



Vidya Prabhakar Sristi
Sristi.vidyap@gmail.com
www.linkedin.com/in/vidya-prabhakar-sristi
(479)-268-8891

SUMMARY:

- 11+ years of experience in **Data warehousing, Informatica, Teradata**, Oracle, SQL, PL/SQL, DB2. Design deploy and maintained **enterprise** class **security**, network and systems management **applications** in **AWS**.
- Expertise in Data Warehousing, **Data Migration, Data Integration** and **Reporting** using Informatica, Informatica Data Quality, Teradata and **Tableau**.
- Implemented various Data warehouse projects and supported them using **Ralph Kimball** and **Bill Inmon** methodologies.
- Designed and implemented secure **Cloud** solutions using **AWS** services like **EC2, S3, IAM** roles and policies, **RDS**, **Auto Scaling** and **Elastic Load Balancing** in **VPC**.
- Created ETL jobs for the data in **Amazon S3** by invoking **AWS Glue ETL jobs** from **AWS Lambda** function
- Experience in Extract, Transform, Load (**ETL**) of data into Data Warehouse using **Informatica 9.x, 10.x**.
- Experience in designing **Workflows, Mappings, Sessions** and **Scheduling** them using Informatica 10.1, 9.6 and 9.5.
- Performed data profiling and analysis using **Informatica Data Explorer (IDE)** and **Informatica Data Quality (IDQ)**.
- **SQL Query Tuning** and Informatica **performance tuning** involving source, target and map level bottlenecks.
- Experience driving in-person **discussions** with **senior personnel** regarding **best practices**, project management and risk mitigation for data warehousing and **cloud services**.
- Extracting and loading data from various sources (**Relational database, Flat Files, XML, Excel**) into data warehouse
- Implemented data warehousing techniques for **Data cleansing**, Slowly Changing Dimension phenomenon's (**SCD**) and Change Data Capture (**CDC**).
- Hands on experience on logical and physical data modelling, performance tuning, code review process, data normalization processes, data integration methods, Agile SDLC and production support.
- Experience **coordination** with Data Integration, Network, DBA, Server Systems and Storage Systems for cross-team projects, resource procurement, system maintenance, server migrations and upgrades and change management
- Worked with Terabyte Size Source Target Databases like **Netezza** & Teradata. Configured **Power Exchange DB2** and **ORACLE CDC** sources
- Experience working in all phases of the Software Development Life Cycle (**SDLC**). Worked in **Kanban** and **Agile** with frequent releases and providing immediate solutions for business needs.
- Knowledge on Microsoft **SQL Business Intelligence** Tools.
- Familiarity working with Relational, **NoSQL**, JavaScript **API**, **Rest API** and **Data Extract API**
- Knowledge on Business intelligence tools like **Business objects, Crystal Reports, OBIEE, Tableau**.
- **Documenting** all aspects of ETL processes, definitions and mappings to support knowledge transfer.
- Excellent oral/written/interpersonal communication skills, quick learner, willing to adapt to the team/organizational environment.

Data Warehousing /ETL Tools	Informatica Power Center 10.2/9.5/9.1/8.6/8.1(Source Analyzer, Data Warehouse Designer, Mapping Designer, Mapplet, Transformations, Workflow Manager, Workflow Monitor, Worklets), AWS, Data cleansing, Autosys, Star Schema, Snowflake Schema, OLTP, SQL*Plus, SQL*Loader
Databases	Oracle 11g/10g/9i, SQL Server 2014/2012/2008, DB2, Cassandra, Netezza, HP Vertica, Teradata, Amazon RDS (PostgreSQL and Oracle)
Programming Languages	SQL, TSQL, PL/SQL, JAVA
Scripting Languages	Shell, Korn, Perl
Big Data Ecosystems	Hadoop, Hive, AWS, CloudWatch, S3, Redshift Spectrum, Athena, Glue, AWS Redshift, Scala, Spark SQL

EDUCATION:

Master of Science in Computer Science, University of Central Missouri, Missouri, United States

PROFESSIONAL EXPERIENCE:

Fannie Mae, Plano TX

April 2020– Till date

Data Engineer

- Contribute to the cloud architecture vision and technical roadmaps and advocate from a strategic perspective to drive cloud transformation and improvements
- Worked in development of applications especially in LINUX environment and familiar with all its commands and worked on Jenkins continuous integration tool for deployment of project and deployed the project into Jenkins using GIT version control system
- Trigger the spark job in AWS lambda script using json parameters which will finally run those in EMR and store the relative output in S3.
- Worked on creating automation scripts for migrating data from onprem to AWS using DMS , SCT, Flyway
- Created CICD pipelines facilitating the end users to use the automation.
- Hands-on experience on implementing **Cloud Solutions** using various AWS Services including **EC2, VPC, S3, Glacier, EFS, AWS Kinesis, Lambda, Directory Services, Cloud Formation, Ops works, CodePipeline, CodeBuild, CodeDeploy, Elastic Beanstalk, RDS, Data Pipeline, DynamoDB**, etc.
- Experience in application lifecycle management (ALM) and DevOps automation including CI/CD pipeline architecture using tools such as CodePipeline, GitHub, and GitLab CI/CD
- Knowledge of serverless & micro-service management using Lambda, ECS, and EKS
- Knowledge of modern system integration & analytics services like API Gateway, Eventbridge, Kinesis, Step Functions, SNS, SQS, and OpenSearch Service (formerly Elasticsearch)
- Knowledgeable in AWS Platform Data services such as **Redshift, Dynamo, S3, and RDS**
- Strong experience in Major AWS services like **Cloud Formation, Cloud Front, Cloud Watch, Cloud Trail, VPC, RDS, DynamoDB, SQS, SNS**.
- Provide technical expertise and ownership in the diagnosis and resolution of an issue, including the determination and provisioning of workaround solutions or escalation to service owners

Environment:	AWS services ,Python ,Airflow, Bitbucket, Urban Code Deploy, Jira, Jenkins, Splunk, Docker, Kubernetes,
--------------	---

Bank Of America, Charlotte NC

Jun 2019– Mar 2020

Application Developer

- Interact with **end-users** and **functional analysts** to identify and develop Business Specification Documents (**BSD**) and transform it into technical requirements.
- Resolve bottlenecks and enhancements, implementing them in Monthly **Production Release Cycles**. Support **deployment** of solution, cut over **planning** and **execution** activities
- Created **reusable expressions**, Mappings for other developers to reuse them accordingly. End to end data movement for coding reusable objects whenever necessary to be used.
- Wrote **Queries to validate** the data loaded which was utilized in building a **Testing Framework** for future use.
- Extensively used parameters, variables at **workflow, session, mapping** levels to dynamically generate data for loading data into database. Used **Autosys** to schedule informatica jobs.
- Thoroughly followed standards, best practices throughout ETL process and naming conventions for various objects in Informatica (Transformations, sessions, maps, workflow names, log files, bad files, input, variable, output ports).
- Debugged **informatica failures** on prod support and implementing fixes in releases. Debugging issues at informatica level (Session logs, Debugger by connecting to prod from DEV), data anomalies in databases.
- Developed Mappings using Transformations like **Expression, Filter, Joiner** and **Lookups** for better data messaging and to migrate clean and consistent data
- Migrated code developed from dev to higher environments using **informatica repository manager**. Validating the migrations after changes and running the code for Business sign off's from BA's and Users.
- To enhance the reconciliation process for BA's and USERS , wrote **reconciliation queries** and regressions test scripts by converting the mapping to SQL for data validation on source to target loads.

- Worked on Creating sessions that would write **LOAD STATISTICS** to a Table in Database after every Run. Extensively used Variables and Parameters in Informatica both at mapping and session Level.
- Extensively used Debugger and enabling Test loads as a part of UNIT TESTING.
- Worked on **Version Control, Performance Tuning** in informatica to have a much better performance and efficient Loads. Created sessions which would analyse the DB Partitions, data and performance when queries are run against it.
- Automating informatica processes where **control** table drives the entire load to what partition data needs to be loaded and rolls the control keys (Dates) to the next day once the load is complete and the job is scheduled to run next day.

Environment:	Informatica 10.2.0, Exadata, Teradata, Netezza, Shell, Python, Perl, Autosys, MS Visio, Toad, SQL Developer, Bitbucket, Urban Code Deploy, Jira
---------------------	---

Nationwide, Columbus OH Application Developer

Nov 2017 – May 2019

- Involved in **requirement** gathering and **analysis** of data from different source systems. Created elevated level **solution design** document and engineered logical and physical models for the new platform
- Technology implementations of **cloud-based** technologies, **data cleansing** and conversions, systems **development life cycle** and interface design/development.
- Development adhering to and complying with all applicable, **federal** and **state laws, regulations** and **guidance**.
- Developed **mappings** and **workflows** in **Informatica**, processed various sources like **Oracle** tables, **Flat Files** (delimited and fixed width), **XML** files and **MQ** source.
- Experience in TIER 1 Production Support, implementing **Change Request, Work Order**, Problem Ticket, and work request. Handled **recovery** process in case of workflow failures.
- **Lead production deployment** of highly-available and **fault-tolerant** enterprise data warehouse applications.
- Created mappings for reconciliation of data for asset management using **Normalizer, Transaction Control** and **Java**
- Improved the **performance** of batch jobs by analysing **explain plans**, implementing indexes, database **partitioning, parallel** hints, Informatica partitioning, pushdown optimization and **caching**
- Developed and executed a migration strategy using **AWS Data Migration service** and **AWS schema conversion Tool** to move Data Warehouse on premise to **AWS Redshift** and use **AWS Cloud Watch** to collect and monitor AWS RDS instances in lower environments.
- Created a **AWS Glue Data Catalog** and used **Athena** to directly query the data from Amazon S3 using the AWS Red Shift Spectrum
- Configured secure file transfer using encryption by **AWS KMS**. Created **Shell** scripts to push and pull files.
- Created an **error framework** using PM Tables in Informatica, PL/SQL blocks/procedures to generate reports **detailing** business and **technical errors** in the batch run.
- Worked on ESP for scheduling jobs. Monitored daily batch jobs proactively and worked for the fixes to stop it from abending.
- Writing Perl scripts for data acquisitions, standardizing the data format and structure and for distinct phases of the **Diff Framework**. And, python scripts for automating some processes across the environment.
- Created Unit test cases in **Ruby** and performed testing of individual code blocks as well as for the batch of workflows.
- Followed **Agile SDLC** method working on everyday tasks. **Planning** stories and **estimations** them, analysed burn down chart to understand the remaining work for the **sprint**.
- Provided **UAT/EBF** (Emergency bug fix) and Production Support for several incremental releases.

Environment:	Informatica 10.2.0.9.6.1, AWS, Oracle12c/11g, Teradata, Netezza, PL/SQL, Shell, Python, Perl, CA work station, Ruby, Data Modelling, MS Visio, Toad, SQL Developer, DB Visualizer, GIT, Urban Code Deploy, ServiceNow
---------------------	---

Blue Cross Blue Shield of North Carolina, Durham NC Application Developer

May 2017 – Oct 2017

- Responsible for migrating ITS host data from **PMHS** to **FACETS**, working as a **Delivery Lead** for an off-shore team.
- Developed ETL for migrating data from **FACETS** to **IDS** and then to **EDW** or **EODS** using **Control-M** jobs. Created data mart for all the rejected and erroneous claims
- Used Informatica Power Centre Workflow manager to create sessions, batches to run with the logic embedded in the mappings.
- Developed PL/SQL and UNIX Shell Scripts for scheduling the sessions in Informatica.
- Responsible for **data validations** across the landscape validating **test strategy** and test **results**
- Developed ETL logics for claims, accounting and finance data for EDW and maintaining claims accounts.

- Performed **data cleansing, massaging, masking**, IDQ and data profiling.
- Developed **reconciliation scripts** as part of unit testing to validate data loaded in the tables. Prepared SQL Queries to validate data in source and target databases.
- Involved in **defect analysis** and **remediation** in UAT environment as per business needs.

Environment:	Informatica 9.6.X, Netezza, Teradata, Shell scripts, Control-M, IBM DB2, AIX UNIX 5.3
---------------------	---

University of Central Missouri
Graduate Assistant

Aug 2016 - Dec 2016

- Worked as a Graduate Student Assistant under **Dr. Ziyuan Meng** and was part of the tech team which developed **Class Attendance Tracking System**
 - **Functionality:** Staff and Faculty can check real-time attendance of classes or events and generate attendance reports for specific student or group of students. Students can check the schedule of their classes, receive notification before class, check-in to class by using GPS signal, scanning NFC/bar code, or providing photo verification using camera.
- Responsible for **Server-Side Coding** and **Database** operations for the project. Maintained project documents with version updates in GIT repository.

Environment:	MySQL, Java, GIT, Servlets, MVC, JavaScript, HTML
---------------------	---

IBM India Pvt Ltd,
Application Architect

April 2012 – Dec 2015

- Responsible for building various subject areas containing information from various upstream and manipulating the data as per business needs and documenting all aspects of ETL processes, definitions and mappings
- Developed well-tuned **Mappings** using Informatica Designer Tool which involves identifying and rectifying the bottlenecks at various levels like Source, Target, Mapping, Session and System Development of Exception Handling and **Error Handling** process.
- Tuning the Informatica mapping, session performance by using Source-Target Optimization. Created dimensions model star schemas using Kimball methodology.
- Created Attributes, Metrics, security filters and other report objects, drilling, defined hierarchies using Micro Strategy.
- Worked on SQL joins, Stored Procedures, Packages, Triggers, Cardinalities, Loops, Aliases, Views, Aggregate conditions, parsing of objects and hierarchies.
- Developed Reusable Maplets and Transformations for standard business units, to use in multiple mappings.
- Created and monitored Sessions using Workflow Manager. Extracted data from Oracle, DB2 and Flat files. Defined Target Load Order Plan for loading data correctly into different Target Tables.
- Defined mapping variables and parameters to capture load history, row counts and to log errors of a session run.

Environment:	Informatica 9.5.1/9.1.0, IBM Maximo, Korn, Shell, DB2, GIT, Oracle 10g/11g, MS SQL Server 2008, DB Visualizer, Autosys, Micro Strategy
---------------------	--