**HAITAO TOm yu**

Lake Forest, IL 60069

haitao.earthenvessel@gmail.com (224)688-0041

**Professional summary**

To pursue a sernior data engineering and data analytics professional.

Highly motivated, team-oriented data, cloud and big data analytics professional with proven record of delivering data analytics projects to meet objectives. Demonstrated success in leading and working with venfors and on-shore and off shore teams, with strong multi-tasking, organizational and team-building skills. Adept at developing variety of data solutions of data warehouse, data marts, data lake to meet rapidly changing data technologies and business environments. Enthusiastic and goal-oriented team player possessing excellent communication, interpersonal skills and leadership capabilities with high level of adaptability. AWS Certified Solution Architect, SysOps, Developer!

Areas of Expertise:

|  |  |
| --- | --- |
|  |  |
| * Cloud/Big Data Architect * ETL Architect/Designer/Developer * AWS Cloud Engineer/DevOps * Business/Data Analyst * Oracle/SQL Server/DB2 * Teradata, Mainframes, Redshift | | * Data Architect/Data Modeler * Team Lead/Proj * Vendor/Outsourcing Management * Project Management * Ab Initio/Informatica Power Center * Python, Shell Scripting, PL/SQL |

**PROFESSIONAL EXPERIENCE**

**SHIETION TECHNOLOGY,** Beijing China                                                                     **2016 – present**

**Cloud and Big Data Analytics Advisor**

Provided the advisory services in data engeering and data analytics to a start up business.

**ComED,** Oakbrook, IL **2014 – 2016**

**ETL Lead**

Led the full life cycle of data warehouse and data marts development. Provided the production support, performance enhancement releases.

* Architected the federated data warehouse based on dimentional model for automatic smart meter readings.
* Worked as the Informatica subject matter expert role(SME) for development projects. Efforts involved review the functional and technical design document, provided ETL design approval for new releases, provided production migration approval for all enhancement projects and bug fixes.
* Conducted the performance tuning for the ETL load batch jobs as well as the data warehouse tuning to support the various business intelligence reports. The tuning efforts involves ETL framework redesign, parallel processing and SQL tuning. The performance tuning efforts result in both daily data loading and reporting to meet SLA.

Technologies and Methodologies:

Informatica 9.5, Oracle 11g, Shell Scripting, SQL, PL/SQL, Visio, ER Win, Agile, Dimenional Modeling

**Capital One,** Rolling Meadows, IL **Jan 2013 - 2014**

**ETL Lead, ETL Designer, Ab Initio Developer**

Led a team of six on shore and off shore to provide data integration design and development using Ab Initio GDE/EME.

* Interacted with business community and conduct the source data analysis and business requirements analysis.
* Developed a standard ETL framework to enable the reusability of similar logic across the board. Involved in System Documentation of Dataflow and methodology
* Utilized all phases of the Software Development Life Cycle (SDLC) – including project analysis, requirements gathering, design documentation, development, unit testing, user acceptance testing, implementation, post implementation support and maintenance

Technologies and Methodologies:

Teradata, Ab Initio 2.16 GDE/EME, Shell Scripting, Perl, SQL

**Citibank,** Columbus, OH  **Nov 2010 – Dec 2012**

**ETL Developer**

* Evaluated business requirements and prepared design and technical specifications within the boundaries of the enterprise architecture and standards.
* Coordinated and facilitated the technical aspects of the solution and was responsible for Ab Initio's application architecture and design; and overseeing the development activities supporting those designs.
* Perform code reviews and ensure compliance with application quality, coding, performance, and corporate standards
* Serve as the Ab Initio technical expert across teams, clients, and the information technology enterprise; providing best practices, definition and positioning
* Provide leadership, manage, develop, and coach a diverse team that can include technical analysts (on shore and offshore), developers and QA teammates.

Technologies and Methodologies:

Teradata, Ab Initio 2.16 GDE/EME, Shell Scripting, Perl, SQL

**Home Depot,** Atlanta GA **2/2010 – 11/2011**

**Data Warehouse Designer**

Led and conducted the data profiling and data integration design efforts for the offshore ETL team

* XML data file generation in the continuous graph, DB2 and maestro ENV.
* Financial data set generation for retail business in DB2 and maestro ENV.
* EDI data set generation for retail business in DB2 and maestro ENV.
* Migration of an existing ab initio application from UNIX environment to LINUX environment.
* Design and Implemented CAR conversion project under the environment of ab Initio 2.15, LINUX, DB2 and     maestro. The implementation involves parallel download from DB2 database and parallel upload to DB2 database.
* Designed and implemented PC Miler application under the environment of ab Initio 2.15, LINUX, DB2 and    maestro. The work involved ab initio transformations and an automated telnet script.
* Checked performance tuning and debugging at different levels such as workflows, mappings, database etc
* Implemented various Performance Tuning techniques on Sources, Targets, Mappings, and Workflows
* Performed System Testing, Regression Testing, Acceptance Testing, Functional Testing and Stress Testing
* Created reports like Master/Detail reports, Cross Tab reports, slice and dice reports, and drill down reports

Technologies and Methodologies:

Oracle 11g, SUN Solaris, Ab Initio, Informatica 9.1, Perl/Shell Scripting, SQL, PL/SQL

**JPMorgan Chase,** Chicago, IL **4/2008 – 2/2010**

**Data Warehouse Developer Ab Initio Developer**

Work involved create business requirement document, create mapping document and create design document. The Graph used the following ab initio components: checksum, generate record, reformat, join, input table, output table, join with DB and update able. The job also involves control m design and implementation.

* Implemented the treasury data mart application in ab initio which are loading about 200 treasury files from 36 different sources daily. Treasure data mart calculates daily transfer pricing on its wholesale and retail CDs.
* Designed and implemented the pre processing graphs which cleansing the in coming source landing files.
* Designed and implemented the main engine which loads the data to 3 financial fact tables (loan, wholesale, swaps).
* Designed and implemented the recovery, reposting process and replication process.
* Designed and implemented load status tracking graph.
* Design and implement the credit card application.  The this application generated the customer monthly credit card
* Generated usage figures and created customer’s profile based on the monthly usage figures. The profile categorizes the customer   into Revolver, Transactor, Borrower and Dormant.
* Design and implemented the monthly account info graph.
* Design and implemented the customer profile graph.

Technologies and Methodologies:

Oracle 11g, SUN Solaris, Ab Initio, Perl/Shell Scripting, SQL, PL/SQL, Control-M, Waterfall

**Allstate,** Northbrook, IL **5/2006 – 2/2008**

**ETL Consultant**

* Analyzed the insurance premium and claim transactional data requirements, complex source data, and the data model, and determined the best methods in extracting, transforming and loading the data into the data staging, warehouse and other system integration projects.
* Designed logical and physical models on Oracle of 40 dimension tables and 5 fact tables.
* Worked on performance tuning for extraction graph.  Reduced the number of components by 30% and the execution time of the new graph is less than 60% of the previous graph.

Technologies and Methodologies:

Mainframe, DB2, Oracle, Ab Initio, Perl/Shell Scripting, SQL, PL/SQL

**EDUCATION**

**Master of Science in Computer Science**

The University of alabama, Tuscallosa, AL

**Master of Engineer in Computer Engineering**

Northwest University, Xi’an, China

**Bachelor of Science in Computer Science**

Northwest University, Xian, China

Designs and develops ETL solutions using data warehouse design best practices and SQL Server Integration Services (SSIS).

• Analyze data requirements, complex source data, and the data model, and determine the best methods in extracting, transforming and loading the data into the data staging, warehouse and other system integration projects.

• Analyze business requirements and outline solutions.

• Provide technical assistance in identifying, evaluating, and developing systems and procedures.

• Document all ETL and data warehouse processes and flows.

• Create or update technical documentation for transition to support teams.

• Develop automated data audit and validation processes.

• Develop and deploy ETL job workflow with reliable error/exception handling and rollback.

• Manage foundational data administration tasks such as scheduling jobs, troubleshooting job errors, identifying issues with job windows, assisting with Database backups and performance tuning.

• Design, Develop, Test, Adapt ETL code & jobs to accommodate changes in source data and new business requirements.

• Provides senior technical leadership to design, architecture, integration and support of the entire data sourcing platform with a focus on high availability, performance, scalability and maintainability

• Manage automation of file processing as well as all ETL processes within a job workflow.

• Develop, Contribute and adhere to the development of standards and sound procedural practices.

• Proactively communicate innovative ideas, solutions, and capabilities over and above the specific task request

• Effectively communicate status, workloads, offers to assist other areas.

• Collaboratively work with a team and independently. Continuously strive for high performing business solutions

• Perform and coordinate unit and system integration testing.

• Participate in design review sessions and ensure all solutions are aligned to pre-defined architectural specifications.

• Ensure data quality throughout entire ETL process, including audits and feedback loops to sources of truth.