, our IT team conceives, develops and delivers impactful technology solutions to support access to quality, affordable health care for our members. We are driven by our collective company purpose: To do everything in our power to stand with our members in sickness and in health. Our IT team unleashes the power of this purpose through technology. We come to work every day to make a difference, and we deliver the highest quality and best solutions to our members

**Responsibilities**

* Involved in gathering requirements from Business users, Business owners and involved in modifying various technicalSpecifications in the Development phase to Maintenance Phase.
* Designing, developing and deploying end-to-end Data Integration solution.
* Designed and Implemented the ETL process using Datastage to load the data from Source to Target Database.
* Designed applications of machine learning, statistical analysis and data visualizations with challenging large data processing problems.
* Developed business logic using python on Django web framework.
* Implemented the workflows using Apache Oozie/Azkaban to automate tasks.
* Have extensively usedboth Oracle, DB2 as source and target databases.
* Integrated large scale data from the retail data base employing Talend and SQL server integration services.
* Involved in Data Extraction from Oracle, Flat files and XML files using Datastage.
* Implemented Dimensions (SCD) Type1, Type 2 to capture the changes using ETL
* Develop ETL mappings for various Sources (.TXT, .CSV, .XML) and also load the data from these sources into relational tables with DataStage.
* Created test harness to enable comprehensive testing utilizing python.
* Tokenization – protegrity Administration managing ESA and DSG appliances, policy/role/data element management.
* Responsible for testing and validating the data at all stages of ETL process.
* Good knowledge on Hadoop cluster architecture and monitoring the cluster.
* Developing a Ruby on rails app that is used by technicians to discover hardware problems on remote systems and guide them through steps to correct those problems.
* Worked on Global Context variables, Context variables, and extensively used over 100+components in Datastage to create jobs.
* Worked on Error handling techniques, optimizing techniques and tuning the ETL flow for better and faster performance.
* Worked Extensively Autosys (Admin Console), toSchedule Jobs in Job Conductor.
* Worked with Oracle SQL Developer, Business Data Analyst while Implementing Unit Testing of ETL Datastage Jobs.
* Scheduling (Autosys) the ETL mappings on daily, weekly, monthly and yearly basis.

Environment: IBM WebsphereDatastage 10.x,9.1, Oracle 10 g/11 g, Toad, Cognos, Tableau Desktop (11.x / 10. x), Tableau Server (9.x/10.x), SAS, Flat flies,UNIX Shell Scripting, ClearCase,DB2.

**AT&T, McLean, Virginia July 2016 – July 2018**

**DataStage Developer**

The aim of the project is to implement Business Intelligence logics to track a missing customer, allocate the customer to appropriate program and to calculate the Cubes, Scorecards, and revenue generated.

**Responsibilities:**

* Extensively involved in different Team review meetings, review & Analysis.
* Create mapping and workflows using Informatica power center to extract, transform and load data into the target environment.
* Monitored continuously and managed the Bigdata (Hadoop) cluster using cloud era manager.
* Interacted extensively with clients during meetings.
* Developed the ETL Development Standards and best practices to design the Data Stage Parallel jobs consistently across the project.
* Created business logic using Python to create planning and tracking functions.
* Did auditing/validation after each load and developed the code to perform actions according to the results.
* Extensively used Parallel Stages like Join, Merge, Lookup, Filter, Aggregator, Copy, Sort,Funnel, Change Data Capture, Remove Duplicates, Surrogate key Generator, Row Generator, and Column Generator for development and de-bugging purposes
* Websense, McAfee, ArcSight, Intrusion prevention systems, firewalls, firemanProofpoint voltage, protegrity.
* Created several Job Sequences for the maintenance of Data Stage jobs using Wait for file, Job Activity,
* Wrote Python Scripts to parse XML documents and load the data in database.
* Provide staging solutions for Data Validation techniques.
* Created smart SQL Queries for minimal resource use during data extraction from source.
* Created a technical document for ETL design for efficient stage use.
* Created universal parameter set for ETLs.
* Assisted in creation of a new database to monitor the growth and performance of different
* Instances.
* Extensively worked on Informatica to extract data from flat flies, excel files and MS access database to load the data into the target database.
* Designed applications of machine learning, statistical analysis and data visualizations with challenging large data processing problems.
* Experienced in managing and reviewing the Hadoop log files.
* ETLs were designed and created based on performance, system availability, and businessrequirements.
* Used shell Scripting and Perl scripting to deploy artifacts that are built by Ant and Maven
* in Unit Testing with the jobs and date loaded in the Target database both in SIT and

UAT.

* Manage scheduling and performance analysis of current jobs in production.

**Environment:** IBM WebsphereDatastage 8.5, Oracle 10g, Quality Stage Designer, Director, Informatica Power Center 9.1.0, SQL, PL/SQL, UNIX Shell Scripting, DataStage Version Control, MS SQL server, Cognos,Framework Manager, QlikView 10/11, Qlik Sense, SSRS, Informatica, Hyperion Reporting, Tableau and analysis 9.x/11. x.

**NY Power authority, New York Aug 2015 – May 2016**

**DataStage Developer**

This project was intended to develop a reporting warehouse. The main objective of the project was to convert data from ITAS to ETPM and to build a complete data warehouse for current, future and Adhoc IBM COGNOS reports. The data warehouse will bring together the multiple sources of data.

**Responsibilities:**

* Responsible for business analysis, requirement collection, design and development of guest data mart.
* Interpreted the information from various operational sources and external systems to document the data stage jobs as per the business requirement.
* Designed the frontend application using Ruby 2.2.2, Rails 4.2.3 CSS3, AJAX and jQuery.
* analyzed the data for inconsistencies, duplicates and bad data.
* Used the Data stage Designer to develop processes for extracting, cleansing, transforming, integrating, and loading data into data warehouse database.
* Used Data stage Manager for importing metadata from repository, new job categories and creating new data elements.
* Used Data stage Director and the runtime engine to schedule running the server jobs, monitoring scheduling and validating its components.
* Used Parallel Extender Development/Debugging stages like Row generator, Column Generator, Head, Tail and Peek.
* Dealt extensively with data from Teradata sources.
* Used Data stage Director to validate schedule run and monitor the data stage jobs.
* Used Before/After Job-Subroutines in Job Properties.
* Used Data stage for developing, programs for scheduling data loading and transformations
* Verified job process logic by preparing test data and conducting unit tests and participated in the systems testing and user acceptance testing.

**Environment:** IBM Info Sphere Data Stage 8.7/8.5 (Designer, Administrator, Director, Manager), Teradata 13.11, Zena, Serena, Windows Server 2003, AIX 6.1, SQL Plus, UNIX Shell Scripting, SQL Loader, SQL, MS SQL Server 2005/2008

**Demas Technologies Pvt Ltd - Hyderabad, India June 2013– July 2014**

**DataStage Consultant**

The project was Affordable Care Act’s 3 Rs Production and Semantic Design. In the wake of the Government’s ongoing project, the aim was to process the daily and monthly source data coming from the DMIH and load it into the Atomic Data Warehouse as per the business rules provided by HHS. Project involved the combination of different modules including Claims, Membership and Income. Project was to design, develop and test ETL programs required for target data warehouse.

**Responsibilities:**

* Participated in discussions with Team leader, Group Members and Technical Manager regarding any Technical and Business Requirement issues
* Validation of Informatica mapping for source compatibility due to version changes at the source.
* Used Partition Parallelism, the same job would effectively run simultaneously by several processing Nodes each handling separate subset of Total data
* Used Data Stage to extract client data from Oracle and mapped the data into a target business warehouse
* Used Data Stage Director to view logs for any errors and warnings
* Used Surrogate Keys to keep track of Slowly Changing Dimensions (SCD)
* Used Quality Stage to automate the conversion of data into standard formats and match variables that are common to records (for example, given name, date of birth, or gender) when unique identifiers are not available
* Used Quality Stage components like Match Designer for designing and testing match passes
* Business-oriented approach: this has task-based user interface, it aided business users in reviewing data for anomalies and changes over time, and provided key functional and design information to developers
* -Adaptable, flexible, and scalable architecture: handled high data volumes with common parallel processing technology, and utilized common services such as connectivity to access a wide range of data sources and targets
* Extensively used Pivot stage to pivot the source data to achieve required table structures like converting data from rows into a column
* Used development/debug stages to test the environment by creating samples of data from given high volume data or by creating mock data
* Imported and exported repositories across projects
* Extensively followed the concepts of Naming Conventions, Parameterizing the variables from Global Environment to stage level
* Performed parallel processing of data extraction and transformation
* Extensively worked with various stages like sequential file, dataset, lookup, peek, transformer, merge, aggregator, row generator, surrogate key generator and many more to design jobs and load the data in to fact and dimension tables
* Implemented Local containers for same job and Shared Containers for multiple jobs, which have the same business logic
* Used stage variables which is a temporary location used to hold the values temporarily at the time of execution for better performance and to reduce the complexity of defining expressions
* Extensively used Data Stage Director and Autosys for Job Scheduling, emailing production support for troubleshooting from LOG files
* Extensively worked on Error Handling

**Environment:** IBM Info Sphere Data Stage 8.5/8.7 (Designer, Administrator, Director, Manager), Quality Stage, Audit Stage, Information Analyzer, Oracle OBIEE 10g/11g 11g/10g, Teradata, DB2 UDB 8.1, Autosys, Windows Server 2003, AIX 6.1, SQL Plus, UNIX Shell Scripting, SQL Loader, MS SQL SERVER 2005/2008, Agile.

**Prime software Technologies, Hyderabad, India May 2012 – May 2013**

**Role: Java Developer**

**Responsibilities:**

* Involved in the development cycle of the project following the design documents provided by the team lead using Swings, Core Java and JPA was used as ORM approach.
* Testing the application and then supporting the testing team.
* Developed the entire project from scratch and had good command on all the functional and technical aspects of the project.
* Was responsible for mentoring the new team members in the project.
* Had developed all the “User Help” documents for the entire product and then linked them to the “Help” Section of it.
* A patent was applied for this project from Wipro.Received 2 “Feather in my Cap” awards for the contribution to this project.